Sums of correlated exponentials: two types of Gaussian correlation structures

vorgelegt von
Diplom-Mathematiker
Anton Klymovskiy (Klimovsky)
aus Charkiw

von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Dirk Ferus
Berichter: Prof. Dr. Anton Bovier
Berichter: Prof. Dr. Erwin Bolthausen

Tag der wissenschaftlichen Aussprache: 24. Juni 2008

Berlin 2008

D83

Abstract

In this thesis, we study the limiting behaviour of the large *sums of strongly correlated exponentials* as the number of their summands and the effective dimension of the correlation structure simultaneously tend to infinity. We consider two types of such sums which are generated by two *a priori* very different Gaussian correlation structures. The first type is a sum of hierarchically correlated random variables which is based on the partition function of *Derrida's generalised random energy model* (GREM) with *external field*. The second type is an infinitesimal sum of genuinely non-hierarchically strongly correlated random variables which is based on the partition function of the *Sherrington-Kirkpatrick (SK) model with multidimensional spins*. We consider the asymptotic behaviour (the thermodynamic limit) of these two sums on a logarithmic scale (i.e., at the level of *free energy*) and also at a more refined level of their *fluctuations* (i.e., at the level of weak limiting laws). Interestingly for the SK model with multidimensional spins, we find traces of a *hierarchical organisation* in the thermodynamic limit. This supports the conjectured in theoretical physics universal behaviour of the sums of such sort.

Concerning the SK model with multidimensional spins, we obtain the following results. We prove upper and lower bounds on the free energy of this model in terms of variational inequalities. The bounds are based on a multidimensional extension of the *Parisi functional*. We generalise and unify the comparison scheme of Aizenman, Sims and Starr and the one of Guerra involving the GREM-inspired processes and Ruelle's probability cascades. For this purpose, an abstract quenched large deviations principle of the Gärtner-Ellis type is obtained. We derive Talagrand's representation of Guerra's remainder term for the SK model with multidimensional spins. The derivation is based on well-known properties of Ruelle's probability cascades and the Bolthausen-Sznitman coalescent. We study the properties of the multidimensional Parisi functional by establishing a link with a certain class of semi-linear partial differential equations. We embed the problem of strict convexity of the Parisi functional in a more general setting and prove the convexity in some particular cases which, however, do not cover the original setup of Talagrand. Finally, we prove the *Parisi formula* for the local free energy in the case of *multidimensional Gaussian a priori distribution* of spins using Talagrand's methodology of a priori estimates.

Concerning the GREM in the presence of uniform external field, we obtain the following results. We compute the *fluctuations of the ground state and of the partition function* in the thermodynamic limit for all admissible values of parameters. We find that the fluctuations are described by a hierarchical structure which is obtained by a certain coarse-graining of the initial hierarchical structure of the GREM with external field. We provide an explicit formula for the free energy of the model. We also derive some large deviation results providing an expression for the free energy in a class of models with Gaussian Hamiltonians and external field. Finally, we prove that the coarse-grained parts of the system emerging in the thermodynamic limit tend to have a certain optimal magnetisation, as prescribed by strength of external field and by parameters of the GREM.

Zusammenfassung

Diese Dissertation behandelt das Grenzwertverhalten der *großen Summen der stark korrelierten Exponenziale*, während die Anzahl ihren Summanden und die effektive Dimension der Korrelationsstruktur gleichzeitig gegen Unendlichkeit gehen. Wir betrachten zwei Arten dieser Summen, die durch die *a priori* sehr unterschiedliche Gauß'schen Korrelationsstrukturen erzeugt werden. Die erste Art ist eine Summe der hierarchisch korrelierten Zufallsvariablen, die auf der Zustandssumme von *Derrida's generalised random energy model* (GREM) mit *externem Feld* basiert. Die zweite Art ist eine infinitesimale Summe der echt nicht-hierarchisch stark korrelierten Zufallsvariablen, die auf der Zustandssumme vom *Sherrington-Kirkpatrick (SK) Modell mit mehrdimensionalen Spins* basiert. Wir betrachten das asymptotische Verhalten (der thermodynamische Limes) dieser Summen auf der logarithmischen Skala (d.h., auf dem Niveau der freien Energie) und außerdem auf dem präziseren Niveau ihrer *Fluktuationen* (d.h., auf dem Niveau der schwachen Grenzwertverteilungen). Interessanterweise finden wir auch im SK Modell mit mehrdimensionalen Spins Spuren der *hierarchischen Organisation* im thermodynamischen Limes. Dies unterstützt das hypothetische universale Verhalten dieser Summen aus der theoretischen Physik.

Bezüglich des SK Modells mit mehrdimensionalen Spins erzielen wir die folgenden Ergebnisse. Wir beweisen die oberen und unteren Schranken für die freie Energie mittels Variationsungleichungen, die auf der mehrdimensionalen Verallgemeinerung des *Parisi-Funktionals* basieren. Wir setzen das Vergleichsschema von Aizenman, Sims und Starr und das von Guerra ein, die die GREM-inspirierten Prozesse und Ruell'schen Wahrscheinlichkeitskaskaden involvieren. Hierfür beweisen wir ein abstraktes "quenched" Prinzip der großen Abweichungen von Gärtner-Ellis Art. Mittels der Eigenschaften der Ruelle'schen Wahrscheinlichkeitskaskaden und des Bolthausen-Sznitman Koaleszents leiten wir die Darstellung von Talagrand des Restterms von Guerra für das SK Modell mit mehrdimensionalen Spins ab. Wir untersuchen die Eigenschaften des mehrdimensionalen Parisi-Funktionals, indem wir eine Verbindung mit einer Kategorie semi-linearer partieller Differentialgleichungen herstellen. Wir betten das Problem der strengen Konvexität des Parisi-Funktionals in einen allgemeineren Kontext ein. Wir zeigen die Konvexität in einigen Fällen, welche jedoch nicht die ursprüngliche Formulierung von Talagrand umfassen. Schließlich beweisen wir die *Parisi-Formel* für die lokale freie Energie im Fall der *mehrdimensionalen Gauß'schen a priori Verteilung* der Spins mit der Methodologie der a priori Abschätzungen von Talagrand.

Bezüglich des GREMs in Anwesenheit des uniformen externen Felds erzielen wir die folgenden Ergebnisse. Wir berechnen die *Fluktuationen des Grundzustandes und der Zustandssumme* im thermodynamischen Limes für alle zulässigen Werte der Parameter. Wir finden, dass im thermodynamischen Limes die Fluktuationen durch eine hierarchische Struktur beschrieben sind. Diese Struktur ist eine Grobkörnung der ursprünglichen hierarchischen Struktur des GREMs mit externem Feld. Wir stellen eine explizite Formel für die freie Energie des Modells zur Verfügung. Wir leiten auch einige Resultate zu den großen Abweichungen, welche einen Ausdruck für die freie Energie einiger Modelle mit Gauß'schen Hamiltonians und externem Feld ermöglichen, her. Schließlich beweisen wir, dass die grobkörnigen Teile des Systems, die in der thermodynamischen Limes auftauchen, dazu neigen eine optimale Magnetisierung zu haben. Diese Magnetisierung ist durch die Stärke des externen Felds und durch die Parameter des GREMs vorgeschrieben.

To my mother and the memory of my father

Contents

Introduction

In this thesis, we study the limiting behaviour of the *large sums of correlated exponentials*

$$Z_M := \sum_{i=1}^{M} \exp(X_M(i)) \tag{0.1}$$

as the number of their summands M tends to infinity (here $\{X_M(i)\}_{i=1}^M$ is an array of correlated random variables). We shall consider the sums with the summands that have the following two characteristic properties: a *wide range of orders of magnitude* and significant *high-dimensional correlations*. Not surprisingly, the sums of this type emerge in a variety of applied contexts, e.g., in models of interacting complex systems being considered in physics, biology, computer science, social sciences, finance, etc.

We should emphasise that, for the purposes of modelling of special magnetic materials called *spin glasses*, study of such sums Z was initiated and brought to very sophisticated (though mathematically unrigorous) levels of analysis in theoretical physics, see, e.g., Mézard *et al.* (1987). For the most recent accounts of mathematical research on such sums Z, we refer to a thorough monograph by Talagrand (2003), very readable introductory text-book of Bovier (2006), and comprehensive collection of review articles (Bolthausen & Bovier, 2007).

In general, the first characteristic feature mentioned above precludes the applicability of such standard tools of probability theory as the law of large numbers and the central limit theorem, since Z may well be (and actually, as we shall see, in many cases is!) dominated by a small number of exceptionally large summands. In these cases we are forced to perform more detailed analysis of such rare events. For this purpose, we shall invoke some techniques of large deviations and also, perhaps more importantly, extreme value theory.

Contrary to extensively studied situations with one-dimensional correlation structures (e.g., Markov processes, where the involved random variables are indexed by a real time variable), we shall deal with the high-dimensional correlation structures, where effective dimension of the index set grows with M.

In the so far rigorously analysed sums Z of the above sort, a new class of limiting objects appears. Moreover, contrary to what might have been expected, the limiting objects possess an important *hierarchical ultrametric structure*. This structure appears to be universal, modulo the number of levels of the hierarchy.

In this thesis, we consider two *a priori* very different instances of such sums of correlated exponentials. These sums are inspired by the following mean-field spin-glass models.

1. **The generalised random energy model with external field.** The generalised random energy model (GREM) was proposed by Derrida (1985) as a model of a random energy landscape with *a priori* ultrametric correlation structure. To define the family of correlated random variables of interest, we first fix the ultrametric structure. Given $N \in \mathbb{N}$, consider the standard *discrete hypercube* $\Sigma_N := \{-1;1\}^N$. It will play the role of the index set. Define the (normalised) *lexicographic overlap* between the configurations $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$ as

$$q_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(2)}) := \begin{cases} 0, & \sigma_1^{(1)} \neq \sigma_1^{(2)} \\ \frac{1}{N} \max\left\{ k \in [1;N] \cap \mathbb{N} : [\sigma^{(1)}]_k = [\sigma^{(2)}]_k \right\}, & \text{otherwise.} \end{cases} \tag{0.2}$$

We equip the index set with the *lexicographic distance* defined as

$$\mathrm{d}_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(2)}) := 1 - q_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(2)}). \tag{0.3}$$

This distance is obviously an *ultrametric*, that is, for all $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)} \in \Sigma_N$, we have

$$\mathrm{d}_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(3)}) \leq \max\left\{ \mathrm{d}_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(2)}), \mathrm{d}_{\mathrm{L}}(\sigma^{(2)}, \sigma^{(3)}) \right\}. \tag{0.4}$$

Let $\mathrm{GREM}_N := \{\mathrm{GREM}_N(\sigma)\}_{\sigma \in \Sigma_N}$ be the Gaussian random process on the discrete hypercube Σ_N with a covariance of the following form

$$\mathbb{E}\left[\mathrm{GREM}_N(\sigma^{(1)})\mathrm{GREM}_N(\sigma^{(2)})\right] = \rho(q_{\mathrm{L}}(\sigma^{(1)}, \sigma^{(2)})), \tag{0.5}$$

where $\rho : [0;1] \to [0;1]$ is the non-decreasing right-continuous function such that $\rho(0) = 0$ and $\rho(1) = 1$. Given $h \in \mathbb{R}_+$, consider the Gaussian process $X := X_N := \{X_N(h,\sigma)\}_{\sigma \in \Sigma_N}$ defined as

$$X_N(h,\sigma) := \mathrm{GREM}_N(\sigma) + \frac{h}{\sqrt{N}} \sum_{i=1}^{N} \sigma_i, \quad \sigma \in \Sigma_N. \tag{0.6}$$

The second summand in (0.6) is called the *external field*. The parameter h represents the *strength of external field*. The GREM with external field is based on the following sum of $M := 2^N$ hierarchically correlated exponentials

$$Z_N(\beta) := \sum_{\sigma \in \Sigma_N} \exp\left[\beta \sqrt{N} X_N(h,\sigma)\right], \tag{0.7}$$

where $\beta > 0$ is the real parameter called the *inverse temperature*. The sum (0.7) is called the *partition function*. Of course, if $\beta \approx 0$ the summands in (0.7) are on the same scale of magnitude and, hence, the law of large numbers applies. As we shall see below, the situation becomes radically different and the law of large numbers breaks down, if β increases above a certain h-dependent threshold $\beta_0(h)$.

2. **The Sherrington-Kirkpatrick model with multidimensional spins.** The Sherrington-Kirkpatrick (SK) model was introduced by Sherrington & Kirkpatrick (1975). This model of a mean-field spin-glass has long been one of the most enigmatic models of statistical mechanics. Mathematically it boils down to the sum of correlated exponentials with a genuine

non-hierarchical correlation structure. The correlations in the SK model are induced by the *Hamming distance* on the discrete hypercube Σ_N. In this thesis, we study the following generalisation of the SK model. Consider arbitrary (not necessarily discrete) subset $\Sigma \subset \mathbb{R}^d$ and the corresponding index set $\Sigma_N := \Sigma^N$. We define the family of Gaussian processes $X := \{X(\sigma)\}_{\sigma\in\Sigma_N}$ as

$$X(\sigma) = X_N(\sigma) := \frac{1}{N}\sum_{i,j=1}^{N} g_{i,j}\langle\sigma_i,\sigma_j\rangle, \tag{0.8}$$

where $G := \{g_{i,j}\}_{i,j=1}^N$ consists of i.i.d. standard normal random variables and, for $x, y \in \mathbb{R}^d$, $\langle x,y\rangle := \sum_{u=1}^d x_u y_u$ is the standard Euclidean scalar product. The SK model with multidimensional spins is based on the following infinitesimal sum of correlated exponentials

$$Z_N(\beta) := \int_{\Sigma_N} \exp\left(\beta\sqrt{N}X(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma), \tag{0.9}$$

where $\mu \in \mathscr{M}_{\mathrm{f}}(\Sigma)$ is some arbitrary (not necessarily uniform or discrete) finite *a priori measure*. The sum (0.9) is called the *partition function*.

To investigate limiting behaviour of the sums (0.7) and (0.9), it is natural to ask the following questions.

As a first approximation, we consider the limiting behaviour of $Z_N(\beta)$ on a logarithmic scale. Consider the random variable

$$p_N(\beta) := \frac{1}{N}\log Z_N(\beta), \tag{0.10}$$

where $Z_N(\beta)$ is given either by (0.9) or by (0.7). Somewhat abusing physical terminology we shall call (0.10) the *free energy*.

Question 0.0.1. *Does p_N converge to some random variable p as $N \uparrow +\infty$? If convergence holds, then in what sense (i.e., in probability, almost surely or L^1)? Finally, can we compute p?*

Assume that we have computed $p(\beta)$, for all $\beta > \beta_0$. It is intuitively clear that being able to compute the limit

$$\lim_{\beta\uparrow+\infty} \beta^{-1}p(\beta) =: m$$

in many cases means that

$$\lim_{N\uparrow+\infty}\left(N^{-1/2}\max_{\sigma\in\Sigma_N} X_N(\sigma)\right) = m. \tag{0.11}$$

Thus, knowing $p(\beta)$ for large β provides very sharp information on extremes of the process X_N as $N \uparrow +\infty$. That is, the constant m not only bounds the extremes, but gives also their *exact* value. This is unusually good.

The second natural more refined approximation is to study the "fluctuations" (weak limiting laws) of the sum Z_N.

Question 0.0.2. *Does there exist some scaling sequence of constants $\{C_N\}_{N=1}^{\infty}$ and a nontrivial random variable Z_0 so that the weak convergence $C_N Z_N \xrightarrow[N\uparrow+\infty]{w} Z_0$ holds? Is the distribution of the random variable Z_0 to some extent universal?*

Question 0.0.2 is, of course, harder than Question 0.0.1, but it seems to be exactly the question that is claimed to be positively answered in theoretical physics.

The main purpose of this thesis is to compute the free energy (Question 0.0.1) and the fluctuations of the partition function (Question 0.0.2) in the $N \uparrow +\infty$ limit for the GREM with external field and the SK model with multidimensional spins. We find that in both models similar hierarchical ultrametric structures do emerge in the $N \uparrow +\infty$ limit. For the GREM with external field we are able to give a complete positive answer to both Questions 0.0.1, 0.0.2. For the general case of the SK model with multidimensional spins we are only able to give upper and lower bounds on the free energy. However, for the special case, where μ (cf. (0.9)) is a Gaussian distribution, we compute the limiting free energy and give a complete answer to Question 0.0.1.

To be more specific, concerning the SK model with multidimensional spins we prove upper and lower bounds on the free energy (0.10) in terms of variational inequalities involving the corresponding multidimensional generalisation of the Parisi functional (Theorems 3.1.1, 3.1.2, 4.1.1, 4.1.2). For this purpose, we generalise and unify the Aizenman-Sims-Starr (AS^2) and Guerra's schemes for the case of multidimensional spins, and employ a quenched large deviations principle (LDP) which may be of independent interest (Theorems 3.3.1 and 3.3.2). Both schemes are formulated in a unifying framework based on the same comparison functional. The functional acts on Gaussian processes indexed by an extended configuration space as in the original AS^2 scheme. As a by-product, we provide also a short derivation of the remainder term in multidimensional Guerra's scheme (Theorem 4.1.4) using well-known properties of the Ruelle probability cascades (RPC) and the Bolthausen-Sznitman coalescent. This gives a clear meaning to the remainder in terms of averages with respect to a measure changed disorder. The change of measure is induced by a reweighting of the RPC using the exponentials of the GREM-inspired process.

We study the properties of the multidimensional Parisi functional by establishing a link between the functional and a certain class of non-linear partial differential equations (PDEs), see Propositions 4.2.1, 4.2.2 and Theorem 4.2.2. We extend the Parisi functional to a continuous functional on a compact space (Theorems 4.2.1, 4.2.2). We show that the class of PDEs corresponds to the Hamilton-Jacobi-Bellman (HJB) equations induced by a linear problem of diffusion control (Proposition 4.2.4). Motivated by a problem posed by Talagrand (2006c), we show the strict convexity of the local Parisi functional in some cases (Theorem 4.2.4).

We partially extend Talagrand's methodology of estimating the remainder term to the multidimensional setting (Theorem 4.1.4, Proposition 4.3.1, Theorem 4.3.1). In the case of multidimensional Gaussian a priori distribution of spins we prove the validity of the Parisi formula (Theorem 5.1.1).

Concerning the GREM in the presence of uniform external field, we find the fluctuations of the ground state and of the partition function in the thermodynamic limit for all admissible values of the model parameters (Theorems 6.1.1, 2.3.1, 6.1.3). An explicit formula for the free energy of this model is given (Theorem 6.1.4). We also derive some large deviation results providing an expression for the free energy of a class of models with external field and Gaussian Hamiltonians. (Theorem 6.2.1). We prove that the coarse-grained parts of the system tend to have a certain optimal magnetisation as prescribed by the strength of external field and by parameters of the GREM.

The thesis is organised as follows (see Figure 0.1 for the chapter dependency graph).

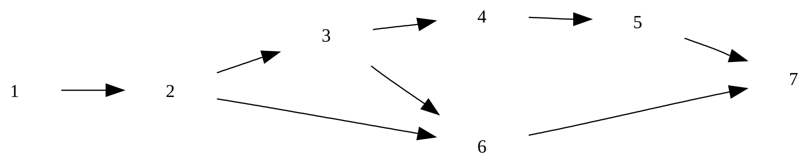

Fig. 0.1. Chapter dependency graph

Chapter 1 contains a description of some modelling ideas of statistical mechanics of disordered systems. We fix some basic notation and terminology, and introduce the models used throughout the work (in particular, the random energy model (REM), the GREM and the SK model).

Chapter 2 is a short account of mathematical results available in the literature which are either directly used in the arguments of this thesis or give a suggestive background and useful examples of what could be achieved. In particular, we discuss in this chapter some interpolation techniques for Gaussian random variables, comparison theorems, concentration of measure, the Poisson-Dirichlet processes, Ruelle's probability cascades and the Bolthausen-Sznitman coalescents.

Chapter 3 is the first chapter of this thesis dealing with the SK model with multidimensional spins. We extend the AS^2 scheme to this model. We record some basic properties of covariance structure of the process (0.8) and establish the relevant concentration of measure results. The chapter contains the tools allowing to compare and interpolate between the free energy-like functionals of two different Gaussian processes. We derive a quenched LDP of the Gärtner-Ellis type under measure concentration assumptions. We conclude the chapter with the derivation (based on the AS^2 scheme) of upper and lower bounds on the free energy in terms of the saddle point of a Parisi-like functional.

Chapter 4 employs the ideas of Guerra's comparison scheme in order to obtain upper and lower bounds on the free energy in the SK model with multidimensional spins. It contains also a useful analytic representation of the remainder term which is used in the next chapter. We also study some properties of the multidimensional Parisi functional such as differentiability and convexity. The chapter is concluded by a partial extension of Talagrand's remainder term estimates to the case of the SK model with multidimensional spins.

Chapter 5 concerns the SK model with multidimensional Gaussian a priori distribution of spins. The main result of this chapter is a derivation of the local Parisi formula for the SK model with multidimensional spins.

Chapter 6 is devoted to study of the fluctuations of the ground state and of the partition function for the REM and GREM with external field. We also provide an explicit formula for limiting free energy of these models. We obtain some large deviation results providing an expression for the free energy of a class of models with Gaussian Hamiltonians and external field.

Chapter 7 lists some open problems related to the models studied in this thesis.

Appendix A contains a proof of the almost superadditivity of the local free energy in the SK model with multidimensional spins. It is an application of the Gaussian comparison results we derived in Chapter 3.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my adviser Prof. Anton Bovier for sharing his vision, erudition and understanding of the fascinating subject of disordered systems with me. Not less important for me were his optimistic approach to resolution of the problems, his skill to understand the subject of matter almost without any words from my side, and his very stimulating habit of dissecting the problem at hand by asking extremely sharp questions.

I would like to thank Prof. Erwin Bolthausen for agreeing to be the coexaminer of this thesis.

I have the pleasure thank the members of the Research Groups on Probability Theory and Mathematical Finance (Technische Universität Berlin), on Interacting Random Systems (Weierstrass Institute for Analysis and Stochastics) and on Probability Theory (Universität Potsdam) for giving me the opportunity to start to understand the beauty and diversity of probability theory.

I am indebted to Prof. Mariya Shcherbina for giving me the initiation in research on disordered systems and for sharing her vision with me. Besides, I would also like to thank my lyceum and university lecturers and professors – the representatives of the Kharkiv mathematical school – for giving me the start in modern mathematics.

I am grateful to Prof. Michel Talagrand and Prof. Dmitry Panchenko for important remarks.

I would like to warmly thank my colleagues Stefan Ankirchner, Andrej Depperschmidt, Alexander Drewitz, Matthias an der Heiden, Ivan Izmestiev, Mesrop Janunts, Pierre-Yves Louis, Martin Slowik, Michel Sortais, Cristian Vormoor, Wiebke Wittmüß for greatly contributing to the cheerful atmosphere I enjoyed both at and after work. I thank Elise Luckfiel for her help with peculiarities of the German language.

This work would totally be unthinkable without the exceptionally wise guide in life and science, who was and is for me my prematurely passed away father. Completion of this thesis would simply be impossible without love and steady support of my mother and professional psychological help of my sister.

Financial support of both the Deutsche Forschungsgemeinschaft (through the Research Training Group "Stochastic Processes and Probabilistic Analysis") and the Helmholz-Gemeinschaft is gratefully acknowledged.

1

Framework

In this chapter, we fix the framework and notations we shall use in the present work. Besides, we try to give some motivational background to the set of problems we shall deal with.

1.1 Physical roots and vision

Starting from the 1970's, an increasing amount of research in statistical mechanics has been targeted at modelling the behaviour of magnetic materials collectively called *spin-glasses*. This led to a group of models usually referred to as *disordered spin systems*. In general, *disordered systems* are complex systems with *spatial inhomogeneity*. On physical grounds, the typical representatives of these systems are alloys, composite materials, porous media, polymers, to name a few. As became clear in recent decades, advances in understanding of the models of disordered systems made in theoretical physics allow for their successful application to systems well beyond the usual scope of physics. For example, we can name the complex systems arising in computer science (combinatorial optimisation, computational complexity, see Hartmann & Weigt (2005); Mézard & Montanari (2007), information processing, see Mézard & Montanari (2007); Nishimori (2001); Opper & Saad (2001)), cognitive science (Amit, 1989; Arbib, 2003; Nishimori, 2001), finance (agent based modelling, see Challet *et al.* (2005); Coolen (2005); risk modelling, see Bouchaud & Potters (2003)), social sciences (Contucci & Graffi, 2007; Durlauf, 1999). Moreover, the models of disordered systems, the vision and heuristics developed in theoretical physics are abstract enough to be viewed as a collection of very sharp statements about basic purely mathematical objects. This explains the ever growing interest in bringing these heuristics on rigorous mathematical grounds in order to understand the real domain of their applicability.

Often the precise mechanisms of inhomogeneity in disordered systems are not known in advance or have a very complicated description. A way to model such situation is to assume that instead of a full deterministic description of the system we have some statistical data on the distribution of inhomogeneities. In that case, the inhomogeneities can be modelled by *random processes*. This randomness is usually referred to as a *quenched disorder*[1]. Disordered systems are often studied in the framework of statistical mechanics which is also the viewpoint adopted in the present thesis.

[1] The term "quenched" emphasises a certain analogy with materials obtained by heat treatment which is called in metallurgy *quenching*. The method amounts to rapid cooling of a heated up sample of metal in order to preserve (or freeze) inhomogeneities in it. In contrast, *annealing* is a process of slow cooling which gives enough time for homogenising diffusion processes.

The main objective of statistical mechanics according to Nishimori (2001) is "the clarification of the macroscopic properties of many-body systems starting from the knowledge of interactions between microscopic elements". In turn, statistical mechanics of disordered systems is a "particularly difficult, but also particularly exciting, branch of the general subject, that is devoted to the same problem in situations when the interactions between the components are very irregular and inhomogeneous, and can only be described in terms of its statistical properties" (Bovier, 2006). For introductory information on disordered systems we refer to Bovier (2006).

Highly nontrivial heuristics such as the *replica symmetry breaking* (RSB) have been developed in theoretical physics to explain the very complex behaviour of even the "simplest" spin-glass models both in dynamics and equilibrium (see, e.g., the landmark monograph by Mézard *et al.* (1987)). Interestingly, these "simple" models are actually very basic but very little studied mathematical objects which starting from the 1980's became the subject of increasingly active mathematical research.

One of the main messages of physical theory of the hierarchical RSB is the following ambitious but unrigorous *universality hypothesis*.

> The main characteristic observables (*order parameters*) of huge disordered spin systems with long range interactions often have ultrametric (tree) structure.

The domain of applicability of the above heuristic seem to be broader than that of the central limit theorem (CLT). Despite considerable efforts in recent years, only rather basic predictions of physical theories were proved rigorously for the most interesting mean-field spin-glass models. Mathematical research on spin-glasses is still concentrated on several most simple cases with more or less ad hoc methods specific to the concrete instance. There is, however, a hope that there exists some general, yet not discovered, unifying mathematical theory explaining all predictions on spin-glasses made in theoretical physics and, especially, the universality hypothesis.

Rigorous understanding of applicability domain of the universality hypothesis would be a very useful result, since the tree structures are dramatically more tractable than the general graphic structures. This allows one not only to get substantially deeper analytical insight, but also to construct faster algorithms solving hard problems at hand.

1.2 Statistical mechanics of spin systems

In this section, we recall some basic objects, terminology and facts used in statistical mechanics of spin systems.

1.2.1 Basic objects

Spin system consists of *sites* indexed by some *finite* index set Λ. Assume $\Sigma \subset \mathbb{R}^d$. To each site we attach an independent copy of the *configuration probability space* $(\Sigma, \mathscr{S}, \mu)$. In this thesis, we shall consider the following configuration spaces: $\Sigma = \{-1; 1\}$, compact subsets of $\mathbb{R}^d$ and, finally, the whole $\mathbb{R}^d$ itself. Compactness of the configuration space Σ is usually assumed. In Chapter 5, we shall, however, treat an example of a non-compact configuration space. The elements $\sigma \in \Sigma$ are traditionally called *spins*[2] or just configurations. A configuration gives complete microscopic description of the site. In applications, the *reference measure* (or *a priori distribution*) μ is often just uniform. The overall configuration space of the whole system is often a product set $\Sigma_\Lambda := \Sigma^\Lambda$. Denote $\mu_\Lambda := \mu^{\otimes \Lambda}$ and also $\mathscr{S}_\Lambda := \mathscr{S}^\Lambda$. Then $(\Sigma_\Lambda, \mathscr{S}_\Lambda, \mu_\Lambda)$ is the (product) *a priori probability space*.

Remark 1.2.1. *Sometimes we shall deal with the a priori measures $\mu \in \mathscr{M}_f(\Sigma)$, i.e., with the measures without probabilistic normalisation. This will not change (modulo a constant shift) the quantities of interest.*

Having defined the configuration space, we would like to have some means to compare the individual configurations. For that purpose we assume that we are given some (measurable) cost functional on the configuration space $H : \Sigma_\Lambda \to \mathbb{R}$ which is called the *Hamiltonian*. We say that the configuration $\sigma^{(1)} \in \Sigma_\Lambda$ is energetically more favourable than the configuration $\sigma^{(2)} \in \Sigma_\Lambda$, if $H(\sigma^{(1)}) > H(\sigma^{(2)})$.

Remark 1.2.2. *Physical convention is that the configurations with smaller energies are more favourable. Mathematically this is only a matter of considering the Hamiltonian $\tilde{H}_\Lambda := -H_\Lambda$.*

The *inverse temperature* is the real parameter $\beta \in \mathbb{R}_+$. A *Gibbs measure* (also called the Boltzmann distribution) $\mathscr{G}_\Lambda(\beta) \in \mathscr{M}_1(\Sigma_\Lambda)$ is a measure satisfying

$$\frac{\mathrm{d}\mathscr{G}_\Lambda(\beta)}{\mathrm{d}\mu_\Lambda} = \frac{1}{Z_\Lambda(\beta)} \exp\left(\beta H_\Lambda\right), \tag{1.1}$$

where necessarily

$$Z_\Lambda(\beta) := \mu_\Lambda\left[\exp\left(\beta H_\Lambda\right)\right]. \tag{1.2}$$

The normalisation constant (1.2) is referred to as the *partition function*.

Remark 1.2.3. *Since we do not assume the compactness of Σ, we have to assume the existence of* (1.2)*, for all $\beta \geq 0$. The partition function* (1.2) *is a very familiar object. It is just the Laplace transform of the random variable H_Λ on the a priori probability space $(\Sigma_\Lambda, \mathscr{S}_\Lambda, \mu_\Lambda)$. This immediately implies that the partition function is increasing and also (by the Hölder inequality) logarithmically convex function of β. Moreover, if Σ is a compact set, then $Z_\Lambda(\beta)$ is an analytic function of β.*

A mathematical pattern behind (1.1) is as follows. At inverse temperature $\beta = 0$, or equivalently, at infinite *temperature* $T := \beta^{-1} = +\infty$, we obviously have $\mathscr{G}_\Lambda(\beta) = \mu_\Lambda$ which is a sound physical fact. A system which is extremely heated up almost ignores its internal energetic barriers. Hence, due to the analyticity of the Gibbs weights (with respect to β), we have in a (sufficiently small) neighbourhood of $\beta = 0$ that irrespective of the structure of the Hamiltonian the system looks just so as there is no energetic barriers at all.

On the other extreme, if we consider the low temperature limit, i.e., let $\beta \uparrow +\infty$ (equivalently $T \downarrow +0$) the system is increasingly exposed to energetic costs. That is, the Gibbs measure is more and more concentrated around the energetically most favourable configurations which are

[2] This terminology is a reminiscent of the fact that initially spin systems were developed as microscopic models of magnetism, where each site models an atom or molecule and a spin state is an orientation of its individual magnetic moment.

the maxima of the Hamiltonian H_Λ. Finally, in the low temperature limit the system "freezes" completely at the maxima of the Hamiltonian H_Λ. That is, the support of the Gibbs measure coincides with the set $\operatorname{argmax} H_\Lambda$ which is called the set of *ground states* (Remark 1.2.2 can make this terminology more transparent). More formally, assume (for simplicity) that Σ is a compact set, $H_\Lambda \in C(\Sigma)$, and $\operatorname{card}\left(\operatorname{argmax} H_\Lambda\right) < \infty$. If we denote the uniform measure on the set $\operatorname{argmax} H_\Lambda$ by $U(\operatorname{argmax} H_\Lambda) \in \mathscr{M}_1(\Sigma_\Lambda)$, then it is easy to show that

$$\|\mathscr{G}_\Lambda(\beta) - U(\operatorname{argmax} H_\Lambda)\|_{\mathrm{TV}} \xrightarrow[\beta\uparrow+\infty]{} 0. \tag{1.3}$$

It is exactly this pattern relating the Gibbs measure and the maxima of the Hamiltonian H_Λ which makes the Gibbs measures attractive in applications. Often the problems of finding the extrema of functions depending on a large number of variables are computationally hard. Gibbs measures provide the whole family of objects indexed by the real parameter $\beta \in \mathbb{R}_+$. This allows to approach (both analytically and algorithmically) the hard problem at $\beta = +\infty$ gradually.

1.2.2 Thermodynamic quantities

The *observables* are the measurable functions of spin configurations $O : \Sigma_\Lambda \to \mathbb{R}$[3]. An important quantity is the *thermodynamic potential* defined as

$$P_\Lambda(\beta) := \log Z_\Lambda(\beta). \tag{1.4}$$

Remark 1.2.3 implies that the thermodynamic potential is convex. This quantity has the following useful property (cf. (1.3))

$$\lim_{\beta\uparrow+\infty} \frac{1}{\beta} P_\Lambda(\beta) = \max_{\sigma\in\Sigma_\Lambda} H_\Lambda(\sigma). \tag{1.5}$$

Another useful property of this quantity is that it allows to compute averages with respect to the Gibbs measure of other observables included in the Hamiltonian. For example, we have

$$\frac{\mathrm{d}}{\mathrm{d}\beta} P_\Lambda(\beta) = \mathscr{G}_\Lambda(\beta)\left[H_\Lambda\right].$$

More generally, given some macroscopic observable $O : \Sigma_\Lambda \to \mathbb{R}$, fix some real parameter $h \in \mathbb{R}$ and define, for $\sigma \in \Sigma_\Lambda$, the modified Hamiltonian $H_\Lambda(h;\cdot) : \Sigma_\Lambda \to \mathbb{R}$ as follows

$$H_\Lambda(h;\sigma) := H_\Lambda(\sigma) + hO(\sigma).$$

Then we have

$$\frac{\mathrm{d}}{\mathrm{d}h} P_\Lambda(\beta) = \mathscr{G}_\Lambda(\beta)\left[O\right].$$

[3] In other words, the observables are the random variables on the configuration space.

1.2.3 The elementary Gibbs variational principle

Another motivation for introducing the thermodynamic potential (and the Gibbs measure) is the fact that they solve a natural variational problem.

The functional $S : \mathscr{M}_1(\Sigma)^2 \to \mathbb{R}_+$ is called the *relative entropy* (or a Kullback-Leibler divergence) of ν_1 with respect to ν_2, if the following holds

$$S(\nu_1 \mid \nu_2) := \begin{cases} \nu_2 \left[\log\left(\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2}\right)\right], & \nu_1 \ll \nu_2, \\ +\infty, & \text{otherwise.} \end{cases} \tag{1.6}$$

Define the functional $\varphi_\Lambda(\beta;\cdot) : \mathscr{M}_1(\Sigma_\Lambda) \to \underline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\}$ as

$$\varphi_\Lambda(\beta;\nu) := \beta\nu\left[H_\Lambda\right] - S(\nu \parallel \mu_\Lambda).$$

The following proposition holds true.

Proposition 1.2.1 (elementary Gibbs variational principle)**.** *For any $\beta \in \mathbb{R}_+$, we have*

$$P_\Lambda = \varphi_\Lambda(\beta;\mathscr{G}_\Lambda(\beta)) = \max_{\nu \in \mathscr{M}_1(\Sigma_\Lambda)} \varphi_\Lambda(\beta;\nu). \tag{1.7}$$

Remark 1.2.4. *See, e.g., Ellis (2006) for more complicated instances of the same proposition for infinite-volume (*$\operatorname{card}\Lambda = \infty$*) spin systems.*

In words, (1.7) means that the Gibbs measure $\mathscr{G}_\Lambda(\beta)$ realises the best possible balance between the relative entropy and energy. Besides, the Gibbs variational principle gives a straightforward way to obtain *lower bounds* on the thermodynamic potential by choosing various trial measures in (1.7).

In Chapter 3, following the seminal work of Aizenman *et al.* (2003), we obtain another variational scheme for the SK model with multidimensional spins. This scheme expresses the free energy of this model as the value of the Parisi functional at its saddle point (Theorems 3.1.1 and 3.1.2).

1.2.4 Thermodynamic limit

It is often hopelessly hard to study *exact* behaviour of thermodynamic quantities of interest for *finite but large* systems due to their enormous dimensionality and the presence of irregularities. The receipt of statistical mechanics is to approximate the quantities of interest for the finite systems by their counterparts from the limiting infinite system. This approximation procedure is called the *thermodynamic limit*.

Limiting thermodynamic potential

Consider the sequence of finite index spaces $\{\Lambda_N\}_{N\in\mathbb{N}}$ and the corresponding family of the Hamiltonians $\{H_{\Lambda_N}\}_{N=1}^\infty$ such that

$$\lim_{N\uparrow+\infty} \operatorname{card}\Lambda_N = +\infty.$$

In what follows, we assume that the index spaces $\{\Lambda_N\}_{N\in\mathbb{N}}$ are organised such that, for all $N \in \mathbb{N}$, $\Lambda_N \subset \Lambda_{N+1} \subset \Lambda_0$, where Λ_0 is some countable set such that $\Lambda_0 = \bigcup_{N\in\mathbb{N}} \Lambda_N$.

Question 1.2.1 (thermodynamic potential)**.** *What is the asymptotic behaviour of thermodynamic quantities such as the thermodynamic potential P_{Λ_N} in the thermodynamic limit, i.e., as $N \uparrow +\infty$? Does the limit exists? What is the behaviour of the limit with respect to β?*

To fix some natural scale, it is by convention assumed in statistical mechanics that thermodynamic quantities grow asymptotically linearly[4] with size of the system, that is,

$$P_{\Lambda_N}(\beta) \underset{N\uparrow+\infty}{\sim} p(\beta)\operatorname{card}\Lambda_N, \tag{1.8}$$

where $p(\beta)$ is some real constant. Therefore, in view of Question 1.2.1 and for notational convenience, we define the density of thermodynamic potential which is called the *free energy* as

$$p_N(\beta) := \frac{1}{\operatorname{card}\Lambda_N} P_{\Lambda_N}(\beta). \tag{1.9}$$

Note that the assumption (1.8) implies that

$$\lim_{N\uparrow+\infty} p_N(\beta) = p(\beta). \tag{1.10}$$

Infinite-volume Gibbs measures

The following question is rather imprecise as it is stated.

Question 1.2.2 (limiting Gibbs measure)**.** *Can we describe the limiting behaviour of the sequence of Gibbs measures $\{\mathscr{G}_{\Lambda_N}(\beta)\}_{N\in\mathbb{N}}$? What are the limiting objects of this sequence?*

A clean way to answer the previous question is based on the concept of the *Dobrushin-Lanford-Ruelle (DLR) states*, see, e.g., Bovier (2006); Simon (1993). This approach requires, however, more structure from the Hamiltonian than it has been assumed previously. In particular, the Hamiltonian should (1) be defined for any $\Lambda \Subset \Lambda_0$, (2) be the function defined on the whole Σ_{Λ_0}, i.e., $H_\Lambda(\cdot) : \Sigma_{\Lambda_0} \to \mathbb{R}$, and (3) have the following form

$$H_\Lambda(\sigma) = \sum_{A: A\cap\Lambda\neq\emptyset} U_A([\sigma]_A), \tag{1.11}$$

where, for $A \Subset \Lambda_0$, the *interaction potential* $U_A : \Sigma_A \to \mathbb{R}$ is the measurable with respect to $\mathscr{S}_A$ function. In this case, given any index subset $\Lambda \Subset \Lambda_0$ and any *external condition* $\sigma^{(c)} \in \Sigma_{\Lambda_0\setminus\Lambda}$, we define the *local specification* $\mathscr{G}_\Lambda(\beta, \sigma^{(c)}) \in \mathscr{M}_1(\Sigma_{\Lambda_0})$, for a measurable set $\Delta \subset \Sigma_{\Lambda_0}$, as

$$\mathscr{G}_\Lambda(\beta, \sigma^{(c)})(\Delta) := \frac{1}{Z_\Lambda(\beta, \sigma^{(c)})} \int_\Lambda \mathbb{1}_\Delta(\sigma \shortparallel \sigma^{(c)}) \exp\Big(\beta H_\Lambda(\sigma \shortparallel \sigma^{(c)})\Big)\, \mathrm{d}\mu_\Lambda(\sigma), \tag{1.12}$$

where

$$Z_\Lambda(\beta, \sigma^{(c)}) := \int_\Lambda \exp\Big(\beta H_\Lambda(\sigma \shortparallel \sigma^{(c)})\Big)\, \mathrm{d}\mu_\Lambda(\sigma).$$

[4] In physics such quantities are called *extensive*.

The set $\mathfrak{G}_0(\beta) \subset \mathscr{M}_1(\Sigma_{\Lambda_0})$ of all *infinite-volume Gibbs measures* (or states) is the set of measures $\mathscr{G}_0(\beta) \in \mathfrak{G}_0(\beta)$ such that, for all index sets $\Lambda \Subset \Lambda_0$, and all *external conditions* $\sigma^{(\mathrm{c})} \in \Sigma_{\Lambda_0 \setminus \Lambda}$, we have

$$\mathscr{G}_0(\beta)\{\cdot \mid [\sigma]_{\Lambda_0 \setminus \Lambda} = \sigma^{(\mathrm{c})}\} = \mathscr{G}_\Lambda(\beta, \sigma^{(\mathrm{c})})\{\cdot\}, \tag{1.13}$$

$\mathscr{G}_0(\beta)$-almost surely. Equation (1.13) is called the *DLR equation*. Under certain regularity conditions (see, e.g., (Bovier, 2006, Corollary 4.2.8)) it can be proved that the set of all infinite-volume measures $\mathfrak{G}_0(\beta)$ is not empty and can be constructed by rather straightforward limiting procedure. Namely, it can be constructed as a convex hull of a set of all weak limiting points (as $N \uparrow +\infty$) for the following sequences of local specifications $\{\mathscr{G}_{\Lambda_N}(\beta, \sigma_N^{(\mathrm{c})})\}_{N\in\mathbb{N}}$ with all possible external conditions $\{\sigma_N^{(\mathrm{c})}\}_{N\in\mathbb{N}} \subset \Sigma_{\Lambda_0} \setminus \Sigma_{\Lambda_N}$.

The above set of ideas and especially the condition (1.13) shows that the Gibbs measures generated by the Hamiltonians (1.11) can be viewed as graphical (spatial) generalisations of Markov chains (see, e.g., Preston (1976)). Such generalisations (e.g., Markov random fields) are important in various applications such as information processing, see, e.g., Winkler (2003).

In this thesis, we shall deal with mean-field systems, where the external conditions are irrelevant. Moreover, the Hamiltonians of such systems are not representable in the form (1.11). Thus, the described above framework of the DLR equations is not applicable in the context of mean-field models. See Section 1.3.4, for a more precise formulation of the questions which will be dealt with in the present work.

Remark 1.2.5 (infinite-volume Gibbs variational principle)**.** *An alternative to the DLR states way of defining the infinite-volume Gibbs measure is by using the infinite-volume analogues of Proposition 1.2.1, see, e.g., Ellis (2006).*

Phase transitions

After passing to the thermodynamic limit, many kinds of interesting effects can occur at the level of free energy (cf. (1.10)). For example, the free energy may become a non-analytic function of β. Such situation is usually referred to as the *phase transition*. In the context of the DLR states, another indicator of the phase transition at the level of the Gibbs measure is the situation, where the set $\mathfrak{G}_0(\beta)$ contains more than one element.

1.3 Spin-glasses

In this section, we introduce disordered spin systems we are going to study in the present work. These are the spin systems with strongly randomised interactions.

1.3.1 Basic mathematical objects

In this subsection, we introduce the deterministic objects that will be randomised afterwards and used as building blocks of spin-glass models.

Deterministic instances

We start from definition of non-disordered Hamiltonians with pair interactions.

The *interaction kernel* is the real matrix $A \in \mathbb{R}^{\Lambda \times \Lambda}$. The *external field* is the real vector $h = (h_i)_{i \in \Lambda} \in \mathbb{R}^{\Lambda}$. The (pair interaction) *Hamiltonian* is the mapping $H_\Lambda(A,h) : \Sigma_\Lambda \to \mathbb{R}$ defined as

$$H_\Lambda(A,h;\sigma) := \sum_{i,j \in \Lambda} A_{i,j} \langle \sigma_i, \sigma_k \rangle + \sum_{i \in \Lambda} \langle h_i, \sigma_i \rangle. \tag{1.14}$$

Remark 1.3.1. *More generally, for $p \in \mathbb{N}$, one can consider the p-*spin interactions *defined as follows. Let $A^{(p)} \in \mathbb{R}^{\Lambda^p}$ be a real tensor of rank $p \in \mathbb{N}$. We also set $d = 1$ for simplicity. Define the p-spin Hamiltonian $H_\Lambda(A^{(p)},h) : \Sigma_\Lambda \to \mathbb{R}$ in the following way*

$$H_\Lambda(A^{(p)},h;\sigma) := \sum_{i_1,\ldots,i_p \in \Lambda} A_{i_1,\ldots,i_p} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p} + \sum_{i \in \Lambda} h_i \sigma_i. \tag{1.15}$$

Remark 1.3.2. *The presence of external field (i.e., $h \neq 0$) in* (1.14) *introduces a bias which makes the mapping* (1.14) *less symmetric by excluding, e.g., the* rotational symmetry

$$H_\Lambda(A,0,O\sigma) = H_\Lambda(A,0,\sigma),$$

where $O \in \mathscr{O}(d)$ is an arbitrary orthogonal matrix.

Remark 1.3.3. *An important difference of the Hamiltonian* (1.14) *from the Hamiltonian* (1.11) *is that the former one does not depend on spins outside Λ. For this reason, we say that the former Hamiltonian does not interact with the exterior of Λ.*

Graph of interacting sites

Note that the Hamiltonian (1.14) naturally induces the graph of interacting sites or *interaction graph* $G = (V,E)$, with vertices $V := \Lambda$ and edges $E := \{(i,j) \in \Lambda^2 \mid A_{i,j} \neq 0\}$.

1.3.2 Quenched disorder and random processes

Disorder

In the sequel, we shall assume that the interaction matrix A is the random matrix on the *probability space of disorder* $(\Omega, \mathscr{F}, \mathbb{P})$. This probability space is always assumed to be large enough to contain all random variables we shall be dealing with (except those explicitly defined on the spin configuration space Σ_Λ). We shall be interested in *typical behaviour* of systems governed by the random Hamiltonians (1.14).

Random processes

Due to the presence of disorder, the spin-glass Hamiltonian (1.14) can be viewed as a *random process* on the probability space of disorder that is indexed by the configuration space Σ_Λ (cf. Definition 2.1.1). This "Hamiltonian-as-a-random-process" point of view we shall adopt in what follows. In the sequel, we shall consider not only the pair interaction random Hamiltonians but also the "abstract" random processes $H_\Lambda := \{H_\Lambda(\sigma)\}_{\sigma \in \Sigma_{\Lambda_N}}$ on the probability space of disorder. Such Hamiltonians induce the random free energy and the random Gibbs measure as specified by (1.9) and (1.1).

To lighten the notation, we shall use the following simplifications. Given some random interaction matrix A, we shall write simply $H_\Lambda(h;\sigma)$ instead $H_\Lambda(A,h;\sigma)$. Given $\beta \in \mathbb{R}_+$ and $h \in \mathbb{R}$, the random Gibbs measure induced by $H_\Lambda(h;\sigma)$ at inverse temperature β will be denoted by $\mathscr{G}_\Lambda$.

1.3.3 Realistic spin-glass models

We start from two motivating examples of the physical relevance.

The Ruderman-Kittel-Kasuya-Yoshida (RKKY) model

Physically, spin-glasses are amorphous magnetic alloys. We can imagine the following mathematical model. Given some dimension $d \in \mathbb{N}$, let $X_N \subset \mathbb{R}^d$ be the box $X_N := [-N;N]^d$. Suppose we have a family of i.i.d. random variables $\{r_i\}_{i=1}^{N^d}$ uniformly distributed on X_N. Realisations of the family r model the quenched positions of impurities in the sample of material X_N. Thus, we are dealing with quenched inhomogeneities[5]. Put $\Lambda_N := \{1, \ldots, N^d\}$. Let further $J : \mathbb{R} \to \mathbb{R}$ be an even function decaying to zero at infinity. Assume that the spin configuration space is the unit sphere $\Sigma := \mathbb{S}^{d-1}$. Given the direction of *external field* $h \in \mathbb{R}^d$, we define the Hamiltonian $H_N^{\mathrm{RKKY}}(h) : \Sigma_{\Lambda_N} \to \mathbb{R}$ as

$$H_N^{\mathrm{RKKY}}(h;\sigma) := \sum_{\substack{i<j \\ i,j\in\Lambda_N}} J(r_i - r_j)\langle \sigma_i, \sigma_j \rangle + \langle h, \sum_{i\in\Lambda} \sigma_i \rangle, \quad \sigma \in \Sigma_{\Lambda_N}. \tag{1.16}$$

A physical example of the function J is the *RKKY exchange potential*

$$J(r) := \frac{1}{\|r\|^3} \cos(\|r\|). \tag{1.17}$$

Physical reasons for this form of the exchange potential J can be found, e.g., in Fischer & Hertz (1991). The Hamiltonian (1.16) becomes a rather complex expression with the above choice of the exchange potential (1.17). The summands involving the exchange potential decay (albeit slowly) as the distance between the interacting sites grows. Hence, the interaction between the spins $i, j \in \Lambda_N$ with large distance $\|r_i - r_j\|$ between them gives only a small contribution to (1.16). Moreover, the exchange potential oscillates taking positive and negative values. Due to these oscillations, there are summands in (1.16) with $J_{i,j} := J(r_i - r_j) > 0$ (*ferromagnetic interaction*). In this case collinear spins σ_i and σ_j are energetically the most favourable. There are, however, nonnegligible summands in (1.17) with $J_{i,j} < 0$ (*antiferromagnetic interaction*). In the latter case anticollinear spins are energetically the most favourable. These two cases may impose contradicting requirements on the spin configuration σ, since, for the triple $i, j, k \in \Lambda_N$ such that $i < j < k$ with non negligible probability we have $J_{i,j} > 0$, $J_{j,k} > 0$ and $J_{i,k} < 0$. This conflicting *non transitive* situation is called *frustration*.

Unfortunately, even on the heuristic qualitative level of theoretical physics the thermodynamic limit of the RKKY model seems to be very hard to analyse, see, however, Zegarliński (1998).

The Edwards-Anderson (EA) model

The EA model (Edwards & Anderson, 1975) is an attempt to simplify the complicated picture of the previous model, while retaining the most important features of it, namely the *frustration*, *quenched disorder* and *locality of interactions*.

Let $\Lambda_N := \mathbb{Z}^d \cap [-N;N]^d$ and also let $\Sigma := \{-1;1\} \subset \mathbb{R}$. Let now $\{g_{i,j}\}_{i,j\in\Lambda_N}$ be a family of i.i.d. standard Gaussian random variables. Consider the following EA Hamiltonian

[5] An example of such material is the Cu-Mg alloy (with 1% of Mg).

$$H_N^{\mathrm{EA}}(h,\sigma) := \sum_{\substack{\|i-j\|_2=1 \\ i,j\in\Lambda_N}} g_{i,j}\sigma_i\sigma_j + h\sum_{i\in\Lambda}\sigma_i. \tag{1.18}$$

At the time of writing, there seems to be no consensus (except for the cases $d = 1$ and "$d = \infty$"[6]) in the physical literature about the qualitative picture of the EA model in the thermodynamic limit (see Newman & Stein (2007) for a recent review).

1.3.4 Thermodynamic limit in spin-glasses

To lighten the notation in what follows, we shall simply write N instead of Λ_N in quantities depending on Λ_N. We shall sometimes drop the subscript index N in occurrences, where it stays fixed and its value is clear from the context.

Given the random interaction matrix A, we can naturally extend Questions 1.2.1 and 1.2.2 to the case of quenched disorder and ask a few additional ones.

The extension of Question 1.2.1 to the case of quenched disorder is Question 0.0.1 from the introduction. This is the first main question being addressed in the present work.

Question 1.3.1 (maximization algorithm for the average-case scenaio)**.** *Can one produce an efficient (for the average-case scenario) algorithm for finding the maximiser of* (1.14) *or at least an "almost maximiser" of it?*

For algorithmic purposes of Question 1.3.1, it would be interesting to know an answer to the quenched version of Question 1.2.2 and, in particular, to the following.

Question 1.3.2 (structure of the maximisers)**.** *What is the typical structure of the (almost) maximisers of* (1.14)*? How many maximisers are there? Are they very different?*

Since the information about the almost maximisers can be gained from the geometry of the Gibbs measure, it is natural to ask also the following.

Question 1.3.3 (structure of the random Gibbs measure)**.** *Does the random Gibbs measure look almost surely uniform or it concentrates around certain random domains in the thermodynamic limit? Can we construct the limiting random object that encodes the structure of the random Gibbs measure in the thermodynamic limit?*

Question 1.3.3 is certainly very informal. The partial excuse is that at the writing time there is no general framework which gives a proper limiting description of the random Gibbs measure. See, however, Bovier (2006); Newman & Stein (2007), and references therein (cf. also Section 1.5.3).

In this thesis, we shall instead of Question 1.3.3 deal with closely related Question 0.0.2 from the introduction.

To make Questions 1.3.2 and 1.3.3 more sensible, we have to specify some metric structure on the configuration space. This metric structure depends, of course, crucially on the choice of the Hamiltonian. In the next section we shall consider some representative examples.

[6] The latter case is the alias for the SK model on which more is said below.

1.4 Some mean-field Gaussian spin-glass models

In this section, we define the main models considered in the present work. They fall into the class of *mean-field* spin-glass models which Hamiltonians are Gaussian processes indexed by high dimensional configuration spaces.

The term *mean-field* refers to the idea to approximate complex interaction geometries (as in the EA model) by certain effective, in some sense averaged, interactions with “simpler” geometries. One way to make this approximation rigorous is Proposition 1.2.1, where one restricts the maximisation to product measures, see, e.g., Opper & Saad (2001); Simon (1993).

1.4.1 The Sherrington-Kirkpatrick model

The celebrated Sherrington-Kirkpatrick model is based on a very simple instance of the Hamiltonian (1.14) on a *fully connected* graph G. The model ignores completely the locality of interactions but retains high degrees of frustration and quenched disorder. A large body of heuristics concerning spin-glasses were developed in theoretical physics during the analysis of the SK model and then were applied to other models. In a sense, the SK model is the most understood spin-glass model in theoretical physics. Some of heuristics concerning Question 0.0.1 have been rigorously confirmed for the SK model in the series of papers by Aizenman *et al.* (2003); Guerra (2003); Guerra & Toninelli (2002); Talagrand (2006b), see also Section 1.5.5 for some details. But still, as of the time of writing, the model contains quite a few puzzles, especially concerning Questions 1.3.1, 1.3.2 and certainly the hardest Question 1.3.3. For a recent short account of mathematical results on this model see Bolthausen (2007) and references therein.

To define the SK model, we make the following assumption.

Assumption 1.4.1 (Ising spins). *Let the index space be $\Lambda_N := [1;N] \cap \mathbb{N}$ and let the spin configuration space be the Ising one, that is $\Sigma := \Sigma^{Ising} := \{-1;1\}$. Therefore, the spin configuration space Σ_{Λ_N} of the whole system is the* discrete hypercube $\{-1;1\}^N$. *We also consider the uniform a priori distribution μ on the configuration space Σ so that the individual spin configurations are the Rademacher random variables on the a priori probability space.*

Let $\{g_{i,j}\}_{i,j=1}^N$ be a family of i.i.d. standard Gaussian random variables (the same as in (1.18)). We set $A_{i,j} := \frac{1}{\sqrt{N}} g_{i,j}$ and $h_i := h \in \mathbb{R}$. Hence, the Hamiltonian (1.14) assumes the form of the *SK Hamiltonian*

$$H_N^{\mathrm{SK}}(\sigma) := \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{i,j} \sigma_i \sigma_j + h \sum_{i \in \Lambda} \sigma_i. \tag{1.19}$$

The reason for the normalisation factor $1/\sqrt{N}$ in (1.19) is to satisfy our assumption (1.8).

Remark 1.4.1. *The Gaussian character of the interaction matrix is not very crucial, at least, if Question 0.0.1 is concerned, see Carmona & Hu (2006). However, it makes the analysis easier.*

The SK Hamiltonian (1.19) is, thus, a Gaussian process on the spin configuration space Σ_N. For convenience, we shall also use the following centred Gaussian process

$$\mathrm{SK}_N(\sigma) := \frac{1}{N} \sum_{i,j=1}^{N} g_{i,j} \sigma_i \sigma_j. \tag{1.20}$$

The distribution of this process is completely determined by its covariance structure

$$\mathbb{E}\left[\mathrm{SK}_N(\sigma^{(1)})\mathrm{SK}_N(\sigma^{(2)})\right] = \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(1)}\sigma_i^{(2)}\right)^2.$$

The *overlap* is the mapping $R : \Sigma_N^2 \to [-1;1]$ defined as

$$R(\sigma^{(1)},\sigma^{(2)}) := \frac{1}{N}\sum_{i=1}^{N}\sigma^{(1)}\sigma^{(2)}. \tag{1.21}$$

In words, the quantity (1.21) measures the number of matching coordinates of the vectors $\sigma^{(1)}$ and $\sigma^{(2)}$. In fact, if we denote the *Hamming distance* between $\sigma^{(1)}$ and $\sigma^{(2)}$ through $\mathrm{d_H}(\sigma^{(1)},\sigma^{(2)})$, then we have

$$R(\sigma^{(1)},\sigma^{(2)}) = 1 - 2\mathrm{d_H}(\sigma^{(1)},\sigma^{(2)}). \tag{1.22}$$

Remark 1.4.2 (canonical L^2-distance)**.** *The Hamming distance can also be obtained (modulo a constant factor) as the canonical L^2-distance induced by the Gaussian process SK_N on the configuration space Σ_N*

$$\mathrm{d}_H(\sigma^{(1)},\sigma^{(2)}) := \frac{1}{\sqrt{2}}\sqrt{\mathbb{E}\left[(SK_N(\sigma^{(1)}) - SK_N(\sigma^{(2)}))^2\right]}.$$

The model (1.20) can be seen as a particular instance of the models which have the covariance structure of the following form

$$\mathbb{E}\left[X_N(\sigma^{(1)})X_N(\sigma^{(2)})\right] = f(R(\sigma^{(1)},\sigma^{(2)})), \tag{1.23}$$

for *some* function $f : [-1;1] \to \mathbb{R}$ is such that $f(0) = 0$ and $f(1) = 1$.

In particular, the Hamiltonian (1.15) with $A_{i_1,\ldots,i_p} := N^{-p/2}g_{i_1,\ldots,i_p}$, where $\{g_{i_1,\ldots,i_p}\}_{i_1,\ldots,i_p=1}^{N}$ is the family of i.i.d. standard Gaussian random variables, has the covariance structure given by (1.23) with $f(q) := q^p$. This model is called *Derrida's p-spin model.*

1.4.2 The random energy model

Perhaps, the simplest possible model in the framework of Section 1.3.2 is the *random energy model*. It was introduced by Derrida (1980, 1981), who considered the Gaussian process $\{H^{\mathrm{REM}}(\sigma)\}_{\sigma\in\Sigma_N}$ without any correlations. More formally, consider

$$H_N^{\mathrm{REM}}(\sigma) := \sqrt{N}\mathrm{REM}_N(\sigma), \tag{1.24}$$

where $\{\mathrm{REM}(\sigma)\}_{\sigma\in\Sigma_N}$ is the family of 2^N i.i.d. standard Gaussian random variables. We work under Assumption 1.4.1.

As opposed to pair interaction Hamiltonians (cf. (1.14)), the REM and, as well as its generalised version from the next section, lack a description in terms of interacting microscopic spin variables σ_i. Nevertheless, while being a "toy model", the REM serves quite surprisingly as a building block of exact hierarchical structure arising in the thermodynamic limit of the SK model (Section 2.2.1). Note that the SK model is a model with the genuine interacting spins description (1.19). It is conjectured that the above mentioned hierarchical structure emerges *universally* also for other mean-field spin-glass models (see Section 1.5.2). Related universality facts on the level of random Hamiltonians (as opposed to the level of free energy and Gibbs measure, cf. Section 1.3.4) have been established in the series of papers by Bauke & Mertens (2004); Ben Arous *et al.* (2006); Borgs *et al.* (2001, 2005a,b); Bovier & Kurkova (2004c, 2006a,b).

Remark 1.4.3. *The covariance of the process $REM_N(\cdot)$ can also be obtained as the limit of the covariance of Derrida's p-spin model as $p \uparrow +\infty$.*

Remark 1.4.4. *The canonical L^2-distance generated by the process $REM_N(\cdot)$ induces, obviously, the discrete topology on Σ_N.*

It turns out that already this simple model without correlations demonstrates a phase transition. Indeed, the limiting free energy is (see Chapter 6) almost surely (and in L^p) given by

$$p(\beta) = \begin{cases} \frac{\beta^2}{2}, & \beta \leq \beta_c := \sqrt{2\log 2} \\ \frac{\beta_c^2}{2} + \beta_c(\beta - \beta_c), & \beta > \beta_c. \end{cases} \tag{1.25}$$

The two cases in (1.25) can easily be explained by the breakdown of the law of large numbers (LLN) in the partition function of the REM for $\beta > \beta_c$.

The questions listed in Section 1.3.4 have been considered and settled for the REM by Bovier *et al.* (2002). The above game of i.i.d. processes on the hypercube in the framework of Section 1.3.2 can also be played for non-Gaussian distributions, see Ben Arous *et al.* (2005).

1.4.3 The generalised random energy model

The *generalised random energy model* was proposed by Derrida (1985). A recent account of mathematical results on the GREM can be found in Bovier & Kurkova (2007). The results are substantially more elaborate than for the SK model. Even the hardest Question 1.3.3 can be answered for this model completely (Bovier & Kurkova, 2003a). However, perhaps most significantly, the importance of the GREM stems from the fact that it emerges in a variational formula for the free energy of the *a priori* much harder SK model. The variational principle was proposed by Aizenman *et al.* (2003) (see also Section 2.3.3). This principle allows, in particular, to obtain upper bounds on the quenched free energy of the SK model. These bounds have been proved to be exact in the thermodynamic limit (Talagrand, 2006b), as prescribed by the Parisi formula.

The GREM is based on the following random Hamiltonian

$$H_N^{\mathrm{GREM}}(\sigma) := \sqrt{N}\mathrm{GREM}_N(\sigma),$$

where $\{\mathrm{GREM}_N(\sigma)\}_{\sigma\in\Sigma_N}$ is the Gaussian random process satisfying (0.5). (The model uses Assumption 1.4.1.)

Consider the space of discrete order parameters

$$\mathscr{Q}'_n := \{q : [0;1] \to [0;1] \quad | \quad q(0) = 0, q(1) = 1, q \text{ is non-decreasing, piece-wise constant with } n \text{ jumps}\}. \tag{1.26}$$

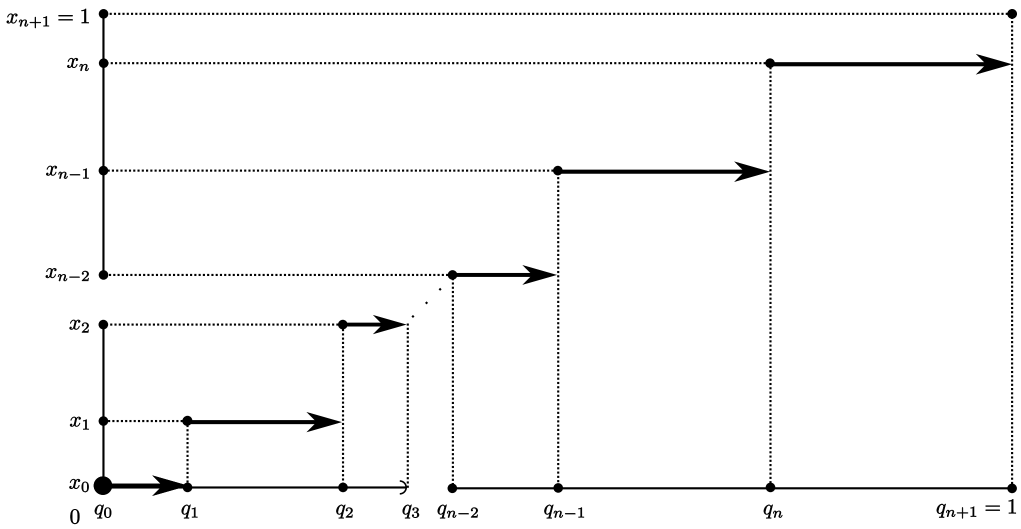

Fig. 1.1. An example of discrete order parameter (overlap distribution function)

Recall the function ρ from (0.5). Assume that $\rho \in \mathscr{Q}'_n$. In what follows, we shall refer to ρ as the *discrete order parameter*. In this case, it is possible to construct the process GREM_N as a finite sum of independent Gaussian processes. Assume that

$$\rho(x) = \sum_{k=1}^{n} q_k \mathbb{1}_{[x_k;x_{k+1})}(x), \tag{1.27}$$

where

$$0 =: x_0 < x_1 < \ldots < x_n = 1, \tag{1.28}$$

$$0 =: q_0 < q_1 < \ldots < q_n = 1. \tag{1.29}$$

See Figure 1.1 for an example graph of discrete order parameter.

Let $\{a_k\}_{k=1}^n \subset \mathbb{R}$ be such that $a_k^2 = q_k - q_{k-1}$. We assume that, for all $k \in [1;n] \cap \mathbb{N}$, we have $x_k N \in \mathbb{N}$[7] and also $a_k \neq 0$. Denote $\Delta x_l := x_l - x_{l-1}$.

Consider the family of i.i.d standard Gaussian random variables

$$\{X(\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(k)}) \mid k \in [1;n] \cap \mathbb{N}, \sigma^{(1)} \in \Sigma_{x_1 N}, \ldots, \sigma^{(k)} \in \Sigma_{x_k N}\}.$$

Using these ingredients, for $\sigma = \sigma^{(1)} \amalg \sigma^{(2)} \amalg \ldots \amalg \sigma^{(n)} \in \Sigma_N$, we have

$$\mathrm{GREM}_N(\sigma) \sim \sum_{k=1}^{n} a_k X(\sigma^{(1)}, \sigma^{(2)}, \ldots, \sigma^{(k)}). \tag{1.30}$$

Equivalence (1.30) is easily verified by computing the covariance of the right hand side. The computation gives, for $\sigma, \tau \in \Sigma_N$

$$\mathrm{Cov}\left[\mathrm{GREM}_N(\sigma)\mathrm{GREM}_N(\tau)\right] = q_{N q_{\mathrm{L}}(\sigma,\tau)}.$$

Figure 1.2 shows the tree structure corresponding to the GREM with the piece-wise constant function ρ such that $n = 3$. This leads to the tree with three "REM levels".

[7] This condition is for notational simplicity. It means that we actually consider instead of N the increasing sequence $\{N_\alpha\}_{\alpha \in \mathbb{N}} \subset \mathbb{N}$ such that $N_\alpha \uparrow +\infty$ as $\alpha \uparrow +\infty$, satisfying $N_\alpha x_k \in \mathbb{N}$, for all $\alpha \in \mathbb{N}$ and all $k \in [1;n] \cap \mathbb{N}$.

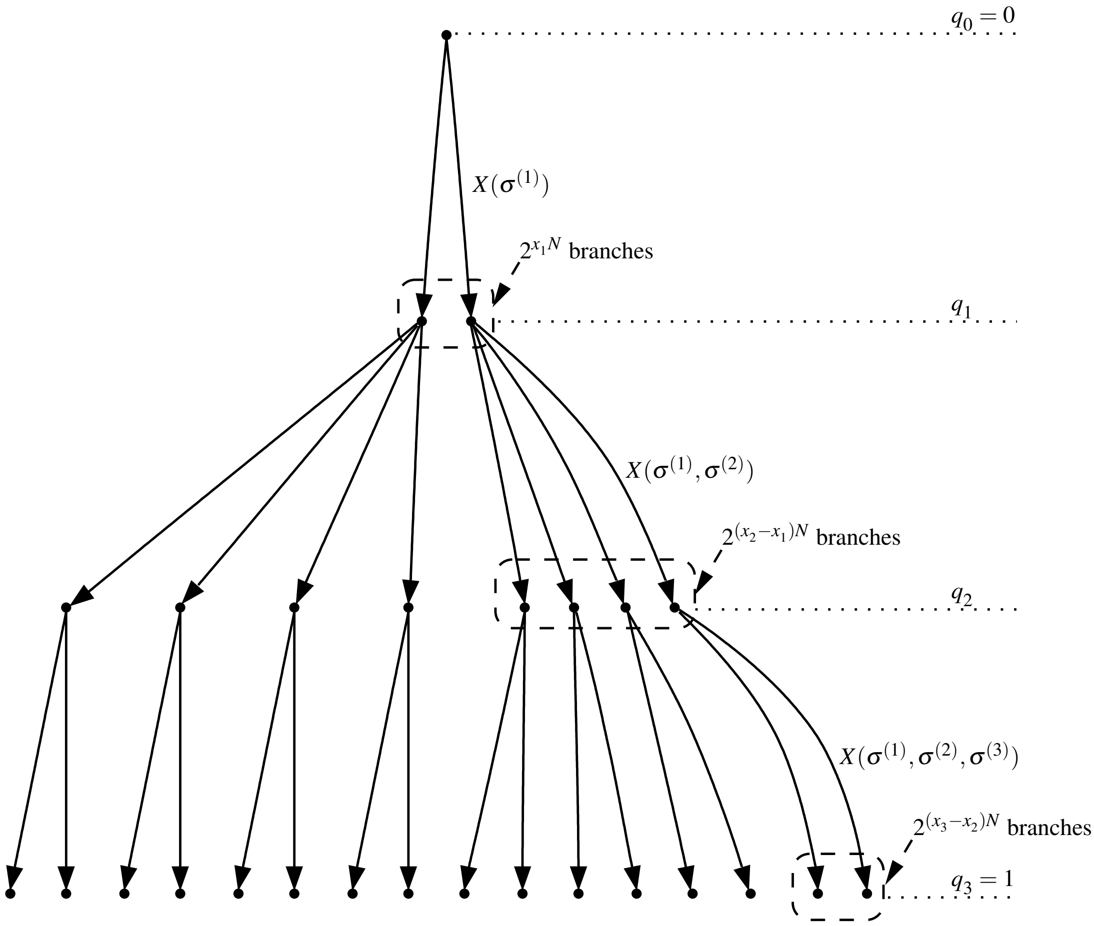

Fig. 1.2. Structure of the GREM with the three levels of hierarchy

1.4.4 Between the GREM and SK model

We now review shortly the results on generalisations of the GREM available in the literature. We classify these results according to the set of questions proposed above.

The GREM is generated by the lexicographic distance (0.3) on the discrete hypercube. Derrida & Gardner (1986) considered the GREM with "lexicographic" external field which is particularly well adapted to the natural lexicographic distance generated by the GREM Hamiltonian. In contrast to the work of Derrida & Gardner (1986), in Chapter 6, we consider the GREM with the uniform (magnetic) field (0.6). This external field is the same as in the SK model, cf. (1.19). We answer both Questions 0.0.1 and 0.0.2 for this model in Chapter 6.

Bolthausen & Kistler (2006) introduced an *a priori* non-hierarchical generalisation of the GREM and analysed it on the level of Question 0.0.1. In a recent preprint, Bolthausen & Kistler (2008) prove that the Gibbs measure of the non-hierarchical GREM hierarchically decomposes in the thermodynamic limit which leads to ultrametricity (Section 1.5.4). This result corresponds to the level of geometry of the Gibbs measure (Question 1.3.3). Due to its *a priori* non-hierarchical character, the model of Bolthausen and Kistler is closer to the SK model than the GREM.

The model of Bolthausen & Kistler (2006) was studied also by Jana (2007); Jana & Rao (2006) at the level of Question 0.0.1 with a generalisation to non-Gaussian distributions of the REM summands in (1.30). Moreover, Jana (2007) analysed the GREM with randomised branching rates (1.28) at the level of Question 0.0.1.

1.4.5 Other mean-field spin-glass models

For the sake of completeness, we mention briefly some mean-field spin-glass models which are not considered in this thesis. Among them are the neural network memory capacity models: the *perceptron* and *Hopfield models*, see Bovier (2006); Shcherbina (2005); Talagrand (2003). Important are also the models of neural network learning, e.g., Klimovsky (2005); Shcherbina & Tirozzi (2002).

Very important for applications are *diluted mean-field spin-glass models*. These models are based on the Hamiltonians with random interaction graphs (e.g., Erdős-Rényi ones). A prominent example is the *satisfiability problem* on random graphs Mézard & Montanari (2007); Talagrand (2003), *random codes* Mézard & Montanari (2007); Nishimori (2001). Naturally, other combinatorial problems such as the *assignment* one can be formulated as mean-field spin-glass models (Talagrand, 2003, Chapter 8). The main distinguishing feature of these models from the ones considered in this thesis is that their Hamiltonians (1.19) are not Gaussian. The important tool in this case seems to be the rigorous versions of the *cavity method*, see, e.g., Bovier (2006); Talagrand (2003). The cavity method is also quite compatible with Gaussian process techniques, as shown by Aizenman *et al.* (2003), see also Section 2.3.3 and Chapter 3.

1.5 Replica symmetry breaking picture

In this section, we shall sketch the heuristic picture of the thermodynamic limit in mean-field spin-glasses developed in theoretical physics (Dotsenko, 2001; Mézard & Montanari, 2007; Mézard *et al.*, 1987; Nishimori, 2001; Parisi, 2007). This picture suggests how the list of questions stated in Section 1.3.4 should be answered.

In landmark series of papers, Parisi and collaborators developed the heuristic Ansatz called the *replica symmetry breaking*, see Mézard *et al.* (1987) and references therein. This Ansatz very plausibly explained effects occurring in a large class of disordered mean-field spin systems. We shall shortly review this Ansatz in the following subsection.

1.5.1 Overlap distribution function and replica symmetry breaking

Assume that $\Sigma_N = \{-1;+1\}^N$. The *overlap* between configurations $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$ is defined as

$$R_N(\sigma^{(1)}, \sigma^{(2)}) := \frac{1}{N} \sum_{i \in \Lambda_N} \sigma_i^{(1)} \sigma_i^{(2)}.$$

The overlap can be viewed as the random variable on the probability space $(\Sigma_N^2, \mathscr{S}_N^{\otimes 2}, \mathscr{G}_N^{\otimes 2})$ (i.e., the overlap is an observable for such a two-times replicated system). To explain the subsequent terminology, we note that one can think of the latter probability space as of two independent *replicas* of the initial spin glass configuration space. However, the marginals of the random measure $\mathscr{G}_N^{\otimes 2}$ are *not* independent random measures on the probability space of disorder, since they depend on the same realisations of disorder.

Let us consider the (finite volume) *overlap distribution function* $F_N(\beta;\cdot) : [0;1] \to [0;1]$ defined as

$$F_N(\beta;q) := \mathbb{E}\left[\mathscr{G}_N^{\otimes 2}(\beta)\left\{R_N(\sigma^{(1)},\sigma^{(2)}) \leq q\right\}\right]. \quad (1.31)$$

We note that the overlap distribution function is indeed a distribution function of some random variable. We shall denote this random variable by $\mathtt{q}$. Assume that the thermodynamic limit

$$F(\beta;q) := \lim_{N\uparrow+\infty} F_N(\beta;q) \quad (1.32)$$

exists. The behaviour of the function (1.32) with respect to β is conjectured (Mézard *et al.*, 1987) to be as follows. There exists $\beta_{\mathrm{c}} \in \mathbb{R}_+$ such that:

1. *Replica symmetric (RS) regime:* for $\beta < \beta_{\mathrm{c}}$, we have $\mathtt{q} \sim \delta_{q(\beta)}$, for some constant $q(\beta) \in \mathbb{R}$.
2. *RSB region:* for $\beta > \beta_{\mathrm{c}}$, we have $\operatorname{card}\operatorname{supp}\mathtt{q} > 1$. Then either of the following two subcases holds
   a) *Discrete RSB:*
      i. *Single-step RSB (1-RSB):* there exist some $\beta_{\mathrm{c}}^{(1)} \in \mathbb{R}$ and $\beta_{\mathrm{c}}^{(2)} \in \overline{\mathbb{R}}_+$ such that $\beta_{\mathrm{c}}^{(2)} > \beta_{\mathrm{c}}^{(1)} > \beta_{\mathrm{c}}$ and, for all $\beta \in (\beta_{\mathrm{c}}^{(1)};\beta_{\mathrm{c}}^{(2)})$, we have

      $$\mathtt{q} \sim w(\beta)\delta_{q_1(\beta)} + (1-w(\beta))\delta_{q_2(\beta)},$$

      for some constants $q_1(\beta), q_2(\beta) \in \mathbb{R}$ and $w(\beta) \in (0;1)$. This regime has been rigorously proven to hold for the REM (Bovier *et al.*, 2002) with $\beta_{\mathrm{c}}^{(2)} = +\infty$, and also for the *p*-spin model (Talagrand, 2000b, 2003), where $\beta_{\mathrm{c}}^{(2)} < +\infty$. Traces of this regime at the level of free energy (Question 0.0.1) were confirmed for the *spherical model* by Talagrand (2006a). In this thesis, we show an analogous result for the SK model with multidimensional Gaussian spins, see Chapter 5.
      ii. *Finite-step RSB (n-RSB):* the 1-RSB regime is a particular case of the following situation

      $$\mathtt{q} \sim \sum_{k=1}^{n+1} w_k(\beta)\delta_{q_k(\beta)},$$

      for some real constants $q_1(\beta) < \ldots < q_{n+1}(\beta)$ and some vector of weights $w(\beta) \in (0;1)^{n+1}$ such that $\sum_{k=1}^{n+1} w_k(\beta) = 1$. See Figure 1.1 for the typical graph of the overlap distribution function. This regime has been rigorously proven for the GREM by Bovier & Kurkova (2004a). In Chapter 6 we prove an analogous result for the GREM with external field at the level of Question 0.0.2.
   b) *Continuous RSB or full RSB (FRSB):* there exists $\beta_{\mathrm{c}}^{(\mathrm{f})} > 0$ such that for $\beta \in (\beta_{\mathrm{c}}^{(\mathrm{f})};+\infty)$, there exist some deterministic weights $w_{\mathrm{m}}, w_{\mathrm{M}} \in [0;1)$, deterministic constants $q_{\mathrm{m}}, q_{\mathrm{M}} \in \mathbb{R}$, and a continuous random variable $\mathtt{q}_{\mathrm{cont}}$ with $\operatorname{supp}\mathtt{q}_{\mathrm{cont}} = [q_{\mathrm{m}}(\beta);q_{\mathrm{M}}(\beta)]$ such that we have the following decomposition

   $$\mathtt{q} \sim w_{\mathrm{m}}(\beta)\delta_{q_{\mathrm{m}}(\beta)} + w_{\mathrm{M}}(\beta)\delta_{q_{\mathrm{M}}(\beta)} + (1-w_{\mathrm{m}}-w_{\mathrm{M}})\mathtt{q}_{\mathrm{cont}}.$$

   This decomposition is believed to hold for the SK model and Derrida's p-spin model for large enough $\beta_{\mathrm{c}}^{(\mathrm{f})}$ (Mézard *et al.*, 1987). Probably the only (at the time of writing) rigorous hint that this regime might hold for the SK model is obtained by Talagrand (2006b) at the level of free energy (which in our classification corresponds to Question 0.0.1). Bovier & Kurkova (2004b) have considered this regime for the continuous version of the GREM (CREM) at the level of Question 1.3.3.

Remark 1.5.1. *The behaviour of the random overlap distribution function*

$$\mathtt{F}_N(\beta;q) := \mathscr{G}_N^{\otimes 2}(\beta)\left\{R_N(\sigma^{(1)},\sigma^{(2)}) \leq q\right\}$$

can be very different from the behaviour of (1.31)*, see Parisi & Talagrand (2004).*

1.5.2 Limiting Gibbs measures and pure states

The RSB described in the previous section is believed to have its roots in the structure of the Gibbs measure.

In classical statistical mechanics of spin systems, Gibbs measures often form a convex compact set $\mathfrak{G}_0(\beta)$, see, e.g., (Simon, 1993, Chapter III.5). The Krein-Milman theorem then assures that there exists the set of extremal points corresponding to the set of all limiting Gibbs measures. These extremal points are called *pure states*.

Somewhat similar but substantially more complicated picture is conjectured (Mézard *et al.*, 1987) to be true for mean-field spin-glass systems in the thermodynamic limit. It is conjectured that in the FRSB case the Gibbs measure decomposes into a countable convex combination of "random pure states", where the coefficients of the convex combination are also random and independent of the pure states. Moreover, it is conjectured that the distribution of these coefficients is universal and the coefficients are hierarchically clustered. See Section 2.2.2, for the random weights which seem to suit this picture. The pure states are claimed (at least for the SK model) to be completely characterised by the "limiting overlaps" which, in turn, are also believed to possess hierarchical structure.

Unfortunately, at the time of writing there is no general mathematical definition of the concept of limiting Gibbs measure for mean-field spin-glass systems which reflects the picture predicted in theoretical physics. However, there are proper definitions and results in several concrete cases, see the next subsection.

Remark 1.5.2 (metastates)**.** *The limiting objects called* metastates *comprehensively encode the behaviour of the limiting Gibbs measures in the case of non-mean-field (lattice) disordered systems (e.g., the EA model). See Bovier (2006); Newman & Stein (2007) and references therein for precise definitions and discussion.*

1.5.3 Higher-level objects

As emphasised, e.g., in (Ellis, 2006), study of higher-level objects such as empirical distribution functions and their large deviations is fruitful for understanding the properties of classical (non-disordered) spin systems. This ideas are closely related to the Gibbs variational principle (Proposition 1.2.1).

Large-deviations of empirical measures

We refer to Bolthausen (2007) for a suggestion of yet another higher-level object approach in the spirit of large deviations of empirical measures in the context of the perceptron model. See, e.g., Comets (1989) for some earlier results in this spirit.

Comparison schemes

Recent breakthroughs in understanding of mean-field spin-glasses can be viewed as rewards on the quest for "proper" higher-level objects. These breakthroughs were initiated by the comparison schemes of Guerra (2003) and Aizenman *et al.* (2003). See Section 2.3.3 for a short review. See Chapters 3, 4 for an extension of the comparison approach to the SK model with multidimensional spins.

Empirical overlap distribution functions

Bovier & Kurkova (2004a,b) have obtained a comprehensive description of the limiting Gibbs measure for the REM, GREM and CREM using the *empirical overlap distribution function*.

To define this object, we need the following ingredients. Let $\mathscr{Q}' \subset [0;1]^{[0;1]}$ be the set of all non-decreasing right-continuous piece-wise constant functions $x : [0;1] \to [0;1]$ with $x(0) = 0$, $x(1) = 1$ such that x has only a finite number of jumps (Figure 1.1). We will refer to the set $\mathscr{Q}'$ as the *set of discrete order parameters*. We equip it with the L^1 topology. The compactification of this space we denote by $\mathscr{Q}$ and call the *set of order parameters*. The latter space is obviously isometric to the space of all non-decreasing càdlàg functions $x : [0;1] \to [0;1]$ such that $x(0) = 0$ and $x(1) = 1$ equipped with the L^1 topology.

Now suppose that the Gibbs measure $\mathscr{G}_N(\beta)$ is induced by the GREM Hamiltonian. Consider the "random overlap distribution field" $x_N(\beta;\cdot) : \Sigma_N \times [0;1] \to [0;1]$ defined as

$$x_N(\beta;\sigma^{(1)},q) := \mathscr{G}_N(\beta)\{\sigma^{(2)} \in \Sigma_N : q_{\mathrm{L}}(\sigma^{(1)},\sigma^{(2)}) \leq q\}. \tag{1.33}$$

Note that given $\sigma^{(1)} \in \Sigma_N$, the above field (1.33) induces the random order parameter

$$x_N(\beta;\sigma^{(1)},\cdot) \in \mathscr{Q}'.$$

In words, (1.33) is the distribution of the overlap between given configuration $\sigma^{(1)} \in \Sigma_N$ and the "equilibrated" (i.e., Gibbs-distributed) configuration $\sigma^{(2)} \in \Sigma_N$. Finally, we define the *empirical overlap distribution* $\mathscr{K}_N(\beta)$ of $\mathscr{M}_1(\mathscr{X}')$ as

$$\mathscr{M}_1([0;1]) \ni \mathscr{K}_N(\beta) := \int_{\Sigma_N} \delta_{x_N(\sigma^{(1)},\cdot)} \mathscr{G}_N(\beta;\mathrm{d}\sigma^{(1)}).$$

The precise relation between the empirical object $\mathscr{K}_N(\beta)$ and the overlap distribution function (1.31) is the following

$$\mathbb{E}\left[\int_{\mathscr{X}} x\mathscr{K}_N(\beta;\mathrm{d}x)\right](q) = F_N(\beta;q).$$

It is shown by Bovier & Kurkova (2003b, 2004b) that

$$\mathscr{K}_N(\beta) \xrightarrow[N\uparrow+\infty]{w} \mathscr{K}_0(\beta), \tag{1.34}$$

where $\mathscr{K}_0(\beta)$ is the random element in $\mathscr{M}_1(\mathscr{Q}')$. Moreover, the explicit distribution of the random measure $\mathscr{K}_0(\beta)$ is provided. In particular, this random measure $\mathscr{K}_0(\beta)$ is the "countable convex combination" (with random weights) of Dirac measures. In turn, the Dirac measures are supported by the random overlap distributions corresponding to the atomic measures of the following form

$$(1-w_1)\delta_0 + (w_2 - w_1)\delta_{\bar{x}_1} + \ldots + w_{l(\beta)}\delta_{\bar{x}_{l(\beta)}},$$

where $\{w_k\}_{k=1}^{l(\beta)} \subset [0;1]$ are certain random weights ($0 \leq w_1 \leq \ldots \leq w_{l(\beta)} \leq 1$) with hierarchical organisation, $\{\bar{x}_k\}_{k=1}^{l(\beta)} \subset [0;1]$ is a certain coarse-graining of $\{x_k\}_{k=1}^{n}$, and $l(\beta) \leq n$. This result gives a precise mathematical meaning to the "convex-combination-of-the-pure-states" heuristics mentioned in Section 1.5.2. Besides, this result comprehensively displays the hierarchical "geometry" of the Gibbs measure in the thermodynamic limit. We refer to Bovier & Kurkova (2007) for the extensive presentation. Essentially, the limiting empirical overlap distribution object $\mathscr{K}_0(\beta)$ is equivalent to the Bolthausen-Sznitman coalescent see Section 2.2.3 and references therein.

Gibbs measure of the SK-like models

We refer to Talagrand (2000b, 2003) for the results on the structure of the Gibbs measure in the thermodynamic limit for Derrida's p-spin model (with large enough p) in the 1-RSB regime. See also Talagrand (2007a) for some related results and conjectures for the SK model in the FRSB regime.

1.5.4 Ultrametricity

In physical literature, it is believed that the canonical metric structure of the SK model weighted by the Gibbs measure becomes ultrametric in the thermodynamic limit. One possible way to formalise this (Talagrand, 2007b) is the following assertion about the three replicas which is formulated quite in the spirit of (1.31) as

$$\lim_{N\uparrow+\infty} \mathbb{E}\left[\mathscr{G}_N^{\otimes 3}(\beta)\left[R(\sigma^{(1)},\sigma^{(3)}) \geq R(\sigma^{(1)},\sigma^{(2)}) \wedge R(\sigma^{(2)},\sigma^{(3)}) - \varepsilon\right]\right] = 1 \tag{1.35}$$

which is required to hold for each $\varepsilon > 0$.

The (unproved) assertion (1.35) is believed to be caused by the hierarchical organisation of the pure states.

Remark 1.5.3. *The definition of ultrametric* (0.4) *implies that the metric space* (Σ, d_U) *has the following property. For any triangle with vertices* $\sigma^{(k)}$, $k \in \{1,2,3\}$, *we have that the lengths of least two of the three sides* $\mathrm{d}_U(\sigma^{(1)},\sigma^{(2)})$, $\mathrm{d}_U(\sigma^{(2)},\sigma^{(3)})$, *and* $\mathrm{d}_U(\sigma^{(1)},\sigma^{(3)})$ *are the same. Using the language of balls this means that, given two generic balls, the balls are either disjoint or one of them contains the other.*

The assertion (1.35) states roughly that the canonical L^2 distance generated by the SK model weighted with the corresponding Gibbs measure is almost ultrametric in the thermodynamic limit.

1.5.5 Free energy in the SK model: the Parisi formula

Using mathematically unrigorous heuristics of the "zero replica limit", Parisi (see, e.g., Mézard *et al.* (1987)) proposed a formula which answers Question 0.0.1 for the SK model. As we shall see below, this formula seems to be in line with the general RSB picture. Hence, the Parisi formula can be treated as an evidence of the RSB picture. This formula was rigorously proved by Talagrand (2006b) for a class of models which essentially reduces to Derrida's p-spin model with even $p \geq 2$. This proof was extended by Panchenko (2005b) to cover the case which in our nomenclature corresponds to the SK model with multidimensional spins with $d = 1$ and a priori measure with bounded support, cf. (0.9). See Section 2.3.2 for a short review of the available results in the case $d > 1$, Chapter 3 for partial results in the case $d > 1$, and Chapter 5 for extension to the case of multidimensional Gaussian spins.

Using the definitions of Section 1.5.3, we may state the following result.

Theorem 1.5.1 (the Parisi formula, Guerra (2003); Talagrand (2006b)). *Given $x \in \mathscr{Q}'$, let the function $f_x(q,y) : [0;1] \times \mathbb{R} \to \mathbb{R}$ be a unique solution of the following backward semilinear terminal value problem*

$$\begin{cases} \partial_q f(q,y) + \frac{1}{2}\left(\partial_y^2 f(q,y) + x(q)(\partial_y f)^2(q,y)\right) = 0, & q \in [0;1), y \in \mathbb{R}, \\ f(1,y) = \log\cosh\left(\beta\sqrt{2}(y+h)\right), & y \in \mathbb{R}. \end{cases} \tag{1.36}$$

Define the Parisi functional *as*

$$\mathscr{P}(\beta,x) := f_x(0,0) - \frac{\beta^2}{2}\int_0^1 qx(q)\mathrm{d}q.$$

Then, for any $\beta > 0$ and any $h \in \mathbb{R}$, the following Parisi formula *holds*

$$p(\beta,h) = \inf_{x\in\mathscr{Q}'} \mathscr{P}(\beta,x), \tag{1.37}$$

where $p(\beta,h)$ is the limiting free energy of the SK model with the Hamiltonian (1.24).

Remark 1.5.4. *It is easy to see (Guerra, 2003; Talagrand, 2006c) that the Parisi functional is Lipschitz continuous with respect to $x \in \mathscr{Q}'$ (recall that the space $\mathscr{Q}'$ is equipped with the L^1 topology). Hence, it can be extended by continuity onto the whole compact space $\mathscr{Q}$. Then the extended Parisi functional attains its infimum on $\mathscr{Q}$.*

The Parisi formula (1.37) can more intuitively be reformulated in terms of the AS^2 scheme. The scheme expresses the Parisi functional in terms of difference between the free energies of two GREM-like models, see Section 2.3.3 for details. This shows explicitly the relevance of the hierarchical structure of the GREM for the SK model.

In this thesis, we partially extend Theorem 1.5.1 to the case of the SK model with multidimensional spins, see Chapters 3, 4 and 5.

2

Often used tools and some existing results

The present chapter consists of three parts. In Section 2.1, we record some tools and techniques that are often used in the context of mean-field spin-glasses. In Section 2.2, we record some known limiting objects and their basic properties that are extensively used in the subsequent chapters. Finally, in Section 2.3, we give a short review of the known results on Gaussian mean-field spin-glasses which are closely related to or are special cases of the results obtained in this thesis.

2.1 Often used ingredients

In this section, we record for the future reference some basic results on Gaussian processes.

In what follows, whenever we speak about Gaussian random variables (or processes), we assume by default that they are centred.

We also accept the following.

Definition 2.1.1 (random process)**.** *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space. Suppose also that* $(S, \mathscr{S})$ *is a measure space. Let* I *be an abstract index space. The collection of the measurable mappings*

$$X := \{X(i) : \Omega \to S\}_{i \in I}$$

is called the random process X.

In particular, we shall be interested in Gaussian random processes indexed by high-dimensional index spaces. For a general discussion of Gaussian random processes we refer, e.g., to monographs by Adler (1990); Adler & Taylor (2007); Bogachev (1998); Dudley (1999); Fernique (1997); Ledoux & Talagrand (1991); Lifshits (1995).

2.1.1 Interpolation

This subsection is devoted to a short account of comparison results between the functionals of Gaussian processes.

Integration by parts

We begin by recalling well-known integration by parts formula which is the source of many comparison results for functionals of Gaussian processes.

Let $F : X \to \mathbb{R}$ be a functional on the linear space X. Given $x \in X$ and $e \in X$, the *directional (Gâteaux) derivative* of F at x along the direction e is

$$\partial_{x\rightsquigarrow e}F(x) := \partial_t F(x+te)\Big|_{t=0}. \tag{2.1}$$

We now ready to state the following.

Lemma 2.1.1. *Let $\{g(i)\}_{i\in I}$ be a real-valued Gaussian process (the set I is an arbitrary index set), and let h be some Gaussian random variable. Define the vector $e \in \mathbb{R}^I$ as $e(i) := \mathbb{E}[hg(i)]$, $i \in I$. Let $F : \mathbb{R}^I \to \mathbb{R}$ be such that, for all $f \in \mathbb{R}^I$, the function*

$$\mathbb{R} \ni t \mapsto F(f+te) \in \mathbb{R} \tag{2.2}$$

is either locally absolute continuous or everywhere differentiable on $\mathbb{R}$. Moreover, assume that the random variables $hF(g)$ and $\partial_{g\rightsquigarrow e}F(g)$ are in L^1.

Then

$$\mathbb{E}[hF(g)] = \mathbb{E}\left[\partial_{g\rightsquigarrow e}F(g)\right]. \tag{2.3}$$

The previous lemma coincides with (Panchenko, 2005b, Lemma 4) (modulo the differentiability condition on (2.2) and the integrability assumptions which are needed, e.g., for (Bogachev, 1998, Theorem 5.1.2)).

Corollary 2.1.1 (integration by parts for finite dimensional Gaussian vectors)**.** *Using the notations of Lemma 2.1.1, suppose that* $\operatorname{card} I < \infty$. *Assume, further, that the function $F : \mathbb{R}^I \to \mathbb{R}$ has the first derivatives of moderate growth.*

Then

$$\mathbb{E}[hF(g)] = \sum_{i\in I} \mathbb{E}[hg_i]\mathbb{E}\left[\partial_i F(g)\right]. \tag{2.4}$$

See (Talagrand, 2003, formula (A.41)).

Quadratic interpolation paths method

The following is a simple but powerful equality which is a consequence of the Gaussian integration by parts. It can be considered as a tool of comparison between two Gaussian processes that are seen through the prism of expected value of some functional of them.

Proposition 2.1.1. *Consider two independent Gaussian random vectors $X = \{X_i\}_{i=1}^n$, $Y = \{Y_i\}_{i=1}^n$ and the n-variate function $F : \mathbb{R}^n \to \mathbb{R}$. We require that F has second-order derivatives of moderate growth. Define*

$$Z(t) := \sqrt{t}X + \sqrt{1-t}Y. \tag{2.5}$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[F(Z(t))] = \frac{1}{2}\sum_{i,j=1}^{n}(\mathbb{E}[X_iX_j] - \mathbb{E}[Y_iY_j])\mathbb{E}\left[\partial^2_{i,j}F(Z(t))\right]. \tag{2.6}$$

In particular,

$$\mathbb{E}[F(X)] = \mathbb{E}[F(Y)] + \frac{1}{2}\sum_{i,j=1}^{n}\int_0^1 (\mathbb{E}[X_iX_j] - \mathbb{E}[Y_iY_j])\mathbb{E}\left[\partial^2_{i,j}F(Z(t))\right]\mathrm{d}t. \tag{2.7}$$

See, e.g., the proof of (Talagrand, 2003, Proposition 2.4.3). Analogous ideas can be found in Kahane (1986). An interesting generalisation of the previous proposition to more general *interpolation paths* than (2.5) is (Aizenman *et al.*, 2007, Corollary A.2).

Remark 2.1.1. *Definition* (2.5) *and the fact that* $Z(0) = Y$ *and* $Z(1) = X$ *may explain the name "quadratic interpolation".*

Corollary 2.1.2 (generalised Slepian's lemma)**.** *Let X and Y be two independent d-dimensional Gaussian vectors. Let* D_1 *and* D_2 *be some subsets of* $\{1,\dots,d\} \times \{1,\dots,d\}$*. Assume that*

$$\begin{cases} \mathbb{E}[X_i X_j] \le \mathbb{E}[Y_i Y_j], & (i,j) \in D_1 \\ \mathbb{E}[X_i X_j] \ge \mathbb{E}[Y_i Y_j], & (i,j) \in D_1 \\ \mathbb{E}[X_i X_j] = \mathbb{E}[Y_i Y_j], & (i,j) \notin (D_1 \cup D_2). \end{cases}$$

Let $F : \mathbb{R}^d \to \mathbb{R}$ *be the function with the second derivatives of moderate growth and such that the following holds*

$$\begin{cases} \partial^2_{i,j} F(x) \ge 0, & (i,j) \in D_1 \\ \partial^2_{i,j} F(x) \le 0, & (i,j) \in D_1. \end{cases}$$

Then

$$\mathbb{E}[F(X)] \le E[F(Y)]. \tag{2.8}$$

See (Ledoux & Talagrand, 1991, Theorem 3.11). A similar result can be extracted from Joag-Dev *et al.* (1983).

2.1.2 Concentration of measure

We refer to the monograph of Ledoux (2001) for an extensive account of results on concentration of measure. We state here a typical result of concentration of measure for i.i.d. Gaussian random variables.

Theorem 2.1.1 (Gaussian concentration measure of bound)**.** *Equip* $\mathbb{R}^n$ *with the Euclidean norm. Suppose* $F : \mathbb{R}^n \to \mathbb{R}$ *is Lipschitzian with some constant* $L > 0$*. Further, let* $\{g_i\}_{i=1}^n$ *be a family of i.i.d. standard normal random variables.*

Then, for any $t \le 0$*, we have*

$$\mathbb{P}\{|F(g) - \mathbb{E}[F(g)]| \ge t\} \le 2\exp\left(-\frac{t^2}{2L^2}\right). \tag{2.9}$$

See, e.g., (Ledoux, 2001, Corollary 2.6) for a proof.

Proposition 2.1.2 (a measure conentration bound for the SK free energy)**.** *Using the notations of Chapter 1, suppose* $\Sigma \subset \mathbb{R}$ *is a bounded set. Further, assume* $\Omega \subset \Sigma_N$*. Let* $\{H(\sigma)\}_{\sigma \in \Sigma_N}$ *be the Gaussian process with the correlation structure satisfying the following condition*

$$\sup_{\sigma^{(1)},\sigma^{(2)} \in \Sigma_N} \left| \frac{1}{N} \mathbb{E}\left[H(\sigma^{(1)}) H(\sigma^{(2)})\right] - \xi(R(\sigma^{(1)},\sigma^{(2)})) \right| \le c(N),$$

where $\xi : \mathbb{R} \to \mathbb{R}$ *is some continuous function and* $c(N) \xrightarrow[N\to+\infty]{} 0$. *Define*

$$X := \log \int_\Omega \exp\left(\beta H_N(\sigma)\right) \mathrm{d}\mu_N(\sigma).$$

Then, for any $t \geq 0$, *we have*

$$\mathbb{P}\{|X - \mathbb{E}[X]| > 2t\sqrt{LN}\} \leq 2\exp\left(-t^2\right),$$

where $L := \max\{\xi(\sigma^2) : \sigma \in \Sigma\} + c(N)$.

See (Panchenko, 2005b, Lemma 12) for a proof.

2.1.3 Superadditivity

The following is a classical theorem which is usually attributed to Fekete with an extension due to Aizenman *et al.* (2003).

Theorem 2.1.2 (superadditivity)**.** *Consider a real sequence* $\{a_n\}_{n=1}^\infty \subset \mathbb{R}$. *Suppose it is superadditive, i.e., for all* $n, m \in \mathbb{N}$, *we have*

$$a_{n+m} \geq a_n + a_n.$$

Then there exists the following limit

$$\lim_{n\to\infty} \frac{a_n}{n} =: a_0 = \sup_{n\in\mathbb{N}} \frac{a_n}{n} \in \overline{\mathbb{R}}.$$

Moreover,

$$a_0 = \inf_{n\in\mathbb{N}} \left(\sup_{n\in\mathbb{N}}\right) \frac{a_n}{n} = \lim_{m\to\infty} \frac{1}{m} \underline{\lim}_{n\to\infty} \left(\overline{\lim}_{n\to\infty}\right) \mathbb{E}[a_{n+m} - a_m]. \tag{2.10}$$

See, e.g., (Aizenman *et al.*, 2007, Lemma B.1) for a proof.

Subadditivity-based arguments were historically used to prove the existence of free energy in the thermodynamic limit, cf. Question 1.2.1. In the context of the SK model, the superadditivity in conjunction with the quadratic interpolation method was used for the first time by Guerra & Toninelli (2002).

2.2 Limiting objects

In this section, we shall shortly review several limiting objects that arise in the context of mean-field spin-glass models with Gaussian Hamiltonians.

In Sections 2.2.1 and 2.2.2, we consider limiting higher-level objects for the exponentials of the REM and GREM Hamiltonians, respectively. In Section 2.2.3, we consider a limiting higher-level object for the GREM overlap structure.

2.2.1 The Poisson-Dirichlet process

In this subsection, we shall deal with a limiting structure for the REM. It was first suggested by Ruelle (1987). The limiting structure is closely related to the *Poisson-Dirichlet process*.

Using the notation of Pitman & Yor (1997), the object we shall deal with is the Poisson point process $\mathrm{PD}(x,0)$. The generic two-parameter Poisson-Dirichlet processes $\mathrm{PD}(x,a)$ can also be obtained within framework of the REM exponentials, see Talagrand (2007a).

Definition 2.2.1 (limiting REM exponentials, Ruelle (1987))**.** *Assume* $x \in (0;1)$. *The* limiting REM exponentials $\xi(x)$ *is the Poisson point process on* $\mathbb{R}_+$ *with the following intensity density*

$$\mathbb{R}_+ \ni t \mapsto xt^{-x-1}. \tag{2.11}$$

We refer, e.g., to Daley & Vere-Jones (2003); Kallenberg (1983); Leadbetter *et al.* (1983); Resnick (1987) for the discussion of random measures and the Poisson point processes.

A limiting higher-level object

To motivate the Definition 2.2.1, we need the following linear scaling function

$$u_N(x) := \frac{x}{a_N} + b_N, \tag{2.12}$$

where

$$\begin{aligned} a_N &:= (2N\log 2)^{1/2}, \\ b_N &:= a_N - \frac{\sqrt{2}}{2a_N}\log(a_N\sqrt{2\pi}). \end{aligned}$$

Consider the *rescaled REM process*

$$\overline{\mathrm{REM}}(\sigma) := \gamma u_N^{-1}(\mathrm{REM}(\sigma)), \tag{2.13}$$

where $\gamma := (2\log 2)^{-1/2}$. The relation between the rescaled REM process (2.13) and the Poisson point process $\xi(x)$ is established through the convergence of the corresponding "higher-level object" (cf. Section 1.5.3). Namely, define the *empirical process of the REM exponentials* $\mathscr{W}_N(\beta)$ – a random element in $\mathscr{M}_{\mathrm{f}}(\mathbb{R}_+)$ – as

$$\mathscr{W}_N(\beta) := \sum_{\sigma\in\Sigma_N} \delta_{\exp\left[\beta\overline{\mathrm{REM}}(\sigma)\right]}. \tag{2.14}$$

The following "low temperature behaviour" then holds.

Theorem 2.2.1 (the limit of the REM exponentials)**.** *Suppose* $\beta > \sqrt{2\log 2}$. *Let* $x(\beta) := \sqrt{2\log 2}/\beta \in (0;1)$. *The point process* (2.14) *converges to* $\xi(x(\beta))$ *in distribution. More formally, we have*

$$\mathscr{W}_N(\beta) \xrightarrow[N\uparrow+\infty]{w} \xi(x(\beta)).$$

See, e.g., Bolthausen & Sznitman (2002); Bovier *et al.* (2002) for a proof.

By a slight abuse of notation, let $\{\xi(\alpha)\}_{\alpha\in\mathbb{N}}$ be the enumeration of all *atom positions* for a given realisation of ξ on $\mathbb{R}_+$.

Properties

The particular form of the density (2.11) implies the following result.

Theorem 2.2.2 (some basic properties of the REM exponentials)**.**

1. *For any $\varepsilon > 0$, we have*

$$\mathbb{E}\left[\operatorname{card}\{\alpha \in \mathbb{N} : \xi(\alpha) > \varepsilon\}\right] = \varepsilon^{-x}. \tag{2.15}$$

2. *Almost surely, it is possible to reorder the atoms $\{\xi(\alpha)\}_{\alpha\in\mathbb{N}} = \{\xi(i)\}_{i\in\mathbb{N}}$ such that, for all $i \in \mathbb{N}$, we have*

$$\xi(i) > \xi(i+1).$$

3. *The atom positions have the following "profile"*

$$\xi(i)i^{1/x} \xrightarrow[i\uparrow+\infty]{} 1.$$

4. *The atoms have the following integrability property. The sum $\sum_{i=1}^{\infty}\xi(i)^{v}$ converges if and only if $v > x$.*
5. *The partition function $Z := \sum_{i=1}^{\infty}\xi(i)$ is almost surely finite and has an infinitely divisible distribution satisfying $Z \sim 2^{-1/x}(Z+Z')$, where Z' is an independent copy of Z.*
6. *The v-th moment of the partition function Z is finite, i.e., $\mathbb{E}[Z^{v}] < \infty$ if and only if $u < x$.*

For the proofs, see (Aizenman *et al.*, 2007, Theorem 5.1).

Remark 2.2.1. *Theorem 2.2.2 implies that the sequence $\{\xi(\alpha)\}_{\alpha\in\mathbb{N}}$ can be treated as a random element of $\mathscr{M}_f(\mathbb{N})$.*

In the sequel, we shall assume that the random sequence $\{\xi(i)\}_{i\in\mathbb{N}}$ is decreasing.

Random permutations and their distributions

To state the next property of the limiting REM we need some additional ingredients. Let Y be the non-negative random variable independent of ξ satisfying the following moment condition

$$\lambda(Y) := \mathbb{E}[Y^{x}]^{1/x} < +\infty.$$

Denote by ν the distribution of Y. Let, further, $\{Y(i)\}_{i\in\mathbb{N}}$ be the i.i.d. copies of Y. Consider the following multiplicative deformation of the limiting REM process

$$\{\xi(i)\}_{i=1}^{\infty} \mapsto \{\xi(i)Y(i)\}_{i=1}^{\infty}. \tag{2.16}$$

Due to Theorem 2.2.2, the resulting sequence $\{\xi(i)Y(i)\}_{i\in\mathbb{N}}$ can also be reordered into the decreasing sequence $\{\tilde{\xi}(i)\}_{i\in\mathbb{N}}$, i.e.,

$$\tilde{\xi}(i) > \tilde{\xi}(i+1), i \in \mathbb{N}.$$

Let $\pi : \mathbb{N} \to \mathbb{N}$ be the corresponding random permutation, i.e., $\tilde{\xi}(i) = \xi(\pi(i))Y(\pi(i))$. Define also $\tilde{Y}(i) := Y(\pi(i))$. Then the following result holds.

Theorem 2.2.3 (the averaging property of the REM, the "quasi-stationarity")**.** *The limiting REM process satisfies*

1. The averaging (or the "quasi-stationarity") property:

$$\sum_{\alpha\in\mathbb{N}} \delta_{\tilde{\xi}(\alpha)} \sim \sum_{\alpha\in\mathbb{N}} \delta_{\lambda(Y)\xi(\alpha)}. \tag{2.17}$$

2. The reordering effect: *the sequence $\{\tilde{Y}(i)\}_{i\in\mathbb{N}}$ consists of i.i.d. random variables which are independent of $\tilde{\xi}$. Moreover, the distribution of, e.g., $\tilde{Y}(1)$ is absolute continuous with respect to ν and has the following density*

$$\mathbb{R}_+ \ni \tilde{t} \mapsto \frac{\tilde{t}^x}{\lambda(Y)}.$$

For a proof see (Ruzmaikina & Aizenman, 2005, Proposition 3.1) and (Aizenman *et al.*, 2007, Theorem 5.2).

Define the normalisation operation $\mathscr{N} : \mathscr{M}_{\mathrm{f}}(\mathbb{N}) \to \mathscr{M}_1(\mathbb{N})$ as follows

$$\mathscr{N}(\eta)\{i\} := \frac{\eta\{i\}}{\eta(\mathbb{N})}. \tag{2.18}$$

Note, that Theorem 2.2.3 implies that

$$\sum_{\alpha\in\mathscr{A}} \delta_{\mathscr{N}(\xi)\{\alpha\}} \sim \sum_{i\in\mathbb{N}} \delta_{\mathscr{N}(\tilde{\xi})\{i\}}.$$

Ruzmaikina & Aizenman (2005) show that under some mild regularity conditions the "quasi-stationarity" property (2.17) characterises the distribution of the point process $\xi(x)$ completely.

The limiting Gibbs measure

The above results are precise enough to give an answer to Question 1.3.3 in the REM case.

It is shown by Bovier (2001); Bovier & Kurkova (2004a) (see also (Talagrand, 2003, Theorem 1.2.1)) that, in particular, for $\beta > \sqrt{2\log 2}$, we have the following weak convergence

$$\sum_{\sigma\in\Sigma_N} \delta_{\mathscr{G}_N(\beta)\{\sigma\}} \xrightarrow[N\uparrow+\infty]{} \mathscr{N}(\xi(x(\beta))),$$

where $x(\beta) := \sqrt{2\log 2}/\beta$.

For $\beta \in [0;\sqrt{2\log 2})$, the situation is substantially easier. Given an arbitrary $M \in \mathbb{N}$ and $\eta \in \Sigma_M$, we have

$$\mathscr{G}_N(\beta)\{\sigma \in \Sigma_N : [\sigma]_M = \eta\} \xrightarrow[N\uparrow+\infty]{} \frac{1}{2^M},$$

almost surely, see, e.g., (Bovier, 2006, Section 9.3) for more details.

2.2.2 The Ruelle probability cascades

After the seminal work of Derrida (1985) who introduced the GREM, the corresponding limiting probability structure – probability cascades – was identified by Ruelle (1987). The RPC is the point process, which atoms possess certain hierarchical correlations.

More formally, we shall need the following index spaces. Given $n \in \mathbb{N}$, define $\mathscr{A}_0 := \emptyset$, $\mathscr{A}_k := \mathbb{N}^k$ and $\mathscr{A} := \mathscr{A}_n$. The latter index space can be seen as an (infinitely wide) tree with n-levels of hierarchy.

Define also the *set of discrete order parameters with n jumps* $\mathscr{Q}'_n \subset \mathscr{Q}'$ which consists of the functions of the following form

$$x(q) = \sum_{i=0}^{n} x_i \mathbb{1}_{[q_i;q_{i+1})}(q), \tag{2.19}$$

where, naturally,

$$\begin{aligned} 0 &=: x_0 < x_1 < \ldots < x_n < x_{n+1} := 1, \\ 0 &=: q_0 < q_1 < \ldots < q_n < q_{n+1} := 1, \end{aligned} \tag{2.20}$$

see Figure 1.1 for a sketch of x.

Definition 2.2.2 (Ruelle's probability cascade, Ruelle (1987))**.** *Given some $x \in \mathscr{Q}'_n$, let, for all $k \in [1;n] \cap \mathbb{N}$ and all $\alpha \in \mathscr{A}_k$, the point processes $\xi_{k,[\alpha]_{k-1}}$ be the independent ones with*

$$\xi_{k,[\alpha]_{k-1}} \sim \mathrm{PPP}(\mathbb{R}_+ \ni q \mapsto x_k q^{-x_k-1} \in \mathbb{R}). \tag{2.21}$$

The Ruelle probability cascade *then is the point process $\xi = \xi(x_1,\ldots,x_n)$ with the atom positions $\{\xi(\alpha)\}_{\alpha\in\mathscr{A}_n}$ defined as follows*

$$\xi(\alpha) := \prod_{k=1}^{n} \xi_{k,[\alpha]_{k-1}}(\alpha_k). \tag{2.22}$$

A limiting higher-level object

Assume that the order parameter $x \in \mathscr{Q}'_n$ is such that, for all $k = [1;n] \cap \mathbb{N}$, the following holds

$$\begin{aligned} x_k &= \frac{k}{n}, \\ \Delta q_k &> \Delta q_{k+1}. \end{aligned} \tag{2.23}$$

Define

$$\gamma_k := \left(\frac{\Delta q_k}{(2\log 2)\Delta x_k} \right)^{1/2}.$$

To motivate Definition 2.2.2, consider the following *rescaled GREM process*

$$\overline{\mathrm{GREM}}(\sigma) := \sum_{k=1}^{n} \gamma_k u^{-1}_{\Delta x_k N}(X(\sigma^{(1)},\sigma^{(2)},\ldots,\sigma^{(k)})). \tag{2.24}$$

Define also the correspondent point process of the rescaled GREM exponentials

$$\mathscr{W}_N^{(n)}(\beta) := \sum_{\sigma \in \Sigma_N} \delta_{\exp\left[\beta \overline{\mathrm{GREM}}(\sigma)\right]}. \tag{2.25}$$

The following theorem then describes the "very low temperature regime" of the limiting GREM.

Theorem 2.2.4 (the limit of the GREM exponentials, Bovier & Kurkova (2004a))**.** *Put* $\{\beta_k := \sqrt{2\log 2\Delta x_k/\Delta q_k}\}_{k=1}^n \subset \mathbb{R}_+$. *Fix an arbitrary* $\beta > \beta_n$. *For* $k \in [1;n] \cap \mathbb{N}$, *define* $x_k(\beta) := \beta_k/\beta \in (0;1)$.

Then the point process (2.25) *converges to* $\xi(x_1(\beta),\ldots,x_n(\beta))$ *in distribution. More formally, we have*

$$\mathscr{W}_N^{(n)}(\beta) \xrightarrow[N\uparrow+\infty]{w} \xi(x_1(\beta),\ldots,x_n(\beta)).$$

See (Bovier & Kurkova, 2004a, Theorem 1.7) for a proof of more elaborate result which covers also all remaining intermediate regimes $\beta \in [0;\beta_n]$.

Remark 2.2.2. *Condition* (2.23) *implies, in particular, that all jumps of x are the extreme points of the their concave hull. See Theorem 2.3.2 for similar results in the case of generic order parameters.*

Properties

Theorem 2.2.5 (some elementary properties of the RPC)**.**

1. *The partition function* $Z := \sum_{\alpha\in\mathscr{A}_n} \xi(\alpha)$ *is almost surely finite and, moreover,*

$$Z \sim Z(x_1) \prod_{k=2}^{n} \mathbb{E}\left[Z(x_k)^{x_{n-1}}\right]^{1/x_{n-1}}.$$

   In particular, we have $\mathbb{E}[\log Z] < \infty$.

2. *Almost surely, it is possible to reorder the atoms* $\{\xi(\alpha)\}_{\alpha\in\mathbb{N}} = \{\xi(i)\}_{i\in\mathbb{N}}$ *such that, for all* $i \in \mathbb{N}$, *we have*

$$\xi(i) > \xi(i+1).$$

   That is, there exists the corresponding random permutation $\pi : \mathbb{N} \to \mathscr{A}_n$ *which satisfies* $\xi(k) = \xi(\pi(k))$.

See, e.g., (Aizenman *et al.*, 2007, Theorem 5.3) for a proof.

The limiting Gibbs measure

As already mentioned above, Question 1.3.3 can completely be answered for the GREM. It is shown by (Bovier & Kurkova, 2004a, Theorem 1.9) that, in particular, for $\beta > \beta_n$, we have

$$\sum_{\sigma\in\Sigma_N} \delta_{\mathscr{G}_N(\beta)\{\sigma\}} \xrightarrow[N\uparrow+\infty]{w} \mathscr{N}(\xi(x_n(\beta))),$$

where $\xi(x_n(\beta)) \sim \mathrm{PD}(x_n(\beta),0)$. Moreover, at the level of limiting structures the normalisation (2.18) of the RPC is equidistributed with the normalisation of the Poisson-Dirichlet process. More precisely,

$$\mathcal{N}(\xi(x_1,\ldots,x_n)) \sim \mathcal{N}(\xi(x_n)),$$

where $\xi(x_n) \sim \mathrm{PD}(x_n, 0)$. See, e.g., (Bolthausen & Sznitman, 1998, Lemma 2.1) for a proof. Hence, interestingly, the hierarchical RPC structure flattens down to the Poisson-Dirichlet structure, if the point process of limiting Gibbs weights is concerned.

The hierarchical structure is, however, important for encoding the RSB picture (cf. Section 1.5). It is this structure which is captured by the limiting object introduced in the following section.

2.2.3 The Bolthausen-Sznitman coalescent and random permutations

A deep insight of Bolthausen & Sznitman (1998) was to identify the dynamical system generated by the RPC with remarkable properties which allow to perform many spin-glass calculations in a more transparent way. In what follows, we shall use this dynamical system in computations involving the replicas of the system (see Chapter 4).

A contribution of Bovier & Kurkova (2004a,b) was, in particular, to relate the RPC, the *continuous time coalescent* process of Bolthausen & Sznitman (1998), and its dual – the *continuous time branching processes* of Bertoin & Le Gall (2000) – with the original setup of Derrida (1985).

Definition

Let $\xi = \xi(x_1,\ldots,x_n)$ be the RPC process. Theorem 2.2.5 guarantees that there exists the rearrangement $\xi = \{\xi(i)\}_{i\in\mathbb{N}}$ of ξ's atoms in a decreasing order. Recall (0.2) and define the (random) *limiting ultrametric overlap* $q_{\mathrm{L}} : \mathbb{N}^2 \to [0;n] \cap \mathbb{Z}$ as

$$q_{\mathrm{L}}(i,j) := \max\{k \in [0;n] \cap \mathbb{Z} : [\pi(i)]_k = [\pi(j)]_k\}, \tag{2.26}$$

where we use the convention that $\max \emptyset = 0$. This overlap valuation (2.26) induces the sequence of *random partitions* of $\mathbb{N}$ into *equivalence classes*. Namely, given a $k \in \mathbb{N} \cap [0;n]$, we define, for any $i, j \in \mathbb{N}$, the *Bolthausen-Sznitman equivalence relation* as follows

$$i \underset{k}{\sim} j \overset{\text{def}}{\Longleftrightarrow} q_{\mathrm{L}}(i,j) \geq k. \tag{2.27}$$

Note the conceptual similarity between the equivalence (2.27) and the limiting empirical overlap distribution $\mathcal{K}_0(\beta)$ from (1.34).

In what follows, we shall, by a slight abuse of the language, not distinguish between the concepts of the *equivalence relation* and the partition of a set into the *equivalence classes*.

Partitions induced by equivalence (2.27) have a nice property that the partitions with smaller indices k are the coarsenings of the partitions with the larger k's. Moreover, these random partitions are induced by a *Markovian structure* and can be seen as the states of the *continuous-time coalescent process* at certain deterministic times introduced by Bolthausen & Sznitman (1998).

Indeed, for an arbitrary index set I, let $\mathcal{E}(I)$ be the set of all equivalence relations on I. The sets $\mathcal{E}(I)$ are compact, if seen as the subsets of the product spaces $\{0,1\}^{I\times I}$ (we equip the set $\{0,1\}$ with the discrete topology). Finally, let Γ be the Markov process

$$\Gamma := \{\Gamma(t) : \Omega \to \mathcal{E}(\mathbb{N})\}_{t\in\mathbb{R}_+}$$

with the following properties.

1. The *initial state* is $\Gamma(0) = \Pi_0$, where $\Pi_0 \in \mathscr{E}(\mathbb{N})$ is the partition of $\mathbb{N}$ into singleton sets.
2. *Coalescence property*: for any $s,t \in \mathbb{R}_+$, whenever $s < t$, we have $\Gamma(t) \preceq \Gamma(s)$, i.e., the partition $\Gamma(s)$ is not finer than $\Gamma(t)$.
3. *Finite-dimensional traces*: for any $I \Subset \mathbb{N}$, let Γ_I be the trace of the process Γ on the set $\mathscr{E}(I)$. Let $a_I := \{a_I(\Pi,\Pi')\}_{\Pi,\Pi' \in \mathscr{E}(I)}$ be the *transition rate matrix* (Q-matrix) of the Markov process $\Gamma(I)$ defined as follows. For $\Pi,\Pi' \in \mathscr{E}(I)$, denoting $N := \operatorname{card} I$, define

$$a_I(\Pi,\Pi') := \begin{cases} \left((N-2)\binom{N-2}{k-2}\right)^{-1}, & \Pi' \text{ is obtained by gluing } k \in [2;N] \cap \mathbb{N} \text{ classes} \\ & \text{of } \Pi \text{ together;} \\ -\sum_{E \neq E'} a_I(E,E'), & \Pi = \Pi'; \\ 0, & \text{otherwise.} \end{cases}$$

Definition 2.2.3 (the Bolthausen-Sznitman coalescent)**.** *As is shown in (Bolthausen & Sznitman, 1998, Theorem 1.2), the above properties uniquely identify the distribution of the continuous-time pure-jump Markov process Γ which is then called the* Bolthausen-Sznitman coalescent*.*

See Figure 2.1 for a realisation of the trace of the Bolthausen-Sznitman coalescent on a finite set.

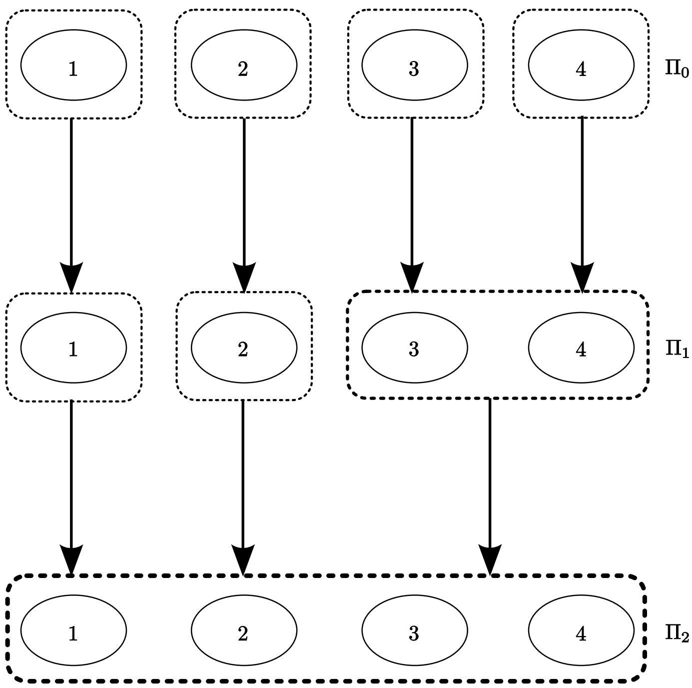

Fig. 2.1. A realisation of the trace of the Bolthausen-Sznitman coalescent on $I := \{1,2,3,4\}$

Remark 2.2.3. *An explicit representation of the Markov semigroup of the Bolhausen-Sznitman coalescent is certainly also available, see (Bolthausen & Sznitman, 1998, Proposition 1.4).*

Properties

For $k \in [0;n] \cap \mathbb{N}$, let the partition $\Pi_k \in \mathscr{E}(\mathbb{N})$ be the partition generated by the Bolthausen-Sznitman equivalence "$\underset{k}{\sim}$". The following theorem motivates Definition 2.2.3 by relating the distributions of the limiting ultrametric overlap distribution function with the Bolthausen-Sznitman coalescent.

Theorem 2.2.6 (limiting ultrametric overlap and the Bolthausen-Sznitman coalescent)**.**

1. *The law of the "limiting ultrametric overlaps" vector* $(\Pi_{n-1}, \ldots, \Pi_1)$ *coincides with the law of the following Bolthausen-Sznitman vector*

$$\left(\Gamma(t_1), \Gamma(t_2), \ldots, \Gamma(t_{n-1})\right),$$

where, for $k \in [1; n-1] \cap \mathbb{N}$

$$t_k := \log\left(\frac{x_n}{x_k}\right) \in \mathbb{R}_+.$$

2. *The coalescent* Γ *and the point process* $\mathscr{N}(\xi(x_1, \ldots, x_n))$ *are independent.*

See (Bolthausen & Sznitman, 1998, Theorem 2.2) for a proof.

The limiting overlaps and certain random permutations

So far, in the present chapter, we were dealing with an important ingredient of the RSB picture, namely, with the scaling limits of the (unnormalised) GREM exponentials at the level of point processes. This lead us to the RPC. The second important ingredient of the Parisi RSB picture are the overlaps (see Section 1.4.5).

Recall (0.2). Given $x \in \mathscr{Q}_n'$, define the *limiting GREM overlap* $q : \mathscr{A}^2 \to [0;1]$ as

$$q(\alpha^1, \alpha^2) := q_{q_{\mathrm{L}}(\alpha^1, \alpha^2)}. \tag{2.28}$$

We shall need also the *randomised limiting GREM overlap* $q : \mathbb{N}^2 \to [0;1]$ as

$$q(i,j) := q_{q_{\mathrm{L}}(i,j)}.$$

Note that the above definitions allow also for another representation of the reordered limiting GREM overlap

$$q(i,j) = q(\pi(i), \pi(j)).$$

Finally, define the *filtered limiting GREM process* $\{Y(\alpha, t) : \alpha \in \mathscr{A}, t \in \mathbb{R}_+\}$ as follows

$$Y(\alpha; t) := \sum_{k=0}^{n} \mathbb{1}_{[q_k;1]}(t) W_k([\alpha]_k; t \wedge q_{k+1} - q_k), \tag{2.29}$$

where $\Delta q_k := q_{k+1} - q_k$ and $\{\{W_k([\alpha]_k; t)\}_{t \in \mathbb{R}_+}\}_{k \in [0;n] \cap \mathbb{N}, \alpha \in \mathscr{A}}$ is the family of independent (for different indices α and k) standard 1-D Wiener processes. Definition (2.29) readily implies

$$\mathrm{Cov}\left[Y(\alpha^{(1)}, t), Y(\alpha^{(2)}, s)\right] = t \wedge s \wedge q(\alpha^{(1)}, \alpha^{(2)}).$$

Now, similarly to (2.16) with the help of the filtered limiting GREM we reweight the initial RPC weights ξ as

$$\tilde{\xi}(\alpha;t) := \xi(\alpha)\exp\left(f(t,Y(\alpha,t))\right), \tag{2.30}$$

where we assume, for simplicity, that $f : [0;1] \times \mathbb{R} \to \mathbb{R}$ is such that, for any $t \in [0;1]$, $f(t,\cdot) \in C^{(2)}(\mathbb{R})$, and any $c > 0$ we have $\int_{\mathbb{R}} \exp\left(f(t,x) - cx^2\right) \mathrm{d}x < \infty$ and, moreover, that

$$\sup_{x\in\mathbb{R}} \left(|\partial_x f(t,x)| + |\partial^2_{xx} f(t,x)|\right) < +\infty.$$

Further, the point process defined in (2.30) can also be reordered in a decreasing way, i.e., there exists the mapping $\tilde{\pi}_t : \mathbb{N} \to \mathscr{A}$ such that, for all $i \in \mathbb{N}$, the following holds

$$\tilde{\xi}(\tilde{\pi}_t(i);t) > \tilde{\xi}(\tilde{\pi}_t(i+1);t). \tag{2.31}$$

In what follows, we shall use the short-hand notations $\tilde{\xi}(i;t) := \tilde{\xi}(\tilde{\pi}_t(i);t)$, $\tilde{Y}(i,t) := Y(\tilde{\pi}_t(i),t)$ and $\tilde{q}_t := \{\tilde{q}_t(i,j) := q(\tilde{\pi}_t(i),\tilde{\pi}_t(j))\}_{i,j\in\mathbb{N}}$.

Theorem 2.2.7 (some limiting GREM properties)**.** *Given a discrete order parameter $x \in \mathscr{Q}'_n$, we have*

1. Independence #1: *the normalised RPC point process $\mathscr{N}(\xi)$ is independent from the corresponding randomised limiting GREM overlaps q.*
2. Independence #2: *the reordered filtered limiting GREM $\tilde{Y}$ is independent from the corresponding reordered weights $\tilde{\xi}$.*
3. The change of measure: *given $I \Subset \mathbb{N}$, let $\mu_I(\cdot|q)$ be the conditional distribution of $\{Y(i,1)\}_{i\in I}$, and $\tilde{\mu}_I(\cdot|q)$ be the conditional distribution of $\{\tilde{Y}(i,1)\}_{i\in I}$ both conditional on q. Then*

$$\frac{\mathrm{d}\tilde{\mu}_I(\cdot|q)}{\mathrm{d}\mu_I(\cdot|q)} = \prod_{k=0}^{n} \prod_{i\in\left(I/\underset{k}{\sim}\right)} \exp\left(x_k\left(f(x_{k+1},Y(i,x_{k+1})) - f(x_k,Y(i,x_k))\right)\right),$$

where the innermost product is taken over all equivalence classes on the index set I induced by the equivalence $\underset{k}{\sim}$.

4. The averaging property: *the function $f : [0;1] \times \mathbb{R} \to \mathbb{R}$ satisfies* (1.36)*, if and only if*

$$\left(\{\xi(\alpha)\exp\left(f(t,Y(\alpha;t))\right)\}_{\alpha\in\mathscr{A}}, \tilde{q}_t\right) \sim \left(\{\xi(\alpha)\exp\left(f(s,Y(\alpha;s))\right)\}_{\alpha\in\mathscr{A}}, \tilde{q}_s\right),$$

for all $s,t \in [0;1]$.

See, e.g., Arguin (2007); Bolthausen & Sznitman (1998) for a proof.

Proposition 2.2.1 (expected mutial overlap distribution, Bolthausen & Sznitman (1998); Ruelle (1987))**.** *For any $k \in [1;n+1] \cap \mathbb{N}$, we have*

$$\mathbb{E}\left[\mathscr{N}(\xi) \otimes \mathscr{N}(\xi)\left\{(\alpha^{(1)},\alpha^{(2)}) \in \mathscr{A}^2 : q_L(\alpha^1,\alpha^2) \leq k\right\}\right] = x_k.$$

See Bolthausen & Sznitman (1998); Ruelle (1987) for a proof.

Some applications

We would like to point out that the Poisson-Dirichlet distribution and Bolthausen-Sznitman coalescents appear in various applied contexts: in combinatorial stochastic structures, see, e.g., Arratia *et al.* (2003); Basdevant (2006); Bertoin (2006); Pitman (2006), in statistics, see, e.g., Teh *et al.* (2006). The structures are recently proved to emerge as equilibrium distributions for certain systems of interacting diffusions, see, e.g., Chatterjee & Pal (2007).

2.3 Some results on Gaussian mean-field spin-glasses

2.3.1 The GREM

Recall the definition of the GREM from Section 1.4.3 and the decomposition (1.30) in particular. Let $\rho \in \mathscr{Q}'_n$. Assume that the process $\{\mathrm{GREM}(\sigma) : \sigma \in \Sigma_N\}$ has ρ as its order parameter. Assume that ρ satisfies (1.27). Define, for $j,k \in [1;n+1]\cap\mathbb{N}$, $j<k$, the "slopes" $A_{j,k}$ of the discrete order parameter x as

$$A_{j,k} := \frac{q_k - q_{j-1}}{2\log 2(x_k - x_{j-1})}.$$

Define, further, the increasing sequence of indices $\{J_l\}_{l=0}^m \subset [0;n+1]\cap\mathbb{N}$ as follows. Start from $J_0 := 0$, and define iteratively

$$J_l := \min\left\{ J \in [J_{l-1};n+1]\cap\mathbb{N} : A_{J_{l-1},J} > A_{J+1,k}, \text{ for all } k > J \right\}.$$

This subsequence of indices induces the following coarse-graining of the initial GREM. Define the following (possibly) "coarse-grained" parameters

$$\begin{aligned} \bar{q}_l &:= q_{J_l} - q_{J_{l-1}}, \\ \bar{x}_l &:= x_{J_l} - x_{J_{l-1}}, \\ \bar{\gamma}_l &:= \sqrt{\frac{\bar{q}_l}{(2\log 2)\bar{x}_l}}. \end{aligned}$$

Recall (2.12) and define the GREM scaling function $u_{J,N} : \mathbb{R} \to \mathbb{R}$ as

$$u_{J,N}(x) := \sum_{l=1}^{m} \left(\sqrt{(2\log 2)N\bar{x}_l\bar{q}_l} - \frac{1}{2\sqrt{N}}\bar{\gamma}_l \log\left((2\log 2)\bar{x}_l\right) \right) + \frac{x}{N^{-1/2}}.$$

Define the rescaled GREM process as

$$\overline{\mathrm{GREM}}(\sigma) := u_{J,N}^{-1}(\mathrm{GREM}(\sigma)). \tag{2.32}$$

Remark 2.3.1. *Note that* (2.32) *is compatible with* (2.24).

Define the point process of the rescaled GREM energies $\mathscr{E}_N$ as

$$\mathscr{E}_N := \sum_{\sigma\in\Sigma_N} \delta_{\overline{\mathrm{GREM}}(\sigma)}.$$

The limiting object we are about to define now is a "logarithm" of the RPC point process. Given $K \in \mathbb{R}_+$, assume that the point process $\mathscr{P}^{(1)}(K)$ on $\mathbb{R}$ satisfies

$$\mathscr{P}^{(1)}(K) \sim \mathrm{PPP}\left(K \exp(-x)\mathrm{d}x\right).$$

Consider the following collection of independent point processes

$$\{\mathscr{P}^{(k)}_{\alpha_1,\ldots,\alpha_{l-1}}(K) \mid \alpha_1,\ldots,\alpha_{l-1} \in \mathbb{N}; l \in [1;m] \cap \mathbb{N}\}$$

such that

$$\mathscr{P}^{(k)}_{\alpha_1,\ldots,\alpha_{k-1}}(K) \sim \mathscr{P}^{(1)}(K).$$

Given $\underline{K} \in \mathbb{R}^m_+$, define the *limiting GREM point process* $\mathscr{P}_m(K)$ on $\mathbb{R}^m$ as follows

$$\mathscr{P}_m(\underline{K}) := \sum_{\alpha \in \mathscr{A}_m} \delta_{\left(\mathscr{P}^{(1)}(K_1;\alpha_1), \mathscr{P}^{(2)}_{\alpha_1}(K_2;\alpha_2), \ldots, \mathscr{P}^{(m)}_{\alpha_1,\alpha_2,\ldots,\alpha_{m-1}}(K_m;\alpha_m)\right)}.$$

Now we are ready to state the following generalisation of Theorem 2.2.4.

Theorem 2.3.1 (the limit of the GREM point process)**.** *Assume that the "slopes" $\{\gamma_l\}_l$ form a decreasing sequence, i.e., for all $l \in [1;m] \cap \mathbb{N}$, we have*

$$\gamma_l > \gamma_{l+1}.$$

Then there exists some $\underline{K} = \underline{K}(x) \in (0;1]^m$ such that

$$\mathscr{E}_N \xrightarrow[N\uparrow\infty]{} \int_{\mathbb{R}^m} \delta_{\gamma_1 e_1 + \ldots + \gamma_m e_m} \mathscr{P}_m(\underline{K}; \mathrm{d}e_1, \ldots, \mathrm{d}e_m).$$

Moreover, $K_l = 1$ if and only if, for all $x \in (x_{l-1}; x_l)$, we have

$$[0;1]^2 \ni (x, \rho(x)) \notin \partial \operatorname{conv} \Gamma(\rho),$$

where $\Gamma(\rho)$ is the sub-graph of the function ρ, i.e.,

$$\Gamma(\rho) := \left\{(x', \rho(x)) \in [0;1]^2 : x' \leq x, x \in [0;1]\right\}.$$

See (Bovier & Kurkova, 2004a, Theorem 1.5) for a proof.

Remark 2.3.2. *Theorem 2.2.4 is a simple particular case of Theorem 2.3.1.*

Theorem 2.3.1 allows for a complete characterisation of the limiting distribution of the GREM's partition function. To formulate the result, we need the β-dependent threshold $l(\beta) \in [0;m] \cap \mathbb{N}$ such that above this threshold ($l > l(\beta)$) all coarse grained levels l of the limiting GREM are in a "frozen state" ("low temperature regime"). Below this threshold ($l \leq l(\beta)$) the levels are in the "high temperature regime". Namely, define

$$l(\beta) := \max\{l \in [1;m] \cap \mathbb{N} : \gamma_l \beta > 1\},$$

and let $l(\beta) := 0$, if $\bar{\gamma}_1 \beta \leq 1$.

Theorem 2.3.2 (the limiting distribution of the GREM partition function). *Given $\beta \in \mathbb{R}_+$, there exists the constant $C(\beta;x) \in (0;1]$ such that one of the following two cases holds.*

1. *If $l(\beta) = 0$, then*

$$\frac{Z_N(\beta)}{2^N \exp(\beta^2 N/2)} \xrightarrow[N\uparrow+\infty]{w} C(\beta,x).$$

2. *If $l(\beta) > 0$, then*

$$\exp\left(\sum_{l=1}^{l(\beta)} \left(-\beta N\sqrt{(2\log 2)\bar{q}_l\bar{x}_l} + \beta\bar{\gamma}_l \log((8\log 2)\pi N\bar{x}_l)\right)\right.$$
$$\left. -N\sum_{k=J_{l(\beta)}+1} \left(\beta^2 q_k/2 + (2\log 2)\bar{x}_l\right)\right) Z_N(\beta) \xrightarrow[N\uparrow+\infty]{w}$$
$$C(\beta,x)\int_{\mathbb{R}^{l(\beta)}} \exp\left[\beta\left(\bar{\gamma}_1 x_1 + \bar{\gamma}_2 x_2 + \ldots + \bar{\gamma}_{l(\beta)} x_{l(\beta)}\right)\right]$$
$$\times \mathscr{P}_{l(\beta)}(\underline{K}(x); \mathrm{d}e_1, \ldots, \mathrm{d}e_{l(\beta)}).$$

Moreover, $C(\beta;x) = 1$, if and only if $\beta\bar{\gamma}_{l(\beta)+1} < 1$.

See (Bovier & Kurkova, 2004a, Theorem 1.7) for a proof.

2.3.2 The SK model with multidimensional spins

Mean-field spin-glass models with multidimensional *(Heisenberg)* spins were considered in the theoretical physics literature, see, e.g., Sherrington (2007) and references therein. Rigorous results are, however, rather scarce. An early example is the result of Fröhlich & Zegarliński (1987), where the bounds on the free energy in the high temperature regime are obtained. Methods of stochastic analysis and large deviations were used by Toubol (1998) to identify the limiting distribution of the partition function and also to obtain some information about the geometry of the Gibbs measure for small β. More recent treatments of the high temperature regime using very different methods are due to Talagrand (2000a), see also (Talagrand, 2003, Section 2.13). The importance of the SK model with multidimensional spins for understanding the ultrametricity of the original model of Sherrington & Kirkpatrick (1975) (which corresponds to $d = 1$ and μ being the Rademacher measure in the above notations) was emphasised by Talagrand (2007b).

The most advanced recent study of spin-glass models with multidimensional spins was attempted by Panchenko & Talagrand (2007b), where the multidimensional spherical spin-glass model was considered. The authors combined the techniques of Panchenko (2005b); Talagrand (2006b) to obtain partial results on the ultrametricity and also get some information on the local free energy for their model.

Multidimensional spins with compact support

Equip $\mathbb{R}^d$ with the Euclidean norm. Assume $\Sigma := B(0,\sqrt{d}) \subset \mathbb{R}^d$. Given $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, define the matrix $R(\sigma^{(1)}, \sigma^{(2)}) \in \mathbb{R}^{d\times d}$ with the following entries

$$R(\sigma^{(1)}, \sigma^{(2)})_{u,v} := \frac{1}{N}\sum_{i=1} \sigma^{(1)}_{i,u}\sigma^{(2)}_{i,v}.$$

Theorem 2.3.3 (existence of the RS order parameter)**.** *There exists $L > 0$ such that if $L\beta d \leq 1$, then there exist matrices $Q^{(1)}, Q^{(2)} \in \mathbb{R}^{d\times d}$ such that*

$$\mathbb{E}\left[\mathscr{G}_N(\beta)\left[\|R(\sigma,\sigma)-Q^{(2)}\|_2^2\right]\right] \leq \frac{K(d)}{N},$$
$$\mathbb{E}\left[\mathscr{G}_N(\beta)\otimes\mathscr{G}_N(\beta)\left[\|R(\sigma^{(1)},\sigma^{(2)})-Q^{(1)}\|_2^2\right]\right] \leq \frac{K(d)}{N}.$$

See (Talagrand, 2003, Theorem 2.13.1) for a proof.

The following theorem provides additional information on the matrices $Q^{(1)}, Q^{(2)}$. Assume $Q^{(2)}$ is a symmetric non-negative definite matrix. There exists the Gaussian family $\{Y_u\}_{u=1}^d$ with covariance

$$\mathbb{E}\left[Y_u Y_v\right] = \beta^2 Q^{(2)}_{u,v}.$$

Define the function $\mathscr{E} : \mathbb{R}^d \to \mathbb{R}$ as follows

$$\mathscr{E}(x) := \exp\left(\sqrt{2}\langle x, Y\rangle + \beta^2\langle (Q^{(2)}-Q^{(1)})x, x\rangle\right).$$

Theorem 2.3.4 (representation of the RS order parameters)**.** *Assume that $\beta L d \leq 1$. We have*

$$Q^{(1)}_{u,v} = \mathbb{E}\left[\frac{1}{Z^2}\int x_u \mathscr{E}(x)\mathrm{d}\mu(x)\int x_v \mathscr{E}\mathrm{d}\mu(x)\right], \tag{2.33}$$
$$Q^{(2)}_{u,v} = \mathbb{E}\left[\frac{1}{Z}\int x_u x_v \mathscr{E}(x)\mathrm{d}\mu(x)\right], \tag{2.34}$$

where

$$Z := \int \mathscr{E}(x)\mathrm{d}\mu(x).$$

Moreover,

$$p(\beta) = \mathbb{E}\left[\log\int_\Sigma \mathscr{E}(x)\mathrm{d}\mu(x)\right] - \frac{\beta^2}{2}\left(\|Q^{(2)}\|_2^2 - \|Q^{(1)}\|_2^2\right).$$

See (Talagrand, 2003, Theorems 2.13.2 and 2.13.3) for proofs.

One-dimensional spins with compact support

Let $d = 1$. Given $u \in \mathbb{R}_+$, let us generalise slightly the objects of Section 1.5.3, namely let $\mathscr{Q}'(u) \subset [0;1]^{[0;u]}$ be the set of all non-decreasing right-continuous piece-wise constant functions $x : [0;u] \to [0;1]$ with $x(0) = 0$, $x(u) = 1$, such that they have only a finite number of jumps. Let also $\mathscr{Q}'_n(u) \subset \mathscr{Q}'(u)$ be the subset of order parameters with exactly $n \in \mathbb{N}$ jumps.

Assume the real sequence $\{c_N\}_{N\in\mathbb{N}}$ satisfies $c_N \xrightarrow[N\uparrow+\infty]{} 0$. Let $\{H_N(\sigma)\}_{\sigma\in\Sigma_N}$ be the family of Gaussian processes such that, for all $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, we have

$$\left|\frac{1}{N}\mathbb{E}\left[H_N(\sigma^{(1)})H_N(\sigma^{(2)})\right] - \xi\left(R(\sigma^{(1)},\sigma^{(2)})\right)\right| \leq c_N, \tag{2.35}$$

where the function $\xi : \mathbb{R} \to \mathbb{R}$, for all $x > 0$, satisfies the following (symmetry and convexity) conditions

$$\xi(0) = 0, \xi(x) = \xi(-x), \xi''(x) > 0.$$

A particular example of the Hamiltonian satisfying (2.35) is Derrida's p-spin interaction Hamiltonian (cf. (1.15), $p \in \mathbb{N}$)

$$H_{N,p}(\sigma) := p^{-1/2} N^{-(p-1)/2} \sum_{i_1,\ldots,i_p=1}^{N} g_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \tag{2.36}$$

where $\{g_{i_1,\ldots,i_p}\}_{i_1,\ldots,i_p=1}^{N}$ are i.i.d. standard normal random variables.

Let $[d;D] \subset \mathbb{R}$ be the smallest interval such that

$$\mu\left\{\sigma^2 \in [d;D]\right\} = 1.$$

Assume $\{\varepsilon_N\}_{N=1}^{\infty} \subset \mathbb{R}_+$ is such that $\varepsilon_N \xrightarrow[N\uparrow+\infty]{} 0$. Define the set of all configurations with approximately the same self-overlap

$$U_N(u) := \left\{\sigma \in \Sigma_N : |R(\sigma,\sigma) - u| \leq \varepsilon_N\right\}.$$

Define the corresponding *local free energy*

$$p_N(u, \varepsilon_N) := \frac{1}{N} \mathbb{E}\left[\log Z_N(u, \varepsilon_N)\right],$$

where

$$Z_N(u, \varepsilon_N) := \int_{U_N(u)} \exp\left(H_N(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma).$$

Given $u \in [d;D]$, consider an arbitrary $x \in \mathscr{Q}_n'(u)$ and $\lambda \in \mathbb{R}$. Consider, further, the family of independent Gaussian random variables $\{z_k\}_{k=0}^{n}$ such that

$$\mathrm{Var}\left[z_k\right] = \xi'(q_{k+1}) - \xi'(q_k).$$

Define the sequence of functionals $\{X_k : \mathscr{Q}_n'(u) \times \mathbb{R} \to \mathbb{R}\}_{k=0}^{n+1}$ iteratively as follows. Start from

$$X_{n+1}(x, \lambda) := \log \int_{\Sigma} \exp\left(\sigma \sum_{k=0}^{n} z_n + \lambda \sigma^2\right) \mathrm{d}\mu(\sigma),$$

and continue, for $k \in \{n, n-1, \ldots, 0\}$, recursively

$$X_k(x, \lambda) := \frac{1}{x_k} \log \mathbb{E}\left[\exp\left(x_k X_{k+1}(x, \lambda)\right)\right].$$

Note that the functional $X_0(x, \lambda)$ is deterministic. Define the functional $\mathscr{P} : \mathscr{Q}_n'(u) \times \mathbb{R} \to \mathbb{R}$ as follows

$$\mathscr{P}(x,\lambda) := -\lambda u + X_0(x,\lambda) - \frac{1}{2}\sum_{k=1}^{n} x_k \left(\theta(q_{k+1}) - \theta(q_k)\right),$$

where

$$\theta(q) := q\xi'(q) - \xi(q). \tag{2.37}$$

Define the *local Parisi free energy* as

$$\mathscr{P}(\xi,u) := \inf_{(x,\lambda)\in\mathscr{Q}'(u)\times\mathbb{R}} \mathscr{P}(x,\lambda).$$

Finally, consider the (global) *Parisi free energy*

$$\mathscr{P}(\xi) := \sup_{u\in[d;D]} \mathscr{P}(\xi,u).$$

Theorem 2.3.5 (the saddle point Parisi formula)**.** *Given $u \in [d;D]$, there exists the vanishing sequence $\{\varepsilon_N\}_{N\in\mathbb{N}}$ such that*

$$\lim_{N\to+\infty} p_N(u,\varepsilon_N) = \mathscr{P}(\xi,u).$$

Moreover,

$$p(\beta) = \mathscr{P}(\xi).$$

See (Panchenko, 2005b, Theorems 1 and 2) for proofs.

Multiple spherical spin-glass models

In this paragraph, deviating from the rest of the thesis, we shall consider the model with spins $\sigma \in \Sigma_N$, where Σ_N is a non-product space (namely, a suitably chosen Euclidean sphere). Consider the Gaussian Hamiltonian of the spherical spin-glass model[1] $\{H_N(\sigma)\}_{\sigma\in S(0,\sqrt{N})}$ satisfying condition (2.35). Recall that (2.36) is a particular example of this situation. Define the *free energy of the spherical spin-glass model* as follows

$$p_N(\beta,h) := \frac{1}{N}\mathbb{E}\left[\log\int_{S(0,\sqrt{N})} \exp\left(\beta H_N(\sigma) + h\sum_{i=1}^{N}\sigma_i\right) \mathrm{d}\lambda_N(\sigma)\right], \tag{2.38}$$

where $\lambda_N \in \mathscr{M}_1\left(S(0,\sqrt{N})\right)$ is the uniform distribution on $S(0,\sqrt{N})$.

Consider $x \in \mathscr{Q}'_n$, where $\mathscr{Q}'_n$ is defined in (1.26). The limiting free energy for (2.38) was computed by Crisanti & Sommers (1992) and then the computation was made rigorous by Talagrand (2006a) using the methods of Talagrand (2006b).

Given a real parameter $b > 1$, define the sequences $\{d_l\}_{l=1}^n$ and $\{D_l\}_{l=1}^n$ as

$$d_k := \sum_{k=l}^{n} x_k \left(\xi'(q_{k+1}) - \xi'(q_k)\right),$$
$$D_l := b - d_l.$$

[1] Sometimes also referred to as “the spherical SK model”

Define the *Parisi functional* $\mathscr{P}(\beta,h) : \mathscr{Q}_n' \times (1;+\infty) \to \mathbb{R}$ for the spherical spin-glass model as

$$\mathscr{P}(\beta,h;x,b) := \frac{1}{2}\left(b - 1 - \log b + \frac{1}{D_1}(h^2 + \xi'(q_1)) + \sum_{k=1}^{n} \frac{1}{x_l} \log \frac{D_{l+1}}{D_l}\right.$$
$$\left. - \sum_{k=1}^{n} x_l \left(\theta(q_{l+1}) - \theta(q_l)\right)\right).$$

(See (2.37) for the definition of $\theta : [0;1] \to \mathbb{R}$.) Also let us introduce the *Crisanti-Sommers functional* $\mathscr{C}\mathscr{S}(\beta,h) : \mathscr{Q}_n' \to \mathbb{R}$ as follows. Set $\delta_k := \sum_{l=k}^{n} x_l(q_{l+1} - q_l)$ and define

$$\mathscr{C}\mathscr{S}(\beta,h;x) := \frac{1}{2}\left(h^2\delta_1 + \frac{q_1}{\delta_1} + \sum_{k=1}^{n-1} \frac{1}{x_k} \log \frac{\delta_k}{\delta_{k+1}} + \log \delta_n + \sum_{k=1}^{n} x_k(\xi(q_{k+1}) - \xi(q_k))\right).$$

The following result is obtained by Talagrand (2006a).

Theorem 2.3.6 (the Parisi formula for the spherical spin-glass model)**.** *We have*

$$p(\beta,h) = \mathscr{P}(\beta,h) := \inf_{x\in\mathscr{Q}',b>1} \mathscr{P}(\beta,h;x,b)$$
$$= \inf_{x\in\mathscr{Q}'} \mathscr{C}\mathscr{S}(\beta,h;x).$$

See Talagrand (2006a) for a proof.

Now, we are ready to consider multiple copies of the spherical model defined above. Let $Q \in \mathrm{Sym}(d)$ be the non-negative definite matrix with elements $Q_{u,v} \in [-1;1]$ and $Q_{u,u} = 1$, for all $u,v \in [1;d]\cap\mathbb{N}$. Given $\varepsilon > 0$, consider the following set of configurations having almost the same self-overlap given by Q, namely

$$U(\varepsilon) := \left\{\sigma \in S(0,\sqrt{N})^d : |R(\sigma,\sigma)_{u,v} - Q_{u,v}| \le \varepsilon, \text{ for all } u,v \in [1;d]\cap\mathbb{N}\right\}.$$

Consider $\beta_1,\dots,\beta_d > 0$ and $h_1,\dots,h_d \in \mathbb{R}$. Denote $\underline{\beta} := \{\beta_u\}_{u=1}^d$ and $\underline{h} := \{h_u\}_{u=1}^d$. Define the *local free energy* on $U(\varepsilon)$ of the d copies of the spherical model as follows

$$p_N^{(d)}(\underline{\beta},\underline{h},U(\varepsilon)) := \frac{1}{N}\mathbb{E}\left[\int_{U(\varepsilon)} \exp\left(\sum_{u=1}^{d} \beta_u H_N(\sigma_{\cdot,u}) + \sum_{u=1}^{d}\left(h_u \sum_{u=1}^{N} \sigma_{i,u}\right)\right)\right].$$

It is easy to see that

$$p_N^{(d)}(\underline{\beta},\underline{h},U(\varepsilon)) \le \sum_{u=1}^{d} p_N(\beta_u,h_u)$$

which implies that

$$\varlimsup_{N\to\infty} p_N^{(d)}(\underline{\beta},\underline{h},U(\varepsilon)) \le \sum_{u=1}^{d} \mathscr{P}(\beta_u,h_u).$$

Consider the following inequality

$$\overline{\lim_{\varepsilon\to+0}}\,\overline{\lim_{N\to\infty}}\, p_N^{(d)}(\underline{\beta},\underline{h},U(\varepsilon)) < \sum_{u=1}^{d} \mathscr{P}(\beta_u,h_u). \tag{2.39}$$

An interesting result of Panchenko & Talagrand (2007b) indicates that, in general, inequality (2.39) does not hold, as one can infer from the following theorem.

Theorem 2.3.7 (local Parisi formula for the multidimensional spherical model (Panchenko & Talagrand, 2007b)). *Assume we are in the situation of the Hamiltonian* (2.36) *with $p = 2$, i.e., the classical SK case. Suppose $\beta_u := \beta > 1$, for all $u \in [1;d]\cap\mathbb{N}$,* and $\min r_u \geq 1$, *where $\{r_u\}_{u=1}^d$ are the eigenvalues of the self-overlap matrix Q.*

Then

$$\lim_{\varepsilon\to+0}\lim_{N\to+\infty} p_N^{(d)}(\underline{\beta},\underline{0},U(\varepsilon)) = d\cdot\mathscr{P}(\beta,0).$$

See (Panchenko & Talagrand, 2007b, Theorem 2) for a proof.

2.3.3 The Aizenman-Sims-Starr comparison scheme for the SK model

The AS^2 scheme (Aizenman *et al.*, 2003, 2007) gives an intrinsic way to obtain variational upper bounds on the free energy in the SK model. The scheme is also based on a comparison between two Gaussian processes. The first process is the sum of the original SK Hamiltonian X and a GREM-inspired process indexed by additional index space $\mathscr{A} := \mathbb{N}^n$. The second one is another GREM-inspired process indexed by the extended configuration space $\Sigma_N \times \mathscr{A}$. The scheme uses a comparison functional defined on Gaussian processes indexed by the extended configuration space equipped with the product measure between the original a priori measure and Ruelle's probability cascade. The role of the comparison functional in the AS^2 scheme is played by a free energy functional acting on the Gaussian processes indexed by the extended configuration space.

It is interesting to note that Panchenko & Talagrand (2007a) have reexpressed Guerra's scheme for the SK model using the RPC.

Increments of the free energy of the SK model

Suppose $H_N(\sigma)$ is the SK Hamiltonian given by (1.19) and $P_N(\beta,h)$ is given by (1.4).

Theorem 2.3.8 (superadditivity of the free energy). *For any $N,M \in \mathbb{N}$,*

$$P_N(\beta,h) + P_M(\beta,h) \leq P_{N+M}(\beta,h). \tag{2.40}$$

See Guerra & Toninelli (2002) for a proof.

Theorem 2.3.8 immediately gives (by Theorem 2.1.2) the existence of the thermodynamic limit (cf. Question 0.0.1). The next theorem builds upon more subtle consequences of superadditivity, namely on relation (2.10).

Theorem 2.3.9 (incremental representation of the free energy). *We have*

$$\lim_{N\to\infty}\underline{\lim_{M\to\infty}}\frac{1}{N}\mathbb{E}\left[\log\left(\frac{Z_{N+M}(\beta,h)}{Z_N(\beta,h)}\right)\right] = p(\beta,h).$$

See, e.g., (Aizenman *et al.*, 2007, Corollary 3.5) for a proof.

Remark 2.3.3. *Note that the free energy $p_N(\beta,h)$ is sharply concentrated around its expectation – a fact that was first established by a martingale argument by Pastur & Shcherbina (1991) and then made more precise by the general machinery of concentration of measure, see, e.g., (Talagrand, 2003, Corollary 2.2.5). Any of these two facts together with Theorem 2.3.8 is sufficient to conclude that*

$$p_N(\beta,h) \xrightarrow[N\uparrow+\infty]{} p(\beta,h), \quad \text{almost surely and in } L^1.$$

Let $\{C(\alpha)\}_{\alpha\in\Sigma_M}$, $\{B(\alpha)\}_{\alpha\in\Sigma_M}$, and $\{A(\alpha,\sigma)\}_{\sigma\in\Sigma_N,\alpha\in\Sigma_M}$ be three independent Gaussian processes with the following correlation structures[2]

$$\begin{aligned}
\mathrm{Cov}\left[C(\alpha^{(1)}),C(\alpha^{(2)})\right] &= \left(\frac{M}{N+M}R_M(\alpha^{(1)},\alpha^{(2)})\right)^2,\\
\mathrm{Cov}\left[A(\sigma^{(1)},\alpha^{(1)}),A(\sigma^{(2)},\alpha^{(2)})\right] &= \frac{2N}{N+M}R_N(\sigma^{(1)},\sigma^{(2)})R_M(\alpha^{(1)},\alpha^{(2)}),\\
\mathrm{Cov}\left[B(\alpha^{(1)}),B(\alpha^{(2)})\right] &= \left(\frac{N}{N+M}R_M(\alpha^{(1)},\alpha^{(2)})\right)^2.
\end{aligned}$$

As a short but crucial computation shows, Theorem 2.3.9 implies the following result.

Theorem 2.3.10 (second incremental representaion of the free energy)**.** *We have*

$$\begin{aligned}
&\lim_{N\to\infty}\varliminf_{M\to\infty}\frac{1}{N}\mathbb{E}\left[\log\left(\frac{\sum_{(\sigma,\alpha)\in\Sigma_N\times\Sigma_M}\eta(\alpha)\exp\left(\beta\sqrt{M}A(\sigma,\alpha)+h\sum_{i=1}^N\sigma_i\right)}{\sum_{\alpha\in\Sigma_N}\eta(\alpha)\exp\left(\beta\sqrt{M}B(\alpha)\right)}\right)\right]\\
&\quad = p(\beta,h),
\end{aligned} \tag{2.41}$$

where

$$\eta(\alpha) := \exp\left(\beta\sqrt{M}C(\alpha)+h\sum_{i=1}^M\sigma_i\right). \tag{2.42}$$

See, e.g., (Bovier, 2006, Section 11.3.1) or (Aizenman *et al.*, 2007, Theorem 4.1) for a proof.

Theorem 2.3.10 suggests that in (2.41) instead of (2.42) one can consider arbitrary random weights $\eta := \{\eta(\alpha)\}_{\alpha\in\mathscr{A}}$, where $\mathscr{A}$ is a countable set and $\sum_{\alpha\in\mathscr{A}}\eta(\alpha)<\infty$ almost surely. Assume also that the Gaussian processes $A := \{A(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma_N\\ \alpha\in\mathscr{A}}}$ and $B := \{B(\alpha)\}_{\alpha\in\mathscr{A}}$ have the following correlation structures

$$\begin{aligned}
\mathrm{Cov}\left[A(\sigma^{(1)},\alpha^{(1)}),A(\sigma^{(2)},\alpha^{(2)})\right] &= 2R_N(\sigma^{(1)},\sigma^{(2)})q(\alpha^{(1)},\alpha^{(2)}),\\
\mathrm{Cov}\left[B(\alpha^{(1)}),B(\alpha^{(2)})\right] &= q(\alpha^{(1)},\alpha^{(2)})^2,
\end{aligned} \tag{2.43}$$

where $q := \{q(\alpha^{(1)},\alpha^{(2)})\}_{\alpha^{(1)},\alpha^{(2)}\in\mathscr{A}}$ is the non-negative definite kernel with $q(\alpha,\alpha)=1$, for all $\alpha\in\mathscr{A}$. Assume, further, that the following functional

$$G_N(\beta,h,\eta,q) := \frac{1}{N}\mathbb{E}\left[\frac{\sum_{(\sigma,\alpha)\in\Sigma_N\times\mathscr{A}}\eta(\alpha)\exp\left(\beta\sqrt{N}A(\sigma,\alpha)+h\sum_{i=1}^N\sigma_i\right)}{\sum_{\alpha\in\mathscr{A}}\eta(\alpha)\exp\left(\beta\sqrt{N}B(\alpha)\right)}\right]$$

is well defined.

[2] It is easy prove the existence of such processes.

Theorem 2.3.11 (the AS^2 variational bound)**.** *For any $N \in \mathbb{N}$ and any η, q as above, we have*

$$p_N(\beta,h) \leq G_N(\beta,h,\eta,q).$$

Moreover,

$$G_N(\beta,h,\eta,q) - p_N(\beta,h) \\ = \frac{\beta^2}{2}\int_0^1 \mathscr{G}(\beta,h,t,\eta,q)\otimes\mathscr{G}(\beta,h,t,\eta,q)\left[\left(q(\alpha^{(1)},\alpha^{(2)}) - R(\sigma^{(1)},\sigma^{(2)})\right)^2\right]\mathrm{d}t,$$

where $\mathscr{G}(\beta,h,t,\eta,q) \in \mathscr{M}_1\left(\Sigma_N \times \mathscr{A}\right)$ is the Gibbs measure induced by the Hamiltonian $H(t)$: $\Sigma \times \mathscr{A} \to \mathbb{R}$ defined as follows

$$H(t;\sigma,\alpha) := \sqrt{t}(H_N(\sigma)+B(\alpha)) + \sqrt{1-t}A(\sigma,\alpha) + h\sum_{i=1}^N \sigma,$$

where $t \in [0;1]$. More precisely, for a measurable $f : \Sigma_N \times \mathscr{A} \to \mathbb{R}$, we have

$$\mathscr{G}(\beta,h,t,\eta,q)\,[f] = \frac{1}{Z}\sum_{(\sigma,\alpha)\in\Sigma\times\mathscr{A}} f(\sigma,\alpha)\eta(\alpha)\exp\left(\beta\sqrt{N}H(t;\sigma,\alpha)\right),$$

where Z is the usual probabilistic normalisation factor.

See (Aizenman *et al.*, 2007, Theorem 4.1) for a proof.

Comparison with the GREM, relation with the Parisi functional

Consider the following *comparison functional*

$$\Phi(\eta)[T] := \mathbb{E}\left[\log\sum_{(\sigma,\alpha)\in\Sigma_N\times\mathscr{A}}\eta(\alpha)\exp\left(\beta\sqrt{N}T(\sigma,\alpha) + h\sum_{i=1}^N\sigma\right)\right], \tag{2.44}$$

where $T := \{T(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma_N\\ \alpha\in\mathscr{A}}}$ is an arbitrary Gaussian process such that (2.44) exists.

We then obviously have

$$G_N(\beta,h,\eta,q) = \Phi(\eta)[A] - \Phi(\eta)[B].$$

Given $x \in \mathscr{Q}'_n$, let $\eta := \xi(x_1,\ldots,x_n)$ be the corresponding RPC. Let, further, $q = q(x)$ be the limiting GREM overlap generated by x (cf. (2.28)). Let $A = A(q)$ and $B = B(q)$ be the corresponding Gaussian processes with covariances given by (2.43).

Theorem 2.3.12 (comaprison with the GREM, relation with the Parisi functional)**.** *Let the function $f_x(q,y) : [0;1]\times\mathbb{R}\to\mathbb{R}$ be the solution of* (1.36).

Then we have

$$\Phi(\eta)[A] = f_x(0,0),$$
$$\Phi(\eta)[B] = \frac{\beta^2}{2}\int_0^1 qx(q)\mathrm{d}q,$$

and consequently

$$G_N(\beta,h,\eta,q) = \mathscr{P}(\beta;x).$$

See, e.g., (Aizenman *et al.*, 2007, Lemma 6.2) for a proof.

Note that Theorems 2.3.12 and 2.3.11 immediately imply the upper bound in (1.37), i.e.,

$$p(\beta,h) \leq \inf_{x\in\mathscr{Q}'}\mathscr{P}(\beta,x). \tag{2.45}$$

3

The Aizenman-Sims-Starr scheme for the SK model with multidimensional spins

In this chapter, we are mainly concerned with the question of the validity of the Parisi formula in the case where spins take values in a d-dimensional Riemannian manifold. We address the issue of extending the approach of Aizenman, Sims and Starr to the multidimensional spins.

3.1 Introduction

The recent rigorous proof of the celebrated *Parisi formula* for the free energy of the SK model, due to Talagrand (2006b), based on the ingenious interpolation schemes of Guerra (2003) and Aizenman *et al.* (2003) constitutes one of the major recent achievements of probability theory. Recently, these results have been generalised to spherical SK-models (Talagrand, 2006a) and to models with spins taking values in a bounded subset of $\mathbb{R}$ (Panchenko, 2005b).

A particular case ($d = 1$, μ with bounded support) of the model we are considering here was treated by Panchenko (2005b). He used the techniques of Talagrand (2006b) to prove that in the case $d = 1$ upper and lower bounds on the free energy coincide (cf. (3.13) and (3.20) in this chapter). However, the results of (Panchenko, 2005b, Section 5 and the proofs of Theorems 2, 5 and 9) are based on relatively detailed differential properties of the optimal Lagrange multipliers in the saddle point optimisation problem of interest. These properties are harder to obtain in multidimensional situations such as that we are dealing with here. In fact, as we show in Theorems 3.1.1 and 3.1.2, one can obtain the same saddle point variational principles without invoking the detailed properties of the optimal Lagrange multipliers. This is achieved using a quenched LDP of the Gärtner-Ellis type.

Definition of the model

We refer to the introduction of this thesis for the definition of the SK model with multidimensional spins. (See, in particular, (0.8) and (0.9).)

Throughout the chapter we assume that we are given a large enough probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that all random variables under consideration are defined on it. Without further notice we shall assume that all Gaussian random variables (vectors and processes) are centred.

We shall be interested mainly in the *free energy*

$$p_N(\beta) := \frac{1}{N} \log \int_{\Sigma_N} \exp\left(\beta \sqrt{N} X(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma), \tag{3.1}$$

where $\beta \geq 0$ is the *inverse temperature* and $\mu \in \mathscr{M}_{\mathrm{f}}(\Sigma)$ is some arbitrary (not necessarily uniform or discrete) finite *a priori measure*. We assume that the a priori measure μ is such that

(3.1) is finite. We shall be interested in proving bounds on the thermodynamic limits of these quantities, e.g., on

$$p(\beta) := \lim_{N\uparrow+\infty} p_N(\beta). \tag{3.2}$$

Remark 3.1.1. *Note that there is no need to include the additional external field terms into the Hamiltonian* (0.8)*, since they could be absorbed into the a priori measure* μ*.*

Main results

Below we introduce the notations, assumptions and formulate the main results of this chapter.

Assumption 3.1.1. *Suppose that the configuration space* Σ *is bounded and such that* $0 \in \operatorname{int}\operatorname{conv}\Sigma$*, where* $\operatorname{conv}\Sigma$ *denotes the convex hull of* Σ*.*

The examples listed below verify this assumption:

1. Multicomponent Ising spins. $\Sigma = \{-1;1\}^d$ – the discrete hypercube.
2. Heisenberg spins. $\Sigma = \left\{\sigma \in \mathbb{R}^d : \|\sigma\|_2 = 1\right\}$ – the unit Euclidean sphere.
3. $\Sigma = \left\{\sigma \in \mathbb{R}^d : \|\sigma\|_2 \leq 1\right\}$ – the unit Euclidean ball.

Remark 3.1.2. *The boundedness assumption can be relaxed and replaced by concentration properties of the a priori measure. In Section 5.2 we will exemplify this in the case of a Gaussian a priori distribution. In general a subgaussian distribution will suffice.*

Consider the space of all *symmetric matrices* $\operatorname{Sym}(d) := \left\{\Lambda \in \mathbb{R}^{d\times d} \mid \Lambda = \Lambda^*\right\}$. Denote

$$\operatorname{Sym}^+(d) := \{\Lambda \in \operatorname{Sym}(d) \mid \Lambda \succeq 0\},$$

where the notation $\Lambda \succeq 0$ means that the matrix Λ is non-negative definite. We equip the space $\operatorname{Sym}(d)$ with the Frobenius (Hilbert-Schmidt) norm

$$\|M\|_{\mathrm{F}}^2 := \sum_{u,v=1}^{d} M_{u,v}^2, \quad M \in \operatorname{Sym}(d).$$

We shall also denote the corresponding (tracial) scalar product by $\langle\cdot,\cdot\rangle$. For $r > \max\{\|\sigma\|_2^2 : \sigma \in \Sigma\}$, define

$$\mathscr{U} := \left\{U \in \operatorname{Sym}(d) \mid U \succeq 0, \|U\|_2 \leq r\right\}.$$

We will call the set $\mathscr{U}$ the *space of the admissible self-overlaps*. In analogy to the usual overlap in the standard SK model, we define, for two configurations, $\sigma^{(i)} = (\sigma_1^{(i)}, \sigma_2^{(i)}, \ldots, \sigma_N^{(i)}) \in \Sigma_N$, $i = 1,2$, the (mutual) *overlap matrix* $R_N(\sigma^{(1)},\sigma^{(2)}) \in \mathbb{R}^{d\times d}$ whose entries are given by

$$R_N(\sigma^{(1)},\sigma^{(2)})_{u,v} := \frac{1}{N}\sum_{i=1}^{N} \sigma_{i,u}^{(1)}\sigma_{i,v}^{(2)}, \quad u,v \in [1;d]\cap\mathbb{N}. \tag{3.3}$$

Fix an *overlap matrix* $U \in \mathscr{U}$. Given a subset $\mathscr{V} \subset \mathscr{U}$, define the set of the *local configurations*,

$$\Sigma_N(\mathscr{V}) := \left\{\sigma \in \Sigma_N \mid R_N(\sigma,\sigma) \in \mathscr{V}\right\}.$$

Next, define the *local free energy*

$$p_N(\mathscr{V}) := \frac{1}{N}\log\int_{\Sigma_N(\mathscr{V})} \mathrm{e}^{\beta\sqrt{N}X(\sigma)}\mathrm{d}\mu^{\otimes N}(\sigma). \tag{3.4}$$

We also define

$$p(\mathscr{V}) := p(\beta,\mathscr{V}) := \lim_{N\uparrow+\infty} p_N(\mathscr{V}), \tag{3.5}$$

where the existence of the limit follows from a result of (Guerra & Toninelli, 2003, Theorem 1). Consider a sequence of matrices $\mathscr{Q} := \{Q^{(k)} \in \mathrm{Sym}(d)\}_{k=0}^{n+1}$ such that

$$0 =: Q^{(0)} \prec Q^{(1)} \prec \ldots \prec Q^{(n+1)} := U, \tag{3.6}$$

where the ordering is understood in the sense of the corresponding quadratic forms. Consider in addition a partition of the unit interval $x := \{x_k\}_{k=0}^{n+1}$, i.e.,

$$0 =: x_0 < x_1 < \ldots < x_{n+1} := 1. \tag{3.7}$$

Let $\{z^{(k)}\}_{k=0}^{n}$ be a sequence of independent Gaussian d-dimensional vectors with

$$\mathrm{Cov}\left[z^{(k)}\right] = Q^{(k+1)} - Q^{(k)}.$$

Given $\Lambda \in \mathrm{Sym}(d)$, define

$$X_{n+1}(x,\mathscr{Q},U,\Lambda) := \log\int_{\Sigma} \exp\left(\sqrt{2}\beta\left\langle\sum_{k=0}^{n} z_k,\sigma\right\rangle + \langle\Lambda\sigma,\sigma\rangle\right)\mathrm{d}\mu(\sigma). \tag{3.8}$$

Define, for $k \in \{n,\ldots,0\}$, by a descending recursion,

$$X_k(x,\mathscr{Q},U,\Lambda) := \frac{1}{x_k}\log\mathbb{E}_{z^{(k)}}\left[\exp\left(x_k X_{k+1}(x,\mathscr{Q},U,\Lambda)\right)\right] \tag{3.9}$$

with

$$X_0(x,\mathscr{Q},U,\Lambda) := \mathbb{E}_{z^{(0)}}\left[X_1(x,\mathscr{Q},U,\Lambda)\right], \tag{3.10}$$

where $\mathbb{E}_{z^{(k)}}[\cdot]$ denotes the expectation with respect to the σ-algebra generated by the random vector $z^{(k)}$.

Remark 3.1.3. *Section 4.1.4 contains the more general framework of dealing with the recursive quantities* (3.10) *which in particular brings to light the links with certain non-linear parabolic PDEs. For these PDEs the recursion* (3.1.2) *is closely related to an iterative application of the well-known Hopf-Cole transformation, see, e.g., Evans (1998).*

Define the *local Parisi functional* as

$$f(x,\mathscr{Q},U,\Lambda) := -\langle\Lambda,U\rangle - \frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right) + X_0(x,\mathscr{Q},U,\Lambda). \tag{3.11}$$

Assumption 3.1.2 (Hadamard squares)**.** *We shall say that a sequence, $\{Q^{(i)}\}_{i=1}^{n}$, of matrices satisfies Assumption 3.1.2, if*

$$\left(Q^{(1)}\right)^{\odot 2} \prec \ldots \prec \left(Q^{(n)}\right)^{\odot 2} \prec \left(Q^{(n+1)}\right)^{\odot 2}. \tag{3.12}$$

Remark 3.1.4. *The above assumption on the matrix order parameters $\mathscr{Q}$ is necessary only to employ the AS^2 scheme. In contrast, Guerra's scheme (Theorems 4.1.1 and 4.16) does not require the above assumption.*

One may verify that the matrices q and ρ in (Talagrand, 2003, Theorems 2.13.1 and 2.13.2) correspond to the matrices $Q^{(1)}$ and $Q^{(2)}$ of this chapter ($n = 1$). (See also (3.23) below.) Furthermore, a straightforward application of the Cauchy-Schwarz inequality shows that the matrices q and ρ actually satisfy Assumption 3.1.2. We also note that in the simultaneous diagonalisation scenario in which the matrices in (3.6) are diagonalisable in the same orthogonal basis (see Sections 4.2.3 and 5.2.2) this assumption is also satisfied.

The first main result of the present chapter uses the AS^2 scheme to establish the upper bound on the limiting free energy $p(\beta)$ in terms of the saddle point problem for the local Parisi functional (3.11).

Theorem 3.1.1. *For any closed set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \leq \sup_{U \in \mathscr{V} \cap \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{3.13}$$

where the infimum runs over all x satisfying (3.7)*, all $\mathscr{Q}$ satisfying both* (3.6) *and Assumption 3.1.2, and all $\Lambda \in \mathrm{Sym}(d)$.*

We were not able to prove in general that the r.h.s. of (3.13) gives also the lower bound to the thermodynamic free energy. See, however, Theorem 5.1.1 for a positive example.

To formulate the lower bound on (3.2) we need some additional definitions.

Let the *comparison index space* be $\mathscr{A} := \mathbb{N}^n$. Given $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}$, define

$$Q(\alpha^{(1)}, \alpha^{(2)}) := Q^{(q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}))}, \tag{3.14}$$

where $q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)})$ is defined in (0.2) Given a $d \times d$-matrix M and $p \in \mathbb{R}$, we denote by $M^{\odot p}$ the $d \times d$-matrix with entries

$$\left(M^{\odot p}\right)_{u,v} := (M_{u,v})^p.$$

The matrix valued lexicographic overlap (3.14) can be used to construct the multidimensional ($d \geq 1$) versions of the GREM (see, e.g., Bovier & Kurkova (2007) and references therein for a review of the results on the one-dimensional case of the model). Here we shall need the following two GREM-inspired real-valued Gaussian processes: $A := \{A(\sigma,\alpha)\}_{\sigma \in \Sigma_N, \alpha \in \mathscr{A}}$ and $B := \{B(\alpha)\}_{\alpha \in \mathscr{A}}$ with covariance structures

$$\begin{aligned} \mathbb{E}\left[A(\sigma^{(1)},\alpha^{(1)})A(\sigma^{(2)},\alpha^{(2)})\right] &= 2\langle R(\sigma^{(1)},\sigma^{(2)}), Q(\alpha^{(1)},\alpha^{(2)})\rangle, \\ \mathbb{E}\left[B(\alpha^{(1)})B(\alpha^{(2)})\right] &= \|Q(\alpha^{(1)},\alpha^{(2)})\|_{\mathrm{F}}^2. \end{aligned}$$

Note that the process A can be represented in the following form:

$$A(\sigma,\alpha)=\left(\frac{2}{N}\right)^{1/2}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle, \tag{3.15}$$

where $\{A_i:=\{A_i(\alpha)\}_{\alpha\in\mathscr{A}}\}_{i=1}^N$ are the i.i.d. (for different indices i) Gaussian $\mathbb{R}^d$-valued processes with the following covariance structure: for $i\in[1;N]\cap\mathbb{N}$, for all $\alpha^{(1)},\alpha^{(2)}\in\mathscr{A}$ and all $u,v\in[1;d]\cap\mathbb{N}$ assume that the following holds

$$\mathbb{E}\left[A_i(\alpha^{(1)})_u A_i(\alpha^{(2)})_v\right]=Q(\alpha^{(1)},\alpha^{(2)})_{u,v}.$$

Given $t\in[0;1]$, we define the *interpolating AS^2 Hamiltonian*

$$H_t(\sigma,\alpha):=\sqrt{t}\,(X(\sigma)+B(\alpha))+\sqrt{1-t}A(\sigma,\alpha). \tag{3.16}$$

Next, we define the random probability measure $\pi_N\in\mathscr{M}_1(\Sigma_N\times\mathscr{A})$ through

$$\pi_N:=\mu^{\otimes N}\otimes\xi,$$

where $\xi=\xi(x)$ is the RPC. We denote by $\{\xi(\alpha)\}_{\alpha\in\mathscr{A}}$ the enumeration of the atom locations of the RPC and consider the enumeration as a random measure on $\mathscr{A}$ (independent of all other random variables around). Define the *local AS^2 Gibbs measure* $\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})$ by

$$\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\,[f]:=\frac{1}{Z_N(t,\mathscr{V})}\int_{\Sigma_N(\mathscr{V})\times\mathscr{A}}f(\sigma,\alpha)\mathrm{e}^{\sqrt{N}\beta H_t(\sigma,\alpha)}\mathrm{d}\pi_N(\sigma,\alpha), \tag{3.17}$$

where $f:\Sigma_N\times\mathscr{A}\to\mathbb{R}$ is an arbitrary measurable function for which the right-hand side of (3.17) is finite. For $\mathscr{V}\subset\mathscr{U}$, define the *$AS^2$ remainder term* as

$$\begin{aligned}&\mathscr{R}_N(x,\mathscr{Q},U,\mathscr{V})\\&:=-\frac{1}{2}\int_0^1\mathbb{E}\left[\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\otimes\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\left[\|R_N(\sigma^{(1)},\sigma^{(2)})-Q(\alpha^{(1)},\alpha^{(2)})\|_{\mathrm{F}}^2\right]\right]\mathrm{d}t.\end{aligned} \tag{3.18}$$

We define also the *limiting AS^2 remainder term*

$$\mathscr{R}(x,\mathscr{Q},U):=\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow\infty}\mathscr{R}_N(x,\mathscr{Q},B(U,\varepsilon))\leq 0, \tag{3.19}$$

where $B(U,\varepsilon)$ is the ball with centre U and radius ε. (The existence of the limiting remainder term is proved in Theorem 3.1.2.)

The second main result of this chapter uses the AS^2 scheme to establish a lower bound on (3.2) in terms of the same saddle point Parisi-type functional as in the upper bound which includes, however, the non-positive remainder term (3.19). In one-dimensional situations Talagrand (2006b) and Panchenko (2005b), respectively, have shown that the corresponding error term vanishes on the optimiser of the Parisi functional.

Theorem 3.1.2. *For any open set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \geq \sup_{U \in \mathscr{V} \cap \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} \left[f(x,\mathscr{Q},\Lambda,U) + \mathscr{R}(x,\mathscr{Q},U) \right], \tag{3.20}$$

where the infimum runs over all x satisfying (3.7)*, all $\Lambda \in \mathrm{Sym}(d)$, and all $\mathscr{Q}$ satisfying both* (3.6) *and Assumption 3.1.2.*

Remark 3.1.5. *The comparison scheme of Guerra (2003) (see also more recent accounts Talagrand (2007a), Guerra (2005) and Aizenman* et al. *(2007)) is also applicable to our model and is covered by our quenched LDP approach, see Theorems 4.1.1 and 4.16 for the formal statements. Guerra's scheme seems to be more amenable (compared to the Aizenman-Sims-Starr one) for Talagrand's remainder estimates (Talagrand, 2006b), see Section 4.3. The scheme is based on the following interpolation*

$$\widetilde{H}_t(\sigma,\alpha) := \sqrt{t}X(\sigma) + \sqrt{1-t}A(\sigma,\alpha) \tag{3.21}$$

which induces the corresponding local Gibbs measure (3.17) *and remainder term* (3.18) *by substituting* (3.16) *with* (3.21)*. Guerra's scheme does not include the process B and, hence, does not require Assumption 3.1.2. Recovering the terms corresponding to $\Phi_N(x,\mathscr{U})[B]$ (see,* (3.105)*) in the Parisi functional requires then a short additional calculation (Lemma 4.1.1).*

Note that the results of (Talagrand, 2003, Theorems 2.13.2 and 2.13.3) imply that at least in the high temperature region (i.e., for small enough β) the Parisi formula for the SK model with multidimensional spins is valid with $n = 1$

$$p(\beta) = f(x,\mathscr{Q}^*,0,U^*) = \sup_{U \in \mathscr{U}} \inf_{(\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{3.22}$$

where the matrices $Q^{*(2)} = U^*$ and $Q^{*(1)}$ solve the following system of equations:

$$\begin{cases} \partial_{Q^{(2)}_{u,v}} f(x,\mathscr{Q}^*,0,U^*) = 0, & u,v \in [1;d] \cap \mathbb{N}, \\ \partial_{Q^{(1)}_{u,v}} f(x,\mathscr{Q}^*,0,U^*) = 0, & u,v \in [1;d] \cap \mathbb{N}. \end{cases} \tag{3.23}$$

Note that the system (3.23) coincides with the mean-field equations obtained in Theorem 2.3.4.

3.2 Some preliminary results

3.2.1 Covariance structure

Our definition of the overlap matrix in (3.3) is motivated by the fact that, as can be seen from a straightforward computation

$$\mathbb{E}\left[X_N(\sigma^{(1)})X_N(\sigma^{(2)})\right] = \sum_{u,v=1}^{d} \left(R_N(\sigma^{(1)},\sigma^{(2)})_{u,v}\right)^2 = \|R_N(\sigma^{(1)},\sigma^{(2)})\|_2^2, \tag{3.24}$$

that is, the the covariance structure of the process $X_N(\sigma)$ is given by the square of the Frobenius (Hilbert-Schmidt) norm of the matrix $R_N(\sigma^{(1)},\sigma^{(2)})$. The basic properties of the overlap matrix are summarised in the following proposition.

Proposition 3.2.1. *We have, for all* $\sigma^{(1)}, \sigma^{(2)}, \sigma \in \Sigma_N$,

1. Matrix representation. $R_N(\sigma^{(1)}, \sigma^{(2)}) = \frac{1}{N}\left(\sigma^{(1)}\right)^* \sigma^{(2)}$.
2. Symmetry #1. $R_N^{u,v}(\sigma^{(1)}, \sigma^{(2)}) = R_N^{v,u}(\sigma^{(2)}, \sigma^{(1)})$.
3. Symmetry #2. $R_N^{u,v}(\sigma, \sigma) = R_N^{v,u}(\sigma, \sigma)$.
4. Non-negative definiteness #1. $R_N(\sigma, \sigma) \succeq 0$.
5. Non-negative definiteness #2.

$$\begin{bmatrix} R_N(\sigma^{(1)}, \sigma^{(1)}) & R_N(\sigma^{(1)}, \sigma^{(2)}) \\ R_N(\sigma^{(1)}, \sigma^{(2)})^* & R_N(\sigma^{(2)}, \sigma^{(2)}) \end{bmatrix} \succeq 0.$$

6. *Suppose* $U := R_N(\sigma^{(1)}, \sigma^{(1)}) = R_N(\sigma^{(2)}, \sigma^{(2)})$, *then*

$$\|R(\sigma^{(1)}, \sigma^{(2)})\|_F^2 \leq \|U\|_F^2.$$

Proof. The proof is straightforward. □

3.2.2 Concentration of measure

The following concentration of measure result for the free energy is standard.

Proposition 3.2.2. *Let* $(\Sigma, \mathfrak{S})$ *be a Polish space. Suppose* μ *is a random finite measure on* Σ. *Suppose, moreover, that* $X(\sigma)$, $\sigma \in \Sigma$ *is the family of Gaussian random variables independent of* μ *which possesses a bounded covariance, i.e.,*

$$\text{there exists } K > 0 \text{ such that } \sup_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma} |\operatorname{Cov}(X(\sigma^{(1)}), X(\sigma^{(2)}))| \leq K. \tag{3.25}$$

Assume that

$$f(X) := \log \int_\Sigma \mathrm{e}^{X(\sigma)} \mathrm{d}\mu(\sigma) < \infty.$$

Then

$$\mathbb{P}\left\{|f(X) - \mathbb{E}[f(X)]| \geq t\right\} \leq 2\exp\left(-\frac{t^2}{4K}\right).$$

Remark 3.2.1. *An analogous result was given in a somewhat more specialised case in Panchenko (2005b).*

Proof. This is an adaptation of the proof of (Talagrand, 2003, Theorem 2.2.4). We can not apply the comparison Theorem 3.2.5 directly, so we resort to the basic interpolation argument as stated in Proposition 2.1.1. For $j = 1, 2$, let the processes $X_j(\cdot)$ be the two independent copies of the process $X(\cdot)$. For $t \in [0; 1]$, let

$$X_{j,t} := \sqrt{t} X_j + \sqrt{1-t} X$$

and

$$F_j(t) := \log \int_\Omega \exp\left(X_{j,t}(\sigma)\right) \mathrm{d}\mu(\sigma).$$

For $s \in \mathbb{R}$, let

$$\varphi_s(t) := \mathbb{E}\left[\exp\left(s(F_1 - F_2)\right)\right].$$

Hence, differentiation gives

$$\dot{\varphi}_s(t) = s\mathbb{E}\left[\exp\left(s(F_1 - F_2)\right)(\dot{F}_1 - \dot{F}_2)\right] \tag{3.26}$$

(the dots indicate the derivatives with respect to t) and also

$$\begin{aligned}\dot{F}_j(t) =& \frac{1}{2}\left(\int_\Sigma \exp\left(X_{j,t}(\sigma)\right) \mathrm{d}\mu(\sigma)\right)^{-1} \\ &\times \int_\Sigma \left(t^{-1/2}X_j(\sigma) - (1-t)^{-1/2}X(\sigma)\right)\exp\left(X_{j,t}(\sigma)\right)\mathrm{d}\mu(\sigma).\end{aligned} \tag{3.27}$$

Now, we substitute (3.27) back to (3.26) and apply Corollary 2.1.1 to the result. After some tedious but elementary calculations we get

$$\begin{aligned}\dot{\varphi}_s(t) =& s^2\mathbb{E}\Bigg[\exp\left(s(F_1 - F_2)\right)\left(\int_\Sigma \exp X_{1,t}(\sigma)\mathrm{d}\mu(\sigma)\int_\Sigma \exp X_{2,t}(\sigma)\mathrm{d}\mu(\sigma)\right)^{-1} \\ &\int_\Sigma \mathrm{Cov}(X(\sigma^{(1)}), X(\sigma^{(2)}))\exp\left(X_{1,t}(\sigma^{(1)}) + X_{2,t}(\sigma^{(2)})\right)\mathrm{d}\mu(\sigma^{(1)})\mathrm{d}\mu(\sigma^{(2)})\Bigg].\end{aligned}$$

Thus, thanks to (3.25), we obtain

$$\dot{\varphi}_s(t) \le Ks^2\varphi_s(t).$$

The conclusion of the theorem follows now exactly as in the proof of (Talagrand, 2003, Theorem 2.2.4).

□

We now apply this general result to the our model and also to the free energy-like functional of the GREM-inspired process A.

Proposition 3.2.3. *Suppose $\Sigma \subset B(0,r)$, for $r > 0$. For $\Omega \subset \Sigma_N$, denote*

$$P_N^{SK}(\beta, \Omega) := \log \int_\Omega \exp\left(\sqrt{N}\beta X_N(\sigma)\right)\mathrm{d}\mu^{\otimes N}(\sigma),$$

and

$$P_N^{GREM}(\beta, \Omega) := \log \int_{\Omega \times \mathscr{A}} \exp\left(\beta\sqrt{2}\sum_{i=1}^N \langle A_i(\alpha), \sigma_i\rangle\right)\mathrm{d}\pi_N(\sigma, \alpha).$$

Then, for all $\Omega \subset \Sigma_N$, we have

1. For any $t > 0$,

$$\mathbb{P}\left\{\left|P_N^{SK}(\beta, \Omega) - \mathbb{E}\left[P_N^{SK}(\beta, \Omega)\right]\right| > t\right\} \le 2\exp\left(-\frac{t^2}{4\beta^2 r^4 N}\right). \tag{3.28}$$

2. *For any $t > 0$,*

$$\mathbb{P}\left\{\left|P_N^{GREM}(\beta,\Omega) - \mathbb{E}\left[P_N^{GREM}(\beta,\Omega)\right]\right| > t\right\} \leq 2\exp\left(-\frac{t^2}{8\beta^2 r^4 N}\right). \tag{3.29}$$

Proof. 1. We would like to use Proposition 3.2.2. By (3.24) and the Cauchy-Bouniakovsky-Schwarz inequality, we have, for all $N \in \mathbb{N}$, $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, that

$$\mathrm{Cov}(X_N(\sigma^{(1)},\sigma^{(2)})) = \|R_N(\sigma^{(1)},\sigma^{(2)})\|_{\mathrm{F}}^2 = \frac{1}{N^2}\sum_{i,j=1}^{N}\langle\sigma_i^{(1)},\sigma_j^{(1)}\rangle\langle\sigma_i^{(2)},\sigma_j^{(2)}\rangle \leq r^4. \tag{3.30}$$

Hence, for all $N \in \mathbb{N}$ and all subsets Ω of Σ_N, we obtain

$$\sup_{\sigma^{(1)},\sigma^{(2)}\in\Sigma}|\mathrm{Cov}(X(\sigma^{(1)}),X(\sigma^{(2)}))| \leq r^4.$$

Thus (3.28) is proved.

2. We fix an arbitrary $N \in \mathbb{N}$, $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}$. We have

$$\begin{aligned}\mathrm{Cov}(A(\sigma^{(1)},\alpha^{(1)}),A(\sigma^{(2)},\alpha^{(2)})) &= \mathbb{E}\left[A(\sigma^{(1)},\alpha^{(1)})A(\sigma^{(2)},\alpha^{(2)})\right] \\ &= \sum_{i=1}^{N}\langle Q(\alpha^{(1)},\alpha^{(2)})\sigma_i^{(1)},\sigma_i^{(2)}\rangle.\end{aligned}$$

Bound (3.30) implies that, for any $U \in \mathscr{U}$, we have $\|U\|_2 \leq r^2$. Since $Q(\alpha^{(1)},\alpha^{(2)}) \in \mathscr{U}$, we obtain

$$\begin{aligned}|\langle Q(\alpha^{(1)},\alpha^{(2)})\sigma_i^{(1)},\sigma_i^{(2)}\rangle| &\leq \|Q(\alpha^{(1)},\alpha^{(2)})\|_2\|\sigma_i^{(1)}\|_2\|\sigma_i^{(2)}\|_2 \\ &\leq \|Q(\alpha^{(1)},\alpha^{(2)})\|_2 r^2 \leq r^4.\end{aligned}$$

Therefore, using Proposition 3.2.2, we obtain (3.29).

□

3.2.3 Gaussian comparison inequalities for free energy-like functionals

We begin by recalling well-known integration by parts formula which is the source of many comparison results for functionals of Gaussian processes.

The following proposition connects the computation of the derivative of the free energy with respect to the parameter that linearly occurs in the Hamiltonian with a certain Gibbs average for a replicated system.

Proposition 3.2.4. *Consider a Polish measure space $(\Sigma,\mathfrak{S})$ and a random measure μ on it. Let $X = \{X(\sigma)\}_{\sigma\in\Sigma}$ and $Y := \{Y(\sigma)\}_{\sigma\in\Sigma}$ be two independent Gaussian real-valued processes. For $u \in \mathbb{R}$, we define*

$$H_u(\sigma) := uX(\sigma) + Y(\sigma).$$

Assume that, for all $u \in [a,b] \Subset \mathbb{R}$, we have

$$\int \exp\left(H_u(\sigma)\right) \mathrm{d}\mu(\sigma) < \infty, \int X(\sigma) \exp\left(H_u(\sigma)\right) \mathrm{d}\mu(\sigma) < \infty$$

almost surely, and also that

$$\mathbb{E}\left[\log \int \exp\left(H_u(\sigma)\right) \mathrm{d}\mu(\sigma)\right] < \infty.$$

Then we have

$$\frac{\mathrm{d}}{\mathrm{d}u}\mathbb{E}\left[\log \int \mathrm{e}^{H_u(\sigma)} \mathrm{d}\mu(\sigma)\right] = u\mathbb{E}\left[\mathscr{G}(u) \otimes \mathscr{G}(u)\left[\operatorname{Var} X(\sigma) - \mathbb{E}\left[X(\sigma), X(\tau)\right]\right]\right],$$

where $\mathscr{G}(u)$ is the random element of $\mathscr{M}_1(\Sigma)$ which, for any measurable $f : \Sigma \to \mathbb{R}$, satisfies

$$\mathscr{G}(u)\left[f\right] = \frac{1}{Z(u)} \int f(\sigma) \exp\left(H_u(\sigma)\right) \mathrm{d}\mu(\sigma).$$

Proof. We write

$$\frac{\mathrm{d}}{\mathrm{d}u} \log \int \mathrm{e}^{H_u(\sigma)} \mathrm{d}\mu(\sigma) = \int X(\sigma) \frac{\mathrm{e}^{H_u(\sigma)}}{Z_u(\beta)} \mathrm{d}\mu(\sigma), \tag{3.31}$$

where $Z_u(\beta) := \int \mathrm{e}^{\beta H_u(\sigma)} \mathrm{d}\mu(\sigma)$. The main ingredient of the proof is the Gaussian integration by parts formula. Denote, for $\tau \in \Sigma$, $e(\tau) := \mathbb{E}\left[X(\sigma) H_u(\tau)\right]$. By (2.3), we have

$$\mathbb{E}\left[X(\sigma) \frac{\mathrm{e}^{H_u(\sigma)}}{Z_u(\beta)}\right] = \mathbb{E}\left[\partial_X \left(\frac{\mathrm{e}^{H_u(\sigma)}}{\int \mathrm{e}^{H_u(\tau)} \mathrm{d}\mu(\tau)}\right)(X; e)\right]. \tag{3.32}$$

Due to the independence, we have

$$\mathbb{E}\left[X(\sigma) H_u(\tau)\right] = u\mathbb{E}\left[X(\sigma), X(\tau)\right].$$

Henceforth, the computation of the directional derivative in (3.32) amounts to

$$\begin{aligned}
&\frac{\partial}{\partial t}\left[\frac{\mathrm{e}^{H_u(\sigma) + tu \operatorname{Var}(\sigma)}}{\int \mathrm{e}^{H_u(\tau) + tu \operatorname{Cov}(\sigma, \tau)} \mathrm{d}\mu(\tau)}\right] \\
&\quad = \left(\int \mathrm{e}^{H_u(\sigma)} \mathrm{d}\mu(\sigma)\right)^{-2} \left(u \operatorname{Var} X(\sigma) \mathrm{e}^{H_u(\sigma)} \int \mathrm{e}^{H_u(\tau)} \mathrm{d}\mu(\tau)\right. \\
&\qquad \left. - \mathrm{e}^{H_u(\sigma)} \int u \operatorname{Cov}\left[X(\sigma), X(\tau)\right] \mathrm{e}^{H_u(\tau)} \mathrm{d}\mu(\tau)\right).
\end{aligned} \tag{3.33}$$

Substituting the r.h.s. of (3.33) into (3.31), we obtain the assertion of the proposition. □

The following proposition gives a short differentiation formula, which is useful in getting comparison results between the (free energy-like) functionals of Gaussian processes.

Proposition 3.2.5. *Let $(X(\sigma))_{\sigma\in\Sigma}$, $(Y(\sigma))_{\sigma\in\Sigma}$ be two independent Gaussian processes as before. Set*

$$H_t(\sigma) := \sqrt{t}X(\sigma) + \sqrt{1-t}Y(\sigma).$$

Assume that

$$\int \mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma) < \infty, \int X(\sigma)\mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma) < \infty,$$
$$\int Y(\sigma)\mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma) < \infty$$

almost surely, and also that, for all $t \in [0;1]$,

$$\mathbb{E}\left[\log \int \mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma)\right] < \infty.$$

Then we have

$$\begin{aligned}\mathbb{E}\left[\log \int \mathrm{e}^{X(\sigma)}\mathrm{d}\mu(\sigma)\right] &= \mathbb{E}\left[\log \int \mathrm{e}^{Y(\sigma)}\mathrm{d}\mu(\sigma)\right] \\ -\frac{1}{2}\int_0^1 \mathscr{G}(t)\otimes\mathscr{G}(t) &\left[\left(\mathrm{Var}\, X(\sigma^{(1)}) - \mathrm{Var}\, Y(\sigma^{(1)})\right)\right. \\ -\left(\mathrm{Cov}\left[X(\sigma^{(1)}), X(\sigma^{(2)})\right] - \mathrm{Cov}\left[Y(\sigma^{(1)}), Y(\sigma^{(2)})\right]\right)&\Big]\,\mathrm{d}t, \end{aligned} \tag{3.34}$$

where $\mathscr{G}(t)$ is the random element of $\mathscr{M}_1(\Sigma)$ which, for all measurable $f : \Sigma \to \mathbb{R}$, satisfies

$$\mathscr{G}(t)\,[f] = \frac{1}{Z(t)}\int_\Sigma f(\sigma)\exp\left(H_t(\sigma)\right)\mathrm{d}\mu(\sigma). \tag{3.35}$$

Proof. Let us introduce the process

$$W_{u,v}(\sigma) := uX(\sigma) + vY(\sigma).$$

Hence,

$$H_t(\sigma) = W_{\sqrt{t},\sqrt{1-t}}(\sigma). \tag{3.36}$$

Thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\log \int \mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma)\right] = \frac{1}{2}&\left(\frac{1}{\sqrt{t}}\frac{\partial}{\partial u}\mathbb{E}\left[\log \int \mathrm{e}^{W_{u,v}(\sigma)}\mathrm{d}\mu(\sigma)\right]\right. \\ &\left.\left.-\frac{1}{\sqrt{1-t}}\frac{\partial}{\partial v}\mathbb{E}\left[\log \int \mathrm{e}^{W_{u,v}(\sigma)}\mathrm{d}\mu(\sigma)\right]\right)\right|_{u=\sqrt{t},v=\sqrt{1-t}}.\end{aligned}$$

Applying Proposition 3.2.4 and $\int_0^1 \cdot \mathrm{d}t$ to the previous formula, we conclude the proof. □

3.3 Quenched Gärtner-Ellis type LDP

In this section, we derive a quenched LDP under measure concentration assumptions. Theorems 3.3.1 and 3.3.2 give the corresponding LDP upper and lower bounds, respectively. The proofs of the LDP bounds will be adapted to get the proofs of the upper and lower bounds on the free energy of the SK model with multidimensional spins. However, they may be of independent interest.

Note that the existing "level-2" quenched large deviation results of Comets (1989) are applicable only to a certain class of mean-field random Hamiltonians which are required to be "macroscopic" functionals of the joint empirical distribution of the random variables representing the disorder and the independent spin variables. The SK Hamiltonian can not be represented in such form, since the interaction matrix consists of i.i.d. random variables. Moreover, it is assumed in Comets (1989) that the Hamiltonian has the form $H_N(\sigma) = NV(\sigma)$, where $\{V(\sigma)\}_{\sigma\in\Sigma_N}$ is a random process taking values in some fixed bounded subset of $\mathbb{R}$. Since the Hamiltonian of our model is a Gaussian process, this assumption is also not satisfied, due to the unboundedness of the Gaussian distribution.

3.3.1 Quenched LDP upper bound

Throughout this section we impose the following.

Assumption 3.3.1. *Suppose $\{Q_N\}_{N=1}^{\infty}$ is a sequence of random measures on a Polish space $(\mathscr{X},\mathfrak{X})$. Assume that there exists some $L>0$ such that for any Q_N-measurable set $A\subset\mathscr{X}$ we have*

$$\mathbb{P}\left\{\left|\log Q_N(A) - \mathbb{E}\left[\log Q_N(A)\right]\right| > t\right\} \leq \exp\left(-\frac{t^2}{LN}\right).$$

Note that this assumption will hold in the cases we are interested in due to Proposition 3.2.2.

Lemma 3.3.1. *Suppose $\{Q_N\}_{N=1}^{\infty}$ is a sequence of random measures on a Polish space $(\mathscr{X},\mathfrak{X})$ and $\{A_r\subset\mathscr{X} : r\in\{1,\dots,p\}\}$ is a sequence of Q_N-measurable sets such that, for some absolute constant $L>0$, we have*

$$\mathbb{P}\left\{\left|\log Q_N(A_r) - \mathbb{E}\left[\log Q_N(A_r)\right]\right| > t\right\} \leq \exp\left(-\frac{t^2}{LN}\right). \tag{3.37}$$

Then we have

$$\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\left|\log Q_N\left(\bigcup_{r=1}^{p} A_r\right) - \max_{r\in\{1,\dots,p\}}\mathbb{E}\left[\log Q_N(A_r)\right]\right|\right] = 0. \tag{3.38}$$

Proof. First, (3.37) gives

$$\mathbb{P}\left\{\max_{r\in\{1,\dots,p\}}\left|\log Q_N(A_r) - \mathbb{E}\left[\log Q_N(A_r)\right]\right| \geq t\right\} \leq 2p\exp\left(-\frac{t^2}{LN}\right).$$

Since, for $a,b\in\mathbb{R}^p$, the following elementary inequality holds

$$\left|\max_r a_r - \max_r b_r\right| \leq \max_r |a_r - b_r|,$$

we get

$$\mathbb{P}\left\{\left|\max_{r\in\{1,\dots,p\}}\log Q_N(A_r)-\max_{r\in\{1,\dots,p\}}\mathbb{E}\left[\log Q_N(A_r)\right]\right|\geq t\right\}\leq 2p\exp\left(-\frac{t^2}{LN}\right).$$

The last equation in turn implies that

$$\frac{1}{N}\mathbb{E}\left[\left|\max_{r\in\{1,\dots,p\}}\log Q_N(A_r)-\max_{r\in\{1,\dots,p\}}\mathbb{E}\left[\log Q_N(A_r)\right]\right|\right]\leq 2p\int_0^{+\infty}\exp\left(-\frac{Nt^2}{L}\right)\mathrm{d}t, \quad (3.39)$$

and the r.h.s. of the previous formula vanishes as $N\uparrow\infty$.

□

Let $Q_N\in\mathscr{M}(\mathscr{X})$, $N\in\mathbb{N}$ be a family of random measures on $(\mathscr{X},\mathfrak{X})$. Define the Laplace transform

$$L_N(\Lambda):=\int_{\mathscr{X}}\mathrm{e}^{N\langle x,\Lambda\rangle}\mathrm{d}Q_N(x).$$

Suppose that, for all $\Lambda\in\mathbb{R}^d$, we have

$$I(\Lambda):=\lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda)\right]\in\overline{\mathbb{R}}=\mathbb{R}\cup\{-\infty,+\infty\}. \quad (3.40)$$

Define the Legendre transform

$$I^*(x):=\inf_{\Lambda}\left[-\langle x,\Lambda\rangle+I(\Lambda)\right]. \quad (3.41)$$

Define, for $\delta>0$,

$$I^*_\delta(x):=\max\left\{I^*(x)+\delta,-\frac{1}{\delta}\right\}. \quad (3.42)$$

Lemma 3.3.2. *Suppose*

$$0\in\operatorname{int}\mathscr{D}(I):=\operatorname{int}\{\Lambda:I(\Lambda)<+\infty\}. \quad (3.43)$$

Then

1. *The mapping $I^*(\cdot):\mathscr{X}\to\mathbb{R}$ is upper semi-continuous and concave.*
2. *For all $M>0$,*

$$\{x\in\mathscr{X}:I^*(x)\leq M\}\text{ is a compact.}$$

Proof. 1. Since, for all $\Lambda\in\mathscr{D}(I)$, the linear mappings

$$x\mapsto-\langle\Lambda,x\rangle+I(\Lambda)$$

are obviously concave, the infimum of this family is upper semi-continuous and concave.

2. See, e.g., den Hollander (2000) for the proof.

□

Theorem 3.3.1. *Suppose that*

1. *The family* $\{Q_N\}$ *satisfies condition* (3.38).
2. *Condition* (3.40) *is satisfied.*
3. *Condition* (3.43) *is satisfied.*

Then, for any closed set $\mathscr{V} \subset \mathbb{R}^d$*, we have*

$$\overline{\lim_{N\uparrow\infty}} \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{V})\right] \le \sup_{x\in\mathscr{V}} I^*(x). \tag{3.44}$$

Proof. 1. Suppose at first that $\mathscr{V}$ is a compact.

Thanks to (3.41), for any $x \in \mathscr{X}$, there exists $\Lambda(x) \in \mathscr{X}$ such that

$$-\langle x, \Lambda(x)\rangle + I(\Lambda(x)) \le I^*_\delta(x). \tag{3.45}$$

For any $x \in \mathscr{X}$, there exists a neighbourhood $A(x) \subset \mathscr{X}$ of x such that

$$\sup_{y\in A(x)} \langle y-x, \Lambda(x)\rangle \le \delta.$$

By compactness, the covering $\bigcup_{x\in\mathscr{Y}} A(x) \supset \mathscr{V}$ has the finite subcovering, say $\bigcup_{r=1}^p A(x_r) \supset \mathscr{V}$. Hence,

$$\frac{1}{N}\log Q_N(\mathscr{V}) \le \frac{1}{N}\log\left(\bigcup_{r=1}^p Q_N(A(x_r))\right). \tag{3.46}$$

Applying condition (3.38), we get

$$\overline{\lim_{N\uparrow\infty}} \frac{1}{N}\mathbb{E}\left[\max_{r\in\{1,\dots,p\}} \log Q_N(A(x_r)) - \max_{r\in\{1,\dots,p\}} \mathbb{E}\left[\frac{1}{N}\log Q_N(A(x_r))\right]\right] \le 0. \tag{3.47}$$

By the Chebyshev inequality,

$$\begin{aligned} Q_N(A(x)) &\le Q_N\{y\in\mathscr{X} : \langle y-x, \Lambda(x)\rangle \le \delta\} \\ &\le \mathrm{e}^{-\delta N}\int_{\mathscr{X}} \mathrm{e}^{N\langle y-x,\Lambda(x)\rangle}\mathrm{d}Q_N(y) \\ &= \mathrm{e}^{-\delta N}\mathrm{e}^{-N\langle x,\Lambda(x)\rangle} L_N(\Lambda(x)). \end{aligned} \tag{3.48}$$

Hence, (3.48) together with (3.45) yields

$$\begin{aligned} \overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log Q_N(A(x_r))\right] &\le \lim_{N\uparrow+\infty}\left[-\langle x_r, \Lambda(x_r)\rangle + \frac{1}{N}\log L_N(\Lambda(x_r))\right] - \delta \\ &= -\langle x_r, \Lambda(x_r)\rangle + I(\Lambda(x_r)) - \delta \\ &\le I^*_\delta(x_r) - \delta. \end{aligned} \tag{3.49}$$

Combining (3.46), (3.47), (3.49), we obtain

$$\begin{aligned} \overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{V})\right] &\le \max_{r\in\{1,\dots,p\}} I^*_\delta(x_r) - \delta \\ &\le \sup_{x\in\mathscr{V}} I^*_\delta(x) - \delta. \end{aligned}$$

Taking $\delta \downarrow +0$ limit, we get the assertion of the theorem.

2. Let us allow now the set $\mathscr{V}$ to be unbounded. We first prove that the family Q_N is quenched exponentially tight. For that purpose, let

$$R_N(M) := \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{X} \setminus [-M;M]^d)\right],$$

and denote

$$R(M) := \overline{\lim_{N\uparrow+\infty}} R_N(M).$$

We want to prove that

$$\lim_{M\uparrow+\infty} R(M) = -\infty. \tag{3.50}$$

Fix some $u \in \{1,\dots,d\}$. Suppose $\delta_{u,p} \in \{0,1\}$ is the standard Kronecker symbol. Let $e_u \in \mathbb{R}^d$ be an element of the standard basis of $\mathbb{R}^d$, i.e., for all $p \in \{1,\dots,d\}$, we have

$$(e_u)_p := \delta_{u,p}.$$

Thanks to the Chebyshev inequality, we have

$$Q_N\{x_u \le -M\} \le \mathrm{e}^{-NM} \int_{\mathbb{R}^d} \mathrm{e}^{-N\langle x,e_u\rangle} \mathrm{d}Q_N(x), \text{ a.s.} \tag{3.51}$$

Now, we get

$$\begin{aligned}\int_{\mathbb{R}^d} \mathrm{e}^{-N\langle x,e_u\rangle} \mathrm{d}Q_N(x) &= \frac{1}{L_N(\Lambda_e)} \int_{\mathbb{R}^d} \mathrm{e}^{N\langle x,\Lambda_e - e_u\rangle} \mathrm{d}Q_N(x)\\ &= \frac{L_N(\Lambda_e - e_u)}{L_N(\Lambda_e)}, \text{ a.s.} \end{aligned}\tag{3.52}$$

Hence, combining (3.51) and (3.52), we obtain

$$\frac{1}{N}\mathbb{E}\left[\log Q_N\{x_u \le -M\}\right] \le -M + I_N(\Lambda_e - e_u) - I_N(\Lambda_e). \tag{3.53}$$

Using the same argument, we also get

$$\frac{1}{N}\mathbb{E}\left[\log Q_N\{x_u \ge M\}\right] \le -M + I_N(\Lambda_e + e_u) - I_N(\Lambda_e). \tag{3.54}$$

We obviously have

$$R_N(M) \le \frac{1}{N}\mathbb{E}\left[\log Q_N\left(\bigcup_{u=1}^{d} (\{x_u \le -M\} \cup \{x_u \ge M\})\right)\right]. \tag{3.55}$$

Applying condition (3.38) to (3.55), we get

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\Big[\log Q_N\Big(\bigcup_{u=1}^{d}\Big(\{x_u \le -M\} \cup \{x_u \ge M\}\Big)\Big)$$

$$-\max_{u\in\{1,\dots,d\}}\max\Big\{\mathbb{E}[\log Q_N(\{x_u\le -M\})],\mathbb{E}[\log Q_N(\{x_u\ge M\})]\Big\}\Big]\le 0. \quad (3.56)$$

Applying (3.53) and (3.54) in (3.56), we get

$$\begin{aligned}&\overline{\lim_{N\uparrow+\infty}}\frac{1}{N}\mathbb{E}\left[\log Q_N\left(\bigcup_{u=1}^{d}(\{x_u\le -M\}\cup\{x_u\ge M\})\right)\right]\\ &\le -M-I(\Lambda_e)+\max_{u\in\{1,\dots,d\}}\max\{I(\Lambda_e-e_u),I(\Lambda_e+e_u)\}.\end{aligned} \quad (3.57)$$

The bound (3.57) assures (3.50). Now, since we have (with the help of (3.38) and (3.44))

$$\begin{aligned}\overline{\lim_{N\uparrow+\infty}}\frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{V})\right]&\le\overline{\lim_{N\uparrow+\infty}}\frac{1}{N}\mathbb{E}\left[\log Q_N((\mathscr{V}\cap[-M;M]^d)\cup(\mathscr{X}\setminus[-M;M]^d))\right]\\ &\le\max\left\{\sup_{x\in(\mathscr{V}\cap[-M;M]^d)}I^*(x),R(M)\right\},\end{aligned} \quad (3.58)$$

the assertion of the theorem follows from (3.50) by taking the $\overline{\lim}_{M\uparrow+\infty}$ in the bound (3.58). □

3.3.2 Quenched LDP lower bound

Suppose that, for some $\Lambda\in\mathbb{R}^d$ and all $N\in\mathbb{N}$, we have

$$\int_{\mathscr{X}} e^{N\langle y,\Lambda\rangle}dQ_N(y)<+\infty.$$

Let $\widetilde{Q}_{N,\Lambda}\in\mathscr{M}(\mathscr{X})$ be the random measure defined by

$$\widetilde{Q}_{N,\Lambda}(A)=\int_A e^{N\langle y,\Lambda\rangle}dQ_N(y), \quad (3.59)$$

for any Q_N measurable $A\subset\mathscr{X}$.

Lemma 3.3.3. *Suppose the family of random measures Q_N satisfies the following assumptions.*

1. *Measure concentration. There exists some $L>0$ such that, for any Q_N-measurable set $A\subset\mathscr{X}$, we have*

$$\mathbb{P}\left\{\left|\log Q_N(A)-\mathbb{E}\left[\log Q_N(A)\right]\right|>t\right\}\le\exp\left(-\frac{t^2}{LN}\right).$$

2. *Tails decay condition. Let*

$$C(M):=\{x\in\mathscr{X}:\|x\|<M\}.$$

There exists $p\in\mathbb{N}$ such that

$$\lim_{K\uparrow+\infty}\overline{\lim_{N\uparrow\infty}}\int_0^{+\infty}\mathbb{P}\left\{\frac{1}{N}\log\widetilde{Q}_{N,\Lambda}(\mathscr{X}\setminus C(N^p))>-K+t\right\}dt=0.$$

3. *Non-degeneracy. The family of the sets* $\left\{B_j \subset \mathscr{X} : j \in \{1,\dots,q\}\right\}$ *satisfies the following condition*

$$\text{there exists some } j_0 \in \{1,\dots,q\} \text{ such that } \varliminf_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_{j_0})\right] > -\infty. \tag{3.60}$$

Then, for any $\Lambda \in \mathbb{R}^d$, *we have*

$$\varlimsup_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}\left(\bigcup_{j=1}^{q} B_j\right) - \max_{j\in\{1,\dots,q\}} \mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_j)\right]\right] \le 0. \tag{3.61}$$

Remark 3.3.1. *The polynomial growth choice of* $M = M_N := N^p$ *made in assumption* (2) *of the above lemma corresponds certainly to the concrete form of the exponential concentration in assumption* (1). *Both assumptions together with the proof can be adapted to the weaker concentration bounds.*

Proof. We fix some $j \in \{1,\dots,q\}$. Take an arbitrary $\varepsilon > 0$, $M > 0$ and denote $J_{M,\varepsilon} := \mathbb{Z} \cap [-\|\Lambda\|M/\varepsilon; \|\Lambda\|M/\varepsilon]$. Consider, for $i \in J_{M,\varepsilon}$, the following closed sets

$$A_{i,j} := \{x \in B_j : (j-1)\varepsilon \le \langle \Lambda, x\rangle \le j\varepsilon\}.$$

We get

$$\begin{aligned}
\frac{1}{N}\log \widetilde{Q}_{N,\Lambda}\left(\bigcup_{j=1}^{q} B_j\right) &\le \frac{1}{N}\log \widetilde{Q}_{N,\Lambda}\left(\left(\bigcup_{j=1}^{q} B_j \cap C(M)\right) \cup (\mathscr{X}\setminus C(M))\right)\\
&\le \frac{1}{N}\max\Big\{\max_{j\in\{1,\dots,q\}} \log \widetilde{Q}_{N,\Lambda}(B_j\cap C(M)),\\
&\qquad \log \widetilde{Q}_{N,\Lambda}(\mathscr{X}\setminus C(M))\Big\} + \frac{\log(q+1)}{N}.
\end{aligned} \tag{3.62}$$

We have

$$\begin{aligned}
\frac{1}{N}\log \widetilde{Q}_{N,\Lambda}(B_j\cap C(M)) &\le \frac{1}{N}\log\Big(\sum_{i\in J_{M,\varepsilon}} \mathrm{e}^{Ni\varepsilon} Q_N(A_{i,j})\Big)\\
&\le \max_{i\in\{1,\dots,p\}}\left[i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right] + \frac{\log(\operatorname{card} J_{M,\varepsilon})}{N}.
\end{aligned} \tag{3.63}$$

Denote

$$\alpha_N(\varepsilon) := \max_{j\in\{1,\dots,q\}} \max_{i\in J_{M,\varepsilon}} \left(i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right),$$

and

$$\beta_N := \max_{j\in\{1,\dots,q\}} \mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_j)\right],$$

$$\widetilde{\beta}_N(\varepsilon) := \max_{j\in\{1,\dots,q\}} \mathbb{E}\left[\max_{i\in J_{M,\varepsilon}}\left(i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right)\right],$$

$$\gamma_N(M) := \frac{1}{N} \log \widetilde{Q}_{N,\Lambda}(\mathscr{X} \setminus C(M)).$$

We also have

$$\begin{aligned}
\frac{1}{N} \log \widetilde{Q}_{N,\Lambda}(B_j) &\geq \frac{1}{N} \log \widetilde{Q}_{N,\Lambda}(B_j \cap C(M)) \\
&\geq \max_{i \in J_{M,\varepsilon}} \left[(i-1)\varepsilon + \frac{1}{N} \log Q_N(A_{i,j}) \right] \\
&= \max_{i \in J_{M,\varepsilon}} \left[i\varepsilon + \frac{1}{N} \log Q_N(A_{i,j}) \right] - \varepsilon.
\end{aligned} \tag{3.64}$$

Due to condition (1), we have

$$\mathbb{P}\left\{ \left| \alpha_N(\varepsilon) - \widetilde{\beta}_N(\varepsilon) \right| > t \right\} \leq \operatorname{card} J_{M,\varepsilon}\, q \exp\left(-\frac{Nt^2}{L} \right). \tag{3.65}$$

We put $M := M_N := N^p$, and we get

$$\begin{aligned}
\operatorname{card} J_{M,\varepsilon} &\leq 2\|\Lambda\| M/\varepsilon + 1 \\
&\leq 2\|\Lambda\| N^p/\varepsilon + 1.
\end{aligned} \tag{3.66}$$

Let

$$X_N(M,\varepsilon) := \max\{\gamma_N(M), \alpha_N(\varepsilon)\} - \beta_N,$$

then we have

$$\mathbb{P}\{X_N(K,\varepsilon) > t\} \leq \mathbb{P}\{\gamma_N(M) > \beta_N + t\} + \mathbb{P}\{\alpha_N(\varepsilon) > \beta_N + t\}. \tag{3.67}$$

Due to property (3.60), there exists $K > 0$ such that we have

$$\mathbb{P}\{\gamma_N(M) > \beta_N + t\} \leq \mathbb{P}\{\gamma_N(M) > -K + t\}. \tag{3.68}$$

Thanks to (3.64), we have

$$\mathbb{P}\{\alpha_N(\varepsilon) > \beta_N + t\} \leq \mathbb{P}\{\alpha_N(\varepsilon) > \widetilde{\beta}_N(\varepsilon) + t - \varepsilon\}. \tag{3.69}$$

For $t > \varepsilon$, we apply (3.65) and (3.66) to (3.69) to obtain

$$\mathbb{P}\{\alpha_N(\varepsilon) > \beta_N + t\} \leq (2\|\Lambda\| N^p/\varepsilon + 1)\, q \exp\left(-\frac{Nt^2}{L} \right). \tag{3.70}$$

Combining (3.62) and (3.63), we get

$$\begin{aligned}
&\mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}\left(\bigcup_{j=1}^{q} B_j \right) - \max_{j \in \{1,\dots,q\}} \mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_j) \right] \right] \\
&\leq \mathbb{E}\left[X_N(M,\varepsilon) \right] + \frac{\log(q+1)}{N} + \frac{\log\left(2\|\Lambda\| N^p/\varepsilon + 1\right)}{N}.
\end{aligned} \tag{3.71}$$

Now, (3.67), (3.68) and (3.70) imply

$$\begin{aligned}\mathbb{E}\left[X_N(M,\varepsilon)\right] &\leq \int_0^{+\infty} \mathbb{P}\left\{X_N(M,\varepsilon) > t\right\} \mathrm{d}t \\ &\leq \int_\varepsilon^{+\infty} \mathbb{P}\left\{X_N(M,\varepsilon) > t\right\} \mathrm{d}t + \varepsilon \\ &\leq \int_\varepsilon^{+\infty} \mathbb{P}\{\gamma_N(M) > -K + t\} \mathrm{d}t \\ &\quad + (2\|\Lambda\| N^p/\varepsilon + 1)\, q \int_\varepsilon^{+\infty} \exp\left(-\frac{Nt^2}{L}\right) \mathrm{d}t + \varepsilon. \end{aligned} \tag{3.72}$$

Therefore, taking sequentially $\overline{\lim}_{N\uparrow+\infty}$, $\lim_{K\uparrow+\infty}$ and $\lim_{\varepsilon\uparrow+0}$ in (3.72), we arrive at

$$\overline{\lim_{N\uparrow\infty}} \mathbb{E}\left[X_N(M,\varepsilon)\right] \leq 0. \tag{3.73}$$

The bound (3.73) together with (3.71) implies the assertion of the lemma.

□

Let $\hat{Q}_{N,\Lambda}$ be the (random) probability measure defined by

$$\hat{Q}_{N,\Lambda} := \frac{\widetilde{Q}_N}{L_N(\Lambda)}.$$

Lemma 3.3.4. *Suppose that the measure Q_N satisfies the assumptions of the previous lemma. Then* (3.61) *is valid also for $\hat{Q}_{N,\Lambda}$.*

Proof. Similar to the one of the previous lemma. □

Remark 3.3.2. *Recall that a point $x \in \mathscr{X}$ is called an exposed point of the concave mapping I^* if there exists $\Lambda \in \mathbb{R}^d$ such that, for all $y \in \mathscr{X} \setminus \{x\}$, we have*

$$I^*(y) - I^*(x) < \langle y - x, \Lambda \rangle. \tag{3.74}$$

Theorem 3.3.2. *Suppose*

1. *The family $\left\{Q_N : N \in \mathbb{N}\right\} \subset \mathscr{M}(\mathbb{R}^d)$ satisfies the assumptions of Lemma 3.3.3.*
2. *$\mathscr{G} \subset \mathscr{X}$ is an open set.*
3. *$\emptyset \neq \mathscr{E}(I^*) \subset \mathscr{D}(I^*)$ is the set of the exposed points of the mapping I^*.*
4. *Condition* (3.43) *is satisfied.*

Then

$$\underline{\lim}_{N\uparrow+\infty} \frac{1}{N} \mathbb{E}\left[\log Q_N(\mathscr{G} \cap \mathscr{E})\right] \geq \sup_{x\in\mathscr{G}} I^*(x). \tag{3.75}$$

Proof. Let $B(x,\varepsilon)$ be a ball of radius $\varepsilon > 0$ around some arbitrary $x \in \mathscr{X}$. It suffices to prove that

$$\lim_{\varepsilon\downarrow+0} \underline{\lim}_{N\uparrow\infty} \frac{1}{N} \mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] \geq I^*(x). \tag{3.76}$$

Indeed, since we have

$$Q_N(\mathscr{G}) \geq Q_N(B(x,\varepsilon)), \tag{3.77}$$

applying $\frac{1}{N}\log(\cdot)$, taking the expectation, taking $\underline{\lim}_{N\uparrow+\infty}$, $\varepsilon \downarrow +0$ and taking the supremum over $x \in \mathscr{G}$ in (3.77), we get (3.75).

Take any $x \in \mathscr{G} \cap \mathscr{E}$. Then we can find the corresponding vector $\Lambda_e = \Lambda_e(x) \in \mathbb{R}^d$ orthogonal to the exposing hyperplane at the point x, as in (3.74). Define the new ("tilted") random probability measure $\hat{Q}_N$ on $\mathbb{R}^d$ by demanding that

$$\frac{\mathrm{d}\hat{Q}_N}{\mathrm{d}Q_N}(y) = \frac{1}{L_N(\Lambda_e)}\mathrm{e}^{N\langle y,\Lambda_e\rangle}. \tag{3.78}$$

Moreover, we have

$$\begin{aligned}\frac{1}{N}\mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] &= \frac{1}{N}\mathbb{E}\left[\log\int_{B(x,\varepsilon)}\mathrm{d}Q_N(y)\right]\\ &= \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda_e)\right] + \frac{1}{N}\mathbb{E}\left[\int_{B(x,\varepsilon)}\mathrm{e}^{-N\langle y,\Lambda_e\rangle}\mathrm{d}\hat{Q}_N(y)\right]\\ &\geq \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda_e)\right] - \langle x,\Lambda_e\rangle - \varepsilon\|\Lambda_e\|_2 + \frac{1}{N}\mathbb{E}\left[\log\hat{Q}_N(B(x,\varepsilon))\right].\end{aligned}$$

Hence,

$$\lim_{\varepsilon\downarrow+0}\underline{\lim}_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] \geq \left[-\langle x,\Lambda_e\rangle + I(\Lambda_e)\right] + \lim_{\varepsilon\downarrow+0}\underline{\lim}_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log\hat{Q}_N(B(x,\varepsilon))\right].$$

Since we have

$$-\langle x,\Lambda_e\rangle + I(\Lambda_e) \geq I^*(x),$$

in order to show (3.76) it remains to prove that

$$\lim_{\varepsilon\downarrow+0}\underline{\lim}_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log\hat{Q}_N(B(x,\varepsilon))\right] = 0. \tag{3.79}$$

The Laplace transform of $\hat{Q}_N$ is

$$\hat{L}_N(\Lambda) = \frac{L_N(\Lambda+\Lambda_e)}{L_N(\Lambda_e)}.$$

Hence, we arrive at

$$\hat{I}(\Lambda) = I(\Lambda+\Lambda_e) - I(\Lambda_e).$$

Moreover, we have

$$\hat{I}^*(x) = I^*(x) + \langle x,\Lambda_e\rangle - I(\Lambda_e). \tag{3.80}$$

By the assumptions of the theorem, the family Q_N satisfies the assumptions of Lemma 3.3.3. Hence, due to Lemma 3.3.4, the family $\hat{Q}_N$ satisfies (3.38). Thus we can apply Theorem 3.3.1 to obtain

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(\mathbb{R}^d \setminus B(U,\varepsilon))\right] \leq \sup_{y\in\mathscr{U}\setminus B(x,\varepsilon)} \hat{I}^*(y). \tag{3.81}$$

Lemma 3.3.2 implies that there exists some $x_0 \in \mathscr{X}\setminus B(x,\varepsilon)$ (note that $x_0 \neq x$) such that

$$\sup_{y\in\mathscr{X}\setminus B(x,\varepsilon)} \hat{I}^*(y) = \hat{I}^*(x_0).$$

Since Λ_e is an exposing hyperplane, using (3.80), we get

$$\begin{aligned} \hat{I}^*(x_0) &= I^*(x_0) + \langle x_0, \Lambda_e\rangle - I(\Lambda_e) \\ &\leq [I^*(x_0) + \langle x_0, \Lambda_e\rangle] - [I^*(x) + \langle x, \Lambda_e\rangle] < 0, \end{aligned} \tag{3.82}$$

and hence, combining (3.81) and (3.82), we get

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon))\right] < 0.$$

Therefore, due to the concentration of measure, we have almost surely

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\log \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon)) < 0$$

which implies that, for all $\varepsilon > 0$, we have almost surely

$$\lim_{N\uparrow+\infty} \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon)) = 0,$$

and (3.79) follows by yet another application of the concentration of measure.

□

Corollary 3.3.1. *Suppose that in addition to the assumptions of previous Theorem 3.3.2 we have*

1. $I(\cdot)$ *is differentiable on* $\operatorname{int}\mathscr{D}(I)$.
2. *Either* $\mathscr{D}(I) = \mathscr{X}$ *or*

$$\lim_{\Lambda\to\partial\mathscr{D}(I)} \|\nabla I(\Lambda)\| = +\infty.$$

Then $\mathscr{E}(I^*) = \mathbb{R}^d$, *consequently*

$$\underline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{G})\right] \geq \sup_{x\in\mathscr{G}} I^*(x).$$

Proof. The proof is the same as in the classical Gärtner-Ellis theorem (see, e.g., den Hollander (2000)).

□

3.4 The Aizenman-Sims-Starr comparison scheme

In this section, we shall extend the AS^2 scheme to the case of the SK model with multidimensional spins and prove the Theorems 3.1.1 and 3.1.2, as stated in Section 3.1. We use the Gaussian comparison results of Section 3.2.3 in the spirit of AS^2 scheme in order to relate the free energy of the SK model with multidimensional spins with the free energy of a certain GREM-inspired model. Comparing to Aizenman *et al.* (2003), due to more intricate nature of spin configuration space, some new effects occur. In particular, the remainder term of the Gaussian comparison non-trivially depends on the variances and covariances of the Hamiltonians under comparison. To deal with this obstacle, we use the idea of localisation to the configurations having a given overlap (cf. (3.4)). This idea is formalised by adapting the proofs of the quenched Gärtner-Ellis type LDP obtained in Section 3.3.

3.4.1 Naive comparison scheme

We start by recalling the basic principles of the AS^2 comparison scheme (see, e.g., (Bovier, 2006, Chapter 11)). It is a simple idea to get the comparison inequalities by adding some additional structure into the model. However, the way the additional structure is attached to the model might be suggested by the model itself. Later on we shall encounter a real-world use of this trick. Let $(\Sigma, \mathfrak{S})$ and $(\mathscr{A}, \mathfrak{A})$ be Polish spaces equipped with measures μ and ξ, respectively. Furthermore, let

$$X := \{X(\sigma)\}_{\sigma\in\Sigma}, A := \{A(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma, \\ \alpha\in\mathscr{A}}}, B := \{B(\sigma)\}_{\alpha\in\mathscr{A}}$$

be independent real-valued Gaussian processes. Define the *comparison functional*

$$\Phi[C] := \mathbb{E}\left[\log \int_{\Sigma\times\mathscr{A}} \mathrm{e}^{C(\sigma,\alpha)} \mathrm{d}\,(\mu\otimes\xi)\,(\sigma,\alpha)\right], \tag{3.83}$$

where $C := \{C(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma \\ \alpha\in\mathscr{A}}}$ is a suitable real-valued Gaussian process. Theorem 4.1 of Aizenman *et al.* (2007) is easily understood as an example of the following observation. Suppose $\Phi[X]$ is somehow hard to compute directly, but $\Phi[A]$ and $\Phi[B]$ are manageable. We always have the following additivity property

$$\Phi\,[X+B] = \Phi\,[X] + \Phi\,[B]. \tag{3.84}$$

Assume now that

$$\Phi\,[X+B] \le \Phi\,[A] \tag{3.85}$$

which we can obtain, e.g., from Proposition 3.2.5. Combining (3.84) and (3.85), we get the bound

$$\Phi\,[X] \le \Phi\,[A] - \Phi\,[B]. \tag{3.86}$$

3.4.2 Free energy upper bound

Let $\mathscr{V} \subset \mathrm{Sym}(d)$ be an arbitrary Borell set.

Remark 3.4.1. *Note that $\mathscr{U}$ is closed and convex.*

Let

$$\begin{aligned}\Sigma_N(\mathscr{V}) &:= \left\{\sigma \in \Sigma_N : R_N(\sigma,\sigma) \in \mathscr{V}\right\} \\ &= \left\{\sigma \in \Sigma_N : R_N(\sigma,\sigma) \in \mathscr{V} \cap \mathscr{U}\right\}.\end{aligned} \tag{3.87}$$

Let us define the *local comparison functional* $\Phi_N(x,\mathscr{V})$ as follows (cf. (3.83))

$$\Phi_N(x,\mathscr{V})[C] := \frac{1}{N}\mathbb{E}\left[\log \pi_N\left[\mathbb{1}_{\Sigma_N(\mathscr{V})}\exp\left(\beta\sqrt{N}C\right)\right]\right], \tag{3.88}$$

where $C := \{C(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma\\ \alpha\in\mathscr{A}}}$ is a suitable Gaussian process. Let us consider the following family ($N \in \mathbb{N}$) of random measures on the Borell subsets of $\mathrm{Sym}(d)$ generated by the SK Hamiltonian,

$$P_N(\mathscr{V}) := \int_{\Sigma_N(\mathscr{V})} \mathrm{e}^{\beta\sqrt{N}X_N(\sigma)}\mathrm{d}\mu^{\otimes N}(\sigma),$$

and consider also the following family of the random measures generated by the Hamiltonian $A(\sigma,\alpha)$

$$\widetilde{P}_N(\mathscr{V}) := \widetilde{P}_N^{x,\mathscr{Q},U}(\mathscr{V}) := \int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} \exp\left(\beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right)\mathrm{d}\pi_N(\sigma,\alpha), \tag{3.89}$$

where the parameters $\mathscr{Q}$ and U are taken from the definition of the process $A(\alpha)$ (cf. (3.6)). The vector x defines the random measure $\xi \in \mathscr{M}(\mathscr{A})$ (cf. (3.7)), and, hence, also the measure $\pi_N \in \mathscr{M}(\Sigma\times\mathscr{A})$.

Remark 3.4.2. *To lighten the notation, most of the time we shall not indicate explicitly the dependence of the following quantities on the parameters x, $\mathscr{Q}$, U.*

Consider (if it exists) the Laplace transform of the measure (3.89)

$$\widetilde{L}_N(\Lambda) := \int_{\mathscr{U}} \mathrm{e}^{N\langle U,\Lambda\rangle}\mathrm{d}\widetilde{P}_N(U). \tag{3.90}$$

Let (if it exists)

$$\widetilde{I}(\Lambda) := \lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log\widetilde{L}_N(\Lambda)\right]. \tag{3.91}$$

Define the following Legendre transform

$$\widetilde{I}^*(U) := \inf_{\substack{x\in\mathscr{Q}'(1,1),\\ \mathscr{Q}\in\mathscr{Q}'(U,d),\\ \Lambda\in\mathrm{Sym}(d)}}\left[-\langle U,\Lambda\rangle - \Phi_N(x,\mathscr{V})[B] + \widetilde{I}(\Lambda)\right]. \tag{3.92}$$

Denote, for $\delta > 0$,

$$\widetilde{I}^*_\delta(U) := \max\left\{\widetilde{I}^*(U) + \delta, -\frac{1}{\delta}\right\}.$$

Let

$$p(\mathscr{V}) := \lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V})\right]. \tag{3.93}$$

Remark 3.4.3. *Note that the result of Guerra & Toninelli (2003) assures the existence of the limit in the previous formula.*

Lemma 3.4.1. *We have*

1. *The Laplace transform* (3.90) *exists. Moreover, for any* $\Lambda \in \mathrm{Sym}(d)$*, we have*

$$\int_{\mathscr{V}} \mathrm{e}^{N\langle U,\Lambda\rangle} \mathrm{d}P_N(U) = \int_{\Sigma_N(\mathscr{V})} \exp\left(N\langle \Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}X(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma), \tag{3.94}$$

$$\int_{\mathscr{V}} \mathrm{e}^{N\langle U,\Lambda\rangle} \mathrm{d}\widetilde{P}_N(U) = \int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} \exp\left(N\langle \Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma,\alpha). \tag{3.95}$$

2. *The quenched cumulant generating function* (3.91) *exists in the* $N\uparrow\infty$ *limit, for any* $\Lambda \in \mathrm{Sym}(d)$*. Moreover, for all* $N\in\mathbb{N}$*, we have*

$$I_N(\Lambda) := \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda)\right] = X_0(x,\mathscr{Q},\Lambda,U), \tag{3.96}$$

that is $I_N(\cdot)$ *in fact does not depend on* N*.*

Proof. 1. We prove (3.95), the proof of (3.94) is similar. Since $\mathscr{U}$ is a compact, it follows that, for arbitrary $\varepsilon>0$, there exists the following ε-partition of $\mathscr{U}$

$$\mathscr{N}(\varepsilon) = \{\mathscr{V}_r \subset \mathscr{U} : r\in\{1,\ldots,K\}\}$$

such that $\bigcup_r \mathscr{V}_r = \mathscr{U}$, $\mathscr{V}_r\cap\mathscr{V}_s=\emptyset$, $\mathrm{diam}\,\mathscr{V}_r \le \varepsilon$ and pick some $V_r\in\mathrm{int}\,\mathscr{V}_r$, for all $r\neq s$. We denote

$$\widetilde{L}_N(\Lambda,\varepsilon) := \sum_{r=1}^{K} \mathrm{e}^{N\langle\Lambda,V_r\rangle}\int_{\Sigma_N(\mathscr{V}_r)\times\mathscr{A}} \exp\left(\beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right)\mathrm{d}\pi_N(\sigma,\alpha).$$

For small enough ε, we have

$$(1-2N\|\Lambda\|\varepsilon)\,\mathrm{e}^{N\langle\Lambda,R_N(\sigma,\sigma)\rangle} \le \mathrm{e}^{N\langle\Lambda,U\rangle} \le \mathrm{e}^{N\langle\Lambda,R_N(\sigma,\sigma)\rangle}\,(1+2N\|\Lambda\|\varepsilon)\,.$$

Therefore, if we denote

$$\widehat{L}_N(\mathscr{V},\Lambda) := \int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} \exp\left(N\langle \Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma,\alpha),$$

we get

$$(1-2N\|\Lambda\|\varepsilon)\sum_{r=1}^{K}\widehat{L}_N(\mathscr{V}_r,\Lambda) \le \widetilde{L}_N(\Lambda,\varepsilon) \le (1+2N\|\Lambda\|\varepsilon)\sum_{r=1}^{K}\widehat{L}_N(\mathscr{V}_r,\Lambda).$$

Hence,

$$(1-2N\|\Lambda\|\varepsilon)\widehat{L}_N(\mathscr{U},\Lambda) \le \widetilde{L}_N(\Lambda,\varepsilon) \le (1+2N\|\Lambda\|\varepsilon)\widehat{L}_N(\mathscr{U},\Lambda). \tag{3.97}$$

Let $\varepsilon\downarrow +0$ in (3.97) and we arrive at

$$\widetilde{L}_N(\Lambda) = \widehat{L}_N(\mathscr{U},\Lambda).$$

That is, the existence of $L_N(\Lambda)$ and the representation (3.95) are proved.

2. For all $N \in \mathbb{N}$, we have, by the RPC averaging property (see, e.g., (Aizenman *et al.*, 2007, Theorem 5.4) or Theorem 4.1.3, property (4) below), that

$$\frac{1}{N}\mathbb{E}\left[\log \widetilde{L}_N(\mathscr{U},\Lambda)\right] = \Phi_N(x,\mathscr{U})\left[A + N\langle \Lambda, R_N(\sigma,\sigma)\rangle\right] = X_0(x,\mathscr{Q},\Lambda,U).$$

□

Proof of Theorem 3.1.1. In essence, the proof follows almost literally the proof of Theorem 3.3.1. The notable difference is that we apply the Gaussian comparison inequality (Proposition 3.2.5) in order to "compute" the rate function in a somewhat more explicit way.

Due to (3.87), we can without loss of generality suppose that $\mathscr{V}$ is compact. For any $\delta > 0$ and $U \in \mathscr{V}$, by (3.92), there exists $\Lambda(U,\delta) \in \mathrm{Sym}(d)$, $x(U,\delta) \in \mathscr{Q}'(1,1)$ and $Q(U,\delta) \in \mathscr{Q}'(U,d)$ such that

$$-\langle U, \Lambda(U)\rangle + \widetilde{I}(\Lambda(U)) \leq \widetilde{I}^*_\delta(U). \tag{3.98}$$

For any $U \in \mathscr{V}$, there exists an open neighbourhood $\mathscr{V}(U) \subset \mathrm{Sym}(d)$ of U such that

$$\sup_{V \in \mathscr{V}(U)} \langle V - U, \Lambda(U)\rangle \leq \delta.$$

Fix some $\varepsilon > 0$. Without loss of generality, we can suppose that all the neighbourhoods satisfy additionally the condition $\operatorname{diam}\mathscr{V}(U) \leq \varepsilon$. By compactness, the covering $\bigcup_{U\in\mathscr{V}} \mathscr{V}(U) \supset \mathscr{V}$ has a finite subcovering, say $\bigcup_{r=1}^{p} \mathscr{V}(U^{(r)}) \supset \mathscr{V}$. We denote the corresponding to this covering approximants in (3.98) by $\{x^{(r)} \in \mathscr{Q}'(1,1)\}_{r=1}^{p}$ and $\{\mathscr{Q}^{(r)} \in \mathscr{Q}'(U^{(r)},d)\}_{r=1}^{p}$. We have

$$\frac{1}{N}\log P_N(\mathscr{V}) \leq \frac{1}{N}\log\left(\bigcup_{r=1}^{p} P_N(\mathscr{V}(U^{(r)}))\right). \tag{3.99}$$

Due to the concentration of measure Proposition 3.2.3, we can apply Lemma 3.3.1 and get

$$\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\left|\log P_N\left(\bigcup_{r=1}^{p}\mathscr{V}(U^{(r)})\right) - \max_{r\in\{1,\dots,p\}}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right]\right|\right] = 0. \tag{3.100}$$

In fact, since we know that (3.93) exists, (3.100) implies that

$$\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N\left(\bigcup_{r=1}^{p}\mathscr{V}(U^{(r)})\right)\right] = \max_{r\in\{1,\dots,p\}}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right]. \tag{3.101}$$

For $U^{(r)}$, $x = x^{(r)}$, $\mathscr{Q} = \mathscr{Q}^{(r)}$, Proposition 3.2.5 gives

$$\begin{aligned}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right] &= \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(\mathscr{V}(U^{(r)}))\right] - \Phi_N(x,\mathscr{U})[B] \\ &\quad + \mathscr{R}_N(x^{(r)},\mathscr{Q}^{(r)},U^{(r)},\mathscr{V}(U^{(r)})) + \mathscr{O}(\varepsilon) \\ &\leq \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(\mathscr{V}(U^{(r)}))\right] - \Phi_N(x,\mathscr{U})[B] + K\varepsilon,\end{aligned} \tag{3.102}$$

where $K > 0$ is an absolute constant.

By the Chebyshev inequality and Lemma 3.4.1, we have

$$\begin{aligned}\widetilde{P}_N(\mathscr{V}(U)) &\le \widetilde{P}_N\{V \in \mathscr{U} : \langle V-U, \Lambda(U)\rangle \le \delta\} \\ &\le \mathrm{e}^{-\delta N}\int_{\mathscr{U}} \mathrm{e}^{N\langle V-U,\Lambda(U)\rangle}\mathrm{d}\widetilde{P}_N(V) \\ &= \mathrm{e}^{-\delta N}\mathrm{e}^{-N\langle U,\Lambda(U)\rangle}\widetilde{L}_N(\Lambda(U)).\end{aligned}$$

Thus, using (3.102) and (3.98), we get

$$\begin{aligned}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right] &\le \lim_{N\uparrow+\infty}\left[-\langle U^{(r)},\Lambda(U^{(r)})\rangle - \Phi[B] + \frac{1}{N}\log\widetilde{L}_N(\Lambda(U^{(r)}))\right] - \delta + K\varepsilon \\ &= -\langle U_r,\Lambda(U_r)\rangle - \Phi[B] + \widetilde{I}(\Lambda(U_r)) - \delta + K\varepsilon \\ &\le \widetilde{I}^*_\delta(U_r) - \delta + K\varepsilon. \end{aligned} \tag{3.103}$$

Combining (3.99), (3.47), (3.103), we obtain

$$\begin{aligned}p(\mathscr{V}) = \lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V})\right] &\le \max_{r\in\{1,\dots,p\}}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right] \\ &\le \max_{r\in\{1,\dots,p\}}\widetilde{I}^*_\delta(U^{(r)}) + K\varepsilon - \delta \\ &\le \sup_{U\in\mathscr{V}}\widetilde{I}^*_\delta(V) + K\varepsilon - \delta.\end{aligned}$$

Taking $\delta \downarrow +0$ and $\varepsilon \downarrow +0$ limits, we get

$$p(\mathscr{V}) \le \sup_{V\in\mathscr{V}}\widetilde{I}^*(U). \tag{3.104}$$

The averaging property of the RPC (see, e.g., (Aizenman *et al.*, 2007, Theorem 5.4) or property (4) of Theorem 4.1.3) gives

$$\Phi_N(x,\mathscr{U})[B] = \frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right). \tag{3.105}$$

To finish the proof it remains to show that, for any fixed $\Lambda \in \mathrm{Sym}(d)$, we have

$$\widetilde{I}(\Lambda) = X_0(x,\mathscr{Q},\Lambda,U)$$

which is assured by Lemma 3.4.1. □

3.4.3 Free energy lower bound

In this subsection, we return to the notations of Section 3.4.2.

Lemma 3.4.2. *For any $\mathscr{B} \subset \mathrm{Sym}(d)$ such that $\mathrm{int}\,\mathscr{B} \cap \mathrm{int}\,\mathscr{U} \neq \emptyset$ there exists $\Delta \subset \Sigma$ with $\mathrm{int}\,\Delta \neq \emptyset$ such that*

$$\begin{aligned}\lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\int_{\Sigma_N(\mathscr{B})\times\mathscr{A}}\exp\left(N\langle\Lambda,R_N(\sigma,\sigma)\rangle + \sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right)\mathrm{d}\pi_N(\sigma,\alpha)\right] \\ \ge \log\int_\Delta \exp\left(\langle(\beta^2U+\Lambda)\sigma,\sigma\rangle\right)\mathrm{d}\mu(\sigma) > -\infty. \end{aligned}\tag{3.106}$$

Proof. In view of (3.9), iterative application of the Jensen inequality with respect to $\mathbb{E}_{z^{(k)}}$ leads to the following

$$\mathbb{E}\left[X_{n+1}(x,\mathscr{Q},\Lambda,U)\right] \leq X_0(x,\mathscr{Q},\Lambda,U).$$

Performing the Gaussian integration, we get

$$\mathbb{E}\left[X_{n+1}(x,\mathscr{Q},\Lambda,U)\right] \geq \log \int_{\Delta} \exp\left(\langle(\beta^2 U+\Lambda)\sigma,\sigma\rangle\right) \mathrm{d}\mu(\sigma),$$

where $\Delta \subset \Sigma$ is such that $\mu(\Delta) > 0$ and $\{R(\sigma,\sigma) : \sigma \in \Delta^N\} \subset \mathscr{B}$. □

Define the following Legendre transform

$$\widehat{I}^*(U) := \inf_{\substack{x\in\mathscr{Q}'(1,1),\\ \mathscr{Q}\in\mathscr{Q}'(U,d),\\ \Lambda\in\mathrm{Sym}(d)}} \left[-\langle U,\Lambda\rangle - \Phi[B] + \widetilde{I}(\Lambda) + \mathscr{R}(x,\mathscr{Q},U)\right]. \tag{3.107}$$

Proof of Theorem 3.1.2. As it is the case with the proof of Theorem 3.1.1, this proof also follows in essence almost literally the proof of Theorem 3.3.2. The notable difference is that we apply the Gaussian comparison in order to “compute” the rate function in a somewhat more explicit way.

In notations of Theorem 3.3.2, we are in the following situation: $\mathscr{X} := \mathrm{Sym}(d)$ and $\mathfrak{X}$ is the topology induced by any norm on $\mathrm{Sym}(d)$.

Let $B(U,\varepsilon)$ be the ball (in the Hilbert-Schmidt norm) of radius $\varepsilon > 0$ around some arbitrary $U \in \mathscr{V}$. Let us prove at first that

$$\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log P_N(B(U,\varepsilon))\right] \geq \widehat{I}^*(U). \tag{3.108}$$

Similarly to (3.102), for any $(x,\mathscr{Q})$, we have

$$\begin{aligned}\mathbb{E}&\left[\frac{1}{N}\log P_N(B(U,\varepsilon))\right]\\ &= \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(B(U,\varepsilon))\right] - \Phi[B] + \mathscr{R}_N(x,\mathscr{Q},U,B(U,\varepsilon)) + \mathscr{O}(\varepsilon).\end{aligned} \tag{3.109}$$

The random measure $\widetilde{P}_N$ satisfies the assumptions of Corollary 3.3.1. Indeed:

1. Due to representation (3.96), mapping $I(\cdot)$ is differentiable with respect to Λ. Henceforth assumption (1) of the corollary is also fulfilled.
2. Let us note at first that, thanks to Proposition 3.2.3, we have $\mathscr{D}(I) = \mathbb{R}^d$. Thus, the assumption (2) of Corollary 3.3.1 is satisfied, as is condition (3.43).

Moreover, the assumptions of Lemma 3.3.3 are satisfied:

1. The concentration of measure condition is satisfied due to Proposition 3.2.3.
2. The tail decay is obvious since the family $\{\widetilde{P}_N : N \in \mathbb{N}\}$ has compact support. Namely, for all $N \in \mathbb{N}$, we have $\operatorname{supp}\widetilde{P}_N = \mathscr{U}$. Thus the measure $\widetilde{Q}_{N,\Lambda}$ (cf. (3.59)) generated by $\widetilde{P}_N$ has the same support. Thus, $\operatorname{supp}\widetilde{Q}_{N,\Lambda} = \mathscr{U}$.

3. The non-degeneracy is assured by Lemma 3.4.2.

Hence, due to (3.109), arguing in the same way as in Theorem 3.3.2, we arrive at (3.108). Note that the $N \uparrow +\infty$ limit of $\mathscr{R}_N(x, \mathscr{Q}, U, B(U, \varepsilon))$ exists, since in (3.109) the limits of the other two N-dependent quantities exist due to Guerra & Toninelli (2003). The subsequent $\varepsilon \downarrow +0$ limit of the remainder term exists due to the monotonicity.

Finally, taking the supremum over $U \in \mathscr{V}$ in (3.108), we get (3.20). □

4

Estimates of Guerra's remainder term in the SK model with multidimensional spins

In this chapter, we extend the comparison scheme of Guerra (2003) to the case of the SK model with multidimensional spins. We provide a short derivation of the remainder term which is a by-product of this scheme (Theorem 4.1.4). We use the well-known properties of the RPC and Bolthausen-Sznitman coalescent. This gives the remainder term a clear meaning in terms of the averages with respect to the disorder induced by a change of measure. The change of measure is induced by the reweighting of the RPC by means of the GREM-inspired process (cf. Theorem 2.2.7).

We study the properties of the multidimensional Parisi functional by establishing a link between the (multidimensional) Parisi functional and a certain class of non-linear PDEs, see Propositions 4.2.1, 4.2.2 and Theorem 4.2.2. We extend the Parisi functional to a continuous functional on a compact space (Theorems 4.2.1, 4.2.2). We show that the class of PDEs corresponds to the Hamilton-Jacobi-Bellman equations induced by a linear problem of diffusion control (Proposition 4.2.4). Motivated by a problem posed by Talagrand (2006c), we show the strict convexity of the local Parisi functional in some cases (Theorem 4.2.4).

We partially extend Talagrand's methodology of estimating the remainder term to the multidimensional setting (Proposition 4.3.1, Theorem 4.3.1).

4.1 Guerra's comparison scheme

For the SK model, Guerra's scheme gave historically the first way to obtain the variational upper bound on the free energy in terms of the Parisi functional. The scheme is based on the comparison between two Gaussian processes: the first one being the original SK Hamiltonian (0.8) and the other one being a carefully chosen GREM inspired process indexed by $\sigma \in \Sigma_N$. The second important ingredient is a recursively defined non-linear comparison functional acting on the Gaussian processes indexed by $\sigma \in \Sigma_N$.

Talagrand (2006b) using Guerra's scheme and the wealth of other ingenious analytical insights showed that the variational upper bound is also the lower bound for the free energy in the SK model. This established, hence, the remaining half of the Parisi formula.

In this section, we shall apply Guerra's comparison scheme (see the recent accounts by Aizenman *et al.* (2007); Guerra (2005); Talagrand (2007a)) to the SK model with multidimensional spins. However, we shall use also the ideas (and the language) of Aizenman *et al.* (2003). In particular, we shall use the same local comparison functional (3.88) as in the AS^2 scheme, see (4.2). The section contains the proofs of the upper (4.14) and lower (4.16) bounds on the free energy without Assumption 3.1.2. The proofs use the GREM-like Gaussian processes, RPCs as in the AS^2 scheme. We also obtain an analytic representation of the remainder term (which is an artifact of this scheme) using the properties of the Bolthausen-Sznitman coalescent.

4.1.1 Multidimensional Guerra's scheme

Given $n \in \mathbb{N}$, assume that x and $\mathscr{Q}$ satisfy (3.7) and (3.6), respectively. Recall the definitions of the Gaussian processes X and A which satisfy (0.8) and (3.15), respectively. We consider, for $t \in [0;1]$, the following interpolating Hamiltonian on the configuration space $\Sigma_N \times \mathscr{A}$

$$H_t(\sigma,\alpha) := \sqrt{t}X(\sigma) + \sqrt{1-t}A(\sigma,\alpha). \tag{4.1}$$

Given $\mathscr{U} \subset \mathrm{Sym}^+(d)$, the Hamiltonian (4.1) in the usual way induces the following *local free energy*

$$\varphi_N(t,x,Q,\mathscr{U}) := \Phi_N(x,\mathscr{U})\,[H_t], \tag{4.2}$$

where we use the same local comparison functional (3.88) as in the AS^2 scheme. Using (3.4), we obtain then

$$\varphi(0,x,Q,\mathscr{U}) = \Phi_N(x,\mathscr{U})[A] \text{ and } \varphi(1,x,Q,\mathscr{U}) = \Phi_N(x,\mathscr{U})[X] = p_N(\mathscr{U}).$$

Now, we are going to disintegrate the Gibbs measure defined on $\mathscr{U} \times \mathscr{A}$ into two Gibbs measures acting on $\mathscr{U}$ and $\mathscr{A}$ separately. For this purpose we define the correspondent (random) *local free energy* on $\mathscr{U}$ as follows

$$\psi(t,x,Q,\alpha,\mathscr{U}) := \log \int_{\Sigma_N(\mathscr{U})} \exp\left[\beta\sqrt{N}H_t(\sigma,\alpha)\right] \mathrm{d}\mu^{\otimes N}(\sigma). \tag{4.3}$$

For $\alpha \in \mathscr{A}$, we can define the (random) *local Gibbs measure* $\mathscr{G}(t,Q,\alpha,\mathscr{U}) \in \mathscr{M}_1(\Sigma_N)$ by demanding that the following holds

$$\frac{\mathrm{d}\mathscr{G}(t,x,Q,\alpha,\mathscr{U})}{\mathrm{d}\mu^{\otimes N}}(\sigma) := \mathbb{1}_{\Sigma_N(\mathscr{U})}(\sigma)\exp\left[\beta\sqrt{N}H_t(\sigma,\alpha) - \psi(t,x,Q,\mathscr{U},\alpha)\right].$$

Let us define a certain reweighting of the RPC ξ with the help of (4.3). We define the random point process $\{\tilde{\xi}\}_{\alpha\in\mathscr{A}}$ in the following way

$$\tilde{\xi}(\alpha) := \xi(\alpha)\exp(\psi(t,x,Q,\mathscr{U},\alpha)).$$

We also define the *normalisation operation* $\mathscr{N} : \mathscr{M}_{\mathrm{f}}(\mathscr{A}) \to \mathscr{M}_1(\mathscr{A})$ as

$$\mathscr{N}(\xi)(\alpha) := \frac{\xi(\alpha)}{\sum_{\alpha'\in\mathscr{A}} \xi(\alpha')}.$$

We introduce the *local Gibbs measure* $\mathscr{G}(t,x,Q,\mathscr{U}) \in \mathscr{M}_1(\mathscr{U} \times \mathscr{A})$, for any $\mathscr{V} \subset \mathscr{U} \times \mathscr{A}$, as follows

$$\mathscr{G}(t,x,Q,\mathscr{U})\,[\mathscr{V}] := \sum_{\alpha\in\mathscr{A}_n} \mathscr{N}(\tilde{\xi})(\alpha)\mathscr{G}(t,x,Q,\alpha,\mathscr{U})\,[\mathscr{V}]. \tag{4.4}$$

Finally, we introduce, what shall call *Guerra's remainder term*:

$$\mathscr{R}(t,x,Q,\mathscr{U}) := -\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|R(\sigma^1,\sigma^2) - Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right]\right]. \tag{4.5}$$

Note that (4.5) coincides with (3.18) after substituting (3.16) with (3.21).

4.1.2 Local comparison

The results of Section 3.4 can be straightforwardly generalised to the comparison scheme based on (4.1). Given $\varepsilon, \delta > 0$ and $\Lambda \in \mathrm{Sym}(d)$, define

$$\mathscr{V}(\Lambda, \mathscr{U}, \varepsilon, \delta) := \{U' \in \mathrm{Sym}(d) : \|U' - U\|_{\mathrm{F}} < \varepsilon, \langle U' - U, \Lambda\rangle < \delta\}. \tag{4.6}$$

We now specialise to the case $\mathscr{U} = \Sigma_N(\mathscr{V}(\Lambda, U, \varepsilon, \delta))$.

Lemma 4.1.1. *We have*

$$\begin{aligned}\frac{\partial}{\partial t}\varphi_N(t,x,Q,\mathscr{V}(\Lambda,U,\varepsilon,\delta)) =& \mathscr{R}(t,x,Q,\Sigma_N(\mathscr{A}(\Lambda,U,\varepsilon,\delta))) \\ & - \frac{\beta^2}{2}\sum_{k=1}^{n} x_k \left(\|Q^{(k+1)}\|_F^2 - \|Q^{(k)}\|_F^2\right) + \mathscr{O}(\varepsilon).\end{aligned} \tag{4.7}$$

Proof. This is an immediate consequence of Proposition 3.2.5. Indeed, recalling that $Q(\alpha^1, \alpha^1) = U$, and setting $\mathscr{U} := \Sigma(B(U,\varepsilon))$, we have

$$\begin{aligned}&\frac{\partial}{\partial t}\varphi(t,x,Q,\mathscr{U}) \\ &= \frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|R(\sigma^1,\sigma^1) - U\|_{\mathrm{F}}^2 - \|R(\sigma^1,\sigma^2) - Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right.\right. \\ &\quad \left.\left. - \left(\|U\|_{\mathrm{F}}^2 - \|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right)\right]\right] \\ &= -\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|R(\sigma^1,\sigma^2) - Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right]\right] \\ &\quad - \frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|U\|_{\mathrm{F}}^2 - \|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right]\right] + \mathscr{O}(\varepsilon).\end{aligned} \tag{4.8}$$

Using the Proposition 2.2.1, we get

$$\begin{aligned}&\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|U\|_{\mathrm{F}}^2 - \|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right]\right] \\ &\quad = \frac{\beta^2}{2}\mathbb{E}\left[\mathscr{N}(\xi)\otimes\mathscr{N}(\xi)\left[\sum_{k=q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})}^{n}\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right)\right]\right] \\ &\quad = \frac{\beta^2}{2}\sum_{k=1}^{n}\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right)\mathbb{E}\left[\mathscr{N}(\xi)\otimes\mathscr{N}(\xi)\{k \geq q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})\}\right] \\ &\quad = \frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right).\end{aligned} \tag{4.9}$$

Combining (4.8) and (4.9), we get (4.7)

□

Lemma 4.1.2. *We have*

$$\begin{aligned}p_N(\Sigma_N(B(U,\varepsilon))) =& \Phi_N(x, \Sigma_N(B(U,\varepsilon)))[A] - \frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_F^2 - \|Q^{(k)}\|_F^2\right) \\ & + \int_0^1 \mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon))\mathrm{d}t + \mathscr{O}(\varepsilon).\end{aligned} \tag{4.10}$$

Remark 4.1.1. *Note that the above lemma also holds if we substitute $B(U,\varepsilon)$ with the smaller set $\mathscr{V}(\Lambda,U,\varepsilon,\delta)$.*

Proof. The claim follows from (4.7) by integration. □

Proposition 4.1.1. *There exists $C = C(\Sigma,\mu) > 0$ such that, for all $U \in \mathrm{Sym}^+(d)$ as above, and all $\varepsilon,\delta > 0$, there exists an δ-minimal Lagrange multiplier $\Lambda = \Lambda(U,\varepsilon,\delta) \in \mathrm{Sym}(d)$ in* (3.11) *such that, for all $t \in [0;1]$, and all $(x,\mathscr{Q})$, we have*

$$p_N(\Sigma_N(\mathscr{V}(\Lambda,U,\varepsilon,\delta))) \le \inf_{\Lambda\in\mathrm{Sym}(d)} f(x,\mathscr{Q},U,\Lambda) + C(\varepsilon+\delta) \tag{4.11}$$

and

$$\begin{aligned}\lim_{N\uparrow+\infty} p_N(\Sigma_N(B(U,\varepsilon))) \ge & \inf_{\Lambda\in\mathrm{Sym}(d)} f(x,\mathscr{Q},U,\Lambda) + \lim_{N\uparrow+\infty}\int_0^1 \mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon)))\mathrm{d}t \\ & - C(\varepsilon+\delta).\end{aligned} \tag{4.12}$$

Remark 4.1.2. *The following upper bound also holds true. There exists $C = C(\Sigma,\mu) > 0$, such that, for any $\Lambda \in \mathrm{Sym}(d)$,*

$$p_N(\Sigma_N(B(U,\varepsilon))) \le f(x,\mathscr{Q},U,\Lambda) + C\|\Lambda\|_F\varepsilon. \tag{4.13}$$

Proof. The result follows from Lemma 4.1.2 by the same arguments as in the proofs of Theorems 3.1.1 and 3.1.2. □

4.1.3 Free energy upper and lower bounds

Similarly to the quenched LDP bounds for the AS^2 scheme in the SK model with multidimensional spins (see Section 3.3), we get the quenched LDP bounds for Guerra's scheme in the same model without Assumption 3.1.2 on Q.

Recall the definition of the local Parisi functional f (3.11).

Theorem 4.1.1. *For any closed set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \le \sup_{U\in\mathscr{V}\cap\mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{4.14}$$

where the infimum runs over all x satisfying (3.7)*, all $\mathscr{Q}$ satisfying* (3.6) *and all $\Lambda \in \mathrm{Sym}(d)$.*

Proof. The proof is identical to the one of Theorem 3.1.1. □

Define the *local limiting Guerra remainder term $\mathscr{R}(x,\mathscr{Q},U)$* as follows

$$\mathscr{R}(x,\mathscr{Q},U) := -\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow+\infty}\int_0^1 \mathscr{R}(t,\Sigma_N(B(U,\varepsilon)))\mathrm{d}t \le 0. \tag{4.15}$$

The existence of the limits in (4.15) is proved similar to the case of the AS^2 scheme, see the proof of Theorem 3.1.2.

Theorem 4.1.2. *For any open set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \geq \sup_{U \in \mathscr{V} \cap \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} \left[f(x,\mathscr{Q},\Lambda,U) + \mathscr{R}(x,\mathscr{Q},U) \right], \tag{4.16}$$

where the infimum runs over all x satisfying (3.7)*; all $\mathscr{Q}$ satisfying* (3.6) *and all $\Lambda \in \mathrm{Sym}(d)$.*

Proof. The proof is identical to the one of Theorem 3.1.2. The only new ingredient is Lemma 4.1.1 needed to recover Guerra's remainder term (4.5). $\square$

4.1.4 The filtered d-dimensional GREM

Given $U \in \mathrm{Sym}^+(d)$ non-negative definite, denote by $\mathscr{Q}(U,d)$ the set of all càdlàg (right continuous with left limits) $\mathrm{Sym}^+(d)$-valued non-decreasing paths which end in matrix U, i.e.,

$$\mathscr{Q}(U,d) := \{\rho : [0;1] \to \mathrm{Sym}^+(d) \mid \rho(0) = 0; \rho(1) = U; \rho(t) \preceq \rho(s), \text{ for } t \leq s; \rho \text{ is cádlág}\}. \tag{4.17}$$

Definition (4.17) is a multidimensional generalisation of (1.26). Define the natural inverse ρ^{-1} : $\mathrm{Im}\,\rho \to [0;1]$ as

$$\rho^{-1}(Q) := \inf\{t \in [0;1] \mid \rho(t) \succeq Q\},$$

where $\mathrm{Im}\,\rho := \rho([0;1])$. Let $x := \rho^{-1} \circ \rho \in \mathscr{Q}(1,1)$.

Let also $\mathscr{Q}'(U,d) \subset \mathscr{Q}(U,d)$ be the space of all piece-wise constant paths in $\mathscr{Q}(U,d)$ with finite (but arbitrary) number of jumps with an additional requirement that they have a jump at $x = 1$. Given some $\rho \in \mathscr{Q}'(U,d)$, we enumerate its jumps and define the finite collection of matrices $\{Q^{(k)}\}_{k=0}^{n+1} := \mathrm{Im}\,\rho \subset \mathbb{R}^d$. This implies that there exist $\{x_k\}_{k=0}^{n+1} \subset \mathbb{R}$ such that

$$0 =: x_0 < x_1 < \ldots < x_n < x_{n+1} := 1,$$
$$0 =: Q^{(0)} \preceq Q^{(1)} \preceq Q^{(2)} \preceq \cdots \preceq Q^{(n+1)} := U,$$

where $\rho(x_k) = Q^{(k)}$. Let us associate to $\rho \in \mathscr{Q}'(U,d)$ a new path $\tilde{\rho} \in \mathscr{Q}(U,d)$ which is obtained by the linear interpolation of the path ρ. Namely, let

$$\tilde{\rho}(t) := Q^{(k)} + (Q^{(k+1)} - Q^{(k)}) \frac{t - x_k}{x_{k+1} - x_k}, t \in [x_k; x_{k+1}).$$

Let $g : \mathbb{R}^d \to \mathbb{R}$ be a function satisfying Assumption 4.1.1. Let us introduce the *filtered d-dimensional GREM process W*. Let

$$W := \left\{ \{W_k(t,[\alpha]_k)\}_{t \in \mathbb{R}_+} : \alpha \in \mathscr{A}, k \in [0;n] \cap \mathbb{N} \right\}$$

be the collection of independent (for different α and k) $\mathbb{R}^d$-valued correlated Brownian motions satisfying

$$W_k(t,[\alpha]_k) \sim (Q^{(k+1)} - Q^{(k)})^{1/2} W \left(\frac{t - x_k}{x_{k+1} - x_k} \right),$$

where $\{W(t)\}_{t\in\mathbb{R}_+}$ is the standard (uncorrelated) $\mathbb{R}^d$-valued Brownian motion. Now, for $k \in [0;n]\cap\mathbb{N}$, we define the $\mathbb{R}^d$-valued process $\{Y(t,\alpha) \mid \alpha \in \mathscr{A}, t \in [0;1]\}$ by

$$Y(t,\alpha) := \sum_{k=0}^{n} \mathbb{1}_{[x_k;1]}(t) W_k(t \wedge x_{k+1}, [\alpha]_k).$$

Lemma 4.1.3. *For $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}$, we have*

$$\mathrm{Cov}\left[Y(t_1,\alpha^{(1)}), Y(t_2,\alpha^{(2)})\right] = \tilde{\rho}\left(t_1 \wedge t_2 \wedge x_{q_L(\alpha^{(1)},\alpha^{(2)})}\right).$$

Proof. The proof is straightforward. □

Assumption 4.1.1. *Suppose that the function $g : \mathbb{R}^d \to \mathbb{R}$ satisfies $g \in C^{(2)}(\mathbb{R}^d)$ and, for any $c > 0$, we have $\int_{\mathbb{R}^d} \exp\left(g(y) - c\|y\|_2^2\right) \mathrm{d}y < \infty$ and also*

$$\sup_{y\in\mathbb{R}^d} \left(\|\nabla g(y)\|_2 + \|\nabla^2 g(y)\|_2\right) < +\infty, \tag{4.18}$$

where $\nabla^2 g(y)$ denotes the matrix of second derivatives of the function g at $y \in \mathbb{R}^d$.

Assume g satisfies the above assumption. Let $f := f_\rho : [0;1] \times \mathbb{R}^d \to \mathbb{R}$ be the function satisfying the following (backward) recursive definition

$$f(t,y) := \begin{cases} g(y), & t = 1, \\ \frac{1}{x_k} \log \mathbb{E}\left[\exp\left\{x_k f(x_{k+1}, y + Y(x_{k+1},\alpha) - Y(t,\alpha))\right\}\right], & t \in [x_k; x_{k+1}), \end{cases} \tag{4.19}$$

where $k \in [0;n]\cap\mathbb{N}$, $\alpha \in \mathscr{A}$ is arbitrary and fixed.

Remark 4.1.3. *It is easy to recognise that the definition of f is a continuous "algorithmisation" of* (3.10). *Namely, $X_k(x,\mathscr{Q},\Lambda,U) = f(x_k,0)$, where*

$$f(1,y) = g(y) := \log \int_{\Sigma} \exp\left(\sqrt{2}\beta\,\langle y,\sigma\rangle + \langle \Lambda\sigma,\sigma\rangle\right) \mathrm{d}\mu(\sigma). \tag{4.20}$$

4.1.5 A computation of the remainder term

Recall the equivalence relation (2.27). In words, the equivalence $i \underset{k}{\sim} j$ means that the atoms of the RPC ξ with ranks i and j have the same ancestors up to the k-th generation. Varying the k in (2.27), we get a family of equivalences on $\mathbb{N}$ which possesses important Markovian properties, see Bolthausen & Sznitman (1998).

Lemma 4.1.4. *For all $k \in [0;n-1]\cap\mathbb{N}$, we have*

$$\mathbb{E}\left[\sum_{\substack{i \underset{k}{\sim} j \\ i \underset{k+1}{\nsim} j}} \mathscr{N}(\xi)(i)\mathscr{N}(\xi)(j)\right] = x_{k+1} - x_k, \tag{4.21}$$

and also

$$\mathbb{E}\left[\sum_{i} \mathscr{N}(\xi)(i)^2\right] = 1 - x_n. \tag{4.22}$$

Proof. 1. To prove (4.21) we notice that

$$\mathbb{E}\left[\sum_{\substack{i \underset{k}{\sim} j \\ i \underset{k+1}{\nsim} j}} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right] = \mathbb{E}\left[\sum_{i \underset{k+1}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j) - \sum_{i \underset{k}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right]$$
$$= x_{k+1} - x_k,$$

where the last equality is due to Proposition 2.2.1.

2. Similarly, (4.22) follows from the following observation

$$\mathbb{E}\left[\sum_i \mathcal{N}^2(\xi)(i)\right] = \mathbb{E}\left[\sum_{i,j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j) - \sum_{i \underset{n}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right]$$
$$= 1 - x_n,$$

where the last equality is due to Proposition 2.2.1.

□

Note that, using the above notations, we readily have

$$A(\sigma,\alpha) \sim \left(\frac{2}{N}\right)^{1/2} \sum_{i=1}^{N} \langle Y^{(i)}(1,\alpha), \sigma_i \rangle,$$

where $\{Y^{(i)} := \{Y^{(i)}(1,\alpha)\}_{\alpha\in\mathscr{A}}\}_{i=1}^N$ are i.i.d. copies of $\{Y(1,\alpha)\}_{\alpha\in\mathscr{A}}$. Consider the following weights

$$\tilde{\xi}^{(t)}(\alpha) := \xi(\alpha)\exp\left(f(t, Y(t,\alpha))\right).$$

As in (Bolthausen & Sznitman, 1998), the above weights induce the permutation $\tilde{\pi}^{(t)} : \mathbb{N} \to \mathscr{A}$ such that, for all $i \in \mathbb{N}$, the following holds

$$\tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i)) > \tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i+1)). \tag{4.23}$$

In what follows, we shall use the short-hand notations $\tilde{\xi}^{(t)}(i) := \tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i))$, $\tilde{Y}^{(t)}(s,i) := Y(s, \tilde{\pi}^{(t)}(i))$ and $\tilde{Q}^{(t)} := \{\tilde{Q}^{(t)}(i,j) := Q(\tilde{\pi}^{(t)}(i), \tilde{\pi}^{(t)}(j))\}_{i,j\in\mathbb{N}}$.

Theorem 4.1.3. *Given a discrete order parameter $x \in \mathscr{Q}'(1,1)$, we have*

1. Independence #1. *The normalised RPC point process $\mathcal{N}(\xi)$ is independent from the corresponding randomised limiting GREM overlaps q.*
2. Independence #2. *The reordered filtered limiting GREM $\tilde{Y}$ is independent from the corresponding reordered weights $\tilde{\xi}$.*
3. The reordering change of measure. *Given $I \in \mathbb{N}$, let $\nu_I(\cdot|Q)$ be the joint distribution of $\{Y(1,i)\}_{i\in I}$, and $\tilde{\nu}_I(\cdot|Q)$ be the joint distribution of $\{\tilde{Y}^{(1)}(1,i)\}_{i\in I}$ both conditional on Q. Then*

$$\frac{\mathrm{d}\tilde{\nu}_I(\cdot|Q)}{\mathrm{d}\nu_I(\cdot|Q)} = \prod_{k=0}^{n} \prod_{i \in \left(I/\underset{k}{\sim}\right)} \exp\left(x_k \left\{ f(x_{k+1}, Y(x_{k+1}, i)) - f_k(x_k Y(x_k, i)) \right\}\right), \tag{4.24}$$

where the innermost product in the previous formula is taken over all equivalence classes on the index set I induced by the equivalence $\underset{k}{\sim}$.

4. The averaging property. *For all* $s, t \in [0;1]$, *we have*

$$\left(\left\{ \xi^{(t)}(\alpha) \right\}_{\alpha \in \mathscr{A}}, \tilde{Q}^{(t)} \right) \sim \left(\left\{ \xi^{(s)}(\alpha) \right\}_{\alpha \in \mathscr{A}}, \tilde{Q}^{(s)} \right). \tag{4.25}$$

Proof. The proof is the same as in the case of the one-dimensional SK model, see Arguin (2007); Bolthausen & Sznitman (1998). □

Keeping in mind (4.24), we define, for $k \in [0; n-1] \cap \mathbb{N}$, the following random variables

$$T_k(\alpha) := \exp\left(x_k \left[f(x_{k+1}, Y(x_{k+1}, \alpha)) - f(x_k, Y(x_k, \alpha)) \right]\right).$$

Given $k \in [1;n] \cap \mathbb{N}$, assume that $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}$ satisfy $q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)}) = k$. We introduce, for notational convenience, the (random) measure $\mu_k(t, \mathscr{U})$ – an element of $\mathscr{M}_1(\Sigma_N)$ – by demanding the following

$$\begin{aligned} \mu_k(t, \mathscr{U})[g] := \mathbb{E}\big[& T_1(\alpha^1) \cdots T_k(\alpha^1) T_{k+1}(\alpha^1) T_{k+1}(\alpha^2) \cdots T_n(\alpha^1) T_n(\alpha^2) \\ & \mathscr{G}(t, \alpha^{(1)}, \mathscr{U}) \otimes \mathscr{G}(t, \alpha^{(2)}, \mathscr{U})[g] \big], \end{aligned} \tag{4.26}$$

where $g : \mathscr{U}^2 \to \mathbb{R}$ is an arbitrary measurable function such that (4.26) is finite. Using this notation, we can state the following lemma.

Lemma 4.1.5. *For any* $i, j \in \mathbb{N}$, *satisfying* $i \underset{k}{\sim} j$, $i \underset{k+1}{\nsim} j$, *we have*

$$\mathbb{E}\left[\mathscr{G}(t, i, \mathscr{U}) \otimes \mathscr{G}(t, j, \mathscr{U}) \left[\| R(\sigma^1, \sigma^2) - Q(i, j) \|_F^2 \right] \right] = \mu_k(t, \mathscr{U}) \left[\| R(\sigma^1, \sigma^2) - Q^{(k)} \|_F^2 \right]. \tag{4.27}$$

Proof. This is a direct consequence of (4.24) and the fact that under the assumptions of the theorem $Q(i, j) = Q^{(k)}$.

□

Remark 4.1.4. *It is obvious from the previous theorem that* μ_k *is a probability measure.*

The main result of this subsection is an "analytic projection" of the probabilistic RPC representation which integrates out the dependence on the RPC. Comparing to (3.18), it has a more analytic flavor which will be exploited in the remainder estimates (Section 4.3). This is also a drawback in some sense, since the initial beauty of the RPCs is lost.

Theorem 4.1.4. *In the case of Guerra's interpolation* (3.21), *we have*

$$\begin{aligned} \mathscr{R}(t, x, Q, \Sigma_N(B(U, \varepsilon))) = & \frac{1}{2} \sum_{k=0}^{n-1} (x_{k+1} - x_k) \mu_k(t, \Sigma_N(B(U, \varepsilon))) \left[\| R(\sigma^1, \sigma^2) - Q^{(k)} \|_F^2 \right] \\ & + \mathscr{O}(\varepsilon) + \mathscr{O}(1 - x_n), \end{aligned} \tag{4.28}$$

as $\varepsilon \to 0$ *and* $x_n \to 1$.

Proof. Recalling (4.5) and (4.4), we write

$$\mathscr{R}(t,x,Q,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\mathbb{E}\Big[\sum_{i,j}\mathscr{N}(\tilde{\xi})(i)\mathscr{N}(\tilde{\xi})(j) \\ \times \mathscr{G}(t,x,Q,i,\mathscr{U})\otimes\mathscr{G}(t,x,Q,j,\mathscr{U})\Big[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\Big]\Big].$$

Using the Theorem 4.1.3, we arrive to

$$\mathscr{R}(t,x,Q,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\sum_{i,j}\mathbb{E}\left[\mathscr{N}(\tilde{\xi})(i)\mathscr{N}(\tilde{\xi})(j)\right] \\ \times \mathbb{E}\left[\mathscr{G}(t,x,Q,i,\mathscr{U})\otimes\mathscr{G}(t,x,Q,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\right]\right].$$

(We can interchange the summation and expectation since all summands are non-negative.) The averaging property (see Theorem 4.1.3) then gives

$$\mathscr{R}(t,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\sum_{i,j}\mathbb{E}\left[\mathscr{N}(\xi)(i)\mathscr{N}(\xi)(j)\right]\mathbb{E}\left[\mathscr{G}(t,i,\mathscr{U})\otimes\mathscr{G}(t,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\right]\right]. \tag{4.29}$$

For each $k\in[1;n-1]\cap\mathbb{N}$, we fix any indexes $i_0, i_0^{(k)}, j_0^{(k)}\in\mathbb{N}$ such that $i \underset{k}{\sim} j$ and $i \underset{k+1}{\nsim} j$. Rearranging the terms in (4.29), we get

$$\mathscr{R}(t,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\sum_{k=1}^{n}\mathbb{E}\left[\mathscr{G}(t,i_0^{(k)},\mathscr{U})\otimes\mathscr{G}(t,j_0^{(k)},\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q^{(k)}\|_{\mathrm{F}}^2\right]\right] \\ \times \sum_{\substack{i\underset{k}{\sim} j\\ i\underset{k+1}{\nsim} j}}\mathbb{E}\left[\mathscr{N}(\xi)(i)\mathscr{N}(\xi)(j)\right] \\ + \frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,i_0,\mathscr{U})\otimes\mathscr{G}(t,i_0,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-U\|_{\mathrm{F}}^2\right]\right]\sum_i\mathbb{E}\left[\mathscr{N}(\xi)(i)^2\right]. \tag{4.30}$$

Finally, applying Lemmata 4.1.4 and 4.1.5 to (4.30), we arrive at (4.28).

□

4.2 The Parisi functional in terms of differential equations

In this section, we study the properties of the multidimensional Parisi functional. We derive the multidimensional version of the Parisi PDE. This allows to represent the Parisi functional as a solution of a PDE evaluated at the origin. We also obtain a variational representation of the Parisi functional in terms of a HJB equation for a linear problem of diffusion control. As a by-product, we arrive at the strict convexity of the Parisi functional in 1-D which settles a problem of uniqueness of the optimal Parisi order parameter posed by Panchenko (2005a); Talagrand (2006c).

Lemma 4.2.1. *Consider the function $B : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ defined as*

$$B(y,t) := \frac{1}{x} \log \mathbb{E}\left[\exp\left\{xf(y+z(t))\right\}\right],$$

where $f : \mathbb{R}^d \to \mathbb{R}$ satisfies Assumption 4.1.1 and $\{z(t)\}_{t\in[0;1]}$ is a Gaussian $\mathbb{R}^d$-valued process with $\operatorname{Cov}[z(t)] := Q(t) \in \operatorname{Sym}(d)$ such that $Q(t)_{u,v}$ is differentiable, for all u,v. Then

$$\partial_t B(y,t) = \frac{1}{2}\sum_{u,v=1}^{d} \dot{Q}_{u,v}(t)\left(\partial^2_{y_u y_v} B(y,t) + x\partial_{y_u} B(y,t)\partial_{y_v} B(y,t)\right), \quad (t,y) \in (0;1)\times\mathbb{R}^d. \tag{4.31}$$

In particular, the function B is differentiable with respect to the t-variable on $(0;1)$ and $C^2(\mathbb{R}^d)$ with respect to the y-variable.

Proof. Denote $Z := \mathbb{E}\left[\mathrm{e}^{xf(y+z(t))}\right]$. By (Aizenman *et al.*, 2007, Lemma A.1), we have

$$\partial_t B(y,t) = \frac{1}{2x}\left(\frac{1}{Z}\mathbb{E}\Big[\sum_{u,v=1}^{d} \dot{Q}_{u,v}(t)\partial^2_{z_u z_v} \mathrm{e}^{xf(z)}|_{z=y+z(t)}\Big]\right).$$

A straightforward calculation then gives

$$\partial_t B(y,t) = \frac{1}{2x}\left(\frac{1}{Z}\mathbb{E}\Big[\sum_{u,v=1}^{d} \dot{Q}_{u,v}(t)\left(x^2\partial_{z_u} f(z)\partial_{z_v} f(z) + x\partial^2_{z_u z_v} f(z)\right)\mathrm{e}^{xf(z)}|_{z=y+z(t)}\Big]\right). \tag{4.32}$$

We also have

$$\partial_{y_u} B(y,t) = \frac{1}{xZ}\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_u} f(z)|_{z=y+z(t)}\right], \tag{4.33}$$

and

$$\begin{aligned}\partial^2_{y_u y_v} B(y,t) = \frac{1}{x}\Big(&\frac{1}{Z}\mathbb{E}\left[\mathrm{e}^{xf(z)}\left(x^2\partial_{z_u} f(z)\partial_{z_v} f(z) + \partial^2_{z_u z_v} f(z)\right)|_{z=y+z(t)}\right] \\ &- \frac{1}{Z^2}\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_u} f(z)|_{z=y+z(t)}\right]\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_v} f(z)|_{z=y+z(t)}\right]\Big).\end{aligned} \tag{4.34}$$

Combining (4.32), (4.33) and (4.34), we get (4.31). □

Proposition 4.2.1. *Denote $D := \bigcup_{k=0}^{n}(x_k;x_{k+1})$. The function $f = f_\rho$ defined in (4.19) satisfies the final-value problem for the controlled semi-linear parabolic Parisi-type PDE*

$$\begin{cases}\partial_t f(y,t) + \frac{1}{2}\sum_{u,v=1}^{d} \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}_{u,v}(t)\left(\partial^2_{y_u y_v} f(y,t) + x(t)\partial_{y_u} f(y,t)\partial_{y_v} f(y,t)\right) = 0, & (t,y)\in D\times\mathbb{R}^d,\\ f(1,y) = g(y), & y\in\mathbb{R}^d,\\ f(y,x_k-0) = f(y,x_k+0), \quad k\in[1;n]\cap\mathbb{N}, & y\in\mathbb{R}^d.\end{cases} \tag{4.35}$$

Note that $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}(t) = \frac{Q^{(k+1)}-Q^{(k)}}{x_{k+1}-x_k}$, for $t\in(x_k;x_{k+1})$.

Proof. A successive application of Lemma 4.2.1 to (4.19) on the intervals D starting from $(x_n;1)$ gives (4.35). $\square$

Remark 4.2.1. *Note that a straightforward inspection of* (4.19)*, using* (4.32), (4.33) *and* (4.34)*, shows that the function f defined in* (4.19) *is $C^1(D)\cap C([0;1])$ with respect to the t-variable and $C^2(\mathbb{R}^d)$ with respect to the y-variable.*

Lemma 4.2.2. *Given $\rho \in \mathscr{Q}'(U,d)$, the function* (4.19) *satisfies the following:*

$$f_\rho(0,0) = \mathbb{E}\Big[\log \sum_{\alpha\in\mathscr{A}} \xi(\alpha)\exp\{g(Y(1,\alpha))\}\Big]. \tag{4.36}$$

Proof. This is an immediate consequence of the RPC averaging property (4.25). $\square$

Lemma 4.2.3.

1. *Given $k \in [1;n]\cap\mathbb{N}$ and a non-negative definite matrix $Q \in \mathrm{Sym}(d)$, we have*

$$\partial_{Q^{(k)}\rightsquigarrow Q} f_\rho(0,0) = -\frac{1}{2}(x_k - x_{k-1})\mathbb{E}\left[\langle Q,M\rangle\right], \tag{4.37}$$

where $M\in\mathbb{R}^{d\times d}$ is defined as

$$\begin{aligned} M_{u,v} :=& T_1(\alpha^{(1)})\cdots T_k(\alpha^{(1)})T_{k+1}(\alpha^{(1)})T_{k+1}(\alpha^{(2)})\cdots T_n(\alpha^{(1)})T_n(\alpha^{(2)})\\ &\partial_{z_u} g(z)|_{z=Y(1,\alpha^{(1)})}\partial_{z_v} g(z)|_{z=Y(1,\alpha^{(2)})} \end{aligned}$$

with $q_L(\alpha^{(1)},\alpha^{(2)}) = k$. Moreover, (4.37) *does not depend on the choice of $\alpha^{(1)},\alpha^{(2)}\in\mathscr{A}$ but only on k.*

2. *Given a non-negative definite matrix $Q\in\mathrm{Sym}(d)$, we have*

$$\partial_{U\rightsquigarrow Q} f_\rho(0,0) = \frac{1}{2}\mathbb{E}\left[\langle Q,M'\rangle\right], \tag{4.38}$$

where $M'\in\mathrm{Sym}(d)$ is satisfies

$$M'_{u,v} = T_1(\alpha)\cdots T_n(\alpha)\Big(\partial^2_{z_u z_v} g(z) + \partial_{z_u} g(z)\partial_{z_v} g(z)\Big)\Big|_{z=Y(1,\alpha)} + \mathscr{O}(1-x_n),$$

as $x_n\to 1$. Note that (4.38) *obviously does not depend on the choice of $\alpha\in\mathscr{A}$.*

Proof. Applying (Aizenman *et al.*, 2007, Lemma A.1) to (4.36), we obtain

$$\begin{aligned} &\partial_s \mathbb{E}\Big[\log\sum_{\alpha\in\mathscr{A}}\exp\{g(Y(1,\alpha))\}\Big|_{Q^{(k)}=Q^{(k)}+sQ}\Big]\\ &= \frac{1}{2}\mathbb{E}\Big[\sum_{u,v=1}^{d}\mathscr{N}(\tilde{\xi})\otimes\mathscr{N}(\tilde{\xi})\Big[\partial_s\Big(Q(\alpha^{(1)},\alpha^{(2)})_{u,v}|_{Q^{(k)}=Q^{(k)}+sQ}\Big)\\ &\quad\Big\{\mathbb{1}_{\alpha^{(1)}=\alpha^{(2)}}(\alpha^{(1)},\alpha^{(2)})\Big(\partial^2_{z_u z_v} g(z)+\partial_{z_u} g(z)\partial_{z_v} g(z)\Big)\Big|_{z=Y(1,\alpha^{(1)})}\\ &\quad - \partial_{z_u} g(z)|_{z=Y(1,\alpha^{(1)})}\partial_{z_v} g(z)|_{z=Y(1,\alpha^{(2)})}\Big\}\Big|_{Q^{(k)}=Q^{(k)}+sQ}\Big]\Big]. \end{aligned}$$

Note that

$$\partial_s\Big(Q(\alpha^{(1)},\alpha^{(2)})_{u,v}|_{Q^{(k)}=Q^{(k)}+sQ}\Big) = \begin{cases} Q_{u,v}, & q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}) = k,\\ 0, & q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}) \neq k. \end{cases}$$

1. Define $M(\alpha^{(1)},\alpha^{(2)}) \in \mathbb{R}^{d\times d}$ as

$$M(\alpha^{(1)},\alpha^{(2)})_{u,v} := \partial_{z_u} g(z)\big|_{z=Y(1,\alpha^{(1)})} \partial_{z_v} g(z)\big|_{z=Y(1,\alpha^{(2)})}.$$

Hence, we arrive at

$$\partial_{Q^{(k)}\rightsquigarrow Q} f_\rho(0,0) = -\frac{1}{2}\mathbb{E}\Big[\sum_{\alpha^{(1)}\alpha^{(2)}\in\mathscr{A}} \mathbb{1}_{q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})=k}\xi(\alpha^{(1)})\xi(\alpha^{(2)})(\alpha^{(1)},\alpha^{(2)})\langle Q, M(\alpha^{(1)},\alpha^{(2)})\rangle\Big].$$

The proof is concluded similarly to the proof of Theorem 4.1.4 by using the properties of the RPC (Theorem 4.1.3 and Lemma 4.1.4).
2. The proof is the same as in (1).

□

The following is a multidimensional version of (Talagrand, 2006b, Lemma 4.3).

Lemma 4.2.4. *For any $\alpha \in \mathscr{A}$, we have*

1.

$$\partial_{x_k} f_\rho(0,0)|_{x_k=x_{k-1}} = \frac{1}{x_{k-1}}\mathbb{E}\Big[T_1(\alpha)\cdots T_{k-2}(\alpha)T_{k-1}(\alpha)|_{x_k=x_{k-1}}$$
$$\Big(\mathbb{E}\Big[f(x_{k+1},Y(x_{k+1},\alpha))T_k(\alpha)|_{x_k=x_{k-1}}\Big] - f(x_k,Y(x_k,\alpha))\Big)\Big].$$

2. Let $M \in \mathrm{Sym}(d)$ with $M_{u,v} := \partial_{z_u} f(x_k,Y(x_k,\alpha))\partial_{z_v} f(x_k,Y(x_k,\alpha))$, then

$$\partial^2_{Q^{(k)}\rightsquigarrow Q, x_k} f_\rho(0,0) = \frac{1}{2}\mathbb{E}\Big[T_1(\alpha)\cdots T_{k-2}(\alpha)\langle Q,M\rangle\Big].$$

Proof. This proof is the same as in (Talagrand, 2006b). □

We now generalise the PDE (4.35). Given a piece-wise continuous $x \in \mathscr{Q}(1,1)$ and $Q \in \mathscr{Q}(U,d)$, consider the following terminal value problem

$$\begin{cases} \partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (0,1), \\ f(y,1) = g(y). \end{cases} \tag{4.39}$$

We say that $f \in C([0;1]\times\mathbb{R}^d \to \mathbb{R})$ is a piece-wise viscosity solution of (4.39), if there exists the partition of the unit segment $0 =: x_0 < x_1 < \ldots < x_{n+1} := 1$ such that, for each $k \in [0,n]\cap\mathbb{N}$, $f : (x_k;x_{k+1})\times\mathbb{R}^d \to \mathbb{R}$ is a viscosity solution (see, e.g., Briand & Hu (2007)) of

$$\begin{cases} \partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (x_k,x_{k+1}), \\ f(y,x_{k+1}+0) = f(y,x_{k+1}-0), \\ f(y,1) = g(y). \end{cases}$$

Proposition 4.2.2. *For any $\rho^{(1)},\rho^{(2)} \in \mathscr{Q}'(U,d)$, we have*

$$|f_{\rho^{(1)}}(0,0) - f_{\rho^{(2)}}(0,0)| \le \frac{C}{2}\int_0^1 \|\rho^{(1)}(t)-\rho^{(2)}(t)\|_F \mathrm{d}t,$$

where $C = C(\Sigma) := \mathbb{E}\left[\|M\|_F\right]$.

Proof. This is an adaptation of the proof of (Talagrand, 2006c, Theorem 3.1) to the multidimensional case. Assume without loss of generality that the paths $\rho^{(1)}$ and $\rho^{(2)}$ have same jump times $\{x_k\}_{k=0}^{n+1}$. Denote the corresponding overlap matrices as $\{Q^{(1,k)}\}_{k=0}^{n+1}$ and $\{Q^{(2,k)}\}_{k=0}^{n+1}$. Given $s \in [0;1]$, define the new path $\rho(s) \in \mathscr{Q}'(U,d)$ by assuming that it has the same jump times $\{x_k\}_{k=0}^{n+1}$ as the paths $\rho^{(1)}, \rho^{(2)}$ and defining its overlap matrices as $Q^{(k)}(s) := sQ^{(1,k)} + (1-s)Q^{(2,k)}$. On the one hand, we readily have

$$\int_0^1 \|\rho^{(1)}(t) - \rho^{(2)}(t)\|_{\mathrm{F}} \mathrm{d}t = \sum_{k=1}^{n} (x_k - x_{k-1}) \|Q^{(1,k)} - Q^{(2,k)}\|_{\mathrm{F}}.$$

On the other hand, using Lemma 4.2.3, we have

$$|\partial_s f_{\rho(s)}(0,0)| \leq \frac{C}{2} \sum_{k=1}^{n} (x_k - x_{k-1}) \|Q^{(1,k)} - Q^{(2,k)}\|_{\mathrm{F}}.$$

Finally, we have

$$|f_{\rho^{(1)}}(0,0) - f_{\rho^{(2)}}(0,0)| \leq \int_0^1 |\partial_s f_{\rho(s)}(0,0)| \mathrm{d}s.$$

Combining the last three formulae, we get the theorem. □

Remark 4.2.2. *Note that using the same argument and notations as in the previous theorem we get that, for any $(y,t) \in \mathbb{R}^d \times [0;1]$,*

$$|f_{\rho^{(1)}}(y,t) - f_{\rho^{(2)}}(y,t)| \leq \frac{C(\Sigma)}{2} \int_t^1 \|\rho^{(1)}(s) - \rho^{(2)}(s)\|_F \mathrm{d}s.$$

Remark 4.2.3. *Note that we can associate to each $\rho \in \mathscr{Q}(U,d)$ a $\mathrm{Sym}^+(d)$-valued countably additive vector measure $\nu_\rho \in \mathscr{M}([0;1], \mathrm{Sym}^+(d))$ by the following standard procedure. Given $[a;b) \subset [0;1]$, define*

$$\nu_\rho([a;b)) := \rho(b) - \rho(a)$$

and then extend the measure, e.g., to all Borell subsets of $[0;1]$.

Theorem 4.2.1. *Given $U \in \mathrm{Sym}^+(d)$, we have*

1. *The set $\mathscr{Q}(U,d)$ is compact under the topology induced by the following norm*

$$\|\rho\| := \int_0^1 \|\rho(t)\|_F \mathrm{d}t, \quad \rho \in \mathscr{Q}(U,d). \tag{4.40}$$

2. *The functional $\mathscr{Q}'(U,d) \ni \rho \mapsto f_\rho(0,0)$ is Lipschitzian and can be uniquely extended by continuity to the whole $\mathscr{Q}(U,d)$.*

Proof. 1. The topology induced by the norm (4.40) coincides with the topology of weak convergence of the above-defined vector measures. Since $\mathscr{Q}(U,d)$ is a bounded set, it is compact in the weak topology.
2. This is an immediate consequence of Proposition 4.2.2.

□

In the next result, we summarise some results on the PDE (4.39) for the non-discrete parameters, cf. Proposition 4.2.1.

Theorem 4.2.2.

1. Existence. *Assume that Q is in $\mathscr{Q}(U,d)$ and is piece-wise $C^{(1)}$. Assume also that x is in $\mathscr{Q}(1,1)$ and is piece-wise continuous. Then the terminal value problem* (4.39) *has a unique continuous, piece-wise viscosity solution $f_{Q,x} \in C([0;1] \times \mathbb{R}^d)$.*
2. Monotonicity with respect to x. *Assume $Q \in \mathscr{Q}(U,d)$. Assume also that $x^{(1)}, x^{(2)} \in \mathscr{Q}(1,1)$ are such that $x^{(1)}(t) \le x^{(2)}(t)$, almost everywhere for $t \in [0;1]$. Let $f_{Q,x^{(1)}}$ and $f_{Q,x^{(2)}}$ be the corresponding solutions of* (4.39). *Then $f_{Q,x^{(1)}} \le f_{Q,x^{(2)}}$.*
3. Monotonicity with respect to g. *Assume $g_1, g_2 : \mathbb{R}^d \to \mathbb{R}$ satisfy Assumption 4.1.1 and also $g_1 \le g_2$ almost everywhere. Let $f_{g_1}, f_{g_2} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ be the corresponding solutions of* (4.39) *with $g = g_1$, $g = g_2$, respectively. Then $f_{g_1} \le f_{g_2}$.*

Proof. 1. Due to the assumptions, the diffusion matrix $\dot{Q}(t) = \dot{\rho}(t)$ in (4.39) is non-negative definite. Applying (Briand & Hu, 2007, Proposition 8) to the PDE (4.39) successively on the intervals $[x_k; x_{k+1})$, where the $\dot{\rho}$ is continuous, gives the existence of the solutions in viscosity sense and, moreover, gives their continuity. Uniqueness is ensured by (Da Lio & Ley, 2006, Theorem 1.1).

2. By the approximation argument (cf. Theorem 4.2.1), it is enough to assume that $x^{(1)}, x^{(2)} \in \mathscr{Q}'(1,1)$ and $Q \in \mathscr{Q}'(U,d)$. Then Proposition 4.2.1 gives the existence of the corresponding piece-wise classical solutions of (4.39): $f_{Q,x^{(1)}}, f_{Q,x^{(2)}}$. These solutions are obviously also the (unique) piece-wise viscosity solutions of (4.39). The comparison result (Briand & Hu, 2007, Theorem 5) and the non-linear Feynman-Kac formula (Briand & Hu, 2007, Proposition 8) give then the claim.
3. This can be seen either from the representation (4.36) and an approximation argument, or exactly as in (2) by invoking the results of Briand & Hu (2007).

□

4.2.1 The Parisi functional

We consider now a specific terminal condition in the system (4.35) given in (4.20).

Given $\rho \in \mathscr{Q}(U,d)$, let $f_\rho : [0;1] \times \mathbb{R}^d \to \mathbb{R}$ be the value of (the continuous extension onto $\mathscr{Q}(U,d)$ of) the solution of (4.35) with the specific terminal condition given by (4.20). Following the ideas in the physical literature, we now define the *Parisi functional* $\mathscr{P}(\beta,\rho,\Lambda) : \mathbb{R}_+ \times \mathscr{Q}'(U,d) \times \mathrm{Sym}^+(d) \times \mathrm{Sym}(d) \to \mathbb{R}$ in as

$$\mathscr{P}(\beta,\rho,\Lambda) := f_\rho(0,0) - \frac{\beta^2}{2} \int_0^1 x(t) \mathrm{d}\left(\|\rho(t)\|_{\mathrm{F}}^2\right) - \langle U, \Lambda \rangle. \tag{4.41}$$

The integral in (4.41) is understood in the usual Lebesgue-Stiltjes sense.

Remark 4.2.4. *Note that the path integral term in* (4.41) *equals $f(0,0)$, where $f(t,y)$ is the solution of* (4.39) *with the following boundary condition*

$$g(y) := \beta \langle y, \mathbb{1} \rangle = \beta \sum_{u=1}^{d} y_u, \quad y \in \mathbb{R}^d.$$

Obviously $\mathscr{Q}'(d)$ is dense in $\mathscr{Q}(d)$.

Theorem 4.2.3. *We have*

$$p(\beta) \leq \sup_{U \in \mathrm{Sym}^+(d)} \inf_{\substack{\rho \in \mathscr{Q}'(U,d) \\ \Lambda \in \mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda). \tag{4.42}$$

Proof. The bound (4.42) is a straightforward consequence of Theorem 4.1.1. □

4.2.2 On strict convexity of the Parisi functional and its variational representation

In this subsection, we derive a variational representation for Parisi's functional. As a consequence, for $d = 1$, we prove that the functional is strictly convex with respect to the $x \in \mathscr{Q}(1,1)$, if the terminal condition g (cf. (4.39)) is strictly convex and increasing. This result is related to the problem of strict convexity of the Parisi functional in the case of the SK model.

Let $W := \{W(s)\}_{s \in \mathbb{R}_+}$ be the standard $\mathbb{R}^d$-valued Brownian motion and let $\{\mathscr{F}_t\}_{t \in \mathbb{R}_+}$ be the correspondent filtration. Define

$$\mathscr{U}[t;T] := \{u : [t;T] \to \mathbb{R}^d \mid u \text{ is } \{\mathscr{F}_t\}_{t \in \mathbb{R}_+} \text{ progressively measurable}\}.$$

Given $u \in \mathscr{U}[t;1]$, $Q \in \mathscr{Q}(U,d)$ and $x \in \mathscr{Q}(1,1)$, consider the following $\mathbb{R}^d$-valued and adapted to $\{\mathscr{F}_t\}_{t \in \mathbb{R}^+}$ diffusion

$$Y^{(Q,x,u,t,y)}(s) := y - \int_t^s \left(x(s)\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^s \left(\dot{Q}(s)\right)^{1/2} \mathrm{d}W(s), \quad s \in [t;1].$$

Given some function $g : \mathbb{R}^d \to \mathbb{R}$ satisfying Assumption 4.1.1, define $f_{Q,x} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ as

$$f_{Q,x}(y,t) := \sup_{u \in \mathscr{U}[t;1]} \mathbb{E}\left[g(Y^{(Q,x,u,t,y)}(1)) - \frac{1}{2}\int_t^1 \|u(s)\|_2^2 \mathrm{d}s\right]. \tag{4.43}$$

Proposition 4.2.3. *Let $d = 1$. If g is strictly convex and increasing, then the functional $\mathscr{Q}(1,1) \ni x \mapsto f_{Q,x}$ is strictly convex.*

Proof. We have

$$Y^{(Q,x,u,t,y)}(1) = y - \int_t^1 \left(x(s)\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s).$$

By an approximation argument, it is enough to prove the strict convexity for the continuous $x_1, x_2 \in \mathscr{Q}(1,1)$ ($x_1 \neq x_2$). For any $\gamma \in (0;1)$, we have

$$\begin{aligned}
Y^{(Q,\gamma x_1 + (1-\gamma)x_2,u,t,y)}(1) &= -\int_t^1 \left(\gamma x_1 + (1-\gamma)x_2\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s) \\
&< -\gamma\int_t^1 \left(x_1\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s - (1-\gamma)\int_t^1 \left(x_2\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s \\
&\quad + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s) \\
&= \gamma Y^{(Q,x_1,u,t,y)}(1) + (1-\gamma)Y^{(Q,x_2,u,t,y)}(1),
\end{aligned} \tag{4.44}$$

where the strict inequality above is due to the strict concavity of the square root function. The strict convexity and monotonicity of g combined with the representation (4.44) implies that (4.43) is strictly convex as a function of x, since a supremum of a family of convex functions is convex. $\square$

Proposition 4.2.4. *Given a piece-wise continuous $x \in \mathscr{Q}(1,1)$ and a $Q \in \mathscr{Q}(U,d)$ which is piece-wise in $C^1(0;1)$, the function $f_{Q,x} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ defined by* (4.43) *is a unique, continuous, piece-wise viscosity solution of the following terminal value problem*

$$\begin{cases} \partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (0,1), \\ f(y,1) = g(y). \end{cases}$$

Proof. In a way similar to the proof of Theorem 4.2.2, we successively use (Da Lio & Ley, 2006, Theorem 2.1) on the intervals $(x_k; x_{k+1})$, where the data of the PDE are continuous. $\square$

Theorem 4.2.4. *Assume $d = 1$. Suppose also that g satisfies the assumptions of Proposition 4.2.3. For any $u \in \mathbb{R}$, the generalised Parisi functional given by* (4.41) *with $f_\rho(0,0)$ corresponding to the terminal condition g is strictly convex on $Q(u,1)$. Consequently, there exists a unique optimising order parameter.*

Proof. In 1-D, we can choose the coordinates such that $Q := Ut$, on $[0;1]$. Consequently, $\dot{Q} := U :=$ const on $[0;1]$. Hence, it is enough check the strict convexity with respect to $x \in \mathscr{Q}(1,1)$. The result follows by approximation in the norm (4.40) of an arbitrary pair of different elements of $\mathscr{Q}(U,d)$ by a pair of elements of $\mathscr{Q}'(U,d)$ and Propositions 4.2.1, 4.2.3 and 4.2.4. $\square$

Remark 4.2.5. *Due to the monotonicity assumption on g, Theorem 4.2.4 does not cover the case of the SK model, where the terminal value g is given by* (4.20).

4.2.3 Simultaneous diagonalisation scenario

In the setups with highly symmetric state spaces Σ_N (such as the spherical spin models of Panchenko & Talagrand (2007b) or the Gaussian spin models, see Section 5.2 below), less complex order parameter spaces as $Q(U,d)$ suffice.

Given some orthogonal matrix $O \in \mathscr{O}(d)$, we briefly discuss the case $\rho \in \mathscr{Q}_{\text{diag}}(U,O,d)$, where

$$\mathscr{Q}_{\text{diag}}(U,O,d) := \{\rho \in \mathscr{Q}(U,d) \mid \text{for all } t \in [0;1], \text{ the matrix } O\rho(t)O^* \text{ is diagonal}\}.$$

The space $\mathscr{Q}_{\text{diag}}(U,O,d)$ is obviously isomorphic to the space of "paths" with the non-decreasing coordinate functions in $\mathbb{R}^d$, starting from the origin and ending at u, i.e.,

$$\bar{\mathscr{Q}}(u,d) := \{\rho : [0;1] \to \mathbb{R}^d \mid \bar{\rho}(0) = 0; \bar{\rho}(1) = u; \bar{\rho}(t) \preceq \bar{\rho}(s), \text{ for } t \leq s; \bar{\rho} \text{ is cádlág}\},$$

where $u = OUO^* \in \mathbb{R}^d$. The isomorphism is then given by

$$\bar{\mathscr{Q}}(u,d) \ni \bar{\rho} \mapsto O\rho O^* \in \mathscr{Q}_{\text{diag}}(U,O,d). \tag{4.45}$$

4.3 Remainder estimates

In this section, we partially extend Talagrand's remainder estimates to the multidimensional setting. Due to Proposition 4.1.1, to prove the validity of Parisi's formula it is enough to show that all the μ_k terms in (4.28) almost vanish for the almost optimal parameters of the optimisation problem in (4.14). This can be done if the free energy of two coupled replicas of the system (4.48) is strictly smaller than twice the free energy of the uncoupled single system (4.2), see inequality (4.3.2). However, the systems involved in (4.3.2) are effectively at least as complex as the SK model itself. In Section 4.3.2, we again apply Guerra's scheme to obtain the upper bounds on (4.48) in terms of the free energy of the corresponding comparison GREM-inspired model. One might then hope that by a careful choice of the comparison model one can prove inequality (4.3.2). In Sections 4.3.3 and 4.3.4, we formulate some conditions on the comparison system which would suffice to get inequality (4.3.2), giving, hence, the conditional proof of the Parisi formula, see Theorem 4.3.1.

4.3.1 A sufficient condition for μ_k-terms to vanish

In this subsection, we are going to establish a sufficient condition for the measures μ_k to vanish. This condition states roughly the following. Whenever the free energy of a certain replicated system uniformly in N strictly less then twice the free energy of the single system, the measure μ_k vanishes in $N \to +\infty$ limit (see Lemma 4.3.2).

Keeping in mind the definition of μ_k (cf. (4.26)) and of the Hamiltonian $H_t(\sigma,\alpha)$ (cf. (4.1)), we define, for $\alpha^{(1)},\alpha^{(2)} \in \mathscr{A}^{(2),k}$, the corresponding replicated Hamiltonian as

$$H_t^{(2)}(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) := H_t(\sigma^{(1)},\alpha^{(1)}) + H_t(\sigma^{(2)},\alpha^{(2)}). \tag{4.46}$$

Remark 4.3.1. *We note here that the distribution of the Hamiltonian $H_t(k,\sigma^{(1)},\sigma^{(2)})$ depends only on k and not on the choice of the indices $\alpha^{(1)},\alpha^{(2)} \in \mathscr{A}^{(2),k}$.*

Remark 4.3.2. *The superscript (2) in* (4.46) *(and in what follows) indicates that the quantity is related to the twice replicated objects.*

Define

$$\mathscr{A}^{(2),k} := \{(\alpha^{(1)},\alpha^{(2)}) \in \mathscr{A}^2 : q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}) = k\}.$$

Additionally, for any $\mathscr{V} \subset \Sigma(B(U,\varepsilon))^2$ and any suitable Gaussian process,

$$\{F(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) : \sigma^{(1)},\sigma^{(2)} \in \Sigma_N, \alpha^{(1)},\alpha^{(2)} \in \mathscr{A}\},$$

we define the *local remainder comparison functional* as

$$\Phi_{\mathscr{V}}^{(2),k,x}[F] := \frac{1}{N}\mathbb{E}\Big[\log \iint_{\mathscr{V}} \iint_{\mathscr{A}^{(2),k}} \exp\Big\{\beta\sqrt{N}F(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)})\Big\} \mathrm{d}\mu^{\otimes N}(\sigma^{(1)})\mathrm{d}\mu^{\otimes N}(\sigma^{(2)})\mathrm{d}\xi(\alpha^{(1)})\mathrm{d}\xi(\alpha^{(2)})\Big]. \tag{4.47}$$

Define

$$\varphi_N^{(2)}(k,t,x,Q,\mathscr{V}) := \Phi_{\mathscr{V}}^{(2),k}\left[H_t^{(2)}\right]. \tag{4.48}$$

Lemma 4.3.1. *Recalling the definition* (4.2)*, for any* $\mathscr{V} \subset \Sigma(B(U,\varepsilon))^2$*, we have*

$$\varphi_N^{(2)}(k,t,x,Q,\mathscr{V}) \le \varphi_N^{(2)}(k,t,x,Q,\Sigma(B(U,\varepsilon))^2) = 2\varphi_N(t,x,Q,\Sigma(B(U,\varepsilon))). \tag{4.49}$$

Proof. The first inequality in (4.49) is obvious, since the expression under the integral in (4.47) is positive. The equality in (4.49) is an immediate consequence of the RPC averaging property (4.25). □

In what follows, we shall be looking for the sharper (in particular, *strict*) versions of the inequality (4.49) because of the following observation due to Talagrand (2006b).

Lemma 4.3.2. *Fix an arbitrary* $\mathscr{V} \subset \Sigma_N(B(U,\varepsilon))^2$. *Suppose that, for some* $\varepsilon > 0$*, the following inequality holds*

$$\varphi_N^{(2)}(k,t,x,Q,\mathscr{V}) \le 2\varphi_N(t,x,Q,\Sigma_N(B(U,\varepsilon))) - \varepsilon. \tag{4.50}$$

Then, for some $K > 0$*, we have*

$$\mu_k(\mathscr{V}) \le K\exp\left(-\frac{N}{K}\right).$$

Proof. The proof is based on Theorem 3.2.2 and follows the lines of (Panchenko, 2005b, Lemma 7). □

4.3.2 Upper bounds on $\varphi^{(2)}$: Guerra's scheme revisited

In this subsection, we shall develop a mechanism to obtain upper bounds on $\varphi^{(2)}$ defined in (4.48). This will be achieved in the full analogy to Guerra's scheme by using a suitable Gaussian comparison system.

Given $U \in \mathrm{Sym}^+(d)$, we say that $V \in \mathbb{R}^{d\times d}$ is an *admissible mutual overlap matrix for* U, if

$$\mathfrak{U} := \begin{bmatrix} U & V \\ V^* & U \end{bmatrix} \in \mathrm{Sym}^+(2d). \tag{4.51}$$

Furthermore, define

$$\mathscr{V}(U) := \{V \in \mathbb{R}^{d\times d} : V \text{ is an admissible mutual overlap matrix for } U\}.$$

Hereinafter without further notice we assume that $\mathfrak{U} \in \mathrm{Sym}^+(2d)$ has the form (4.51), where V is some admissible mutual overlap matrix for U.

Let $\mathfrak{Q} \in \mathscr{Q}(\mathfrak{U},2d)$. Let $\mathfrak{x} := \{\mathfrak{x}_l \in [0;1]\}_{l=1}^{\mathfrak{n}}$ be the "jump times" of the path ρ. We assume that the "times" are increasingly ordered, i.e.,

$$0 = \mathfrak{x}_0 < \mathfrak{x}_1 < \ldots < \mathfrak{x}_{\mathfrak{n}} < \mathfrak{x}_{\mathfrak{n}+1} = 1.$$

Consider the following collection of matrices

$$\mathfrak{Q} := \{\mathfrak{Q}_l := \mathfrak{Q}(\mathfrak{x}_l) \subset \mathrm{Sym}^+(2d)\}_{l=0}^{\mathfrak{n}+1}.$$

We obviously then have

$$0 = \mathfrak{Q}^{(0)} \prec \mathfrak{Q}^{(1)} \prec \ldots \prec \mathfrak{Q}^{(\mathfrak{n})} \prec \mathfrak{Q}^{(\mathfrak{n}+1)} = \mathfrak{U}. \tag{4.52}$$

Such a path $\mathfrak{Q}$ induces in the usual way the "doubled" GREM overlap kernel $\mathfrak{Q} := \{\mathfrak{Q}(\alpha^{(1)}, \alpha^{(2)}) \in \mathrm{Sym}^+(2d) \mid \alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}_\mathfrak{n}\}$, defined as

$$\mathfrak{Q}(\alpha^{(1)}, \alpha^{(2)}) := \mathfrak{Q}^{(q_\mathrm{L}(\alpha^{(1)}, \alpha^{(2)}))}.$$

We also need the $d \times d$ submatrices of the above overlap such that

$$\mathfrak{Q}(\alpha^{(1)}, \alpha^{(2)}) = \begin{bmatrix} \mathfrak{Q}|_{11}(\alpha^{(1)}, \alpha^{(2)}) & \mathfrak{Q}|_{12}(\alpha^{(1)}, \alpha^{(2)}) \\ \mathfrak{Q}|_{12}(\alpha^{(1)}, \alpha^{(2)})^* & \mathfrak{Q}|_{22}(\alpha^{(1)}, \alpha^{(2)}) \end{bmatrix}. \tag{4.53}$$

Remark 4.3.3. *For $\sigma^{(1)}\sigma^{(2)} \in \Sigma_N$, we shall use the notation $\sigma^{(1)} \shortparallel \sigma^{(2)} \in \left(\mathbb{R}^{2d}\right)^N$ to denote the vector obtained by the following concatenation of the vectors $\sigma^{(1)}$ and $\sigma^{(2)}$*

$$\sigma^{(1)} \shortparallel \sigma^{(2)} := \left(\sigma_i^{(1)}\sigma_i^{(2)} \in \Sigma \times \Sigma \subset \mathbb{R}^{2d}\right)_{i=1}^N.$$

Let us observe that the process

$$X^{(2)} := \left\{X^{(2)}(\tau) = X(\sigma^{(1)}) + X(\sigma^{(1)}) \mid \tau = \sigma^{(1)} \shortparallel \sigma^{(2)}; \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N\right\}$$

is actually an instance of the $2d$-dimensional Gaussian process defined in (0.8). Hence, it has the following correlation structure, for $\tau^1, \tau^2 \in \Sigma_N^{(2)}$,

$$\mathrm{Cov}\left[X^{(2)}(\tau^1), X^{(2)}(\tau^2)\right] = \|R^{(2)}(\tau^1, \tau^2)\|_\mathrm{F}^2.$$

The path ρ induces also the following two new (independent of everything before) comparison process $Y^{(2)} := \left\{Y^{(2)}(\alpha) \in \mathbb{R}^{2d} \mid \alpha \in \mathscr{A}_\mathfrak{n}\right\}$, with the following correlation structures

$$\mathrm{Cov}\left[Y^{(2)}(\alpha^{(1)}), Y^{(2)}(\alpha^{(2)})\right] = \mathfrak{Q}(\alpha^{(1)}, \alpha^{(2)}) \in \mathrm{Sym}^+(d).$$

As usual, let $\{Y_i^{(2)}\}_{i=1}^N$ be the independent copies of $Y^{(2)}$. For the purposes of new Guerra's scheme we define a GREM-like process (cf. (3.15))

$$A^{(2)} = \{A^{(2)}(\tau, \alpha) : \tau = \sigma^{(1)} \shortparallel \sigma^{(2)}; \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N; \alpha \in \mathscr{A}_\mathfrak{n}\}$$

as

$$A^{(2)}(\tau, \alpha) := \left(\frac{2}{N}\right)^{1/2} \sum_{i=1}^N \langle Y_i^{(2)}(\alpha), \tau_i \rangle.$$

We fix some $t \in [0; 1]$. We would now like to apply Guerra's scheme to the comparison functional (4.47) and the following two processes

$$\left\{H_t^{(2)}(\sigma^{(1)}, \sigma^{(2)}, \alpha)\right\}_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N, \alpha \in \mathscr{A}}, \left\{\sqrt{t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha)\right\}_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N, \alpha \in \mathscr{A}}.$$

These two processes are, respectively, the counterparts of the processes $X(\sigma)$ and $A(\sigma,\alpha)$ in Guerra's scheme.

Consider a path $\widetilde{Q} \in \mathscr{Q}'(U,d)$ with the following jumps

$$0 =: \widetilde{Q}^{(0)} \prec \widetilde{Q}^{(1)} \prec \ldots \prec \widetilde{Q}^{(\mathfrak{n})} \prec \widetilde{Q}^{(\mathfrak{n}+1)}.$$

Let $\widetilde{A} := \left\{ \widetilde{A}(\sigma,\alpha) : \sigma \in \Sigma_N; \alpha \in \mathscr{A}_{\mathfrak{n}} \right\}$ be a Gaussian process (independent of all random objects around) with the following covariance structure

$$\mathbb{E}\left[\widetilde{A}(\sigma^{(1)},\alpha^{(1)}) \widetilde{A}(\sigma^{(2)},\alpha^{(2)}) \right] = 2\langle R(\sigma^{(1)},\sigma^{(2)}), \widetilde{Q}(\alpha^{(1)},\alpha^{(2)}) \rangle.$$

For notational convenience, we introduce also the following process

$$\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) := \widetilde{A}(\sigma^{(1)},\alpha^{(1)}) + \widetilde{A}(\sigma^{(2)},\alpha^{(2)}). \tag{4.54}$$

Recalling the replicated Hamiltonian (4.46) and following Guerra's scheme, we introduce, for $s \in [0;1]$, the following interpolating Hamiltonian

$$\begin{aligned} H_{t,s}^{(2)}(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) :=& \sqrt{st} X^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}) + \sqrt{(1-s)t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) \\ &+ \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}). \end{aligned} \tag{4.55}$$

Given $\varepsilon, \delta > 0$ and $\mathfrak{L} \in \mathrm{Sym}(2d)$, define (cf. (4.6))

$$\mathscr{V}^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta) := \{ \mathfrak{U}' \in \mathrm{Sym}^+(2d) : \|\mathfrak{U}' - \mathfrak{U}\|_{\mathrm{F}} < \varepsilon, \langle \mathfrak{U}' - \mathfrak{U}, \mathfrak{L} \rangle < \delta \}.$$

We consider the following set of the local configurations

$$\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta) := \left\{ (\sigma^{(1)},\sigma^{(2)}) \in \Sigma_N \times \Sigma_N : R_N^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \sigma^{(1)} \shortparallel \sigma^{(2)}) \in \mathscr{V}^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta) \right\}. \tag{4.56}$$

Note that $\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta) \subset \Sigma_N(B(U,\varepsilon))^2$. We consider also the RPC $\zeta = \zeta(\mathfrak{x})$ generated by the vector $\mathfrak{x}$ and, for any suitable Gaussian process

$$F := \{ F(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) \mid \sigma^{(1)},\sigma^{(2)} \in \Sigma_N; \alpha^{(1)},\alpha^{(2)} \in \mathscr{A}_n \},$$

define the corresponding local comparison functional (cf. (4.47)) as follows

$$\begin{aligned} \Phi_{\mathscr{V}}^{(2),k,\mathfrak{x}}[F] := \frac{1}{N}\mathbb{E}\Big[\log \iint_{\mathscr{V}} \iint_{\mathscr{A}^{(2),k}} & \exp\left\{ \beta\sqrt{N} F(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) \right\} \\ & \mathrm{d}\mu^{\otimes N}(\sigma^{(1)}) \mathrm{d}\mu^{\otimes N}(\sigma^{(2)}) \mathrm{d}\zeta(\alpha^{(1)}) \mathrm{d}\zeta(\alpha^{(2)}) \Big]. \end{aligned}$$

Define the corresponding local free energy-like quantity as (cf. (4.2))

$$\chi(s,t,k,\mathfrak{x},\mathfrak{Q},\widetilde{\mathfrak{Q}},\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) := \Phi^{(2),k,\mathfrak{x}}_{\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)}\left[H_{t,s}^{(2)} \right]. \tag{4.57}$$

To lighten the notation, we indicate hereinafter only the dependence of χ on s. Denote

$$B^{\mathfrak{x},\mathfrak{Q}} := \frac{t\beta^2}{2} \sum_{l=1}^{\mathfrak{n}} \mathfrak{x}_l \left(\|\mathfrak{Q}^{(l+1)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(l)}\|_{\mathrm{F}}^2 \right).$$

Lemma 4.3.3. *There exists $C = C(\Sigma) > 0$ such that, for any $\mathfrak{U}$ as above, we have*

$$\frac{\partial}{\partial s}\chi(s,t,k,\mathfrak{x},\mathfrak{Q},\widetilde{\mathfrak{Q}},\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq -B^{\mathfrak{x},\mathfrak{Q}} + C\varepsilon, \tag{4.58}$$

Consequently,

$$\begin{aligned}\varphi_N^{(2)}(k,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq & \Phi^{(2),k,\mathfrak{x}}_{\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)}\Big[\sqrt{t}A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)},\alpha^{(1)}) \\ & + \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)},\alpha^{(1)},\alpha^{(2)})\Big] - B^{\mathfrak{x},\mathfrak{Q}} + C\varepsilon. \end{aligned}\tag{4.59}$$

Proof. The idea is the same as in the proof of Theorem 4.1.1 and is based on Proposition 3.2.5. Since we are considering the localised free energy-like quantities (4.57), the variance terms induced by the interpolation (4.55) in (3.34) cancel out (up to the correction $\mathcal{O}(\varepsilon)$) and we are left with the non-positive contribution of the covariance terms. □

Given $\mathfrak{L} \in \mathrm{Sym}(2d)$, we consider the following stencil of the Legendre transform

$$\begin{aligned}\widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}}[F] := -\langle\mathfrak{L},\mathfrak{U}\rangle - B^{\mathfrak{x},\mathfrak{Q}} + \frac{1}{N}\mathbb{E}\Big[\log \iint_{\Sigma_N^2}\iint_{\mathscr{A}^{(2),k}} & \exp\{\beta\sqrt{N}F(\sigma^{(1)},\sigma^{(2)},\alpha^{(1)},\alpha^{(2)}) \\ & + \langle\mathfrak{L}(\sigma^{(1)} \shortparallel \sigma^{(2)}),\sigma^{(1)} \shortparallel \sigma^{(2)}\rangle\} \\ & \mathrm{d}\mu^{\otimes N}(\sigma^{(1)})\mathrm{d}\mu^{\otimes N}(\sigma^{(2)})\mathrm{d}\zeta(\alpha^{(1)})\mathrm{d}\zeta(\alpha^{(2)})\Big]. \end{aligned}\tag{4.60}$$

Definition 4.3.1. *Let $F : \mathrm{Sym}(2d) \to \mathbb{R}$. Given $\delta > 0$, we call $\mathfrak{L}^{(0)} \in \mathrm{Sym}(2d)$ δ-minimal for F, if*

$$F(\Lambda^{(0)}) \leq \inf_{\Lambda\in\mathrm{Sym}(2d)} F(\Lambda) + \delta.$$

Lemma 4.3.4. *There exists $C = C(\Sigma) > 0$ such that, for all $\mathfrak{U}$ and $\mathfrak{Q} \in \mathscr{Q}'(\mathfrak{U},2d)$ as above, all $\varepsilon,\delta > 0$, there exists a δ-minimal Lagrange multiplier $\mathfrak{L} = \mathfrak{L}(\mathfrak{U},\varepsilon,\delta) \in \mathrm{Sym}(2d)$ for (4.60) such that, for all $k \in [1;n] \cap \mathbb{N}$, all $t \in [0;1]$, and all $(x,\mathscr{Q})$, we have*

$$\begin{aligned}\varphi_N^{(2)}(k,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq \inf_{\mathfrak{L}\in\mathrm{Sym}(2d)} & \widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}}\Big[\sqrt{t}A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)},\alpha^{(1)}) \\ & + \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)},\alpha^{(1)},\alpha^{(2)})\Big] \\ & + C(\varepsilon+\delta). \end{aligned}\tag{4.61}$$

Proof. The argument is the same as in the proof of Theorem 3.1.1. □

Consider the family of matrices $\widetilde{\mathfrak{Q}} := \left\{\widetilde{\mathfrak{Q}}^{(l)} \in \mathrm{Sym}^+(2d) \mid l \in [0;\mathfrak{n}+1] \cap \mathbb{N}\right\}$, defined as

$$\widetilde{\mathfrak{Q}}^{(l)} := \begin{bmatrix}\widetilde{Q}^{(l)} & \widetilde{Q}^{(l)} \\ \widetilde{Q}^{(l)} & \widetilde{Q}^{(l)}\end{bmatrix}, \tag{4.62}$$

for $l \in [0;k] \cap \mathbb{N}$, and as

$$\widetilde{\mathfrak{Q}}^{(l)} := \begin{bmatrix} \widetilde{Q}^{(l)} & \widetilde{Q}^{(k)} \\ \widetilde{Q}^{(k)} & \widetilde{Q}^{(l)} \end{bmatrix}, \tag{4.63}$$

for $l \in [k+1; \mathfrak{n}+1] \cap \mathbb{N}$. Additionally we define, for $l \in [0; \mathfrak{n}+1]$, the matrices

$$\widehat{\mathfrak{Q}}^{(l)}(t) := t\mathfrak{Q} + (1-t)\widetilde{\mathfrak{Q}}.$$

Let $\widehat{Z}^{(l)} \in \mathbb{R}^{2d\times 2d}$, for $l \in [0; \mathfrak{x}]$, be independent Gaussian vectors with

$$\operatorname{Cov}\left[\widehat{Z}^{(l)}\right] = 2\beta^2\left(\widehat{\mathfrak{Q}}^{(l+1)}(t) - \widehat{\mathfrak{Q}}^{(l)}(t)\right).$$

Given $\widehat{y} \in \mathbb{R}^{2d}$, $\mathfrak{L} \in \operatorname{Sym}(2d)$, consider the random variable

$$X^{(2)}_{\mathfrak{n}+1}(\widehat{y}, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) := \log \int_\Sigma \int_\Sigma \exp\left(\langle \widehat{y}, \sigma^{(1)} \shortparallel \sigma^{(2)}\rangle + \langle \mathfrak{L}(\sigma^{(1)} \shortparallel \sigma^{(2)}), \sigma^{(1)} \shortparallel \sigma^{(2)}\rangle\right) \mathrm{d}\mu(\sigma^{(1)})\mathrm{d}\mu(\sigma^{(2)}). \tag{4.64}$$

Define recursively, for $l \in [\mathfrak{n}; 0] \cap \mathbb{N}$, the following quantities

$$X^{(2)}_l(\widehat{y}, k, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) := \frac{1}{\mathfrak{x}_l} \log \mathbb{E}^{\widehat{Z}^{(l)}}\left[\exp\left(\mathfrak{x}_l X^{(2)}_{l+1}(\widehat{y} + \widehat{Z}^{(l)}, k, \mathfrak{x}, \widehat{\mathfrak{Q}}^{(l)}(t), \mathfrak{L})\right)\right]. \tag{4.65}$$

Lemma 4.3.5. *We have*

$$\begin{aligned}&\widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}}\left[\sqrt{t}A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) + \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\right] \\ &= -\langle \mathfrak{L}, \mathfrak{U}\rangle + X^{(2)}_0(0, \mathfrak{x}, \widehat{\mathfrak{Q}}^{(l)}(t), \mathfrak{L}).\end{aligned} \tag{4.66}$$

Proof. This is an immediate consequence of the RPC averaging property (4.25). □

Proposition 4.3.1. *Under the conditions of Lemma 4.3.4, we have*

$$\varphi^{(2)}_N(k, t, x, Q, \Sigma^{(2)}_N(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)) \le \inf_{\mathfrak{L}\in\operatorname{Sym}(2d)} \left(-\langle \mathfrak{L}, \mathfrak{U}\rangle + X^{(2)}_0(0, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L})\right) - B^{\mathfrak{x},\mathfrak{Q}} + C(\varepsilon + \delta).$$

Remark 4.3.4. *Similarly to* (4.13)*, there exists $C = C(\Sigma, \mu) > 0$, such that, for any $\mathfrak{L} \in$ Sym$(2d)$,*

$$\varphi^{(2)}_N(k, t, x, Q, \Sigma^{(2)}_N(B(\mathfrak{U}, \varepsilon)) \le -\langle \mathfrak{L}, \mathfrak{U}\rangle - B^{\mathfrak{x},\mathfrak{Q}} + X^{(2)}_0(0, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L})\Big) + C\|\mathfrak{L}\|_F\varepsilon.$$

Proof. Immediately follows from Lemmata 4.3.4 and 4.3.5. □

4.3.3 Adjustment of the upper bounds on $\varphi^{(2)}$

Proposition 3.2.1 implies that there exists $r \in [1; n] \cap \mathbb{N}$ such that

$$\|Q^{(r-1)}\|^2_{\mathrm{F}} < \|V\|^2_{\mathrm{F}} < \|Q^{(r)}\|^2_{\mathrm{F}}. \tag{4.67}$$

Assume $r = k$. (Other cases are similar or easier as shown for 1-D in Talagrand (2006b).) We make the following tuning of the upper bounds of the previous subsection. Set $\mathfrak{n} := n+1$. Let $w \in [x_{r-1}/2; x_r]$. Define

$$\mathfrak{x}_l := \mathfrak{x}_l(w) := \begin{cases} \frac{x_l}{2}, & l \in [0;k-1] \cap \mathbb{N}, \\ w, & l = k, \\ x_l, & l \in [k+1;n+1] \cap \mathbb{N}. \end{cases} \tag{4.68}$$

Let

$$\widetilde{Q}^{(l)} := \begin{cases} Q^{(l)}, & l \in [0;k-1] \cap \mathbb{N}, \\ Q^{(l-1)}, & l \in [k;n+2] \cap \mathbb{N}. \end{cases}$$

Moreover, suppose $\mathfrak{Q} := \{\mathfrak{Q}^{(l)}\}_{l=0}^{n+2}$ satisfy

$$\|\mathfrak{Q}^{(l)}\|_{\mathrm{F}}^2 = \begin{cases} 4\|Q^{(l)}\|_{\mathrm{F}}^2, & l \in [0;k-1] \cap \mathbb{N}, \\ 4\|V\|_{\mathrm{F}}^2, & l = k, \\ 2\left(\|Q^{(l-1)}\|_{\mathrm{F}}^2 + \|V\|_{\mathrm{F}}^2\right), & l \in [k+1;n+2] \cap \mathbb{N}. \end{cases} \tag{4.69}$$

Such $\mathfrak{Q}$ exists due to (4.67). Moreover, if $d \geq 2$, then it is obviously non-unique.

Lemma 4.3.6. *In the above setup, we have*

$$B^{\mathfrak{x},\mathfrak{Q}} := t\beta^2 \Big\{ (w - x_{l-1}) \left(\|Q^{(k)}\|_F^2 - \|V\|_F^2 \right) + \sum_{l=1}^{n} x_l \left(\|Q^{(l+1)}\|_F^2 - \|Q^{(l)}\|_F^2 \right) \Big\}.$$

Proof. The claim is a straightforward consequence of (4.68) and (4.69). □

Define the matrix $\mathfrak{D}^{(n+1)} \in \mathrm{Sym}^+(2d)$ block-wise as

$$\begin{aligned}
\mathfrak{D}^{(n+1)}|_{11} &:= \beta^2 t(U - \mathfrak{Q}^{(n+1)}|_{11}) + \beta^2(1-t)(U - Q^{(n)}) + \mathfrak{L}|_{11}, \\
\mathfrak{D}^{(n+1)}|_{12} &:= \beta^2 t(V - \mathfrak{Q}^{(n+1)}|_{12}) + \mathfrak{L}|_{12}, \\
\mathfrak{D}^{(n+1)}|_{21} &:= \beta^2 t(V - \mathfrak{Q}^{(n+1)}|_{12})^* + \mathfrak{L}|_{12}^*, \\
\mathfrak{D}^{(n+1)}|_{22} &:= \beta^2 t(U - \mathfrak{Q}^{(n+1)}|_{22}) + \beta^2(1-t)(U - Q^{(n)}) + \mathfrak{L}|_{22}.
\end{aligned}$$

Furthermore, we define

$$\mathrm{Sym}^+(2d) \ni \widetilde{\mathfrak{D}}^{(n+1)} := \begin{bmatrix} \beta^2(U - Q^{(n)}) + \Lambda & 0 \\ 0 & \beta^2(U - Q^{(n)}) + \Lambda \end{bmatrix}.$$

Lemma 4.3.7. *We have*

$$X_{n+1}^{(2)}(\widehat{y}, k, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) := \log \int_\Sigma \int_\Sigma \exp\Big(\langle \widehat{y}, \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle + \langle \mathfrak{D}^{(n+1)}(\sigma^{(1)} \shortparallel \sigma^{(2)}), \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle \Big) \\ \times \mathrm{d}\mu(\sigma^{(1)}) \mathrm{d}\mu(\sigma^{(2)}).$$

Proof. Since $\mathfrak{x}_{n+2} = 1$, the result follows from a straightforward calculation of the Gaussian integrals in (4.65) for $l = n+1$. □

Define

$$\widetilde{\mathfrak{L}} := \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix}, \widetilde{\mathfrak{U}} := \begin{bmatrix} U & Q^{(k)} \\ Q^{(k)} & U \end{bmatrix}.$$

Lemma 4.3.8. *For any* $y \in \mathbb{R}^d$, $l \in [0;n+2] \cap \mathbb{N}$, *we have*

$$X_l^{(2)}(y \shortparallel y, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w=x_{k-1}} = \begin{cases} 2X_{l-1}(y,x,\mathscr{Q},U,\Lambda), & l \in [k;n+2] \cap \mathbb{N}, \\ 2X_l(y,x,\mathscr{Q},U,\Lambda), & l \in [0;k-1] \cap \mathbb{N}. \end{cases}$$

Proof. A straightforward (decreasing) induction argument on l gives the result. Indeed: for $l = n+2$, an inspection of (4.64) and (3.8) immediately yields

$$X_{n+2}^{(2)}(y^{(1)} \shortparallel y^{(2)}, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) = X_{n+1}(y^{(1)},x,\mathscr{Q},U,\Lambda) + X_{n+1}(y^{(2)},x,\mathscr{Q},U,\Lambda),$$

where $y^{(1)}, y^{(2)} \in \mathbb{R}^d$. Let $\widehat{Z}^{(l)}$ be a Gaussian $2d$-dimensional vector with

$$\operatorname{Cov}\left[\widetilde{Z}^{(l)}\right] = 2\beta^2(\widetilde{\mathfrak{Q}}^{(l+1)} - \widetilde{\mathfrak{Q}}^{(l)}).$$

Define two Gaussian d-dimensional vectors $\widetilde{Z}^{(l),1}$ and $\widetilde{Z}^{(l),2}$ by demanding that

$$\widetilde{Z}^{(l)} = \widetilde{Z}^{(l),1} \shortparallel \widetilde{Z}^{(l),2}.$$

Due to (4.62) and (4.63), the vectors $\widetilde{Z}^{(l),1}$ and $\widetilde{Z}^{(l),2}$ are independent, for $l \in [k;n+1]$. We have $\widetilde{Z}^{(l),1} \sim \widetilde{Z}^{(l),2}$, for $l \in [0;k-1]$. Assume that $l \in [k;n+1] \cap \mathbb{N}$ and

$$X_{l+1}^{(2)}(y^{(1)} \shortparallel y^{(2)}, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) = X_l(y^{(1)},x,\mathscr{Q},U,\Lambda) + X_l(y^{(2)},x,\mathscr{Q},U,\Lambda).$$

By definition (4.64), we have

$$\begin{aligned} X_l^{(2)}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) &= \frac{1}{\mathfrak{x}_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \left[\exp\left(\mathfrak{x}_l X_{l+1}^{(2)}(y^{(1)} \shortparallel y^{(2)} + \widehat{Z}^{(l)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \mathfrak{L})\right)\right] \\ &= \frac{1}{x_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \left[\exp\left\{x_l\left(X_l(y^{(1)} + \widetilde{Z}^{(l),1},x,\mathscr{Q},U,\Lambda)\right.\right.\right. \\ &\quad \left.\left.\left. + X_l(y^{(2)} + \widetilde{Z}^{(l),2},x,\mathscr{Q},U,\Lambda)\right)\right\}\right] \\ &= X_{l-1}(y^{(1)},x,\mathscr{Q},U,\Lambda) + X_{l-1}(y^{(2)},x,\mathscr{Q},U,\Lambda). \end{aligned}$$

By the construction and previous formula, for $l = k-1$, we have

$$\begin{aligned} X_{k-1}^{(2)}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w=x_{k-1}} &= X_k^{(2)}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) \\ &= X_{k-1}(y^{(1)},x,\mathscr{Q},U,\Lambda) + X_{k-1}(y^{(2)},x,\mathscr{Q},U,\Lambda). \end{aligned}$$

Finally, for $l \in [0;k-2]$, we recursively obtain

$$\begin{aligned} X_l^{(2)}(y^{(1)} \shortparallel y^{(1)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w=x_{k-1}} &= \frac{1}{\mathfrak{x}_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \left[\exp\left(\mathfrak{x}_l X_{l+1}^{(2)}(y^{(1)} \shortparallel y^{(1)} + \widehat{Z}^{(l)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \mathfrak{L})|_{w=x_{k-1}}\right)\right] \\ &= \frac{2}{x_l} \log \mathbb{E}^{\widetilde{Z}^{(l),1}} \left[\exp\left\{\frac{x_l}{2}\left(X_{l+1}(y^{(1)} + \widetilde{Z}^{(l),1},x,\mathscr{Q},U,\Lambda)\right.\right.\right. \\ &\quad \left.\left.\left. + X_{l+1}(y^{(1)} + \widetilde{Z}^{(l),1},x,\mathscr{Q},U,\Lambda)\right)\right\}\right] \\ &= 2X_l(y^{(1)},x,\mathscr{Q},U,\Lambda). \end{aligned}$$

□

Remark 4.3.5. *Motivated by Lemmata 4.3.2 and 4.3.8 (see also Section 4.3.4), we pose the following problem. Is it true that, as in 1-D (see Panchenko (2005b); Talagrand (2006b)), there exists* $\mathfrak{Q} \in \mathscr{Q}'(\mathfrak{U}, 2d)$ *satisfying the assumption* (4.69) *such that the following inequality holds*

$$\inf_{\mathfrak{L}\in\mathrm{Sym}(2d)} \Big(-\langle \mathfrak{L}, \mathfrak{U} \rangle + X_0^{(2)}(0, \mathfrak{x}(w), \widehat{\mathfrak{Q}}(t), \mathfrak{L})|_{w=x_{k-1}} \Big)$$
$$\overset{?}{\leq} 2 \inf_{\Lambda\in\mathrm{Sym}(d)} \Big(-\langle \Lambda, U \rangle + X_0(0, x, \mathscr{Q}, U, \Lambda) \Big)? \tag{4.70}$$

Similar problems have at first been posed in Talagrand (2007b). The resolution of the above problem seems to require more detailed information on the behaviour of the Parisi functional (4.41) *or, equivalently, of the solution of* (4.39) *as a function of* $Q \in \mathscr{Q}(U,d)$.

4.3.4 Talagrand's a priori estimates

We start from defining a class of the almost optimal paths for the optimisation problem in (4.42). Recall the following convenient definition from Panchenko (2005b).

Definition 4.3.2. *Given* $U \in \mathrm{Sym}^+(d)$*, we shall call the triple* $(n, \rho^*, \Lambda^*) \in \mathbb{N} \times \mathscr{Q}'_n(U,d) \times \mathbb{R}^d$ *a* θ-optimiser *of the Parisi functional* (4.41)*, if it satisfies the following two conditions*

$$\mathscr{P}(\beta, \rho^*, \Lambda^*) \leq \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda) + \theta. \tag{4.71}$$

$$\mathscr{P}(\beta, \rho^*, \Lambda^*) = \inf_{\substack{\rho\in\mathscr{Q}'_n(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda). \tag{4.72}$$

Remark 4.3.6. *It is obvious that for any* $\theta > 0$ *such a* θ*-optimiser exists. The main convenient feature of this definition (as pointed out in Talagrand (2006b)) is that n (the number of jumps of* ρ^**) is finite and fixed.*

Recalling (4.11), we set

$$\phi^{(x,\mathscr{Q},\Lambda)}(t) := -\langle U, \Lambda \rangle - \frac{t\beta^2}{2} \sum_{k=1}^{n} x_k \left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2 \right) + X_0(x, \mathscr{Q}, U, \Lambda). \tag{4.73}$$

Under the following assumption (at first proposed in 1-D in Talagrand (2006b)), we shall effectively prove that remainder term almost vanishes on the θ minimisers of (4.41), see Theorem 4.3.1.

Assumption 4.3.1. *Let* $\mathfrak{U} \in \mathrm{Sym}^+(2d)$ *be defined by* (4.51)*. We fix arbitrary* $t_0 \in [0;1)$*,* $\varepsilon > 0$ *and* $\delta > 0$*. There exists* $K = K(t_0, \varepsilon, \delta, \mathfrak{U}) > 0$*,* $\theta(t_0, \varepsilon, \delta, \mathfrak{U}) > 0$*, and* $N_0 = N_0(t_0, \varepsilon, \delta, \mathfrak{U}) \in \mathbb{N}$ *and* $\mathfrak{L}^* \in \mathrm{Sym}(2d)$ *with the following property:*

If (n, ρ^*, Λ^*) *is a* θ*-optimiser, for some* $\theta \in (0; \theta(t_0, \varepsilon, \delta, \mathfrak{U})]$*, then uniformly, for all* $t \in [0; t_0)$*,* $N > N_0$ *and all* $k \in [1;n] \cap \mathbb{N}$*, we have*

$$\varphi_N^{(2)}(k, t, x^*, Q^*, \Sigma_N^{(2)}(\mathfrak{L}^*, \mathfrak{U}, \varepsilon, \delta)) \leq 2\phi^{(x^*, \mathscr{Q}^*, \Lambda)}(t) - \frac{1}{K} \|Q^{*(k)} - V\|_F^2 + C(\varepsilon + \delta). \tag{4.74}$$

Remark 4.3.7. *The validity of the above assumption for general a priori measures is an open problem. However, in the particular case of the Gaussian a priori distribution the assumption is indeed effectively satisfied. See Section 5.2 and Theorem 5.2.1, in particular. This gives a complete proof of the Parisi formula for the case of Gaussian spins.*

Remark 4.3.8. *If the bound* (4.70) *holds then Lemma 4.3.6 with* $w = x_{r-1}$ *would imply that*

$$\varphi_N^{(2)}(k,t,\Sigma_N^{(2)}(\mathfrak{L}^*,\mathfrak{U},\varepsilon,\delta)) \overset{?}{\leq} 2\phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(t) + C(\varepsilon+\delta). \tag{4.75}$$

The above inequality would then be a starting point for the a priori estimates in the spirit of Talagrand (2006b) which might lead to the proof of Assumption 4.3.1.

4.3.5 Gronwall's inequality and the Parisi formula

Theorem 4.3.1. *Suppose Assumption 4.3.1 holds.*
Then we have

$$\lim_{N\uparrow+\infty} p_N(\beta) = \sup_{U\in\mathrm{Sym}^+(d)} \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda).$$

Proof. The proof follows the argument of Talagrand (2006b) (see also Panchenko (2005b)) with the adaptations to the case of multidimensional spins. The main ingredients are the Gronwall inequality and Lemma 4.3.2. Theorem 4.1.1 implies that

$$\lim_{N\uparrow+\infty} p_N(\beta) \leq \sup_{U\in\mathrm{Sym}^+(d)} \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda).$$

We now turn to the proof of the matching lower bound. As in the proof of Theorem 3.1.2, it is enough to show that

$$\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow+\infty} \varphi_N(1,x,Q,B(U,\varepsilon)) \geq \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda). \tag{4.76}$$

1. We fix an arbitrary $U \in \mathrm{Sym}^+(d)$. Fix also some $t_0 \in [0;1)$. By Assumption 4.3.1, we can find the corresponding $\theta(t_0,V,U) > 0$ with the properties listed in the assumption. We pick any $\theta \in (0;\theta(t_0,V,U)]$ and let (n,ρ^*,Λ^*) be a correspondent θ-optimiser. Note that, by definition (4.73), we have

$$\phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(1) = \mathscr{P}(\beta,\rho^*,U,\Lambda^*)$$

   and, by Definition 4.3.2,

$$|\phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(1) - \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,U,\Lambda)| \leq \theta. \tag{4.77}$$

2. We denote

$$\Delta_N(t) := \phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(t) - \varphi_N(t,x^*,Q^*,B(U,\varepsilon)).$$

Note that, due to (4.10), we obviously have

$$\Delta_N(t) \geq -C\varepsilon. \tag{4.78}$$

Define

$$\Delta(t) := \lim_{N\uparrow+\infty} \Delta_N(t).$$

The definition (4.73) and Theorem 4.1.4 yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t) \leq \frac{1}{2}\sum_{k=0}^{n-1}(x_{k+1}-x_k)\mu_k\left[\|R_N(\sigma^{(1)},\sigma^{(2)})-Q^{(k)}\|_{\mathrm{F}}^2\right]+C\varepsilon. \tag{4.79}$$

3. Let us set $D := \sup_{\sigma\in\Sigma}\|\sigma\|_2$. We note that, for any $\sigma^{(1)},\sigma^{(2)} \in \Sigma_N$, we have

$$R(\sigma^{(1)},\sigma^{(2)}) \in [-D^2;D^2]^{d\times d}.$$

Given the constant K from (4.74), for any $c>0$, we define the set

$$\Sigma_N^{(2),k}(U,\varepsilon) := \left\{(\sigma^{(1)},\sigma^{(2)}) \in \Sigma_N(B(U,\varepsilon))^2 : \|R(\sigma^{(1)},\sigma^{(2)})-Q^{(k)}\|_{\mathrm{F}}^2 \geq 2K\left(\Delta_N(t)+c\right)\right\}. \tag{4.80}$$

It is easy to see that by compactness we can find a finite covering of $\Sigma_N^{(2),k}(U,\varepsilon)$ by the neighbourhoods (4.56) with centres, e.g., in the corresponding set of admissible overlap matrices

$$\mathscr{V}_N^{(k)}(U,\varepsilon) := \left\{R(\sigma^{(1)},\sigma^{(1)}) \in [-D^2;D^2]^{d\times d} : (\sigma^{(1)},\sigma^{(2)}) \in \Sigma_N^{(2),k}(U,\varepsilon)\right\}.$$

That is, there exists $M = M(\varepsilon,\delta) \in \mathbb{N}$ and the finite collections of matrices $\{V(i)\}_{i=1}^M \subset \mathscr{V}_N^{(k)}(U,\varepsilon)$ and $\{U(i)\}_{i=1}^M \subset B(U,\varepsilon)\cap \mathrm{Sym}^+(d)$ such that

$$\Sigma_N^{(2),k}(U,\varepsilon) \subset \bigcup_{i=1}^M \Sigma_N^{(2)}(\mathfrak{L}^*(i),\mathfrak{U}(i),\varepsilon,\delta), \tag{4.81}$$

where

$$\mathfrak{U}(i) := \begin{bmatrix} U(i) & V(i) \\ V^*(i) & U(i)\end{bmatrix} \in \mathrm{Sym}^+(2d),$$

and $\mathfrak{L}^*(i)$ is the corresponding δ-minimal Lagrange multiplier.

4. Given $i \in [1;M]\cap\mathbb{N}$, let $(n(i),x^*(i),Q^*(i),\Lambda^*(i))$ be the corresponding to $U(i)$ $\theta(i)$-optimisers. Due to Lipschitzianity of the Parisi functional (Proposition 4.2.2) and the fact that $U(i) \in B(U,\varepsilon)$ we can assume that $n(i) = n$. Using the bound (4.74) and the definition (4.80), we obtain

$$\varphi_N^{(2)}(k,t,x_i^*,Q_i^*,\Sigma_N^{(2)}(\mathfrak{L}^*(i),\mathfrak{U}(i),\varepsilon,\delta)) \leq 2\phi^{(x^*(i),\mathscr{Q}^*(i),\Lambda^*(i))}(t)-\frac{1}{K}\|Q^{(k)}-V(i)\|_{\mathrm{F}}^2+C(\varepsilon+\delta)$$

$$\leq 2\varphi_N(t,x^*,Q^*,B(U,\varepsilon)) - c + C(\varepsilon+\delta),$$

where the last inequality is again due to Lipschitzianity of the Parisi functional (Proposition 4.2.2) which allows to approximate functional's value at $(x^*(i),Q^*(i),\Lambda^*(i))$ by the value at (x^*,Q^*,Λ^*) paying the cost of at most $C\varepsilon$. Choose $c > C(\varepsilon+\delta)$. Then Lemma 4.3.2 implies that there exists $L = L(\varepsilon,\delta,c) > 0$ such that

$$\mu_k\Big(\Sigma_N^{(2)}(\mathfrak{L}^*,\mathfrak{U},\varepsilon,\delta)\Big) \leq L\exp\left(-\frac{N}{L}\right).$$

Therefore, the inclusion (4.81) gives

$$\mu_k\Big(\Sigma_N^{(2),k}(U,\varepsilon)\Big) \leq LM\exp\left(-\frac{N}{L}\right). \tag{4.82}$$

Hence, for each $k \in [1;n]\cap\mathbb{N}$, we have

$$\begin{aligned}\mu_k\Big[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2\Big] &= \mu_k\Big[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2 \mathbb{1}_{\Sigma_N^{(2),k}(U,\varepsilon)}(\sigma^{(1)},\sigma^{(2)})\Big] \\ &\quad + \mu_k\Big[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2\Big(1 - \mathbb{1}_{\Sigma_N^{(2),k}(U,\varepsilon)}(\sigma^{(1)},\sigma^{(2)})\Big)\Big] \\ &=: \mathrm{I} + \mathrm{II}. \end{aligned} \tag{4.83}$$

For all $(\sigma^{(1)},\sigma^{(2)}) \in \Big(\Sigma_N(B(U,\varepsilon))^2 \setminus \Sigma_N^{(2),k}(U,\varepsilon,\delta)\Big)$, we have by definition

$$\|R(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2 < 2K\left(\Delta_N(t) + c\right).$$

Therefore, using Remark 4.1.4, we arrive to

$$\mathrm{II} \leq 2K\left(\Delta_N(t) + c\right). \tag{4.84}$$

The bound (4.82) assures that

$$\mathrm{I} \leq LM\exp\left(-\frac{N}{L}\right). \tag{4.85}$$

5. Combining (4.84) and (4.85) with (4.83) and (4.79), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t) \leq 2K\left(\Delta_N(t)+c\right) + LM\exp\left(-\frac{N}{L}\right) + C(\varepsilon+\delta).$$

Hence,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big((\Delta_N(t)+c)\exp(-2Kt)\Big) &= \exp(-2Kt)\Big(\frac{\mathrm{d}}{\mathrm{d}t}(\Delta_N(t)+c) - 2K(\Delta_N(t)+c)\Big) \\ &\leq \exp(-2Kt)\Big(\frac{\mathrm{d}}{\mathrm{d}t}(LM\exp\left(-\frac{N}{L}\right) + C(\varepsilon+\delta)\Big).\end{aligned}$$

Integrating the above inequality and noting that due to (4.10) $|\Delta_N(0)| \leq C\varepsilon$, we arrive to

$$\begin{aligned}\varDelta_N(t)+c \leq &(C\varepsilon+c)\exp(-2Kt)+LM\exp\left(-\frac{N}{L}\right)\\ &+C(\varepsilon+\delta)(\exp(-2Kt)-1)+C(\varepsilon+\delta).\end{aligned}$$

Passing consequently to the limits $N\uparrow+\infty$, $\varepsilon\downarrow+0$, $\delta\downarrow+0$ and finally $c\downarrow+0$ in the above inequality, we get

$$\lim_{\varepsilon\downarrow+0}\varDelta(t)\leq 0, \quad \text{for all } t\in[0;t_0].$$

The existence of the $N\uparrow+\infty$ limits is guaranteed by the general result of Guerra & Toninelli (2003). The limits $\varepsilon\downarrow+0$, $\delta\downarrow+0$ exist due to monotonicity. Finally, combining the above inequality with (4.78), we get

$$\lim_{\varepsilon\downarrow+0}\varDelta(t)=0, \quad \text{for all } t\in[0;t_0]. \tag{4.86}$$

6. Now it is easy to extend the validity of (4.86) onto the whole interval $[0;1]$. Indeed, due to the boundedness of the derivatives of φ_N and ϕ, we have, for any $t\in[0;1]$,

$$\begin{aligned}\varDelta_N(t) &\leq \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\varDelta_N(t)\mathrm{d}t\\ &= \left(\int_0^{t_0}+\int_{t_0}^1\right)\frac{\mathrm{d}}{\mathrm{d}t}\varDelta_N(t)\mathrm{d}t\\ &\leq \big(\varDelta_N(t_0)-\varDelta_N(0)\big)+\int_{t_0}^1\left|\frac{\mathrm{d}}{\mathrm{d}t}\varDelta_N(t)\right|\mathrm{d}t\\ &\leq \varDelta_N(t_0)+L(1-t_0).\end{aligned} \tag{4.87}$$

Passing to the $N\uparrow+\infty$ limit, applying (4.86), and then to $t_0\to 1$ limit in (4.87), we get

$$\lim_{\varepsilon\downarrow+0}\varDelta(t)=0, \quad \text{for all } t\in[0;1].$$

7. In particular, the previous formula yields

$$0=\lim_{\varepsilon\downarrow+0}\varDelta(1)=\phi^{(x^*,\mathscr{Q}^*,\varLambda^*)}(1)-\lim_{\varepsilon\downarrow+0}\varphi_N(1,x^*,Q^*,B(U,\varepsilon)).$$

Note that $\varphi_N(1,x,Q,B(U,\varepsilon))$ does not depend on the choice of x and Q. Hence, by (4.77), we obtain

$$|\lim_{\varepsilon\downarrow+0}\varphi_N(1,x^*,Q^*,B(U,\varepsilon))-\inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \varLambda\in\mathrm{Sym}(d)}}\mathscr{P}(\beta,\rho,U,\varLambda)|\leq\theta.$$

The proof of (4.76) is finished by noticing that the θ can be made arbitrary small.

□

5

The SK model with multidimensional Gaussian spins

In this chapter, we prove the local Parisi formula (Theorem 2.3.10) for the SK model with multidimensional Gaussian spins.

5.1 Introduction

Let $\Sigma := \mathbb{R}^d$ and fix some vector $h \in \mathbb{R}^d$. Let $\mu \in \mathscr{M}_{\mathrm{f}}(\Sigma)$ be the finite measure with the following density (with respect to the Lebesgue measure λ on Σ)

$$\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}(\sigma) = \left(\frac{\det C}{(2\pi)^d}\right)^{1/2} \exp\left(-\frac{1}{2}\langle C\sigma, \sigma\rangle + \langle h, \sigma\rangle\right), \tag{5.1}$$

where $C \in \mathrm{Sym}^+(d)$. Note that, given $m \in \mathbb{R}^d$ and $C \in \mathrm{Sym}^+(d)$ such that $\det C \neq 0$, the density (5.1) with $h := Cm$ coincides (up to the constant factor $\exp\left(-\frac{1}{2}\langle Cm, m\rangle\right)$) with the Gaussian density with covariance matrix C^{-1} and mean m.

Remark 5.1.1. *It turns out that only matrices C with sufficiently large eigenvalues will result in finite global free energy, cf. Lemma 5.2.8. The local free energy is, in contrast, always finite, see Lemma 5.2.7 and Theorem 5.1.1.*

Consider the function $f : (0 : +\infty)^2 \to \mathbb{R}$ given by

$$f(c,u) = \begin{cases} \beta^2 u^2 + \log cu - cu + 1, & u \in (0; \frac{\sqrt{2}}{2\beta}], \\ (2\sqrt{2}\beta - c)u + \log \frac{c}{\beta} - \frac{1}{2}(1 + \log 2), & u \in (\frac{\sqrt{2}}{2\beta}; +\infty]. \end{cases} \tag{5.2}$$

The following result shows that, at least, in the highly symmetric situation (5.1) with $h = 0$ the multidimensional Parisi formula indeed holds true (see Lemma 5.2.7 for an explanation why the result is indeed a Parisi-type formula).

Theorem 5.1.1. *Let μ satisfy* (5.1) *with $h = 0$. Assume that the matrices U and C are simultaneously diagonalisable in the same basis. Denote by $\{c_v \in \mathbb{R}_+\}_{v=1}^d$ and $\{u_v \in \mathbb{R}_+\}_{v=1}^d$ the eigenvalues of the matrices C and U, respectively. Moreover, assume that* $\min_v u_v > 0$ *and* $\min_v c_v > 0$.

Then we have

$$\lim_{\varepsilon \downarrow +0} \lim_{N \uparrow +\infty} p_N(\Sigma_N(B(U, \varepsilon))) = \sum_{v=1}^d f(c_v, u_v).$$

Remark 5.1.2. *Close results have previously been obtained in the case of the spherical model by Panchenko & Talagrand (2007b), from where we borrow the general methodology of the proof of the above theorem. As noted by Panchenko & Talagrand (2007b), another more straightforward way to obtain the Theorem 5.1.1 is to diagonalise the interaction matrix G and use the properties of the corresponding random matrix ensemble.*

5.2 Proof of the local Parisi formula

In this section, we prove Theorem 5.1.1. The rich symmetries of the Gaussian a priori distribution allow rather explicit computations of the X_0 terms (see (3.10)). This allows us to prove that the analogon of Assumption 4.3.1 is satisfied, implying the Parisi formula for the local free energy (Theorem 5.1.1).

Remark 5.2.1. *The case of Gaussian spins is very tractable due to the (unusually) good symmetry (i.e., the rotational invariance) of the Gaussian measure. Therefore, it is not surprising that in this case the calculus resembles the one for the spherical SK model, cf. Panchenko & Talagrand (2007b); Talagrand (2006a).*

We start from the estimates under a generic (i.e., no simultaneous diagonalisation, cf. Section 4.2.3) scenario.

5.2.1 The case of positive increments

Let, for $k \in [0;n] \cap \mathbb{N}$,

$$\Delta Q^{(k)} := Q^{(k+1)} - Q^{(k)}.$$

We define, for $\Lambda \in \mathrm{Sym}(d)$, a family of matrices $\left\{D^{(l)} \in \mathbb{R}^{d\times d}\right\}_{l=0}^{n+1}$ as follows

$$D^{(n+1)} := C,$$

and, further, for $k \in [0;n] \cap \mathbb{N}$,

$$D^{(k)} := C - \Lambda - 2\beta^2 \sum_{l=k}^{n} x_l \Delta Q^{(l)}. \tag{5.3}$$

We assume that the matrices Λ and C are such that, for all $l \in [1;n+1] \cap \mathbb{N}$, we have

$$D^{(l)} \succ 0.$$

We need the following two small (and surely known) technical Lemmata which exploit the symmetries of our Gaussian setting. We include their statements for reader's convenience.

Lemma 5.2.1. *Fix some vector $h \in \mathbb{R}^d$ and a Gaussian random vector $z \in \mathbb{R}^d$ with $\mathrm{Var}\, z = C^{-1} \in \mathbb{R}^{d\times d}$.*

Then we have

$$\mathbb{E}^z\left[\exp\left(\langle z,h\rangle + \langle \Lambda\sigma,\sigma\rangle\right)\right] = \left(\det\left[C\left(C-\Lambda\right)^{-1}\right]\right)^{1/2} \times \exp\left(\frac{1}{2}\left\langle (C-\Lambda)^{-1}h,h\right\rangle\right).$$

Proof. This is a standard Gaussian averaging argument.

□

Lemma 5.2.2. *For a positive definite matrix $\Delta Q \in \mathrm{Sym}(d)$, let $z \sim \mathcal{N}(0, \Delta Q)$. We fix also another positive definite matrix $D \in \mathrm{Sym}(d)$ such that $\Delta Q^{-1} \succ D^{-1}$.*

Then we have

$$\mathbb{E}^z\left[\exp\left(\frac{1}{2}\langle D^{-1}(z+h), z+h\rangle\right)\right] = \left(\det\left[D(D-\Delta Q)^{-1}\right]\right)^{-1/2} \times \int_{\mathbb{R}^d} \exp\left(\frac{1}{2}\langle (D-\Delta Q)^{-1}h, h\rangle\right).$$

Proof. This is a standard Gaussian averaging argument. See, e.g., Talagrand (2006a) for an argument in 1-D.

□

Now we are ready to compute the term $X_0(x, \mathscr{Q}, U, \Lambda)$ (see (3.10)) corresponding to the a priori distribution (5.1) in a rather explicit way.

Lemma 5.2.3. *We have*

$$X_0(x, \mathscr{Q}, U, \Lambda) = \frac{1}{2}\left(\langle [D^{(1)}]^{-1}, \Delta Q^{(0)}\rangle + \langle [D^{(1)}]^{-1}h, h\rangle + \sum_{l=1}^{n}\frac{1}{x_l}\log\left(\frac{\det D^{(l+1)}}{\det D^{(l)}}\right)\right).$$

Proof. 1. We start from computing the following quantity

$$X_{n+1} := \log\int_{\mathbb{R}^d}\exp\left(\sum_{l=0}^{n}\langle Y^{(l)}, \sigma\rangle + \langle \Lambda\sigma, \sigma\rangle\right) \mathrm{d}\mu(\sigma), \tag{5.4}$$

where $Y^{(l)} \in \mathbb{R}^d$ are independent Gaussian vectors with variance

$$\mathrm{Var}\left[Y^{(l)}\right] = 2\beta^2 \Delta Q^{(l)}.$$

We denote

$$\widetilde{h} := h + \sum_{l=0}^{n} Y^{(l)}.$$

Lemma 5.2.1 gives

$$\int_{\mathbb{R}^d}\exp\left(\sum_{l=0}^{n}\langle Y^{(l)}, \sigma\rangle + \langle \Lambda\sigma, \sigma\rangle\right)\mathrm{d}\mu(\sigma) = \left(\det\left[C(C-\Lambda)^{-1}\right]\right)^{1/2} \times \exp\left(\frac{1}{2}\left\langle (C-\Lambda)^{-1}\widetilde{h}, \widetilde{h}\right\rangle\right).$$

2. Next, we define, for $l \in [0;n] \cap \mathbb{N}$, recursively the following quantities

$$X_l := \frac{1}{x_l} \log \mathbb{E}^{Y_l} \left[\exp\left(x_l X_{l+1}\right)\right].$$

Applying the Lemma 5.2.2 to (5.4) recursively, we obtain

$$X_1 := \frac{1}{2}\langle [(D^{(1)}]^{-1}\left(Y^{(0)}+h\right), Y^{(0)}+h\rangle + \frac{1}{2}\sum_{l=1}^{n}\frac{1}{x_l}\log\left(\frac{\det D^{(l+1)}}{\det D^{(l)}}\right). \tag{5.5}$$

Recall that we have

$$\begin{aligned} X_0 &= \lim_{x\to+0}\frac{1}{x}\log \mathbb{E}^{Y_0}\left[\exp\left(xX_1\right)\right] \\ &= \mathbb{E}^{Y_0}\left[X_1\right] \end{aligned} \tag{5.6}$$

and note that

$$\mathbb{E}^{Y_0}\left[\langle [D^{(1)}]^{-1}(Y^{(0)}+h), Y^{(0)}+h\rangle\right] = 2\beta^2\langle [D^{(1)}]^{-1}, \Delta Q^{(0)}\rangle + \langle [D^{(1)}]^{-1}h, h\rangle. \tag{5.7}$$

Hence, combining (5.6) and (5.7) with (5.5), we obtain the theorem.

□

5.2.2 Simultaneous diagonalisation scenario

In what follows, we employ the simultaneous diagonalisation scenario introduced in Section 4.2.3. Suppose that, for $l \in [0;n+1] \cap \mathbb{N}$, and some matrix $O \in \mathscr{O}(d)$, we have

$$D^{(l)} := O^* d^{(l)} O,$$

where the vectors $d^{(l)} \in \mathbb{R}^d$, for $l \in [0;n] \cap \mathbb{N}$, satisfy

$$0 \prec d^{(l)} \prec d^{(l+1)}.$$

That is, the vectors $d^{(l)}$ are (component-wise) increasingly ordered and non-negative.

Lemma 5.2.4. *We have*

$$X_0(x, \mathscr{Q}, U, \Lambda) = \frac{1}{2}\sum_{v=1}^{d}\left(\frac{2\beta^2 q_v^{(1)} + h_v^2}{d_v^{(1)}} + \sum_{l=1}^{n}\frac{1}{x_l}\log\left(\frac{d_v^{(l+1)}}{d_v^{(l)}}\right)\right), \tag{5.8}$$

$$\frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_F^2 - \|Q^{(k)}\|_F^2\right) = \frac{\beta^2}{2}\sum_{k=1}^{n} x_l\left(\|q^{(k+1)}\|_2^2 - \|q^{(k)}\|_2^2\right). \tag{5.9}$$

Proof. This is a standard argument which relies on the standard invariance properties of the determinant and the matrix trace.

□

Define the 1-D Parisi functional for the case (5.1) as

$$\mathscr{P}(\rho,\lambda) := -\lambda u + \frac{2\beta^2 q^{(1)} + h^2}{d^{(1)}} + \sum_{l=1}^{n} \frac{1}{x_l} \log\left(\frac{d^{(l+1)}}{d^{(l)}}\right)$$
$$- \beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2\right). \tag{5.10}$$

Proposition 5.2.1. *There exists $C = C(\Sigma) > 0$ such that, for all $u \in \mathbb{R}_+^d$ and all $\varepsilon, \delta > 0$, there exists an δ-minimal Lagrange multiplier $\lambda = \lambda(U,\varepsilon,\delta) \in \mathbb{R}^d$ in* (3.11) *such that, for all $t \in [0;1]$ and all (x,ρ), we have*

$$p_N(\Sigma_N(\mathscr{V}(\Lambda,U,\varepsilon,\delta))) \leq \frac{1}{2} \inf_{\rho,\lambda} \left(\sum_{v=1}^{d} \mathscr{P}(\rho_v,\lambda_v)\right) + C(\varepsilon+\delta) \tag{5.11}$$

and

$$\lim_{N\uparrow+\infty} p_N(\Sigma_N(B(U,\varepsilon))) \geq \frac{1}{2} \inf_{\rho,\lambda} \left(\sum_{v=1}^{d} \mathscr{P}(\rho_v,\lambda_v) + \lim_{N\uparrow+\infty} \int_0^1 \mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon)))\mathrm{d}t\right)$$
$$+ C(\varepsilon+\delta). \tag{5.12}$$

Proof. We combine (5.8) and (5.9) and the Proposition 4.1.1 to get (5.11) and (5.12). □

5.2.3 The Crisanti-Sommers functional in 1-D

In this subsection, we adapt the proof of Talagrand (2006a) to obtain the equivalence between the (very tractable) Crisanti-Sommers functional (Crisanti & Sommers, 1992) and the Parisi one (5.10) in the case of the Gaussian a priori measure (5.1). Similar ideas based on the symmetry of the a priori measure were exploited in the case of the spherical models by Ben Arous *et al.* (2001); Panchenko & Talagrand (2007b).

We restrict the consideration to 1-D situation for a moment. Given $u \geq 0$, consider $\rho \in \mathscr{Q}_n'(u,1)$, $\lambda \in \mathbb{R}$, $h \in \mathbb{R}$ and let $\{d^{(l)} \in \mathbb{R}\}_{l=1}^{n+1}$ be the scalars playing the role of matrices $D^{(l)}$ (cf. (5.3)). That is,

$$d^{(l)} := c - \lambda - 2\beta^2 \sum_{k=l}^{n} x_k \left(q^{(k+1)} - q^{(k)}\right),$$
$$d^{(n+1)} := c.$$

We define, for $k \in [1;n] \cap \mathbb{N}$, the family of vectors $\{s^{(k)} \in \mathbb{R}^d\}_{k=0}^{n}$ by

$$s^{(k)} := \sum_{l=k}^{n} x_l \left(q^{(l+1)} - q^{(l)}\right). \tag{5.13}$$

We also define the Crisanti-Sommers functional as follows

$$\mathscr{C}\mathscr{S}(\rho) := 1 - cu + h^2 s^{(1)} + \frac{q^{(1)}}{s^{(1)}} + \sum_{l=1}^{n-1} \frac{1}{x_l} \log\left(\frac{s^{(l)}}{s^{(l+1)}}\right) + \log\left[c(u - q^{(n)})\right]$$
$$+ \beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2\right). \tag{5.14}$$

Lemma 5.2.5. *If (ρ,λ) is an optimiser for* (5.10)*, that is,*

$$\mathscr{P}(\rho,\lambda) = \inf_{(\rho',\lambda')} \mathscr{P}(\rho',\lambda'), \tag{5.15}$$

then, for all $k \in [1;n] \cap \mathbb{N}$, the pair (ρ,λ) satisfies

$$q^{(k)} = \frac{h^2 + 2\beta^2 q^{(1)}}{[d^{(1)}]^2} + \sum_{l=1}^{k-1} \frac{1}{x_l} \left(\frac{1}{d^{(l)}} - \frac{1}{d^{(l+1)}} \right). \tag{5.16}$$

Moreover,

$$\lambda = c - 2\beta^2 (u - q^{(n)}) - (u - q^{(n)})^{-1}, \tag{5.17}$$

and, for all $k \in [1;n] \cap \mathbb{N}$, we have

$$\frac{1}{s^{(k+1)}} - \frac{1}{s^{(k)}} = 2\beta^2 x_k \left(q^{(k+1)} - q^{(k)} \right), \tag{5.18}$$

and also

$$s^{(k)} = \frac{1}{d^{(k)}}. \tag{5.19}$$

Remark 5.2.2. *In the formulation of the theorem (as well as elsewhere), it is implicit that $d^{(k)} = d^{(k)}(\rho,\lambda)$ and $s^{(k)} = s^{(k)}(\rho,\lambda)$.*

Proof. 1. Rearranging the terms in (5.10), we observe that

$$\begin{aligned} \mathscr{P}(\rho',\lambda') = & -\lambda u + \frac{2\beta^2 q^{(1)} + h^2}{d^{(1)}} + \sum_{l=2}^{n} \log d^{(l)} \left(\frac{1}{x_{l-1}} - \frac{1}{x_l} \right) + \frac{1}{x_n} \log d^{(n+1)} - \frac{1}{x_1} \log d^{(1)} \\ & - \beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right). \end{aligned} \tag{5.20}$$

We compute, for $k,l \in [1;n] \cap \mathbb{N}$,

$$\frac{\partial d^{(l)}}{\partial q^{(k)}} = \begin{cases} 0, & k < l, \\ 2\beta^2 x_k, & l = k, \\ 2\beta^2 \left(x_k - x_{k-1} \right), & k > l. \end{cases} \tag{5.21}$$

Using (5.21) and the representation (5.20), we compute the necessary condition for (q,λ) satisfy (4.72), for $k \in [2;n] \cap \mathbb{N}$,

$$\begin{aligned} 0 = \frac{\partial}{\partial q^{(k)}} \mathscr{P}(q,\lambda) = & 2\beta^2 \left(x_k - x_{k-1} \right) \left[-\frac{2\beta^2 q^{(1)} + h^2}{[d^{(1)}]^2} + \sum_{l=2}^{k-1} \frac{1}{d^{(l)}} \left(\frac{1}{x_{l-1}} - \frac{1}{x_l} \right) \right. \\ & \left. + \frac{1}{d^{(k)} x_{k-1}} - \frac{1}{x_1 d^{(1)}} + q_k \right]. \end{aligned} \tag{5.22}$$

We also have (for $k=1$)

$$0=\frac{\partial}{\partial q^{(1)}}\mathscr{P}(q,\lambda)=2\beta^2\left[\frac{d^{(1)}-x_1\left(q^{(1)}+h^2\right)}{[d^{(1)}]^2}-\frac{x_1}{x_1 d^{(1)}}+x_1 q^{(1)}\right]$$

$$=2\beta^2 x_1\left[q^{(1)}-\frac{\left(q^{(1)}+h^2\right)}{[d^{(1)}]^2}\right]. \tag{5.23}$$

Relations (5.22) and (5.23) then imply (5.16).

2. Using the fact that

$$\frac{\partial d^{(l)}}{\partial \lambda}=-1,$$

we obtain

$$\frac{\partial}{\partial \lambda}\mathscr{P}(q,\lambda)=-u+\frac{h^2+2\beta^2 q^{(1)}}{[d^{(1)}]^2}+\sum_{l=1}^{n-1}\frac{1}{x_l}\left(\frac{1}{d^{(l)}}-\frac{1}{d^{(l+1)}}\right)+\frac{1}{d^{(n)}}. \tag{5.24}$$

Applying (5.16) with $k=n$ in (5.24), we obtain that the necessary condition for λ to satisfy (5.15) is as follows

$$0=\frac{\partial}{\partial \lambda}\mathscr{P}(q,\lambda)=-u+q^{(n)}+\frac{1}{x_n}\left(\frac{1}{d^{(n)}}-\frac{1}{d^{(n+1)}}\right)$$

$$=-u+q^{(n)}+\frac{1}{d^{(n)}}=-u+q^{(n)}+\left(c-\lambda-2\beta^2(u-q^{(n)})\right)^{-1} \tag{5.25}$$

which implies (5.17).

3. Relation (5.18) is proved as follows. Subtracting the relations (5.16), we obtain, for $k\in[1;n-1]\cap\mathbb{N}$,

$$x_k\left(q^{(k+1)}-q^{(k)}\right)=\frac{1}{d^{(k)}}-\frac{1}{d^{(k+1)}}. \tag{5.26}$$

By (5.25), we have

$$x_n\left(q^{(n+1)}-q^{(n)}\right)=u-q^{(n)}=\frac{1}{d^{(n)}}.$$

(That is, (5.26) is valid also for $k=n$.) Combining the previous two relations, we get, for $k\in[1;n]\cap\mathbb{N}$,

$$s^{(k)}=\frac{1}{d^{(k)}}. \tag{5.27}$$

Using (5.27) and (5.26), we get

$$2\beta^2 x_k\left(q^{(k+1)}-q^{(k)}\right)=d^{(k+1)}-d^{(k)}$$

$$\begin{aligned}
\text{(by (5.26))} \ &= d^{(k+1)}d^{(k)}x_k\left(q^{(k+1)}-q^{(k)}\right) = d^{(k+1)}d^{(k)}\left(s^{(k)}-s^{(k+1)}\right) \\
\text{(by (5.27))} \ &= \frac{1}{s^{(l+1)}}-\frac{1}{s^{(l)}}
\end{aligned}$$

which is (5.18).

□

Lemma 5.2.6. *If ρ is an optimiser of* (5.14)*, that is,*

$$\mathscr{CS}(\rho) = \inf_{\rho'} \mathscr{CS}(\rho'),$$

then, for all $l \in [1;n] \cap \mathbb{N}$, (5.18) *holds.*

Proof. The strategy is the same as in the previous lemma. We rearrange the summands in (5.14) to get

$$\begin{aligned}
\mathscr{CS}(\rho) =& h^2 s^{(1)} + \frac{q^{(1)}}{s^{(1)}} + \frac{\log s^{(1)}}{x_1} - \frac{\log s^{(n)}}{x_{n-1}} + \sum_{l=2}^{n-1}\left(\frac{1}{x_l}-\frac{1}{x_{l+1}}\right)\log s^{(l)} \\
&+ \log\left(c(u-q^{(n)})\right) + \beta^2 \sum_{l=1}^{n} x_l\left([q^{(l+1)}]^2-[q^{(l)}]^2\right).
\end{aligned} \tag{5.28}$$

We have, for $k,l \in [1;n] \cap \mathbb{N}$,

$$\frac{\partial s^{(l)}}{\partial q^{(k)}} = \begin{cases} 0, & k<l, \\ -x_k, & k=l, \\ x_{k-1}-x_k, & k>l. \end{cases} \tag{5.29}$$

1. Relation (5.29) implies, for $k \in [2;n-1] \cap \mathbb{N}$,

$$\begin{aligned}
\frac{\partial}{\partial q^{(k)}}\mathscr{CS}(\rho) =& h^2(x_{k-1}-x_k) - \frac{q^{(1)}}{[s^{(1)}]^2}(x_{k-1}-x_k) + \frac{x_{k-1}-x_k}{x_1 s^{(1)}} \\
&+ \sum_{l=2}^{k-1}\frac{x_{k-1}-x_k}{s^{(l)}}\left(\frac{1}{x_l}-\frac{1}{x_{l-1}}\right) - \frac{x_k}{s^{(k)}}\left(\frac{1}{x_k}-\frac{1}{x_{k-1}}\right) \\
&+ 2\beta^2 q^{(k)}\left(x_{k-1}-x_k\right) = 0.
\end{aligned}$$

   Hence,

$$\begin{aligned}
2\beta^2 q^{(k)} &= -h^2 + \frac{q^{(1)}}{[s^{(1)}]^2} - \frac{1}{x_1 s^{(1)}} + \frac{1}{x_{k-1}s^{(k)}} - \sum_{l=2}^{k-1}\frac{1}{s^{(l)}}\left(\frac{1}{x_l}-\frac{1}{x_{l-1}}\right) \\
&= -h^2 + \frac{q^{(1)}}{[s^{(1)}]^2} - \sum_{l=1}^{k-1}\frac{1}{x_l}\left(\frac{1}{s^{(l)}}-\frac{1}{s^{(l+1)}}\right).
\end{aligned} \tag{5.30}$$

2. To handle the case $k=n$, we note that

$$\log\left(1+c(u-q^{(n)})\right) = \frac{1}{x_n}\log\left(\frac{s^{(n)}}{s^{(n+1)}}\right),$$

   and, hence, the argument in the previous item shows that (5.30) is also valid for $k=n$.

3. Differentiating the representation (5.28) with respect to $q^{(1)}$ and using (5.29), we obtain

$$\frac{\partial}{\partial q^{(1)}}\mathscr{CS}(\rho) = -x_1 h^2 + \frac{1}{s^{(1)}} + \frac{x_1 q^{(1)}}{[s^{(1)}]^2} - \frac{x_1}{x_1 s^{(1)}} - 2\beta^2 x_1 q^{(1)} = 0.$$

Therefore,

$$2\beta^2 q^{(1)} = -h^2 + \frac{q^{(1)}}{[s^{(1)}]^2}$$

which is (5.30), for $k = 1$.

4. Subtracting equations (5.30), we arrive to (5.18), for all $k \in [1;n] \cap \mathbb{N}$.

□

Proposition 5.2.2. *The functionals* (5.14) *and* (5.10) *are equivalent in the following sense*

$$\inf_{\rho',\lambda'} \mathscr{P}(\rho',\lambda') = \inf_{\rho'} \mathscr{CS}(\rho').$$

Proof. 1. Let (ρ,λ) be the solutions of equations (5.18) and (5.17). Lemma 5.2.6 guarantees that ρ is the optimiser of the Crisnati-Sommers functional and Lemma 5.2.5 assures that (ρ,λ) is the optimiser of the Parisi functional.

2. We have

$$\begin{aligned}\mathscr{P}(\rho,\lambda) - \mathscr{CS}(\rho) = &-\lambda u + 2\beta^2 q^{(1)} s^{(1)} - \frac{q^{(1)}}{s^{(1)}} + cu - 1 \\ &- 2\beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right).\end{aligned} \tag{5.31}$$

We can simplify the $\Phi[B]$-like term (that is the summation) in (5.31), using (5.18) and (5.17). Indeed,

$$\begin{aligned}2\beta^2 \sum_{l=1}^{n-1} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right) &= 2\beta^2 \sum_{l=1}^{n-1} x_l \left(q^{(l+1)}[q^{(l+1)} - q^{(l)}] + q^{(l)}[q^{(l+1)} - q^{(l)}] \right) \\ \text{(by (5.18) and (5.13))} \ &= \sum_{l=1}^{n-1} \left(2\beta^2 q^{(l+1)} \left[s^{(l)} - s^{(l+1)} \right] + q^{(l)} \left[\frac{1}{s^{(l+1)}} - \frac{1}{s^{(l)}} \right] \right).\end{aligned} \tag{5.32}$$

Regrouping the summands in (5.32), we get

$$\begin{aligned}(5.32) \ =& 2\beta^2 \sum_{l=1}^{n-1} s^{(l)} \left(q^{(l+1)} - q^{(l)} \right) + 2\beta^2 \left(q^{(1)} s^{(1)} - q^{(n)} s^{(n)} \right) \\ &+ \sum_{l=1}^{n-1} \frac{q^{(l)} - q^{(l+1)}}{s^{(l+1)}} + \left(\frac{q^{(n)}}{s^{(n)}} - \frac{q^{(1)}}{s^{(1)}} \right).\end{aligned} \tag{5.33}$$

Due to (5.18), we have

$$2\beta^2\left(q^{(l+1)}-q^{(l)}\right)=\frac{s^{(l)}-s^{(l+1)}}{x_l s^{(l)}s^{(l+1)}}=\frac{q^{(l+1)}-q^{(l)}}{s^{(l)}s^{(l+1)}}.$$

Applying the previous relation, we get that the both summations in (5.33) cancel out and we end up with

$$(5.33)\ =2\beta^2\left(q^{(1)}s^{(1)}-q^{(n)}s^{(n)}\right)+\frac{q^{(n)}}{s^{(n)}}-\frac{q^{(1)}}{s^{(1)}}.$$

Now, turning back to (5.31), we get

$$\begin{aligned}\mathscr{P}(\rho,\lambda)-\mathscr{C}\mathscr{S}(\rho)&=-\lambda u-2\beta^2\left(u^2-[q^{(n)}]^2\right)+2\beta^2 q^{(n)}s^{(n)}-\frac{q^{(n)}}{s^{(n)}}+cu-1\\ \text{(by (5.17)) and (5.13)}&=-u\left(c-2\beta^2(u-q^{(n)})-(u-q^{(n)})^{-1}\right)-2\beta^2\left(u^2-[q^{(n)}]^2\right)\\ &\quad-\frac{q^{(n)}}{u-q^{(n)}}+2\beta^2 q^{(n)}\left(u-q^{(n)}\right)+cu-1\\ &=0.\end{aligned}$$

□

5.2.4 Replica symmetric calculations

In this subsection, we shall consider the one dimensional case of the a priori measure (5.1) with $h=0$. We shall also restrict the computations to the case $n=1$ which is often referred to in physical literature as the replica symmetric scenario. It is indeed the right scenario under the above assumptions, as shows Theorem 5.1.1.

Lemma 5.2.7. *Let μ satisfy* (5.1) *with $h=0$. Assume $d=1$, $n=1$ and $c>0$. Given $u\geq 0$, we have*

$$\inf_{\rho\in\mathscr{Q}(u,1)}\mathscr{C}\mathscr{S}(\rho)=\inf_{q\in[0;u]}\left(1-cu+\log\left(c(u-q)\right)+\frac{q}{u-q}+\beta^2\left(u^2-q^2\right)\right)=f(c,u),\quad (5.34)$$

where $f(c,u)$ is defined in (5.2).

Proof. Using the definitions, we obtain

$$\frac{\partial}{\partial q}\mathscr{C}\mathscr{S}(\rho)=\frac{\partial}{\partial q}\left[\log\left(u-q\right)+\frac{q}{u-q}+\beta^2\left(u^2-q^2\right)\right]=\frac{q}{(u-q)^2}-2\beta^2 q.$$

Hence, the critical points of $q\mapsto\mathscr{C}\mathscr{S}(q,u)$ are

$$q_0=0, q_{1,2}=u\pm\frac{\sqrt{2}}{2\beta}.$$

Furthermore, we also have

$$\frac{\partial^2}{\partial q^2}\mathscr{C}\mathscr{S}(q,u)=\frac{1}{(u-q)^2}+\frac{2q}{(u-q)^3}-2\beta^2.$$

Hence, as a simple calculation shows, the infima in (5.34) are attained on

$$q^* = \begin{cases} 0, & u \leq \frac{\sqrt{2}}{2\beta}, \\ u - \frac{\sqrt{2}}{2\beta}, & u > \frac{\sqrt{2}}{2\beta} \end{cases} \tag{5.35}$$

which implies (5.34). □

Lemma 5.2.8. *Under the assumptions of Lemma 5.2.7, we have*

1. *For $c \geq 2\sqrt{2}\beta$, we have*
$$\sup_{u\geq 0} \inf_{q\in[0;u]} \mathscr{CS}(q,u) = \mathscr{CS}(0,u^*) = \beta^2(u^*)^2 + \log cu^* - cu^* + 1,$$
where
$$u^* := \frac{1}{4\beta^2}\left(c - \sqrt{c^2 - 8\beta^2}\right).$$
2. *For $c < 2\sqrt{2}\beta$, we have*
$$\sup_{u\geq 0} \inf_{q\in[0;u]} \mathscr{CS}(q,u) = +\infty.$$

Remark 5.2.3. *Under the assumptions, the above theorem says that from the point of view of the global free energy, the system can only exist in the "high temperature" scenario, cf.* (5.2). *The threshold at $c_0 = 2\sqrt{2}\beta$ could be easily understood from the perspective of the norms of random matrices.*

Proof. 1. Suppose $c \geq 2\sqrt{2}\beta$. Recalling (5.2), for $u \in (0; \frac{\sqrt{2}}{2\beta}]$, we introduce the following function
$$f(u) := \log(cu) + \beta^2 u^2 - cu + 1.$$
We have
$$\frac{\partial}{\partial u} f(u) = \frac{1}{u} + 2\beta^2 u - c.$$
Hence, the critical points of the function f are
$$u_{1,2} = \frac{c \pm \sqrt{c^2 - 8\beta^2}}{4\beta^2}.$$
Furthermore, we have
$$\frac{\partial^2}{\partial u^2} f(u) = 2\beta^2 - \frac{1}{u^2}.$$
We notice that $u^* \leq \frac{\sqrt{2}}{2\beta}$ and, hence, due to (5.2)
$$\mathscr{CS}(0,u^*) = \beta^2(u^*)^2 + \log cu^* - cu^* + 1.$$

2. If $c < 2\sqrt{2}\beta$, then the function
$$u \mapsto (2\sqrt{2}\beta - c)u + \log\frac{c}{\beta} - \frac{1}{2}(1 + \log 2)$$
is unbounded on $(\frac{\sqrt{2}}{2\beta}; +\infty)$.

□

5.2.5 The multidimensional Crisanti-Sommers functional

Recall the definition (4.73).

Proposition 5.2.3. *Assume $d = 1$. Given $u > 0$, we have*

$$2\phi^{(x*,\mathscr{Q}*,\Lambda*)}(t) = \begin{cases} \left(3\sqrt{2}\beta - c\right)u + \log\frac{c}{\beta} - 1 - \frac{\log 2}{2} - t\left(\sqrt{2}u\beta - \frac{1}{2}\right), & u > \frac{\sqrt{2}}{2\beta}, \\ 2\beta^2(u)^2 + \log(cu) - cu + 1 - t\beta^2(u)^2, & u \leq \frac{\sqrt{2}}{2\beta}. \end{cases} \tag{5.36}$$

Proof. Combining (5.10), (5.14) with Lemma 5.2.7 and Proposition 5.2.2, we get the claim. □

5.2.6 Talagrand's a priori estimates

In this subsection, we prove that Assumption 4.3.1 is satisfied in the case of the Gaussian a priori distribution (5.1) with $h = 0$.

Theorem 5.2.1. *Let μ satisfy* (5.1) *with $h = 0$, assume $U \in \mathrm{Sym}^+(d)$ is such that $\min_v u_v > \frac{\sqrt{2}}{2\beta}$ and suppose $C \succ 0$. Let $Q = Q^*$ and $\Lambda = \Lambda^*$.*

Then, for any $t_0 \in (0;1)$ and any $t \in (0;t_0]$, we have (cf. (4.74) *with $k = 1$)*

$$\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq 2\phi^{(x,\mathscr{Q},\Lambda)}(t) - \frac{1}{K}\|Q^{(1)} - V\|_F^2 + \mathscr{O}(\varepsilon+\delta). \tag{5.37}$$

Proof. 1. We employ the notations of Section 4.3.2. Let $\mathfrak{n} = 1$. Given $\mathfrak{U} \in \mathrm{Sym}(2d)$ (cf. (4.51)), choose arbitrary matrices $\left\{\mathfrak{Q}^{(l)} \in \mathrm{Sym}(2d) \mid l \in [0;2] \cap \mathbb{N}\right\}$ satisfying (4.52). Define $\mathfrak{x} := x$ which, in particular, implies that $\zeta = \xi$. Finally, we set, for $l \in [0;n+1] \cap \mathbb{N}$, $\widetilde{Q}^{(l)} := Q^{(l)}$.

2. Proposition 4.3.1 implies that, for any δ-minimal $\mathfrak{L} \in \mathbb{R}^{2d\times 2d}$, we have

$$\begin{aligned}\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq -\langle\mathfrak{L},\mathfrak{U}\rangle - \frac{t\beta^2}{2}\left(\|\mathfrak{Q}^{(2)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(1)}\|_{\mathrm{F}}^2\right) \\ + X_0^{(2)}(1,\mathfrak{x},\widehat{\mathfrak{Q}}^{(l)}(t),\mathfrak{L}) + \mathscr{O}(\varepsilon+\delta).\end{aligned} \tag{5.38}$$

3. We define a matrix $\mathfrak{C} \in \mathbb{R}^{2d\times 2d}$ as follows

$$\mathfrak{C} := \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}.$$

Recalling (5.3), we define also the following matrices $\mathfrak{D}^{(2)} := \mathfrak{C}$ and

$$\mathfrak{D}^{(1)} := \mathfrak{C} - \mathfrak{L} - \left(\widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t)\right). \tag{5.39}$$

Applying the Proposition 5.2.1 to (5.38), we get

$$\begin{aligned}\varphi_N^{(2)}(1,t,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) &\leq \frac{1}{2}\left[-\langle\mathfrak{L},\mathfrak{U}\rangle - t\beta^2\left(\|\mathfrak{Q}^{(2)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(1)}\|_{\mathrm{F}}^2\right)\right. \\ &\quad \left. + 2\beta^2\langle[\mathfrak{D}^{(1)}]^{-1},\widehat{\mathfrak{Q}}^{(1)}(t)\rangle + \log\left(\frac{\det\mathfrak{D}^{(2)}}{\det\mathfrak{D}^{(1)}}\right)\right] + \mathscr{O}(\varepsilon) \\ &=: \widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}} + \mathscr{O}(\varepsilon).\end{aligned} \tag{5.40}$$

4. Assume that the matrices

$$\mathfrak{Q}^{(1)}, \mathfrak{Q}^{(2)}, \mathfrak{D}^{(1)} \in \mathbb{R}^{2d\times 2d} \tag{5.41}$$

are simultaneously diagonalisable in the same basis which is given by the orthogonal matrix $\mathfrak{O} \in \mathbb{R}^{2d\times 2d}$. Let the vectors

$$\mathfrak{q}^{(1)}, \mathfrak{q}^{(2)}, \mathfrak{d}^{(1)} \in \mathbb{R}^{2d} \tag{5.42}$$

be the corresponding spectra of the matrices (5.41). That is, we assume that

$$\begin{aligned}\mathfrak{Q}^{(1)} = \mathfrak{O}^* \operatorname{diag} \mathfrak{q}^{(1)}\mathfrak{O}, \mathfrak{Q}^{(2)} = \mathfrak{O}^* \operatorname{diag} \mathfrak{q}^{(2)}\mathfrak{O},\\ \mathfrak{D}^{(1)} = \mathfrak{O}^* \mathfrak{d}^{(1)}\mathfrak{O}, \widetilde{\mathfrak{Q}}^{(1)} = \mathfrak{O}^* \operatorname{diag} \widetilde{\mathfrak{Q}}'^{(1)}\mathfrak{O},\end{aligned}$$

where we have introduced the matrix $\widetilde{\mathfrak{Q}}'^{(1)}(t) \in \mathrm{Sym}^+(2d)$. By (5.35), we have, $Q^{(2)} - Q^{(1)} = \frac{\sqrt{2}}{2\beta} I$, where I denotes the unit matrix of the suitable dimension. The definitions (4.62) and (4.63) then imply

$$\widetilde{\mathfrak{Q}}^{(2)} - \widetilde{\mathfrak{Q}}^{(1)} = \frac{\sqrt{2}}{2\beta} I. \tag{5.43}$$

Using the definitions and the above relation, we obtain

$$\begin{aligned}\widehat{\mathfrak{Q}}^{(1)}_{v}(t) &= \mathfrak{O}^* \Big(t \operatorname{diag} \mathfrak{q}^{(1)} + (1-t)\widetilde{\mathfrak{Q}}'^{(1)} \Big) \mathfrak{O},\\ \widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t) &= \mathfrak{O}^* \Big(t \operatorname{diag}(\mathfrak{q}^{(2)} - \mathfrak{q}^{(1)}) + (1-t)\frac{\sqrt{2}}{2\beta} I \Big) \mathfrak{O}.\end{aligned} \tag{5.44}$$

Motivated by (5.19), we set

$$\mathfrak{d}^{(1)}_{v} := \Big(\mathfrak{u}_v - \mathfrak{q}^{(1)}_{v} \Big)^{-1}. \tag{5.45}$$

In view of (5.39), the above choice necessarily yields (cf. (5.17))

$$\begin{aligned}\mathfrak{L} &= \mathfrak{C} - \mathfrak{O}^* \operatorname{diag}(\mathfrak{u}_v - \mathfrak{q}^{(1)}_{v})^{-1}\mathfrak{O} - \Big(\widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t) \Big)\\ &= \mathfrak{C} - \mathfrak{O}^* \Big(\operatorname{diag}(\mathfrak{u}_v - \mathfrak{q}^{(1)}_{v})^{-1} + t \operatorname{diag}(\mathfrak{q}^{(2)} - \mathfrak{q}^{(1)}) + (1-t)\frac{\sqrt{2}}{2\beta} I \Big) \mathfrak{O}.\end{aligned} \tag{5.46}$$

Applying Lemma 5.2.4 to (5.40) and using (5.46), (5.45), (5.44), we get the following diagonalised representation of (5.38)

$$\begin{aligned}\varphi^{(2)}_N(1,t,x,Q,\Sigma^{(2)}_N(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq& \frac{1}{2}\log\det\mathfrak{C} - \frac{1}{2}\langle \mathfrak{C},\mathfrak{U}\rangle\\ &+ \frac{1}{2}\sum_{v=1}^{2d}\Big\{ \mathfrak{u}_v \Big[(\mathfrak{u}_v - \mathfrak{q}^{(1)}_{v})^{-1} + 2\beta^2 \Big(t(\mathfrak{q}^{(2)}_{v} - \mathfrak{q}^{(1)}_{v}) + (1-t)\frac{\sqrt{2}}{2\beta} \Big) \Big]\\ &+ 2\beta^2 (\mathfrak{u}_v - \mathfrak{q}^{(1)}_{v}) \Big(t\mathfrak{q}^{(1)}_{v} + (1-t)\widetilde{\mathfrak{q}}^{(1)}_{v} \Big) + \log(\mathfrak{u}_v - \widetilde{\mathfrak{Q}}'^{(1)}_{v,v})\end{aligned}$$

$$- t\beta^2 \Big((\mathfrak{q}_v^{(2)})^2 - (\mathfrak{q}_v^{(1)})^2 \Big) \Big\} + \mathscr{O}(\varepsilon). \tag{5.47}$$

Using the definitions, we get

$$\langle \mathfrak{C}, \mathfrak{U} \rangle = 2\langle C, U \rangle = 2 \sum_{v=1}^{d} c_v u_v,$$
$$\log\det \mathfrak{C} = 2 \log\det C = 2 \sum_{v=1}^{d} \log c_v. \tag{5.48}$$

Motivated by (5.43) (or by (5.35)), we define

$$\mathfrak{q}_v^{(1)} := \mathfrak{u}_v - \frac{\sqrt{2}}{2\beta}. \tag{5.49}$$

In this case, as a straightforward calculation shows, the expression in the curly brackets in (5.47) equals

$$2\sqrt{2}\beta \mathfrak{u}_v + \beta\sqrt{2}\widetilde{\mathfrak{Q}}'^{(1)}_{v,v}(1-t) - \log\beta - \frac{1}{2}(\log 2 - t). \tag{5.50}$$

By the definitions and the general properties of matrix trace, we have

$$\sum_{v=1}^{2d} \widetilde{\mathfrak{Q}}'^{(1)}_{v,v} = \sum_{v=1}^{2d} \widetilde{\mathfrak{Q}}^{(1)}_{v,v} = 2 \sum_{v=1}^{d} Q^{(1)}_{v,v},$$
$$\sum_{v=1}^{2d} \mathfrak{u}_v = 2 \sum_{v=1}^{d} U_{v,v}. \tag{5.51}$$

Combining (5.47) with (5.50), (5.51) and (5.48), we obtain

$$\begin{aligned} \varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) &\le \sum_{v=1}^{d} \Big(-c_v u_v + \log c_v + 3\sqrt{2}u\beta \\ &\qquad - \frac{1}{2}(\log 2 - t) - \sqrt{2}\beta t u - \log\beta - 1 \Big) + \mathscr{O}(\varepsilon) \\ &= 2 \sum_{v=1}^{d} \phi(t)|_{\substack{c=c_v,\\ u=u_v}} + \mathscr{O}(\varepsilon), \end{aligned} \tag{5.52}$$

where in the last line we have used the relation (5.36).

5. To get the version of the a priori bound (5.52) with the quadratic correction term as stated in (5.37), we perturb the r.h.s of (5.38) around our choice of $\mathfrak{D}^{(1)}$ in (5.45), i.e.,

$$\mathfrak{D}^{(1)} = \Big(\mathfrak{U}_v - \mathfrak{Q}_v^{(1)} \Big)^{-1} = \sqrt{2}\beta I,$$

where in the last equality we used (5.49).

□

5.2.7 The local low temperature Parisi formula

Proof of Theorem 5.1.1. The result follows from Theorem 5.2.1 and Theorem 4.3.1. Note that the proof of Theorem 4.3.1 requires a minor modification to cope with the fact that the a priori distribution (5.1) is unbounded. This minor problem can be fixed by considering the pruned Gaussian distribution and using the elementary estimates to bound the tiny Gaussian tails. □

6

The GREM in the presence of uniform external field

In this chapter, we find the fluctuations of the ground state and of the partition function for the GREM with external field. We provide an explicit expression for the free energy of the model. We also obtain some large deviation results providing an expression for the free energy for a class of models with Gaussian Hamiltonians and external field.

6.1 Introduction

Despite the recent substantial progress due to Guerra (2003), Aizenman *et al.* (2003, 2007), and Talagrand (2006b) in establishing rigorously the Parisi formula for the free energy of the celebrated SK model, understanding of the corresponding limiting Gibbs measure is still very limited.

Due to the above mentioned works, it is now rigorously known that Derrida's GREM is closely related to the SK model at the level of free energy, see, e.g., (Bovier, 2006, Section 11.3). Recently Bovier & Kurkova (2003a, 2004a,b, 2007) have performed a detailed study of the geometry of the Gibbs measure for the GREM. This confirmed the predicted in the theoretical physics literature hierarchical decomposition of the Gibbs measure in rigorous terms.

As pointed out in Bovier & Kurkova (2004a) (see also Ben Arous *et al.* (2005)), the GREM-like models may represent an independent interest in various applied contexts, where correlated heavy-tailed inputs play an important role, e.g., in risk modelling.

One of the key steps in the results of Bovier & Kurkova (2004a) is the identification of the fluctuations of the GREM partition function in the thermodynamic limit with Ruelle's probability cascades. In this chapter we also perform this step and study the effect of external field on the fluctuations (i.e., the weak limit laws) of the partition function of the GREM in the thermodynamic limit. We find that the main difference introduced by the presence of external field, comparing to the system without external field, is that the coarse graining mechanism should be altered. This change reflects the fact that the coarse-grained parts of the system tend to have a certain optimal magnetisation as prescribed by the strength of external field and by parameters of the GREM. We use the general line of reasoning suggested by Bovier & Kurkova (2004a), i.e., we consider the point processes generated by the scaling limits of the GREM Hamiltonian. We streamline the proof of the weak convergence of these point processes to the corresponding Poisson point process by using the Laplace transform.

Organisation of the chapter

In the following subsections of the introduction we define the model of interest and formulate our main results on the fluctuations of the partition function of the REM and GREM with external field and also on their limiting free energy (Theorems 6.1.1, 2.3.1, 6.1.3 and 6.1.4). Their proofs are provided in the subsequent sections. Section 6.2 is devoted to the large deviation results providing an expression for the free energy for a class of models with Gaussian Hamiltonians and external field (Theorem 6.2.1). In Section 6.3 we resort to more refined analysis and perform the calculations of the fluctuations of the ground state and of the partition function in the REM with external field in the thermodynamic limit. Section 6.4 contains the proofs of the results on the fluctuations of the ground state and of the partition function for the GREM with external field.

Definition of the model

In contrast to the work of Derrida & Gardner (1986), we consider here the GREM with external field which depends linearly on the total magnetisation (i.e., we consider the uniform magnetic field, cf. the definitions (0.6) and (0.7)). Derrida & Gardner (1986) considered the "lexicographic" external field which is particularly well adapted to the natural lexicographic distance generated by the GREM Hamiltonian.

The important quantities are the *free energy* defined as

$$p_N(\beta,h) := \frac{1}{N}\log Z_N(\beta,h), \tag{6.1}$$

and the *ground state energy*

$$M_N(h) := N^{-1/2} \max_{\sigma\in\Sigma_N} X_N(h,\sigma). \tag{6.2}$$

In what follows, we shall think of β and h as fixed parameters. We shall occasionally lighten our notation by not indicating the dependence on these parameters explicitly.

We denote the *total magnetisation* (cf. the second summand in (0.6)) by

$$m_N(\sigma) := \frac{1}{N}\sum_{i=1}^{N}\sigma_i, \quad \sigma\in\Sigma_N. \tag{6.3}$$

In this chapter, we shall mainly be interested in the weak limit theorems (i.e., fluctuations) of the partition function (0.7) and of the ground state as $N\uparrow+\infty$. To be precise, the general results on Gaussian concentration of measure imply that (6.2) and (0.10) are self-averaging. By the fluctuations of the ground state, we mean the weak limiting behaviour of the rescaled point process generated by the Gaussian process (0.6). This behaviour is studied in Theorems 6.1.1 and 6.1.2 below. These theorems readily imply the formulae for the limiting free energy (0.10) and the ground state (6.2). A recent account of mathematical results on the GREM without external field and, in particular, on the behaviour of the limiting Gibbs measure can be found in Bovier & Kurkova (2007). The GREM with external field was previously considered by Jana & Rao (2006) (see also Jana (2007)), where its free energy was expressed in terms of a variational problem induced by an application of Varadhan's lemma. In this chapter, we apply very different methods to obtain precise control of the fluctuations of the partition function for the GREM with external field. As a simple consequence of these results, we also get a rather explicit[1] formula for the limiting free energy in the GREM with external field (see Theorem 6.1.4).

Main results

Limiting objects

We now collect the objects which appear in weak limit theorems for the GREM partition function and for the ground states. We denote by $I : [-1;1] \to \mathbb{R}_+$ Cramér's entropy function, i.e.,

$$I(t) := \frac{1}{2}[(1-t)\log(1-t) + (1+t)\log(1+t)]. \tag{6.4}$$

Define

$$\mu(t) := \sqrt{2(\log 2 - I(t))},$$

$$M(h) := \max_{t\in[-1;1]} (\mu(t) + ht). \tag{6.5}$$

Suppose that the maximum in (6.5) is attained at $t = t_* = t_*(h)$. (The maximum exists and is unique, since $\mu(t) + ht$ is strictly concave.) Consider the following two real sequences

$$A_N(h) := \left(\mu(t_*)\sqrt{N}\right)^{-1}, \tag{6.6}$$

$$B_N(h) := M(h)\sqrt{N} + \frac{A_N(h)}{2}\log\left(\frac{A_N(h)^2(I''(t_*)+h)}{2\pi(1-t_*^2)}\right). \tag{6.7}$$

Define the *REM scaling function* $u_{N,h}(x) : \mathbb{R} \to \mathbb{R}$ as

$$u_{N,h}(x) := A_N(h)x + B_N(h). \tag{6.8}$$

Given $f : D \subset \mathbb{R} \to \mathbb{R}_+$, we denote by $\mathrm{PPP}(f(x)\mathrm{d}x, x \in D)$ the Poisson point process with intensity f. We start from a basic limiting object. Assume that the point process $\mathscr{P}^{(1)}$ on $\mathbb{R}$ satisfies

$$\mathscr{P}^{(1)} \sim \mathrm{PPP}\left(\exp(-x)\mathrm{d}x, x \in \mathbb{R}\right), \tag{6.9}$$

and is independent of all random variables around. The point process (6.9) is the limiting object which appears in the REM.

Theorem 6.1.1 (fluctuations of the ground state of the REM with external field)**.** *If* $n = 1$ *(the REM case), then, using the above notations, we have*

$$\sum_{\sigma\in\Sigma_N} \delta_{u_{N,h}^{-1}(X_N(h,\sigma))} \xrightarrow[N\to\infty]{w} \mathscr{P}^{(1)}, \tag{6.10}$$

where the convergence is the weak one of the random probability measures equipped with the vague topology.

[1] In contrast to (Jana & Rao, 2006, Theorem 5.1) and (Jana, 2007, Corollary 4.3.5), who stop at the level of variational problem.

To formulate the weak limit theorems for the GREM (i.e., for the case $n > 1$), we need a limiting object which is a point process closely related to the *Ruelle probability cascade*, (Ruelle, 1987). Define, for $j,k \in [1;n+1] \cap \mathbb{N}$, $j < k$, the "slopes" corresponding to the function ρ in (0.5) as

$$\theta_{j,k} := \frac{q_k - q_{j-1}}{x_k - x_{j-1}}.$$

Define also the following h-dependent "modified slopes"

$$\widetilde{\theta}_{j,k}(h) := \theta_{j,k}\mu(t_*(\theta_{j,k}^{-1/2}h))^{-2}.$$

Define the increasing sequence of indices $\{J_l(h)\}_{l=0}^{m(h)} \subset [0;n+1] \cap \mathbb{N}$ by the following algorithm. Start from $J_0(h) := 0$, and define iteratively

$$J_l(h) := \min\left\{J \in [J_{l-1};n+1] \cap \mathbb{N} : \widetilde{\theta}_{J_{l-1},J}(h) > \widetilde{\theta}_{J+1,k}(h), \text{ for all } k > J\right\}. \tag{6.11}$$

Note that $m(h) \leq n$. The subsequence of indices (6.11) induces the following coarse-graining of the initial GREM

$$\bar{q}_l(h) := q_{J_l(h)} - q_{J_{l-1}(h)}, \tag{6.12}$$

$$\bar{x}_l(h) := x_{J_l(h)} - x_{J_{l-1}(h)}, \tag{6.13}$$

$$\bar{\theta}_l(h) := \theta_{J_{l-1},J_l}. \tag{6.14}$$

The parameters (6.12) induce the new order parameter $\rho^{(J(h))} \in \mathscr{Q}'_m$ in the usual way

$$\rho^{(J(h))}(q) := \sum_{l=1}^{m(h)} q_{J_l(h)} \mathbb{1}_{[x_{J_l(h)};x_{J_{l+1}(h)})}(x).$$

Define the GREM scaling function $u_{N,\rho,h} : \mathbb{R} \to \mathbb{R}$ as

$$u_{N,\rho,h}(x) := \sum_{l=1}^{m(h)} \left[\bar{q}_l(h)^{1/2} B_{\bar{x}_l(h)N}\left(\bar{\theta}_l(h)^{-1/2}h\right)\right] + N^{-1/2}x.$$

Define the rescaled GREM process as

$$\overline{\mathrm{GREM}}_N(h,\sigma) := u_{N,\rho,h}^{-1}(\mathrm{GREM}_N(h,\sigma)).$$

Define the point process of the rescaled GREM energies $\mathscr{E}_N$ as

$$\mathscr{E}_N(h) := \sum_{\sigma \in \Sigma_N} \delta_{\overline{\mathrm{GREM}}_N(h,\sigma)}. \tag{6.15}$$

Consider the following collection of independent point processes (which are also independent of all random objects introduced above)

$$\{\mathscr{P}^{(k)}_{\alpha_1,\dots,\alpha_{l-1}} \mid \alpha_1,\dots,\alpha_{l-1} \in \mathbb{N}; l \in [1;m]\cap\mathbb{N}\}$$

such that

$$\mathscr{P}^{(k)}_{\alpha_1,\dots,\alpha_{k-1}} \sim \mathscr{P}^{(1)}.$$

Define the *limiting GREM cascade point process* $\mathscr{P}_m$ on $\mathbb{R}^m$ as follows

$$\mathscr{P}_m := \sum_{\alpha\in\mathbb{N}^m} \delta_{(\mathscr{P}^{(1)}(\alpha_1),\mathscr{P}^{(2)}_{\alpha_1}(\alpha_2),\dots,\mathscr{P}^{(m)}_{\alpha_1,\alpha_2,\dots,\alpha_{m-1}}(\alpha_m))}, \tag{6.16}$$

Consider the following constants

$$\bar{\gamma}_l(h) := \left(\widetilde{\theta}_{J_{l-1},J_l}\right)^{1/2},$$

and define the function $E_{h,f} : \mathbb{R}^m \to \mathbb{R}$ as

$$E^{(m)}_{h,\rho}(e_1,\dots,e_m) := \bar{\gamma}_1(h)e_1 + \ldots + \bar{\gamma}_m(h)e_m.$$

Note that due to (6.11), the constants $\{\bar{\gamma}_l(h)\}_{l=1}^m$ form a decreasing sequence, i.e., for all $l \in [1;m]\cap\mathbb{N}$, we have

$$\bar{\gamma}_l(h) > \bar{\gamma}_{l+1}(h). \tag{6.17}$$

The cascade point process (6.16) is the limiting object which describes the fluctuations of the ground state in the GREM.

Theorem 6.1.2 (fluctuations of the ground state of the GREM with external field)**.** *We have*

$$\mathscr{E}_N(h) \xrightarrow[N\uparrow+\infty]{w} \int_{\mathbb{R}^m} \delta_{E^{(m)}_{h,\rho}(e_1,\dots,e_m)} \mathscr{P}_m(\mathrm{d}e_1,\dots,\mathrm{d}e_m) \tag{6.18}$$

and

$$M_N(h) \xrightarrow[N\uparrow+\infty]{} \sum_{l=1}^{m(h)} \left[\left(\bar{q}_l(h)\bar{x}_l(h)\right)^{1/2} M\left(\bar{\theta}_l(h)^{-1/2}h\right)\right], \tag{6.19}$$

almost surely and in L^1.

Theorem 6.1.2 allows for complete characterisation of the limiting distribution of the GREM partition function. To formulate the result, we need the β-dependent threshold $l(\beta,h) \in [0;m]\cap\mathbb{N}$ such that above it all coarse-grained levels $l > l(\beta,h)$ of the limiting GREM are in the "high temperature regime". Below this threshold the levels $l \le l(\beta,h)$ are in the "frozen state". Given $\beta \in \mathbb{R}_+$, define

$$l(\beta,h) := \max\{l \in [1;n]\cap\mathbb{N} : \beta\bar{\gamma}_l(h) > 1\}.$$

We set $l(\beta,h) := 0$, if $\beta\bar{\gamma}_1(h) \le 1$. The following gives full information about the limiting fluctuations of the partition function at all temperatures.

Theorem 6.1.3 (fluctuations of the partition function of the GREM with external field)**.** *We have*

$$\exp\left[-\beta\sqrt{N}\sum_{l=1}^{l(\beta,h)}\left(\bar{q}_l(h)^{1/2}B_{\bar{x}_l(h)N}\left(\bar{\theta}_l^{-1/2}h\right)\right)\right]$$
$$\times\exp\left[-N\left(\log 2+\log\operatorname{ch}\left(\beta h(1-x_{J_{l(\beta,h)}})\right)+\frac{1}{2}\beta^2\left(1-q_{J_{l(\beta,h)}}\right)\right)\right]\operatorname{ch}^{2/3}\left(\beta h(1-x_{J_{l(\beta,h)}})\right)$$
$$\times Z_N(\beta,h)\xrightarrow[N\uparrow+\infty]{w}K(\beta,h,\rho)\int_{\mathbb{R}^{l(\beta,h)}}\exp\left[\beta E_{h,\rho}^{(l(\beta,h))}(e_1,\ldots,e_{l(\beta,h)})\right]\mathscr{P}_{l(\beta,h)}(\mathrm{d}e_1,\ldots,\mathrm{d}e_{l(\beta,h)}),\tag{6.20}$$

where the constant $K(\beta,h,\rho)$ depends on β, h and ρ only. Moreover, $K(\beta,h,\rho)=1$, if $\beta\gamma_{l(\beta,h)+1}<1$ and $K(\beta,h,\rho)\in(0;1)$, if $\beta\gamma_{l(\beta,h)+1}=1$.

The above theorem suggests that the increasing sequence of the constants $\{\beta_l:=\bar{\gamma}^{-1}\}_{l=1}^{m(h)}\subset\mathbb{R}_+$ can be thought of as the sequence of the inverse temperatures at which the phase transitions occur: at β_l the corresponding coarse-grained level l of the GREM with external field "freezes".

As a simple consequence of the fluctuation results of Theorem 6.1.3, we obtain the following formula for the limiting free energy of the GREM.

Theorem 6.1.4 (free energy of the GREM with external field)**.** *We have*

$$\lim_{N\uparrow+\infty}p_N(\beta,h)=\beta\sum_{l=1}^{l(\beta,h)}\left[(\bar{x}_l\bar{q}_l)^{1/2}\mu(t_*(\bar{\theta}_l^{-1/2}h))+h\bar{x}_l t_*(\bar{\theta}_l^{-1/2}h)\right]$$
$$+\log 2+\log\operatorname{ch}\left(\beta h(1-x_{J_{l(\beta,h)}})\right)+\frac{1}{2}\beta^2\left(1-q_{J_{l(\beta,h)}}\right),\tag{6.21}$$

almost surely and in L^1.

6.2 Partial partition functions, external fields and overlaps

In this section, we propose a way to compute the free energy of disordered spin systems with external field using the restricted free energies of systems without external field. The computation involves a large deviations principle. For gauge invariant systems, we also show that the partition function of the system with external field induced by the total magnetisation has the same distribution as the one induced by the overlap with fixed but arbitrary configuration. This section is based on the ideas of Derrida & Gardner (1986).

Fix $p\in\mathbb{N}$. Given some finite *interaction p-hypergraph* $(V_N,E_N^{(p)})$, where $V_N=[1;n]\cap\mathbb{N}$ and $E_N^{(p)}\subset(V_N)^p$, define the *p-spin interaction Hamiltonian* as

$$X_N(\sigma):=\sum_{i\in E_N^{(p)}}J_i^{(N,p)}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p},\quad\sigma\in\Sigma_N,\tag{6.22}$$

where $J^{(N,p)}:=\left\{J_i^{(N,p)}\right\}_{i\in E_N^{(p)}}$ is the collection of random variables having the symmetric joint distribution. That is, we assume that, for any $\varepsilon^{(1)},\varepsilon^{(2)}\in\{-1;+1\}^{E_N^{(p)}}$, and any $t\in\mathbb{R}^{E_N^{(p)}}$,

$$\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{r\in E_N^{(p)}} t_r\varepsilon_r^{(1)}J_i^{(N,p)}\right)\right] = \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{r\in E_N^{(p)}} t_r\varepsilon_r^{(2)}J_i^{(N,p)}\right)\right], \tag{6.23}$$

where $\mathrm{i} \in \mathbb{C}$ denotes the imaginary unit.

A particular important example of (6.22) is *Derrida's p-spin Hamiltonian* given by

$$\mathrm{SK}_N^{(p)}(\sigma) := N^{-p/2} \sum_{i_1,\ldots,i_p=1}^{N} g_{i_1,\ldots,i_p}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p},$$

where $\{g_{i_1,\ldots,i_p}\}_{i_1,\ldots,i_p=1}^N$ is a collection of i.i.d. standard Gaussian random variables. Note that the condition (6.23) is obviously satisfied.

Given $\mu \in \Sigma_N$, define the corresponding *gauge transformation* $T_\mu : \Sigma_N \to \Sigma_N$ as

$$T_\mu(\sigma)_i = \mu_i\sigma_i, \quad \sigma \in \Sigma_N. \tag{6.24}$$

Note that the gauge transformation (6.24) is obviously an involution. We say that a d-variate random function $f : \Sigma_N^d \to \mathbb{R}$ is *gauge invariant*, if, for any $\mu \in \Sigma_N$ and any $(\sigma^{(1)},\ldots,\sigma^{(d)}) \in \Sigma_N^d$,

$$f(T_\mu(\sigma^{(1)}),\ldots,T_\mu(\sigma^{(d)})) \sim f(\sigma^{(1)},\ldots,\sigma^{(d)}),$$

where $\sim$ denotes equality in distribution. Define the *overlap* between the configurations $\sigma,\sigma' \in \Sigma_N$ as

$$R_N(\sigma,\sigma') := \frac{1}{N}\sum_{i=1}^N \sigma_i\sigma_i'. \tag{6.25}$$

Note that the overlap (6.25) and the lexicographic overlap (0.2) are gauge invariant.

Given a bounded function $F_N : \Sigma_N \to \mathbb{R}$, define the *partial partition function* as

$$Z_N^{(p)}(\beta,q,\varepsilon,X_N,F_N) := \sum_{\sigma:|F_N(\sigma)-q|\le\varepsilon} \exp\left(\beta\sqrt{N}X_N(\sigma)\right). \tag{6.26}$$

Denote

$$U_N := F_N(\Sigma_N), \quad U := \overline{\left(\bigcup_{N=1}^\infty U_N\right)}. \tag{6.27}$$

(The bar in (6.27) denotes closure in the Euclidean topology.) Note that for the case $F_N = R_N$ we obviously have

$$U_N = \left\{1 - \frac{2k}{N} : k \in [0;N]\cap\mathbb{Z}\right\}, \quad U = [-1;1].$$

Proposition 6.2.1 (Derrida & Gardner (1986))**.** *Assume that X_N is given either by* (6.22) *or $X_N \sim \mathrm{GREM}_N$. Fix some gauge invariant bivariate function $F_N : \Sigma_N^2 \to \mathbb{R}$, and $q \in \mathbb{R}$.*

Then, for all $\sigma',\tau' \in \Sigma_N$, we have

$$Z_N^{(p)}(\beta, q, \varepsilon, X_N, F_N(\cdot, \sigma')) \sim Z_N^{(p)}(\beta, q, \varepsilon, X_N, F_N(\cdot, \tau')). \tag{6.28}$$

In particular, the partial partition function (6.26) *with* $F_N := R_N(\cdot, \sigma')$ *has the same distribution as the partial partition function which corresponds to fixing the total magnetisation* (6.3)*, i.e.,*

$$Z_N^{(p)}(\beta, q, \varepsilon, X_N, R_N(\cdot, \sigma')) \sim Z_N^{(p)}(\beta, m, \varepsilon,) := \sum_{\sigma : |m(\sigma) - q| < \varepsilon} \exp\left(\beta \sqrt{N} X_N(\sigma)\right).$$

Remark 6.2.1. *The proposition obviously remains valid for the Hamiltonians* X_N *given by the linear combinations of the p-spin Hamiltonians* (6.22) *with varying* $p \in \mathbb{N}$*.*

Proof. 1. If X_N is defined by (6.22), then (6.28) follows due to the gauge invariance of (6.22) and F_N. Indeed, there exists $\mu \in \Sigma_N$ such that $\sigma' = T_\mu(\tau')$. Define

$$J_i^{(N,p,\mu)} := J_i^{(N,p)} \mu_{i_1} \cdots \mu_{i_p}.$$

Due to the symmetry of the joint distribution of $J^{(N,p)}$, we have

$$\{X_N(\sigma)\}_{\sigma \in \Sigma_N} \sim \{X_N(\sigma)|_{J^{(N,p)} = J^{(N,p,\mu)}}\}_{\sigma \in \Sigma_N}$$

which implies (6.28).

2. If $X_N = \mathrm{GREM}_N$, then, since X_N is a Gaussian process, to prove the equality in distribution, it is enough to check that the covariance of X_N is gauge symmetric. Equivalence (6.28) follows, due to (0.5) and the fact that the lexicographic overlap (0.2) is gauge invariant.

□

The partial partition function (6.26) induces the *restricted free energy* in the usual way:

$$p_N^{(p)}(\beta, q, \varepsilon, X_N, F_N) := \frac{1}{N} \log Z_N^{(p)}(\beta, q, \varepsilon, X_N, F_N). \tag{6.29}$$

Given $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, let

$$C_N(\sigma^{(1)}, \sigma^{(2)}) := \mathbb{E}\left[X_N(\sigma^{(1)}) X_N(\sigma^{(2)})\right], \quad \widetilde{C}_N(\sigma^{(1)}) := C_N(\sigma^{(1)}, \sigma^{(1)}).$$

Define

$$V_N := \{C_N(\sigma, \sigma) : \sigma \in \Sigma_N\}, \quad V := \overline{\left(\bigcup_{N=1}^{\infty} V_N\right)}.$$

The following result establishes a large deviations type relation between the partial free energy and the full one.

Theorem 6.2.1. *Assume* $X_N = \{X_N(\sigma)\}_{\sigma \in \Sigma_N}$ *is a centred Gaussian process and* $F_N : \Sigma_N \to \mathbb{R}$ *are such that, for all* $N, M \in \mathbb{N}$*, all* $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$*, and all* $\tau^{(1)}, \tau^{(2)} \in \Sigma_M$*,*

$$C_{N+M}(\sigma^{(1)} \amalg \tau^{(1)}, \sigma^{(2)} \amalg \tau^{(2)}) \leq \frac{N}{N+M} C_N(\sigma^{(1)}, \sigma^{(2)}) + \frac{M}{N+M} C_M(\tau^{(1)}, \tau^{(2)}), \tag{6.30}$$

$$F_{N+M}(\sigma^{(1)} \amalg \tau^{(1)}) \leq \frac{N}{N+M} F_N(\sigma^{(1)}) + \frac{M}{N+M} F_M(\tau^{(1)}). \tag{6.31}$$

Assume that C_N *and* F_N *are bounded uniformly in N.*

Then

1. *The following holds*

$$p_N(\beta, X_N, F_N) := \frac{1}{N} \log \sum_{\sigma \in \Sigma_N} \exp\left(\beta \sqrt{N} X_N(\sigma) + N F_N(\sigma)\right)$$
$$\xrightarrow[N\uparrow+\infty]{} p(\beta, X, F), \quad \textit{almost surely and in } L^1. \tag{6.32}$$

2. *The limiting free energy $p(\beta, X, F)$ is almost surely deterministic.*
3. *We have*

$$\lim_{\varepsilon\downarrow+0} \lim_{N\uparrow+\infty} p_N^{(p)}(\beta, q, \varepsilon, X_N, F_N) =: p^{(p)}(\beta, q, X, F)$$
$$= \sup_{v\in V} \inf_{\lambda\in\mathbb{R}, \gamma\in\mathbb{R}} \left(-\lambda q - \gamma v + p(\beta, X, \lambda F + \gamma \widetilde{C})\right),$$
$$\textit{almost surely and in } L^1. \tag{6.33}$$

4. *Finally,*

$$p(\beta, X, F) = \sup_{q\in U} \left(p^{(p)}(\beta, q, X, F) + q\right). \tag{6.34}$$

Remark 6.2.2.

1. *If there exists $\{\mathrm{const}_N \in \mathbb{R}_+\}_{N=1}^{\infty}$ such that, for all $\sigma \in \Sigma_N$,*

$$\widetilde{C}_N(\sigma) = \mathrm{const}_N, \tag{6.35}$$

then (6.33) *simplifies to*

$$p^{(p)}(\beta, q, X, F) = \inf_{\lambda\in\mathbb{R}} \left(-\lambda q + p(\beta, X, \lambda F)\right), \quad \textit{almost surely and in } L^1. \tag{6.36}$$

2. *Inequality* (6.31) *can alternatively be substituted by the assumption (see (Guerra & Toninelli, 2003, Theorem 1)) that $F_N(\sigma) = f(S_N(\sigma))$, where $f : \mathbb{R} \to \mathbb{R}$, $f \in C^1(\mathbb{R})$, and $S_N : \Sigma_N \to \mathbb{R}$ is the bounded function such that, for all $\sigma \in \Sigma_N$, $\tau \in \Sigma_M$,*

$$S_{N+M}(\sigma \amalg \tau) = \frac{N}{N+M} S_N(\sigma) + \frac{M}{N+M} S_M(\tau).$$

3. *It is easy to check that the assumptions of Proposition 6.2.1 are fulfilled, e.g., for*

$$X_N := c_1 \mathrm{GREM}_N + c_2 \mathrm{SK}_N^{(p)},$$

and

$$F_N(\cdot) := f_1(R_N(\cdot, \sigma^{(N)})) + f_2(q_L(\cdot, \sigma^{(N)})),$$

where $\sigma^{(N)} \in \Sigma_N$, $c_1, c_2 \in \mathbb{R}$, and $f_1, f_2 : \mathbb{R} \to \mathbb{R}$, such that $f_1 \in C^1(\mathbb{R})$, f_2 is convex. Note that in this case, due to Proposition 6.2.1, the free energies (6.32) *and* (6.33) *does not depend on the choice of the sequence $\{\sigma^{(N)}\}_{N=1}^{\infty} \subset \Sigma_N$.*

Proof. Similarly to (Contucci *et al.*, 2003, Theorem 1) and (Guerra & Toninelli, 2003, Theorem 1) we obtain (6.32). Then (6.32) implies that

$$p(\beta, X_N, \lambda F_N + \gamma \widetilde{C}_N) \xrightarrow[N\uparrow+\infty]{} p(\beta, X, \lambda F + \gamma \widetilde{C}), \quad \text{almost surely and in } L^1.$$

Hence, we can apply the quenched large deviation results (Theorems 3.3.1 and 3.3.2) which readily yield (6.34) and (6.33) (or (6.36), in the case of (6.35)). □

Remark 6.2.3. *Derrida & Gardner (1986) sketched a calculation of the free energy defined in* (6.32) *in the following case*

$$F_N = q_L \text{ and } X_N = \mathrm{GREM}_N. \tag{6.37}$$

This case is easier than the case (0.10) *we are considering here, since both q_L and* GREM_N *have lexicographic nature, cf.* (0.5) *and* (0.2).

6.3 The REM with external field revisited

In this section, we recall some known results on the limiting free energy of the REM with external field. However, we give some new proofs of these results which illustrate the approach of Section 6.2. Moreover, we prove the weak limit theorem for the ground state and for the partition function of the REM with external field.

Recall that the REM corresponds to the case $n = 1$ in (1.30). This implies that the process X is simply a family of 2^N i.i.d. standard Gaussian random variables. To emphasise this situation we shall write $\mathrm{REM}(\sigma)$ instead of $\mathrm{GREM}(\sigma)$.

6.3.1 Free energy and ground state

Let us start by recalling the following well-known result on the REM.

Theorem 6.3.1 (Derrida (1980); Eisele (1983); Olivieri & Picco (1984))**.** *Assume that $n = 1$ and let $p(\beta, h)$ be given by* (6.1). *The following assertions hold*

1. *We have*

$$\lim_{N\to\infty} p_N(\beta, 0) = \begin{cases} \frac{\beta^2}{2} + \log 2, & \beta \le \sqrt{2\log 2} \\ \beta\sqrt{2\log 2}, & \beta \ge \sqrt{2\log 2} \end{cases}, \quad \textit{almost surely and in } L^1. \tag{6.38}$$

2. *For all $\beta \ge \sqrt{2\log 2}$ and $N \in \mathbb{N}$, we have*

$$0 \le \mathbb{E}[p_N(\beta, 0)] \le \beta\sqrt{2\log 2}. \tag{6.39}$$

See, e.g., (Bovier, 2006, Theorem 9.1.2) for a short proof.

Given $k \in [0; N] \cap \mathbb{N}$, define the set of configurations having a given magnetisation

$$\Sigma_{N,k} := \{\sigma \in \Sigma_N : \sum_{i=1}^{N} \sigma_i = N - 2k\}. \tag{6.40}$$

Lemma 6.3.1. *Set* $t_{k,N} := \frac{N-2k}{N}$. *Given any* $\varepsilon > 0$, *uniformly in* $k \in [0;N] \cap \mathbb{N}$ *such that*

$$t_{k,N} \in [-1+\varepsilon; 1-\varepsilon],$$

we have the following asymptotics

$$\binom{N}{k} \underset{N\uparrow+\infty}{=} \sqrt{\frac{2}{\pi}} \frac{2^N \mathrm{e}^{-NI(t_{k,N})}}{\sqrt{N(1-t_{k,N}^2)}} \left(1 + \frac{1}{N}\left(\frac{1}{12} + \frac{1}{3(1-t_{k,N}^2)}\right) + \mathscr{O}\left(\frac{1}{N^2}\right)\right). \tag{6.41}$$

Proof. A standard exercise on Stirling's formula. □

Theorem 6.3.2 (Dorlas & Wedagedera (2001))**.** *Assume that* $n = 1$ *(the REM case) and let* $p(\beta,h)$ *be given by* (6.1). *We have*

$$\begin{aligned} p(\beta,h) &:= \lim_{N\to\infty} p_N(\beta,h) \\ &= \begin{cases} \log 2 + \log \operatorname{ch} \beta h + \frac{\beta^2}{2}, & \beta \le \sqrt{2(\log 2 - I(t_*))} =: \beta_0 \\ \beta(\sqrt{2(\log 2 - I(t_*))} + ht_*), & \beta \ge \sqrt{2(\log 2 - I(t_*))} \end{cases}, \quad \textit{almost surely and in } L^1, \end{aligned} \tag{6.42}$$

and $t_* \in (-1;1)$ *is a unique maximiser of the following concave function*

$$(-1;1) \ni t \mapsto ht + \sqrt{2(\log 2 - I(t))}.$$

Proof. For the sake of completeness, we give a short proof based on (the ideas of) Theorem 6.2.1. Put

$$M_{k,N} := \begin{cases} \lfloor \log_2 \binom{N}{k} \rfloor, & k \in [1;N-1] \cap \mathbb{N} \\ 1, & k \in \{0,N\} \end{cases},$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than x. Consider the free energy (cf. (6.29)) of the REM of volume $M_{k,N}$

$$p_{k,N}(\beta) := \frac{1}{M_{k,N}} \log \sum_{\sigma \in \Sigma_N^k} \exp\Big(\beta M_{k,N}^{1/2} \mathrm{REM}(\sigma)\Big),$$

where $\mathrm{REM} := \{\mathrm{REM}(\sigma)\}_{\sigma\in\Sigma_N}$ is the family of standard i.i.d. Gaussian random variables. Let

$$\widetilde{p}_{k,N}(\beta) := \frac{M_{k,N}}{N} p_{k,N}\Big(\Big(\frac{N}{M_{k,N}}\Big)^{\frac{1}{2}} \beta\Big). \tag{6.43}$$

Note that (6.43) is the restricted free energy (cf. (6.29)) of the REM, where the restriction is imposed by the total magnetisation (6.3) given by $t_{k,N}$.

We claim that the family of functions $\mathscr{P} := \{\mathbb{E}[p_N(\cdot)]\}_{N\in\mathbb{N}}$ is uniformly Lipschitzian. Indeed, uniformly in $\beta \ge 0$, we have

$$\partial_\beta \mathbb{E}\left[p_N(\beta)\right] = N^{-1/2}\mathbb{E}\left[\mathscr{G}_N(\beta,0)\left[X_N(\sigma)\right]\right] \le N^{-1/2}\mathbb{E}\left[\max_{\sigma\in\Sigma_N} X(\sigma)\right] \xrightarrow[N\uparrow+\infty]{} \sqrt{2\log 2}.$$

Hence, the family $\mathscr{P}$ has uniformly bounded first derivatives.

Given $t \in (-1;1)$ and $t_{k_N,N} \in U_N$ (cf. (6.27)) such that $\lim_{N\uparrow+\infty} t_{k_N,N} = t$, using (6.41), we have

$$\lim_{N\uparrow+\infty} \frac{M_{k_N,N}}{N} = 1 - I(t)\log_2 \mathrm{e}. \tag{6.44}$$

Using (6.44) and the uniform Lipschitzianity of the family $\mathscr{P}$, we get

$$\lim_{N\uparrow+\infty} \widetilde{p}_{k_N,N}(\beta) = (1-I(t)\log_2 \mathrm{e})p\left(\frac{\beta}{\sqrt{1-I(t)\log_2 \mathrm{e}}}\right). \tag{6.45}$$

Combining (6.45) with (6.34) and (6.36), we get

$$p(\beta,h) = \max_{t\in[-1;1]}\left\{t\beta h + (1-I(t)\log_2 \mathrm{e})p\left(\frac{\beta}{\sqrt{1-I(t)\log_2 \mathrm{e}}}\right)\right\}. \tag{6.46}$$

To find the maximum in (6.46), we consider two cases.

1. If $\beta \le \sqrt{2(\log 2 - I(t_*))}$, then according to (6.38), we have

$$p\left(\frac{\beta}{\sqrt{1-I(t)\log_2 \mathrm{e}}}\right) = \log 2 + \frac{\beta^2}{2(1-I(t)\log_2 \mathrm{e})}.$$

Hence, (6.46) implies

$$p(\beta,h) = \max_{t\in[-1;1]}\left\{t\beta h + \frac{\beta^2}{2} + \log 2 - I(t)\right\} = \log 2 + \log \operatorname{ch} \beta h + \frac{\beta^2}{2}, \tag{6.47}$$

where the last equality is due to the fact that the expression in the curly brackets is concave and, hence, the maximum is attained at a stationary point. The stationarity condition reads

$$t = t_0(\beta,h) := \tanh \beta h. \tag{6.48}$$

It is easy to check that the following identity holds

$$I(t) = t\tanh^{-1} t - \log \operatorname{ch} \tanh^{-1} t. \tag{6.49}$$

Combining (6.49) and (6.48), we get (6.47).

2. If $\beta \ge \sqrt{2(\log 2 - I(t_*))}$, then again by (6.38), we have

$$p\left(\frac{\beta}{\sqrt{1-I(t)\log_2 \mathrm{e}}}\right) = \frac{\beta\sqrt{2\log 2}}{\sqrt{1-I(t)\log_2 \mathrm{e}}}.$$

Hence, (6.46) transforms to

$$p(\beta,h) = \max_{t\in[-1;1]}\left\{t\beta h + \beta\sqrt{2(\log 2 - I(t))}\right\} = \beta\left(\sqrt{2(\log 2 - I(t_*))} + ht_*\right), \tag{6.50}$$

where the last equality is due to the concavity of the expression in the curly brackets.

Combining (6.47) and (6.50), we get (6.42). □

Remark 6.3.1. *We note that due to the continuity of the free energy as a function of β, we have at the freezing temperature β_0*

$$t_0(\beta_0, h) = t_*(h). \tag{6.51}$$

Theorem 6.3.2 suggests that the following holds.

Theorem 6.3.3. *Under the assumptions of Theorem 6.3.2, we have*

$$\lim_{N\uparrow+\infty} \frac{1}{\sqrt{N}} \max_{\sigma\in\Sigma_N} X_N(h,\sigma) = \sqrt{2(\log 2 - I(t_*))} + ht_*, \quad \textit{almost surely and in } L^1. \tag{6.52}$$

Proof. We have

$$\frac{1}{\beta} p_N(\beta) \le \frac{1}{N} \log\left(N\beta\sqrt{N} \max_{\sigma\in\Sigma_N} X_N(h,\sigma)\right) = \frac{\log N}{\beta N} + \frac{1}{\sqrt{N}} \max_{\sigma\in\Sigma_N} X_N(h,\sigma). \tag{6.53}$$

In view of (6.42), relation (6.53) readily implies that

$$\sqrt{2(\log 2 - I(t_*))} + ht_* \le \varliminf_{N\uparrow+\infty} N^{-1/2} \max_{\sigma\in\Sigma_N} X_N(h,\sigma). \tag{6.54}$$

We also have

$$\frac{1}{\beta} p_N(\beta) \ge \frac{1}{\sqrt{N}} \max_{\sigma\in\Sigma_N} X_N(h,\sigma)$$

which combined again with (6.42) implies that

$$\sqrt{2(\log 2 - I(t_*))} + ht_* \ge \varlimsup_{N\uparrow+\infty} N^{-1/2} \max_{\sigma\in\Sigma_N} X_N(h,\sigma). \tag{6.55}$$

Due to the standard concentration of Gaussian measure (e.g., (Ledoux, 2001, (2.35))) and the fact that the free energy (6.1) is Lipschitzian with the constant $\beta\sqrt{N}$ as a function of $X_N(h,\cdot)$ with respect to the Euclidean topology, the bounds (6.54) and (6.55) combined with the Borell-Cantelli lemma give the convergence (6.52). □

6.3.2 Fluctuations of the ground state

In this subsection, we shall study the limiting distribution of the point process generated by the properly rescaled process of the energy levels, i.e. (6.15).

Proof of Theorem 6.1.1. Let us denote

$$\mathscr{E}_N(h) := \sum_{\sigma\in\Sigma_N} \delta_{u_{N,h}^{-1}(X_N(h,\sigma))}. \tag{6.56}$$

We treat $\mathscr{E}_N(h)$ as a random pure point measure on $\mathbb{R}$. Given some test function $\varphi \in C_0^+(\mathbb{R})$ (i.e., a non-negative function with compact support), consider the Laplace transform of (6.56) corresponding to φ

$$
\begin{aligned}
L_{\mathscr{E}_N(h)}(\varphi) &:= \mathbb{E}\left[\exp\left\{-\sum_{\sigma\in\Sigma_N}\varphi\left(u_{N,h}^{-1}(X_N(h,\sigma))\right)\right\}\right] \\
&= \prod_{k=0}^{N}\left\{\frac{1}{2\pi}\int_{\mathbb{R}}\exp\left(-\varphi\left(u_{N,h}^{-1}(x+\frac{h}{\sqrt{N}}(N-2k))\right)-\frac{x^2}{2}\right)\mathrm{d}x\right\}^{\binom{N}{k}}. \qquad (6.57)
\end{aligned}
$$

Introduce the new integration variables $y = u_{N,h}^{-1}(x+\frac{h}{\sqrt{N}}(N-2k))$. We have

$$
\begin{aligned}
(6.57) &= \prod_{k=0}^{N}\left\{\frac{A_N(h)}{2\pi}\int_{\mathbb{R}}\exp\left(-\varphi(y)-\frac{1}{2}\left(u_{N,h}(y)-\frac{h}{\sqrt{N}}(N-2k)\right)^2\right)\mathrm{d}y\right\}^{\binom{N}{k}} \\
&= \exp\left\{\sum_{k=0}^{N}\binom{N}{k}\log\left(1-\frac{A_N(h)}{\sqrt{2\pi}}\int_{\mathbb{R}}(1-\mathrm{e}^{-\varphi(y)})\exp\left[-\frac{1}{2}\left(u_{N,h}(y)-\frac{h}{\sqrt{N}}(N-2k)\right)^2\right]\right)\right\}. \\
&\qquad (6.58)
\end{aligned}
$$

Note that the integration in (6.58) is actually performed over $y \in \operatorname{supp}\varphi$, since the integrand is zero on the complement of the support. It is easy to check that uniformly in $y \in \operatorname{supp}\varphi$ the integrand in (6.58) and, hence, the integral itself are exponentially small (as $N \uparrow +\infty$). Consequently, we have

$$
\begin{aligned}
(6.58) \underset{N\uparrow+\infty}{=} \exp\Bigg\{ &-\int_{\operatorname{supp}\varphi}(1-\mathrm{e}^{-\varphi(y)}) \\
&\sum_{k=0}^{N}\binom{N}{k}\frac{A_N(h)}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(u_{N,h}(y)-\frac{h}{\sqrt{N}}(N-2k)\right)^2\right]\Bigg)(1+o(1))\Bigg\}. \qquad (6.59)
\end{aligned}
$$

Denote $t_{k,N} := \frac{N-2k}{N}$. Using Lemma 6.3.1, we get

$$
\begin{aligned}
(6.59) \underset{N\uparrow+\infty}{=} \exp\Bigg\{ &-(1+o(1))\int_{\operatorname{supp}\varphi}(1-\mathrm{e}^{-\varphi(y)}) \\
&\times\sum_{k=0}^{N}\frac{A_N(h)}{\pi\left(N(1-t_{k,N}^2)\right)^{1/2}}\exp\left[N\left(\log 2 - I(t_{k,N})\right)-\frac{1}{2}\left(u_{N,h}(y)-ht_{k,N}\sqrt{N}\right)^2\right]\Bigg)\Bigg\}. \\
&\qquad (6.60)
\end{aligned}
$$

Note that despite the fact that Lemma 6.3.1 is valid only for $t_{k,N} \in [-1+\varepsilon; 1-\varepsilon]$, we can still write (6.60), since the both following sums are negligible:

$$
\begin{aligned}
0 \le \sum_{k:t_{k,N}\in([-1;-1+\varepsilon]\cup[1-\varepsilon,1])} &\frac{A_N(h)}{\pi\left(N(1-t_{k,N}^2)\right)^{1/2}}\exp\Big[N\left(\log 2 - I(t_{k,N})\right) \\
&-\frac{1}{2}\left(u_{N,h}(y)-ht_{k,N}\sqrt{N}\right)^2\Big] \le KN\exp(-LN),
\end{aligned}
$$

and

$$
0 \le \sum_{k:t_{k,N}\in([-1;-1+\varepsilon]\cup[1-\varepsilon,1])}\binom{N}{k}\frac{A_N(h)}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(u_{N,h}(y)-ht_{k,N}\sqrt{N}\right)^2\right] \le KN\exp(-LN).
$$

Consider the sum appearing in (6.60)

$$S_N(h,y) := \sum_{k=0}^{N} \frac{A_N(h)}{\pi\left(N(1-t_{k,N}^2)\right)^{1/2}} \exp\left[N\left(\log 2 - I(t_{k,N})\right) - \frac{1}{2}\left(u_{N,h}(y) - ht_{k,N}\sqrt{N}\right)^2\right]. \tag{6.61}$$

Introduce the functions $f_N, g_N : [-1;1] \to \mathbb{R}$ as

$$f_N(t) := I(t) + \frac{1}{2}\left(\frac{u_{N,h}(y)}{\sqrt{N}} - ht\right)^2 - \log 2,$$
$$g_N(t) := \frac{A_N(h)}{\pi\,(N(1-t^2))^{1/2}}.$$

Note that definition (6.5) implies

$$I'(t_*) = h\mu(t_*). \tag{6.62}$$

A straightforward computation using (6.6), (6.7) and (6.62) gives

$$f_N''(t) = I''(t) + h > 0, \tag{6.63}$$
$$f_N'(t_*) = -\frac{h}{(2\mu(t_*)N)}\left[2y + \log\left(\frac{I''(t_*)+h}{4\pi(1-t_*^2)(\log 2 - I(t_*))N}\right)\right] = O\left(\frac{\log N}{N}\right), \tag{6.64}$$
$$f_N(t_*) = -\frac{1}{N}\left[y + \frac{1}{2}\log\left(\frac{I''(t_*)+h}{4\pi(1-t_*^2)(\log 2 - I(t_*))N}\right)\right] + o\left(\frac{1}{N}\right). \tag{6.65}$$

Hence, since (6.64) vanishes even after being multiplied by $\sqrt{N}$, (6.64) is negligible for the purposes of the asymptotic Laplace principle. This readily implies that uniformly in $y \in \operatorname{supp}\varphi$

$$S_N(h,y) \underset{N\uparrow+\infty}{\sim} \frac{N g_N(t_*)}{2}\left(\frac{2\pi f_N''(t_*)}{N}\right)^{1/2} \exp\left[N f_N(t_*)\right]. \tag{6.66}$$

Using (6.63), (6.64) and (6.65) in the r.h.s. of (6.66), we obtain that uniformly in $y \in \operatorname{supp}\varphi$

$$S_N(h,y) \underset{N\uparrow+\infty}{\sim} \exp(-y). \tag{6.67}$$

Finally, combining (6.67) and (6.60), we obtain

$$\lim_{N\uparrow+\infty} L_{\mathscr{E}_N(h)}(\varphi) = \exp\left\{-\int_{\mathbb{R}}\left(1-\mathrm{e}^{-\varphi(y)}\right)\mathrm{e}^{-y}\mathrm{d}y\right\}. \tag{6.68}$$

The r.h.s. of (6.68) is the Laplace transform of PPP $(\mathrm{e}^{-x}\mathrm{d}x, x \in \mathbb{R})$. Then a standard result implies the claim (6.10). $\square$

6.3.3 Fluctuations of the partition function

In this subsection, we compute the weak limiting distribution of the partition function under the natural scaling induced by (6.8). Define

$$C_N(\beta,h) := \exp\left[\beta M(h)N + \frac{\beta}{2\mu(t_*)}\log\left(\frac{I''(t_*)+h}{4\pi(1-t_*^2)(\log 2 - I(t_*))N}\right)\right], \tag{6.69}$$

$$D_N(\beta,h) := \mathrm{ch}^{-2/3}(\beta h)\exp\left[N\left(\log 2 + \log\mathrm{ch}\,\beta h + \frac{\beta^2}{2}\right)\right], \tag{6.70}$$

$$\alpha(\beta,h) := \frac{\beta}{\mu(t_*)}. \tag{6.71}$$

Theorem 6.3.4. *If $\beta > \mu(t_*)$, then*

$$\frac{Z_N(\beta,h)}{C_N(\beta,h)} \xrightarrow[N\uparrow+\infty]{w} \int_{\mathbb{R}} \mathrm{e}^{\alpha(\beta,h)x}\mathrm{d}\mathscr{P}^{(1)}(x). \tag{6.72}$$

If $\beta < \mu(t_*)$, then

$$\frac{Z_N(\beta,h)}{D_N(\beta,h)} \xrightarrow[N\uparrow+\infty]{w} 1. \tag{6.73}$$

Proof. This is a specialisation of Theorem 6.1.3 which is proved in Section 6.4.2. □

6.4 The GREM with external field

In this section, we obtain the main results of this chapter concerning the GREM with external field. We prove the limit theorems for the distribution of the partition function and that of the ground state.

6.4.1 Fluctuations of the ground state

As in the REM, we start from the ground state fluctuations (cf. Theorem 6.1.1). The following is the main technical result of this section which shows exactly in which situations the GREM with external field has the same scaling limit behaviour as the REM with external field.

Proposition 6.4.1. *Either of the following two cases holds*

1. *If, for all $l \in [2;n]\cap\mathbb{N}$,*

$$\frac{\log 2 - I(t_*(\theta_{l,n}^{-1/2}h))}{\log 2 - I(t_*(h))} < \theta_{l,n}, \tag{6.74}$$

then we have

$$\sum_{\sigma\in\Sigma_N} \delta_{u_{N,h}^{-1}(X_N(h,\sigma))} \xrightarrow[N\uparrow+\infty]{w} \mathrm{PPP}(\mathrm{e}^{-x}, x\in\mathbb{R}^d). \tag{6.75}$$

2. *If, for all $l \in [2,\dots,n] \cap \mathbb{N}$,*

$$\frac{\log 2 - I(t_*(\theta_{l,n}^{-1/2}h))}{\log 2 - I(t_*(h))} \leq \theta_{l,n}, \tag{6.76}$$

and there exists (at least one) $l_0 \in [2;n] \cap \mathbb{N}$

$$\frac{\log 2 - I(t_*(\theta_{l_0,n}^{-1/2}h))}{\log 2 - I(t_*(h))} = \theta_{l_0,n}, \tag{6.77}$$

then there exits the constant $K = K(\rho,h) \in (0;1)$ such that

$$\sum_{\sigma\in\Sigma_N} \delta_{u_{N,h}^{-1}(X_N(h,\sigma))} \xrightarrow[N\uparrow+\infty]{w} \mathrm{PPP}(K\mathrm{e}^{-x}, x\in\mathbb{R}). \tag{6.78}$$

Remark 6.4.1. *If condition (6.76) is violated, i.e., there exists $l_0 \in [2;n] \cap \mathbb{N}$ such that*

$$\frac{\log 2 - I(t_*(\theta_{l,n}^{-1/2}h))}{\log 2 - I(t_*(h))} > \theta_{l,n}, \tag{6.79}$$

then the REM scaling (cf. (6.10), (6.8)*) is too strong to reveal the structure of the ground state fluctuations of the GREM. Theorem 6.1.2 shows how the scaling and the limiting object should be modified to capture the fluctuations of the GREM in this regime.*

Proof. 1. Denote $N_l := \Delta x_l N$, for $l \in [1;n]$. We fix arbitrary test function $\varphi \in C_{\mathrm{K}}^{+}(\mathbb{R})$, i.e., a nonnegative function with compact support. Consider the Laplace transform $L_{\mathscr{E}_N(h)}(\varphi)$ of the random measure $\mathscr{E}_N(h)$ evaluated on the test function φ.

$$\begin{aligned} L_{\mathscr{E}_N(h)}(\varphi) &:= \mathbb{E}\left[\exp\Big(-\sum_{\sigma\in\Sigma_N}(\varphi\circ u_{N,h}^{-1})(X_N(h,\sigma))\Big)\right] \\ &= \mathbb{E}\left[\prod_{\sigma\in\Sigma_N}\exp\Big(-(\varphi\circ u_{N,h}^{-1})(X_N(h,\sigma))\Big)\right]. \end{aligned} \tag{6.80}$$

Consider also the family of i.i.d standard Gaussian random variables

$$\{X(\sigma^{(l)},\sigma^{(2)},\dots,\sigma^{(n)}) \mid l\in[1;n]\cap\mathbb{N}, \sigma^{(l)}\in\Sigma_{N_l},\dots,\sigma^{(n)}\in\Sigma_{N_n}\}.$$

Given $l \in [1;n] \cap \mathbb{N}$ and $y \in \mathbb{R}$, define

$$\begin{aligned} L_N(l,v) := \mathbb{E}\Big[\prod_{\sigma^{(l)}\sqcup\dots\sqcup\sigma^{(n)}\in\Sigma_{(1-x_{l-1})N}} &\exp\Big(-\varphi\circ u_{N,h}^{-1}(v + a_l X(\sigma^{(l)}) + \\ &\dots + a_n X(\sigma^{(l)},\dots,\sigma^{(n)}) + h(1-x_{l-1})\sqrt{N} m(\sigma^{(l)},\dots,\sigma^{(n)}))\Big)\Big]. \end{aligned} \tag{6.81}$$

We readily have

$$L_{\mathscr{E}_N(h)}(\varphi) = L_N(1,0). \tag{6.82}$$

Due to the tree-like structure of the GREM, for $l \in [1;n-1] \cap \mathbb{N}$, we have the following recursion

$$L_N(l,v) = \prod_{\sigma^{(l)} \in \Sigma_{N_l}} \mathbb{E}\left[L_N(l+1, v + a_l X + h\Delta x_l \sqrt{N} m(\sigma^{(l)}))\right], \tag{6.83}$$

where X is a standard Gaussian random variable. Introduce the following quantities

$$Y_N(h,y,v,t,l) := u_{N,h}(y) - h(1-x_{l-1})\sqrt{N}t - v.$$

We claim that, for any $l \in [1;n] \cap \mathbb{N}$, uniformly in $v \in \mathbb{R}$ satisfying

$$v \le \sqrt{N}\left(M(h) - \delta - (1-q_{l-1})\,\mu(h) - h(1-x_{l-1})t_*(\theta_{l,n}^{-1/2}h)\right),$$

we have

$$\begin{aligned}\log L_N(l,v) \underset{N\uparrow+\infty}{\sim} & -\frac{A_N(h)}{\sqrt{2\pi(1-q_{l-1})}} \sum_{k=0}^{(1-x_{l-1})N} \left(\binom{(1-x_{l-1})N}{k}\right. \\ & \times \left.\int_{\mathbb{R}} (1-\mathrm{e}^{-\varphi(y)}) \exp\Big[-\frac{1}{2(1-q_{l-1})} Y_N(h,y,v,t_{k,(1-x_{l-1})N}, l)^2\Big] \mathrm{d}y\right).\end{aligned} \tag{6.84}$$

We shall prove (6.84) by a decreasing induction in l starting from $l=n$.

2. The base of induction is a minor modification of the proof of Theorem 6.1.1. By the definition (6.81) and independence, we have

$$L_N(n,v) = \prod_{k=0}^{N_n} \left(\mathbb{E}\exp\Big[-(\varphi \circ u_{h,N}^{-1})(a_n X + h\Delta x_n \sqrt{N} t_{k,N_n} + v)\Big]\right)^{\binom{N_n}{k}}. \tag{6.85}$$

For fixed $k \in [0;N_n] \cap \mathbb{Z}$,

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-(\varphi \circ u_{N,h}^{-1})(a_n X + h\Delta x_n \sqrt{N} t_{k,N} + v)\right)\right] \\ &\quad = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \mathrm{d}x \exp\Big[-x^2/2 - (\varphi \circ u_{N,h}^{-1})(a_n x + h\Delta x_n \sqrt{N} t_{k,N} + v)\Big].\end{aligned} \tag{6.86}$$

We introduce in (6.86) the new integration variable

$$y := u_{N,h}^{-1}\left(a_n x + h\Delta x_n \sqrt{N} t_{k,N_n} + v\right). \tag{6.87}$$

Using the change of variables (6.87), we get that the r.h.s. of (6.86) is equal to

$$\frac{A_N(h)}{\sqrt{2\pi}a_n} \int_{\mathbb{R}} \mathrm{d}y \exp\left[-\frac{1}{2a_n^2} Y_N(h,y,v,t_{k,N_n},n)^2 - \varphi(y)\right]. \tag{6.88}$$

Combining (6.85) and (6.88), we get

$$L_N(n,v) = \prod_{k=0}^{N_n}\Big(\frac{A_N(h)}{\sqrt{2\pi}a_n}\int_{\mathbb{R}}\mathrm{d}y\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t_{k,N_n},n)^2-\varphi(y)\Big]\Big)^{\binom{N_n}{k}}$$
$$=\prod_{k=0}^{N_n}\Big(1-\frac{A_N(h)}{\sqrt{2\pi}a_n}\int_{\mathbb{R}}\mathrm{d}y\big(1-\mathrm{e}^{-\varphi(y)}\big)\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t_{k,N_n},n)^2\Big]\Big)^{\binom{N_n}{k}}.$$

Define

$$V_N(h,v,t,n) := \frac{A_N(h)}{\sqrt{2\pi}a_n}\int_{\mathbb{R}}\mathrm{d}y\big(1-\mathrm{e}^{-\varphi(y)}\big)\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t,n)^2\Big].$$

Given any small enough $\delta>0$, it straightforward to show that uniformly in $v\in\mathbb{R}$ such that

$$v\le\sqrt{N}\left(M(h)-h\Delta x_n t_*(h\theta_{n-1,n}^{-1/2})-\delta\right),$$

we have

$$L_N(n,v) \underset{N\uparrow+\infty}{=} \prod_{k=0}^{N_n}\left(1-\binom{N_n}{k}V_N(h,v,t_{k,N_n},n)\right)\big(1+\mathscr{O}(\mathrm{e}^{-CN})\big). \tag{6.89}$$

Indeed, we have

$$\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t_{k,N_n},n)^2\Big]\le\exp\Big[-N_n\big(\log 2-I(t_n)\big)\Big].$$

Next, using the fact that $(1-\mathrm{e}^{-\varphi(\cdot)})\in C_K^+(\mathbb{R})$, we get for some $C>0$

$$\int_{\mathbb{R}}\mathrm{d}y\big(1-\mathrm{e}^{-\varphi(y)}\big)\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t_{k,N_n},n)^2\Big]\le C\exp\Big[-N_n\big(\log 2-I(t_n)\big)\Big].$$

Applying the elementary bounds

$$x-x^2\le\log(1+x)\le x,\ \text{for } |x|<\frac{1}{2} \tag{6.90}$$

to

$$x := -\frac{A_N(h)}{\sqrt{2\pi}a_n}\int_{\mathbb{R}}\mathrm{d}y\big(1-\mathrm{e}^{-\varphi(y)}\big)\exp\Big[-\frac{1}{2a_n^2}Y_N(h,y,v,t_{k,N_n},n)^2\Big],$$

and using the fact that, due to (6.41), there exists $C>0$ such that uniformly in $k\in[1;N_n]\cap\mathbb{N}$

$$x^2\le\exp\Big[-2N_n\big(\log 2-I(t_n)\big)\Big]\binom{N_n}{k}\le\mathrm{e}^{-CN},$$

we get (6.89) and, consequently, (6.84) holds for $l=n$.

3. For simplicity of presentation, we prove only the induction step $l=n\rightsquigarrow l=n-1$. Due to (6.83), we have

$$L_N(n-1,v)=\prod_{k_{n-1}=0}^{N_{n-1}}\mathbb{E}\Big[L_N(n,v+a_{n-1}X+h\Delta x_{n-1}\sqrt{N}t_{k_{n-1},N_{n-1}})\Big]^{\binom{N_{n-1}}{k_{n-1}}}. \tag{6.91}$$

Define

$$t(k_n,k_{n-1}) := \frac{1}{1-x_{l-2}}\left(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}}\right).$$

Fix an arbitrary $\delta > 0$ and $\varepsilon > 0$. Due to (6.84) with $l = n$, there exists some $C > 0$, such that uniformly for all k_n, k_{n-1} with

$$t_{k_n,k_{n-1}} \in \{t \in [-1;1] : |t_*(\theta_{n-1,n}^{-1/2} h) - t_{k_n,k_{n-1}}| \le \varepsilon\},$$

and uniformly for all $v, x \in \mathbb{R}$ satisfying

$$\Delta x_n(\log 2 - I(t_{k_n,N_n})) \le \frac{1}{2a_n^2}\Big(M(h) - \delta - a_{n-1}x - N^{-1/2}v \\ - h(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}})\Big)^2, \tag{6.92}$$

we obtain

$$\left|\log L_N(n, v + a_{n-1}x + h\Delta x_{n-1}\sqrt{N} t_{k,N_{n-1}})\right| \le CN\exp(-N/C). \tag{6.93}$$

Define

$$x_N(v) := \frac{\sqrt{N}}{a_{n-1}}\Big(M(h) - \delta - vN^{-1/2} - a_n(2\Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2} \\ - h(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}})\Big). \tag{6.94}$$

Using the elementary bounds

$$1 + x \le \mathrm{e}^x \le 1 + x + x^2, \text{ for } |x| < 1, \tag{6.95}$$

and the bound (6.93), we obtain

$$\mathbb{E}\left[\mathbb{1}_{\{X \le x_N(v)\}} L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}})\right] \\ \underset{N\uparrow+\infty}{=} \mathbb{P}\{X \le x_N(v)\} + \mathbb{E}\left[\mathbb{1}_{\{X \le x_N(v)\}} \log L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}})\right] \\ + \mathscr{O}(N\exp(-N/C)). \tag{6.96}$$

Given $k_{n-1} \in [1; N_{n-1}] \cap \mathbb{N}$, we have

$$\mathbb{E}\left[\mathbb{1}_{\{X \le x_N(v)\}} \exp\left(-\frac{1}{2a_n^2} Y_N(h, y, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}}, t_{k_n,N_n}, n)^2\right)\right] \\ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_N(v)} \mathrm{d}x \exp\left[-\frac{x^2}{2} - \frac{1}{2a_n^2}\left(u_{N,h}(y) - a_{n-1}x - h\sqrt{N}(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}}) - v\right)^2\right] \\ = \exp\left(-\frac{1}{1-q_{n-2}} Y_N(h, y, v, t(k_n, k_{n-1}), n-1)^2\right)$$

$$\times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_N(v)} \exp\left[-\frac{a_n^2+a_{n-1}^2}{2a_n^2} \left(x - \frac{a_{n-1}}{a_n^2+a_{n-1}^2} Y_N(h,y,v,t(k_n,k_{n-1}),n-1) \right)^2 \right] \mathrm{d}x. \tag{6.97}$$

We claim that due to the strict inequalities (6.74), we have

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_N(v)} \exp\left[-\frac{a_n^2+a_{n-1}^2}{2a_n^2} \left(x - \frac{a_{n-1}}{a_n^2+a_{n-1}^2} Y_N(h,y,v,t(k_n,k_{n-1}),n-1) \right)^2 \right] \mathrm{d}x \xrightarrow[N\uparrow+\infty]{} \frac{a_n}{\left(a_n^2+a_{n-1}^2\right)^{1/2}}, \tag{6.98}$$

uniformly in $v \in \mathbb{R}$ such that

$$v \le \sqrt{N}\left(M(h) + \delta' - h(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}}) - (a_n^2+a_{n-1}^2)\mu(h) \right) =: v_N^{\max}, \tag{6.99}$$

where $0 < \delta'$ exists due to strict inequality (6.74), for $l = n$. Indeed, due to the standard bounds on Gaussian tails, to show (6.98) it is enough to check that

$$\frac{a_{n-1}}{a_n^2+a_{n-1}^2} Y_N(h,y,v,t(k_n,k_{n-1}),n-1) + \delta\sqrt{N} \le x_N(v), \tag{6.100}$$

for v satisfying (6.99). Due to (6.74) with $l = n$, there exists $\delta_3 > 0$ such that we have

$$(2\Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2} \le \mu(h) - \delta_3. \tag{6.101}$$

Choosing a small enough $\delta' > 0$, we have

$$\begin{aligned}
&x_N(v) - \frac{a_{n-1}}{a_n^2+a_{n-1}^2} Y_N(h,y,v,t(k_n,k_{n-1}),n-1) + \delta\sqrt{N} \\
&= a_n^2(M(h) - vN^{-1/2} - h\left(\Delta x_n t_{k_n,N_n} + \Delta x_{n-1} t_{k_{n-1},N_{n-1}})\right) \\
&\quad - (a_n^2+a_{n-1}^2)\left(a_n(2\Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2} - \delta\right) \\
&\underset{(6.99)}{\ge} a_n^2\left((a_n^2+a_{n-1}^2)\mu(h) - \delta'\right) - (a_n^2+a_{n-1}^2)\left(a_n(2\Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2} - \delta\right) \\
&\underset{(6.101)}{\ge} a_n^2\left((a_n^2+a_{n-1}^2)\mu(h) - \delta'\right) - (a_n^2+a_{n-1}^2)\left(a_n^2\mu(h) - \delta\right) \\
&= (a_n^2+a_{n-1}^2)(\delta_3 a_n^2 + \delta) - a_n^2\delta' > 0
\end{aligned}$$

which proves (6.100).
We claim that there exists $C > 0$ such that uniformly in $k_{n-1} \in [1; N_{n-1}] \cap \mathbb{N}$ and in $v \in \mathbb{R}$ satisfying (6.99) we have

$$\binom{N_{n-1}}{k_{n-1}} \mathbb{P}\{X \ge x_N(v)\} \le \exp(-N/C). \tag{6.102}$$

Indeed, in view of (6.41) and due to the classical Gaussian tail asymptotics, to obtain (6.102) it is enough to show that

$$N_{n-1}(\log 2 - I(t_{k_{n-1},N_{n-1}})) \leq \frac{1}{2} x_N^2(v_N^{\max}). \tag{6.103}$$

Using (6.99) and (6.94), we obtain

$$x_N(v_N^{\max}) = \frac{N^{1/2}}{a_{n-1}} \Big((a_n^2 + a_{n-1}^2)\mu(h) - a_n(2\Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2} + \delta' - \delta \Big). \tag{6.104}$$

If $n > 2$, then due to strict inequality (6.74), for $l = n-2$, there exists $\delta'' > 0$ such that we have

$$\begin{aligned}(a_n^2 + a_{n-2}^2)\mu(h) - \delta'' &> \Big((\log 2 - I(t_*(\theta_{l,n}^{-1/2} h)))(a_n^2 + a_{n-2}^2)(\Delta x_n + \Delta x_{n-1}) \Big)^{1/2} \\ &\geq (2a_{n-1}^2 \Delta x_{n-1}(\log 2 - I(t_{k_{n-1},N_{n-1}})))^{1/2} \\ &\quad + (2a_n^2 \Delta x_n(\log 2 - I(t_{k_n,N_n})))^{1/2},\end{aligned} \tag{6.105}$$

where the last inequality may be obtained as a consequence of Slepian's lemma (Slepian, 1962). If $n = 2$, then (6.105) follows directly from Slepian's lemma. Combining (6.104) and (6.105), we get (6.103). Note that (6.102), in particular, implies that

$$\mathbb{P}\{X \geq x_N(v)\} \leq \exp(-N/C). \tag{6.106}$$

Given $k_{n-1} \in [1; N_{n-1}] \cap \mathbb{N}$, denote

$$L_N(n-1, v, k_{n-1}) := \mathbb{E}\left[L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}}) \right]^{\binom{N_{n-1}}{k_{n-1}}}$$

Due to (6.106) and (6.96), we have

$$\begin{aligned}L_N(n-1, v, k_{n-1}) &= \mathbb{E}\left[(\mathbb{1}_{\{X \leq x_N(v)\}} + \mathbb{1}_{\{X > x_N(v)\}}) L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}}) \right]^{\binom{N_{n-1}}{k_{n-1}}} \\ &= \Big(1 + \mathbb{E}\left[\mathbb{1}_{\{X \leq x_N(v)\}} L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}}) \right] \\ &\quad + \mathscr{O}\left(\mathbb{P}\{X \geq x_N(v)\} + N\exp(-N/C) \right) \Big)^{\binom{N_{n-1}}{k_{n-1}}}.\end{aligned}$$

Using (6.102) and the standard bounds (6.90) and (6.95), we get

$$\begin{aligned}L_N(n-1, v, k_{n-1}) = \exp\Big\{ \binom{N_{n-1}}{k_{n-1}} \mathbb{E}\left[\mathbb{1}_{\{X \leq x_N(v)\}} \log L_N(n, v + a_{n-1}X + h\Delta x_{n-1}\sqrt{N} t_{k_{n-1},N_{n-1}}) \right] \\ + \mathscr{O}(N\exp(-N/C)) \Big\}.\end{aligned}$$

Applying (6.98), (6.97), (6.84), for $l = n$, we obtain

$$\log L_N(n-1, v, k_{n-1}) = -\frac{A_N(h)}{\sqrt{2\pi(a_n^2 + a_{n-1}^2)}} \sum_{k_n=0}^{N_n} \left(\binom{N_n}{k_n} \binom{N_{n-1}}{k_{n-1}} \right.$$

$$\times \int_{\mathbb{R}} \left(1-\mathrm{e}^{-\varphi(y)}\right) \exp\left[-\frac{1}{2(a_n^2+a_{n-1}^2)} Y_N(h,y,v,t_{k_n,k_{n-1}},n-1)^2\right] \mathrm{d}y\Big) + \mathscr{O}(N\exp(-N/C)).$$

Finally, we arrive at

$$\begin{aligned}
\log L_N(n-1,v) &= \sum_{k_{n-1}=0}^{N_{n-1}} \log L_N(n-1,v,k_{n-1}) \\
&= -\frac{A_N(h)}{\sqrt{2\pi(a_n^2+a_{n-1}^2)}} \sum_{k_n=0}^{N_n} \sum_{k_{n-1}=0}^{N_{n-1}} \Bigg(\binom{N_n}{k_n} \binom{N_{n-1}}{k_{n-1}} \\
&\quad \times \int_{\mathbb{R}} \left(1-\mathrm{e}^{-\varphi(y)}\right) \exp\left[-\frac{1}{2(a_n^2+a_{n-1}^2)} Y_N(h,y,v,t_{k_n,k_{n-1}},n-1)^2\right] \mathrm{d}y\Bigg) \\
&\quad + \mathscr{O}(N^2\exp(-N/C)) \\
&= -\frac{A_N(h)}{\sqrt{2\pi(a_n^2+a_{n-1}^2)}} \sum_{k=0}^{N_n+N_{n-1}} \Bigg(\binom{N_n+N_{n-1}}{k} \\
&\quad \times \int_{\mathbb{R}} \left(1-\mathrm{e}^{-\varphi(y)}\right) \exp\left[-\frac{1}{2(a_n^2+a_{n-1}^2)} Y_N(h,y,v,t_{k,N_n+N_{n-1}},n-1)^2\right] \mathrm{d}y\Bigg) \\
&\quad + \mathscr{O}(N^2\exp(-N/C)).
\end{aligned}$$

4. Combining (6.82) and (6.84) for $l=1$, we obtain

$$L_{\mathscr{E}_N(h)}(\varphi) = \exp\left(-\int_{\mathbb{R}} \left(1-\mathrm{e}^{-\varphi(y)}\right) S_N(h,y)\mathrm{d}y + o(1)\right), \tag{6.107}$$

where $S_N(h,y)$ is given by (6.61). Invoking the proof of Theorem 6.1.1, we get that

$$\begin{aligned}
L_{\mathscr{E}_N(h)}(\varphi) &\xrightarrow[N\uparrow+\infty]{} \exp\left(-\int_{\mathbb{R}} \left(1-\mathrm{e}^{-\varphi(y)}\right)\mathrm{e}^{-y}\mathrm{d}y\right) \\
&= L_{\mathscr{P}(\mathrm{e}^{-x})}(\varphi).
\end{aligned}$$

This establishes (6.75).

5. The proof of (6.78) is very similar to the above proof of (6.75). The main difference is that (6.98) does not hold. Instead, if (6.77) holds for $l_0 = n$, then we have

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_N(v)} \exp\left[-\frac{a_n^2+a_{n-1}^2}{2a_n^2}\left(x - \frac{a_{n-1}}{a_n^2+a_{n-1}^2} Y_N(h,y,v,t(k_n,k_{n-1}),n-1)\right)^2\right] \mathrm{d}x \\
&\xrightarrow[N\uparrow+\infty]{} \frac{a_n}{\left(a_n^2+a_{n-1}^2\right)^{1/2}} \mathbb{P}\Bigg\{X < \frac{\sqrt{N}}{a_{n-1}\sqrt{a_n^2+a_{n-1}^2}} \Big[M(h) - vN^{-1/2} - (1-x_{n-2})ht_*(h\theta_{n,n}^{-1/2}) \\
&\quad - (a_n^2+a_{n-1}^2)\mu(h)\Big]\Bigg\},
\end{aligned} \tag{6.108}$$

uniformly in

$$v \le \sqrt{N}\left(M(h) - (1-x_{n-2})ht_*(h\theta_{n,n}^{-1/2}) - (a_n^2 + a_{n-1}^2)\mu(h)\right) - \delta'.$$

The subsequent applications of the recursion (6.83) to (6.108) give rise to the constant $K(h,\rho) \in (0;1)$ in (6.78).

□

Proof of Theorem 6.1.2. The existence of the r.h.s. of (6.18) follows from (Bovier & Kurkova, 2004a, Theorem 1.5 (ii)). It remains to show convergence (6.18) itself. We apply Proposition 6.4.1 to each coarse-grained block. Note that the assumption (6.74) of Proposition 6.4.1 is fulfilled, due to the construction of the blocks, cf. (6.11), (6.17). The result then follows from (Bovier & Kurkova, 2004a, Theorem 1.2).

The representation of the limiting ground state (6.19) is proved exactly as in (Bovier & Kurkova, 2004a, Theorem 1.5 (iii))). □

6.4.2 Fluctuations of the partition function

In this subsection we compute the limiting distribution of the GREM partition function under the scaling induced by (6.8). The analysis amounts to handling both the low and high temperature regimes. The low temperature regime is completely described by the behaviour of the ground states which is summarised in Theorem 6.1.2. The high temperature regime is considered in Lemma 6.4.1 below.

Lemma 6.4.1. *Assume* $l(\beta,h) = 0$. *Then*

$$\begin{aligned}\exp\left[-N\left(\log 2 + \log \operatorname{ch}\beta h + \frac{\beta^2}{2}\right)\right] \operatorname{ch}^{2/3}(\beta h) Z_N(\beta,h) \\ \xrightarrow[N\uparrow+\infty]{w} K(\beta,h),\end{aligned} \tag{6.109}$$

where $K(\beta,h) = 1$, *if* $\beta\bar{\gamma}_1(h) < 1$, *and* $K(\beta,h) \in (0;1)$, *if* $\beta\bar{\gamma}_1(h) = 1$.

Proof. We follow the strategy of (Bovier & Kurkova, 2004a, Lemma 3.1). By the very construction of the coarse graining algorithm (6.11), we have

$$\begin{aligned}&\widetilde{\theta}_{1,k} \le \widetilde{\theta}_{1,J_1} = \bar{\gamma}_1(h)^2, \quad k \in [1;J_1] \cap \mathbb{N},\\ &\widetilde{\theta}_{1,k} < \widetilde{\theta}_{1,J_1}, \quad k \in (J_1;n] \cap \mathbb{N}.\end{aligned} \tag{6.110}$$

Assume $\beta\bar{\gamma}_1(h) < 1$. Hence, due to (6.110), we have

$$\beta\widetilde{\theta}_{1,k}^{1/2} < 1, \quad k \in [1;n] \cap \mathbb{N}. \tag{6.111}$$

Strict inequality (6.111) implies that there exists $\varepsilon > 0$ such that, for all $k \in [1;n] \cap \mathbb{N}$,

$$\left(\beta^2 - \frac{1}{2}(\beta-\varepsilon)^2\right) q_k < x_k \left(\log 2 - I(t_*(h(x_k/q_k)^{1/2}))\right). \tag{6.112}$$

We have

$$\mathbb{E}\left[Z_N(\beta,h)\right] = \sum_{k=0}^{N} \binom{N}{k} \exp\left(\beta h t_{k,N} N + \frac{\beta^2 N}{2}\right) =: S_N(\beta,h). \tag{6.113}$$

Note that due to (6.41)

$$S_N(\beta,h) \underset{N\uparrow+\infty}{\sim} \sum_{k=0}^{N} g_N(t_{k,N}) \exp\left(N f(t_{k,N})\right), \tag{6.114}$$

where

$$f(t) := \log 2 - I(t) + \beta h t + \beta^2/2,$$
$$g_N(t) := \left(\frac{2}{\pi N(1-t^2)}\right)^{1/2}.$$

A straightforward computation gives

$$f'(t_0) = \beta h - \tanh^{-1}(t_0(\beta,h)) = 0$$
$$f''(t_0) = -(1-t_0^2)^{-1} = -\operatorname{ch}^2(\beta h),$$
$$g_N(t_0) = \left(\frac{2}{\pi N(1-t^2)}\right)^{1/2} = \left(\frac{2}{\pi N}\right)^{1/2} \operatorname{ch}(\beta h).$$

The asymptotic Laplace method then yields

$$S_N(\beta,h) \underset{N\uparrow+\infty}{\sim} \operatorname{ch}^{-2/3}(\beta h) \exp\left[N\left(\log 2 + \log \operatorname{ch}\beta h + \frac{\beta^2}{2}\right)\right]. \tag{6.115}$$

For $p \leq q$, define

$$\mathrm{GREM}_N^{(p,q)}(\sigma^{(1)},\ldots,\sigma^{(q)}) := \sum_{k=p}^{q} a_k X(\sigma^{(1)},\ldots,\sigma^{(k)}).$$

Consider the event

$$E_N(\sigma) := \Big\{\mathrm{GREM}_N^{(1,k)}(\sigma^{(1)},\ldots,\sigma^{(k)}) < (\beta+\varepsilon) q_k \sqrt{N},$$
$$\text{for all } k \in [1;n] \cap \mathbb{N}\Big\}.$$

Define the truncated partition function as

$$Z_N^{(\mathrm{T})}(\beta,h) := \sum_{\sigma\in\Sigma_N} \mathbb{1}_{E_N(\sigma)} \exp\left[\beta\sqrt{N} X_N(h,\sigma)\right]. \tag{6.116}$$

The truncation (6.116) is mild enough in the following sense

$$\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right] = S_N(\beta,h)\mathbb{P}\left\{\mathrm{GREM}_N^{(1,k)}(\sigma^{(1)},\ldots,\sigma^{(k)}) < \varepsilon q_k \sqrt{N}, \text{for all } k \in [1;n] \cap \mathbb{N}\right\}$$

$$\underset{N\uparrow+\infty}{\sim} \mathbb{E}\left[Z_N(\beta,h)\right]. \tag{6.117}$$

We write

$$\begin{aligned}\frac{Z_N(\beta)}{\mathbb{E}\left[Z_N(\beta)\right]} &= \frac{Z_N^{(\mathrm{T})}(\beta)}{\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right]} \times \frac{\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right]}{\mathbb{E}\left[Z_N(\beta)\right]} + \frac{Z_N(\beta)-Z_N^{(\mathrm{T})}(\beta)}{\mathbb{E}\left[Z_N(\beta)\right]}\\ &=: (\mathrm{I})\times(\mathrm{II})+(\mathrm{III}).\end{aligned}$$

Due to (6.117), we get

$$(\mathrm{II}) \underset{N\uparrow+\infty}{\sim} 1, \quad (\mathrm{III}) \xrightarrow[N\uparrow+\infty]{L^1} 0.$$

To estimate (I), we fix any $\delta>0$, and use the Chebyshev inequality

$$\mathbb{P}\left\{|(\mathrm{I})-1|>\delta\right\} \leq \left(\delta\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right]\right)^{-2}\operatorname{Var}\left[Z_N^{(\mathrm{T})}(\beta)\right]. \tag{6.118}$$

Expanding the squares, we have

$$\begin{aligned}\operatorname{Var}\left[Z_N^{(\mathrm{T})}(\beta)\right] &= \mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)^2\right]-\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right]^2\\ &= \sum_{p=1}^{n}\sum_{\sigma^{(1)}\amalg\ldots\amalg\sigma^{(k)}\in\Sigma_{x_kN}}\mathbb{E}\Bigg[\exp\Big\{2\beta\sqrt{N}\Big(\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})\\ &\quad +2\beta h x_p m_{x_pN}(\sigma^{(1)},\ldots,\sigma^{(p)})\sqrt{N}\Big)\Big\}\\ &\quad\times \sum_{\substack{\sigma^{(p+1)}\amalg\ldots\amalg\sigma^{(n)},\\ \tau^{(p+1)}\amalg\ldots\amalg\tau^{(n)}\in\Sigma_{(1-x_p)N},\\ \sigma^{(p+1)}\neq\tau^{(p+1)}}}\exp\Big\{\beta\sqrt{N}\\ &\quad\times\Big(\mathrm{GREM}_N^{(p+1,n)}(\sigma^{(1)},\ldots,\sigma^{(n)})+\mathrm{GREM}_N^{(p+1,n)}(\tau^{(1)},\ldots,\tau^{(n)})\\ &\quad+h(1-x_p)\sqrt{N}(m_{(1-x_p)N}(\sigma^{(p+1)}\amalg\ldots\amalg\sigma^{(n)})+m_{(1-x_p)N}(\tau^{(p+1)}\amalg\ldots\amalg\tau^{(n)})\Big)\Big\}\\ &\quad\times\mathbb{1}_{E_N(\sigma^{(1)}\amalg\ldots\amalg\sigma^{(n)})}\mathbb{1}_{E_N(\tau^{(1)}\amalg\ldots\amalg\tau^{(n)})}\Bigg].\end{aligned} \tag{6.119}$$

Hence, due to the independence, we arrive at

$$\begin{aligned}\operatorname{Var}\left[Z_N^{(\mathrm{T})}(\beta)\right] &\leq \sum_{p=1}^{n}\sum_{k=0}^{x_kN}\binom{N}{k}\mathbb{E}\Bigg[\exp\Big\{2\beta\sqrt{N}\Big(\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})\\ &\quad +h x_p t_{k,N}\sqrt{N}\Big)\Big\}\mathbb{1}_{\left\{\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})<(\beta+\varepsilon)q_p\sqrt{N}\right\}}\Bigg]\\ &\quad\times\left(\sum_{k=0}^{(1-x_p)N}\binom{(1-x_p)N}{k}\mathbb{E}\Bigg[\exp\Big(\beta\sqrt{N}(\mathrm{GREM}_N^{(p+1,n)}(\sigma^{(1)},\ldots,\sigma^{(n)})\right.\end{aligned}$$

$$+h(1-x_p)t_{k,(1-x_p)N}\sqrt{N}\Big)\Big]\Big]\Big)^2. \tag{6.120}$$

Assume that X is a standard Gaussian random variable. Using the standard Gaussian tail bounds, we have

$$\begin{aligned}
&\mathbb{E}\Bigg[\exp\Big(2\beta\sqrt{N}(\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})+hx_pt_{k,N}\sqrt{N})\Big)\\
&\qquad \mathbb{1}_{\left\{\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})<(\beta+\varepsilon)q_p\sqrt{N}\right\}}\Bigg]\\
&= \exp\left\{N\left(2\beta^2q_p+\beta ht_{k,N}\right)\right\}\mathbb{P}\left\{X\geq(\beta-\varepsilon)\sqrt{q_pN}\right\}\\
&\underset{N\uparrow+\infty}{\leq} C\exp\left\{N\left(2\beta^2q_p+\beta ht_{k,N}-\frac{1}{2}(\beta-\varepsilon)^2q_p\right)\right\}.
\end{aligned} \tag{6.121}$$

Similarly to (6.114), using (6.41) and (6.121), we have

$$\begin{aligned}
&\sum_{k=0}^{x_kN}\binom{N}{k}\mathbb{E}\Bigg[\exp\Big(2\beta\sqrt{N}(\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})\\
&\qquad +hx_pt_{k,N}\sqrt{N})\Big)\mathbb{1}_{\left\{\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)})<(\beta+\varepsilon)q_p\sqrt{N}\right\}}\Bigg]\\
&\underset{N\uparrow+\infty}{\leq} C\sum_{k=0}^{x_kN}\exp\left\{N\left(x_p(\log 2-I(t_{k,x_pN}))+2\beta^2q_p+2\beta hx_pt_{k,x_pN}-\frac{1}{2}(\beta-\varepsilon)^2q_p\right)\right\}=:P_N(p).
\end{aligned} \tag{6.122}$$

Using (6.41), we also obtain

$$\begin{aligned}
&\sum_{k=0}^{(1-x_p)N}\binom{(1-x_p)N}{k}\mathbb{E}\Bigg[\exp\Big(\beta\sqrt{N}(\mathrm{GREM}_N^{(p+1,n)}(\sigma^{(1)},\ldots,\sigma^{(n)})\\
&\qquad +h(1-x_p)\sqrt{N}t_{k,(1-x_p)N}\Big)\Bigg]\\
&\underset{N\uparrow+\infty}{\leq} C\sum_{k=0}^{(1-x_p)N}\exp\Big\{N\Big((1-x_p)(\log 2-I(t_{k,(1-x_p)N}))+\frac{1}{2}(1-q_p)\beta^2\\
&\qquad +\beta h(1-x_p)t_{k,(1-x_p)N}\Big)\Big\}=:\widetilde{P}_N(p).
\end{aligned} \tag{6.123}$$

Combining (6.120), (6.122) and (6.123), we get

$$\mathrm{Var}\left[Z_N^{(\mathrm{T})}(\beta)\right]\underset{N\uparrow+\infty}{\leq}\sum_{p=1}^{N}P_N(p)\widetilde{P}_N^2(p). \tag{6.124}$$

For any $p\in[1;n]\cap\mathbb{N}$, we have the following factorisation

$$\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right] = \sum_{k=0}^{x_pN} \binom{x_pN}{k} \mathbb{E}\left[\exp\left(\beta\sqrt{N}(\mathrm{GREM}_N^{(1,p)}(\sigma^{(1)},\ldots,\sigma^{(p)}) + hx_pt_{k,x_pN}\sqrt{N})\right)\right.$$
$$\times \sum_{k=0}^{(1-x_p)N} \binom{(1-x_p)N}{k} \exp\Big(\beta\sqrt{N}(\mathrm{GREM}_N^{(p+1,n)}(\sigma^{(1)},\ldots,\sigma^{(n)})$$
$$\left.+h(1-x_p)Nt_{k,(1-x_p)N}\sqrt{N})\Big)\mathbb{1}_{E_N(\sigma^{(1)}\|\ldots\|\sigma^{(n)})}\right]. \tag{6.125}$$

Hence, again similarly to (6.114), we obtain

$$\mathbb{E}\left[Z_N^{(\mathrm{T})}(\beta)\right] \underset{N\uparrow+\infty}{\sim} C\sum_{k=0}^{x_pN} \exp\left\{N\left(x_p(\log 2 - I(t_{k,x_pN})) + \frac{1}{2}q_p\beta^2 + \beta hx_pt_{k,x_pN}\right)\right\}$$
$$\times \sum_{k=0}^{(1-x_p)N} \exp\Big\{N\Big((1-x_p)(\log 2 - I(t_{k,(1-x_p)N})) + \frac{1}{2}(1-q_p)\beta^2$$
$$+\beta h(1-x_p)t_{k,(1-x_p)N}\Big)\Big\} =: Q_N(p) \times \widetilde{P}_N(p). \tag{6.126}$$

Denote

$$R_N(p) := q_p\beta^2 + 2x_p \max_{t\in[-1;1]} \{\log 2 - I(t) + \beta ht\}.$$

We observe that similarly to (6.115) we have

$$\frac{Q_N^2(p)}{\exp(NR(p))} \underset{N\uparrow+\infty}{\sim} C. \tag{6.127}$$

Combining (6.124), (6.126), (6.127) and (6.51), we get

$$(6.118) \underset{N\uparrow+\infty}{\leq} C\sum_{p=1}^{n} \frac{P_N(p)}{Q_N^2(p)} = C\sum_{p=1}^{n} \frac{P_N(p)/\exp(NR(p))}{Q_N^2(p)/\exp(NR(p))} \underset{N\uparrow+\infty}{\leq} C\sum_{p=1}^{n} \frac{P_N(p)}{\exp(NR(p))}$$
$$\underset{N\uparrow+\infty}{\leq} C\sum_{p=1}^{n} \exp\left\{N\left((\beta^2 - \frac{1}{2}(\beta-\varepsilon)^2)q_p - (\log 2 - I(t_0))x_p\right)\right\}$$
$$= C\sum_{p=1}^{n} \exp\left\{N\left((\beta^2 - \frac{1}{2}(\beta-\varepsilon)^2)q_p - (\log 2 - I(t_*(h(x_p/q_p)^{1/2})))x_p\right)\right\} \xrightarrow[N\uparrow+\infty]{} 0, \tag{6.128}$$

where the convergence to zero in the last line is assured by the choice of ε in (6.112). Finally, combining (6.118) and (6.128), we get

$$(\mathrm{I}) \xrightarrow[N\uparrow+\infty]{\mathbb{P}} 1.$$

This finishes the proof of (6.109) in the case $\beta\bar{\gamma}_1(h) < 1$.

The case $\beta\bar{\gamma}_1(h) = 1$ is a little bit more tedious and uses the information about the low temperature regime obtained in Theorem 6.1.2 in the spirit of the proof of (Bovier & Kurkova, 2004a, Lemma 3.1). The lemma follows. □

Proof of Theorem 6.1.3. The proof is verbatim the one of (Bovier & Kurkova, 2004a, Theorem 1.7), where the analysis of the high temperature regime (Bovier & Kurkova, 2004a, Lemma 3.1) is substituted by Lemma 6.4.1. The low temperature regime is governed by the fluctuations of the ground state which are summarised in Theorem 6.1.2.

□

6.4.3 Formula for the free energy of the GREM

Proof of Theorem 6.1.4. The L^1 convergence follows immediately from Theorem 6.1.3. Almost sure convergence is a standard consequence of Gaussian measure concentration, e.g., (Ledoux, 2001, (2.35)), and the Borell-Cantelli lemma. □

7

Some open problems and outlook

In this thesis, we studied two types of the large sums of strongly correlated exponentials. The first type is a sum of hierarchically correlated random variables (the GREM with external field). The second type is an infinitesimal sum of genuine non-hierarchically strongly correlated random variables (the SK model with multidimensional spins). We were interested in the limiting behaviour of such sums as the number of their summands and the effective dimension of the correlation structure simultaneously tend to infinity.

For the GREM with external field, we provided an explicit expression for the free energy (Theorem 6.1.4) and even obtained precise information about the fluctuations of the partition function (Theorem 6.1.3). Our understanding of the SK model with multidimensional spins is only at the level of free energy (cf. Question 0.0.1) and is substantially less elaborate. Below we list some open questions concerning the latter model.

In view of the bounds on the free energy (of the SK model with multidimensional spins) obtained using the generalised AS^2 scheme (Theorems 3.1.1 and 3.1.2, cf. also (3.22)), it is natural to pose the following question.

Question 7.0.1 (Parisi-type formula, the AS^2 scheme version)**.** *Using the notations of Chapter 3 and, in particular, of Theorem 3.1.1, for which a priori distributions of spins do we have*

$$p(\beta) = \sup_{U \in \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{7.1}$$

where the infimum runs over all x satisfying (3.7)*, all* $\Lambda \in \mathrm{Sym}(d)$*, all* $\mathscr{Q}$ *satisfying both* (3.6) *and "Hadamard squares" Assumption 3.1.2?*

In view of the bounds on the free energy obtained using generalised Guerra's scheme (see Theorems 4.1.1 and 4.1.2), we pose the following version of Question 7.0.1.

Question 7.0.2 (Parisi-type formula, Guerra's scheme version)**.** *Using the notation of Theorem 4.1.1, for which a priori distributions of spins do we have* (7.1)*, where the infimum in* (7.1) *runs over all x satisfying* (3.7)*, all* $\mathscr{Q}$ *satisfying* (3.6)*, and all* $\Lambda \in \mathrm{Sym}(d)$*?*

In view of the results on the Parisi-type formula in the case of multidimensional Gaussian a priori distribution (Chapter 5), we pose the following question.

Question 7.0.3 (simultaneous diagonalisation scenario)**.** *Using the notations of Section 4.2.3, for which a priori distributions of spins does* (7.1) *hold, where the infimum in* (7.1) *runs over all* $x \in \mathscr{Q}'_n(1,1)$*, all* $\mathscr{Q} \in \mathscr{Q}_{diag}(U,O,d) \cap \mathscr{Q}'_n(U,d)$*, all* $O \in \mathscr{O}(d)$*, all* $n \in \mathbb{N}$ *and all* $\Lambda \in \mathrm{Sym}(d)$*?*

In view of the conditional positive answer to Question 7.0.2 which we obtained by Talagrand's methodology of a priori estimates (cf. Theorem 4.3.1), we pose the following question.

Question 7.0.4. *For which a priori distributions of spins is Assumption 4.3.1 satisfied?*

A related question (see Remark 4.3.8) seems to be the following one.

Question 7.0.5. *For which a priori distributions of spins does the condition* (4.75) *hold?*

We embedded an open problem of strict convexity of the Parisi functional posed by Panchenko (2005a); Talagrand (2006c) into a more general setting in Section 4.2. In view of our partial result on strict convexity (see Theorem 4.2.4), this problem can be generalised and paraphrased in the following way.

Question 7.0.6 (multidimensional Parisi functional)**.** *Given a piece-wise continuous $x \in \mathscr{Q}(1,1)$ and $Q \in \mathscr{Q}(U,d)$, assume $f_{x,\mathscr{Q}}$ is the solution of* (4.39)*. For which terminal conditions g and parameters Q is the mapping $\mathscr{Q}'(1,1) \ni x \mapsto f_{x,\mathscr{Q}}(0,0) \in \mathbb{R}$ strictly convex?*

Remark 7.0.2. *The case of Question 7.0.6 with $d = 1$ essentially corresponds to the problem posed by Panchenko (2005a); Talagrand (2006c).*

Remark 7.0.3. *It would, perhaps, also be useful to understand the behaviour of the mapping $\mathscr{Q}(U,d) \ni Q \mapsto f_{x,\mathscr{Q}}(0,0) \in \mathbb{R}$.*

A

Appendix

The general result of Guerra & Toninelli (2003) implies that the thermodynamic limit of the local free energy (3.5) exists almost surely and in L^1. The following existence of the limiting average overlap is an immediate consequence of this.

Proposition A.0.1. *We have*

$$\mathbb{E}\left[\mathscr{G}_N(\beta)\otimes\mathscr{G}_N(\beta)\left[\operatorname{Var}H_N(\sigma)-\mathbb{E}\left[H_N(\sigma)H_N(\sigma')\right]\right]\right]\xrightarrow[N\uparrow+\infty]{}C(\beta)\geq 0,$$

where $C:\mathbb{R}_+\to\mathbb{R}_+$.

Proof. The free energy is a convex function of β (a consequence of the Hölder inequality). Hence, by a result of Griffiths (1964) the following holds

$$\lim_{N\uparrow\infty}\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p_N(\beta)\right]=\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p(\beta)\right].$$

Proposition 3.2.4 implies

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p_N(\beta)\right]=\beta\mathbb{E}\left[\mathscr{G}_N(\beta)\otimes\mathscr{G}_N(\beta)\left[\operatorname{Var}H_N(\sigma)-\mathbb{E}\left[H_N(\sigma)H_N(\sigma')\right]\right]\right].$$

□

The following super-additivity result is an application of the Gaussian comparison inequalities obtained in Subsection 3.2.3. Note that the result does not provide enough information for the cavity-like argument of Aizenman *et al.* (2003).

Proposition A.0.2. *For any* $\mathscr{V}\equiv B(U,\varepsilon)\subset\mathscr{U}$*, we have*

$$N\mathbb{E}\left[p_N(\mathscr{V})\right]+M\mathbb{E}\left[p_M(\mathscr{V})\right]\leq(N+M)\mathbb{E}\left[p_{N+M}(\mathscr{V})\right]+(N+M)\mathscr{O}(\varepsilon),$$

as $\varepsilon\downarrow+0$.

Proof. Define the process $Y_{N,M}:=\{Y(\sigma):\sigma=\alpha\sqcup\tau;\alpha\in\Sigma_N,\tau\in\Sigma_M\}$ as follows

$$Y(\alpha\sqcup\tau):=\left(\frac{N}{N+M}\right)^{1/2}X_N^{(1)}(\alpha)+\left(\frac{M}{N+M}\right)^{1/2}X_M^{(2)}(\tau),$$

where $X^{(1)}$ and $X^{(2)}$ are two independent copies of the process X. Given some Gaussian process $\{C(\sigma)\}_{\sigma\in\Sigma_N}$, let us introduce the functional $\Phi_N(\beta)[C]$ as follows

$$\Phi_{N,M}(\beta)[C] := \mathbb{E}\left[\log \mu^{\otimes(N+M)}\left[\mathbb{1}_{\Sigma_N(\mathscr{V})}\mathbb{1}_{\Sigma_M(\mathscr{V})}\exp(\beta\sqrt{N+M}C)\right]\right].$$

Now, set $\varphi(t) := \Phi_{N+M}(\beta)\left[\sqrt{t}X_{N+M}+\sqrt{1-t}Y_{N,M}\right]$. Applying Proposition 3.2.5, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \frac{\beta^2(N+M)}{2}\mathbb{E}\,[\mathscr{G}(t)\otimes\mathscr{G}(t)\,[\\
&\Big(\mathrm{Var}\,X_{N+M}(\sigma^{(1)}) - \mathrm{Var}\,Y_{N,M}(\sigma^{(1)})\Big)\\
&-\Big(\mathrm{Cov}\left[X_{N+M}(\sigma^{(1)}),X_{N+M}(\sigma^{(2)})\right] - \mathrm{Cov}\left[Y_{N,M}(\sigma^{(1)}),Y_{N,M}(\sigma^{(2)})\right]\Big)\Big]\Big].\end{aligned}\tag{A.1}$$

Note that we have

$$\begin{aligned}\varphi(0) &= N\mathbb{E}\left[p_N(\mathscr{V})\right] + M\mathbb{E}\left[p_M(\mathscr{V})\right],\\
\varphi(1) &\le (N+M)\mathbb{E}\left[p_{N+M}(\mathscr{V})\right],\end{aligned}\tag{A.2}$$

where the last inequality is due to the fact that, for all $\alpha\in\Sigma_N(\mathscr{V})$ and all $\tau\in\Sigma_N(\mathscr{V})$, we have

$$\alpha \sqcup \tau \in \Sigma_{N+M}(\mathscr{V}).$$

Moreover, for $\sigma = \alpha\sqcup\tau$ with $\alpha\in\Sigma_N(\mathscr{V})$ and $\sigma\in\Sigma_M(\mathscr{V})$ we have

$$\begin{aligned}\mathrm{Var}\,X_{N+M}(\sigma) - \mathrm{Var}\,Y_{N,M}(\sigma) &= \left\|\frac{N}{N+M}R_N(\alpha,\alpha) + \frac{M}{N+M}R_M(\tau,\tau)\right\|_2^2 - \frac{N}{N+M}\|R_N(\alpha,\alpha)\|_2^2\\
&\quad - \frac{M}{N+M}\|R_M(\tau,\tau)\|_2^2 = \mathscr{O}(\varepsilon).\end{aligned}$$

Also, due to convexity of the norm, we have

$$\begin{aligned}&\mathrm{Cov}\left[X_{N+M}(\sigma^{(1)}),X_{N+M}(\sigma^{(2)})\right] - \mathrm{Cov}\left[Y_{N,M}(\sigma^{(1)}),Y_{N,M}(\sigma^{(2)})\right]\\
&= \left\|\frac{N}{N+M}R_N(\alpha^{(1)},\alpha^{(2)}) + \frac{M}{N+M}R_M(\tau^{(1)},\tau^{(2)})\right\|_2^2 - \frac{N}{N+M}\|R_N(\alpha^{(1)},\alpha^{(2)})\|_2^2\\
&\quad - \frac{M}{N+M}\|R_M(\tau^{(1)},\tau^{(2)})\|_2^2 \le 0.\end{aligned}$$

Applying $\int_0^1 \mathrm{d}t$ to (A.1) and using the previous two formulae, we get the claim. □

Notation Index

Sets, spaces

$\mathbb{N}$	natural numbers, i.e., $\{1,2,\dots\}$
$\mathbb{R}$	real axis, i.e., $(-\infty;+\infty)$
$\mathbb{R}_+$	nonnegative reals, i.e., $[0;+\infty)$
$\overline{\mathbb{R}}_+$	compactified nonnegative reals, i.e., $[0;+\infty]$
$\underline{\mathbb{R}}$	$[-\infty;+\infty)$
$\mathbb{S}^{d-1}$	unit Euclidean sphere in $\mathbb{R}^d$ with centre at $0 \in \mathbb{R}^d$
$B(x,r)$	Euclidean ball with centre at x and radius $r \in \mathbb{R}_+$
$\mathbb{R}^{d\times d}$	real $d \times d$ matrices
$\mathrm{Sym}(d)$	symmetric $d \times d$ matrices
$\mathrm{Sym}^+(d)$	symmetric non-negative-definite $d \times d$ matrices
$\mathscr{O}(d)$	orthogonal $d \times d$ matrices
$C^{(k)}(D)$	k-times continuously differentiable functions $f : D \subset \mathbb{R}^d \to \mathbb{R}$
$C^+_{\mathrm{K}}(D)$	nonnegative functions $f : D \subset \mathbb{R}^d \to \mathbb{R}_+$ with compact support
$\mathscr{M}_{\mathrm{f}}(D)$	finite measures on the metric space D
$\mathscr{M}_1(D)$	probability measures on the metric space D
Λ	a finite index set
Λ_0	an infinite index set
Σ	a single spin configuration space
Σ_Λ (or Σ_N)	whole system configuration space
Σ^{Ising}	Ising spins, $\Sigma^{\mathrm{Ising}} := \{-1;1\}$
$(\Omega,\mathscr{F},\mathbb{P})$	probability space of disorder
$\mathfrak{G}_0(\beta)$	set of all infinite-volume (DLR) Gibbs measures
$\mathscr{Q}(U,d)$	set of all càdlàg $\mathrm{Sym}^+(d)$-valued non-decreasing (in the sense of quadratic forms) "paths" which start in 0 and end in matrix U, (4.17)
$\mathscr{Q}'(U,d)$	set of all piece-wise constant paths in $\mathscr{Q}(U,d)$ with finite (but arbitrary) number of jumps
$\mathscr{Q}'_n(U,d)$	set of all piece-wise constant paths in $\mathscr{Q}'(U,d)$ with exactly $n \in \mathbb{N}$ jumps
$\mathscr{E}(I)$	set of all equivalence relations on the set I

Relations

$\Lambda \Subset \Lambda_0$	set Λ is the finite subset of the infinite set Λ_0

$x_\alpha \xrightarrow[(\alpha)]{w} y$	(x_α) weakly converges to y along the direction α
$x_\alpha \xrightarrow[(\alpha)]{\mathbb{P}} y$	(x_α) converges to y in probability along the direction α
$x_\alpha \underset{(\alpha)}{\sim} y_\alpha$	(x_α) and (y_α) are asymptotically equivalent along the direction α, i.e., $\lim_\alpha \frac{y_\alpha}{x_\alpha} = 1$
$x_\alpha \underset{(\alpha)}{=} \mathscr{O}(y_\alpha)$	"the big-O notation", rough asymptotic domination along the direction α, i.e., $\overline{\lim}_\alpha \lvert\frac{x_\alpha}{y_\alpha}\rvert < +\infty$
$A \succeq 0$	matrix $A \in \mathbb{R}^{d\times d}$ is nonnegative definite
$A \preceq B$	$B - A \succeq 0$
$X \sim Y$	random variables X and Y are equidistributed
$i \underset{k}{\sim} j$	Bolthausen-Sznitman equivalence relation, (2.27)
$\Pi_1 \preceq \Pi_2$	partition Π_1 is not finer than Π_2

Operations

$x := y$ $(y =: x)$	element x equals y by definition
$x \shortparallel y$	concatenation of the vectors x and y
$[x]_k$	projection of the vector $x \in \mathbb{R}^n$ on its first k (with $k \le n$) coordinates, i.e., $(x_1, \dots, x_k) \in \mathbb{R}^k$
$[\sigma]_A$	projection of the vector $\sigma \in \Sigma_{\Lambda_0}$ on Σ_A, where $A \subset \Lambda$
$\lfloor x \rfloor$	largest integer smaller than x
$x \vee y$	maximum of x and y
$x \wedge y$	minimum of x and y
$\mathscr{N}(\cdot)$	normalisation operation, (2.18)
$\operatorname{card} \Lambda$	cardinality of the set Λ
$\operatorname{diag} x$	diagonal matrix in $\mathbb{R}^{d\times d}$ induced by the vector $x \in \mathbb{R}^d$
$\det A$	determinant of the matrix $A \in \mathbb{R}^{d\times d}$
$\nabla f(x)$	gradient of the function $f : D \subset \mathbb{R}^d \to \mathbb{R}$, i.e., $(\partial_1 f(x), \dots, \partial_d f(x))$
$\nabla^2 f(x)$	Hessian of the function $f : D \subset \mathbb{R}^d \to \mathbb{R}$, i.e., $(\partial^2_{u,v} f(x))^d_{u,v=1}$
$\partial_{x \rightsquigarrow e} F(x)$	directional (Gâteaux) derivative of $F : X \to \mathbb{R}$ at the point x along the direction e
$\langle x, y\rangle$	Euclidean scalar product between the vectors $x, y \in \mathbb{R}^d$, i.e., $\sum_{u=1}^d x_u y_u$
$\langle A, B\rangle$	tracial scalar product between the matrices $A, B \in \mathbb{R}^{d\times d}$, i.e., $\sum_{u,v=1}^d A_{u,v} B_{u,v}$
$\mu \otimes \nu$	product measure
$\mu^{\otimes \Lambda}$	product measure, i.e., $\bigotimes_{i\in\Lambda} \mu$
$x^{\odot a}$	coordinate-wise raising of the vector $x \in \mathbb{R}^d$ to the power $a \in \mathbb{R}$
$A^{\odot a}$	entry-wise raising of the matrix $x \in \mathbb{R}^{d\times d}$ to the power $a \in \mathbb{R}$
$\operatorname{int} D$	interior of the set D
$\operatorname{conv} D$	convex hull of the set D
∂D	border of the set D

Norms, metrics, valuations

$\mathbb{1}_D(x)$	indicator function of the set D
$\mathbb{E}[X]$	expectation of the random variable X

$\mu[X]$	average of the function X with respect to the measure μ, i.e., $\int X d\mu$
$\mathrm{Var}[X]$	variance of the random variable X
$\mathrm{Cov}[X,Y]$	covariance between the random variables X and Y
$\lVert x \rVert_2$	Euclidean norm of the vector $x \in \mathbb{R}^d$, i.e., $\sqrt{x_1^2 + \cdots + x_n^2}$
$\lVert A \rVert_{\mathrm{F}}$	Frobenius (Hilbert-Schmidt) norm of the matrix $A \in \mathbb{R}^{d\times d}$
$\lVert \mu \rVert_{\mathrm{TV}}$	total variation norm of the finite (signed) measure μ, i.e., $\lVert \mu \rVert_{\mathrm{TV}} = \sup_D \lvert \mu(D) \rvert$
$\mathrm{d_H}(\sigma^{(1)}, \sigma^{(2)})$	Hamming distance,
$\mathrm{d_L}(\alpha^{(1)}, \alpha^{(2)})$	lexicographic distance, (0.3)
$R(\sigma^{(1)}, \sigma^{(2)})$	overlap matrix, (3.3)
$Q(\alpha^{(1)}, \alpha^{(2)})$	ultrametric overlap matrix, (3.14)
$q(i,j)$	limiting ultrametric overlap of $i,j \in \mathbb{N}$, see (2.26)
$S(\nu \mid \mu)$	relative entropy, (1.6)
$I(\cdot)$	Cramèr Entropy, (6.4)
$I^*(\cdot)$	large deviations (good) rate function, (3.41)
p_Λ	free energy
$\mathscr{P}(\cdot)$	Parisi functional, (5.10)
$\mathscr{C}\mathscr{S}(\cdot)$	Crisanti-Sommers functional, (5.14)

Processes, random variables, measures

μ	a priori measure on Σ
μ_Λ (or μ_N)	a priori product measure on Σ_Λ (or Σ_N)
$\mathscr{G}_\Lambda$ (or $\mathscr{G}_N$)	finite volume Gibbs measure on Σ_Λ (or Σ_N), (1.1)
$\mathscr{G}_0(\beta)$	infinite-volume (DLR) Gibbs measure
$\mathscr{G}_\Lambda(\beta, \sigma^{(c)})$	Gibbsian local specification on Σ_Λ, given the external condition $\sigma^{(c)} \in \Lambda_0 \setminus \Lambda$, (1.12)
$\mathbb{P}$	probability on the space of disorder
$U(D)$	uniform distribution on the set $D \subset \mathbb{R}^d$
$\mathrm{REM}(\cdot)$	Derrida's REM process, (1.24)
$\mathrm{GREM}(\cdot)$	Derrida's GREM process, (0.5)
$\mathrm{SK}(\cdot)$	Sherrington-Kirkpatrick process, (1.20)
$\mathrm{PPP}(D \ni x \mapsto f(x) \in \mathbb{R})$	Poisson point process with the density $f : D \subset \mathbb{R}^d \to \mathbb{R}$
$\mathrm{PD}(x,a)$	Poisson-Dirichlet point process
ξ	Ruelle's probability cascade, (2.22)
$\mathscr{E}_N$	point process of the GREM energy levels, (6.15)
$H_\Lambda(A,h;\sigma)$	spin glass Hamiltonian, (1.14)

Abbreviations

i.i.d.	independent identically distributed
k-D	k-dimensional
LLN	law of large numbers
LDP	large deviations principle
CLT	central limit theorem
PDE	partial differential equation

BSDE	backward stochastic differential equation
HJB equation	Hamilton-Jacobi-Bellman equation
càdlàg	continue à droite, limite à gauche (right continuous with left limits)
RS	replica symmetric
RSB	replica symmetry breaking
FRSB	full replica symmetry breaking
RPC	Ruelle’s probability cascade
REM	random energy model
GREM	generalised random energy model
CREM	continuous GREM
SK model	Sherrington-Kirkpatrick model
EA model	Edwards-Anderson model
RKKY model	Ruderman-Kittel-Kasuya-Yoshida model
AS^2 scheme	Aizenman-Sims-Starr scheme

References

ADLER, ROBERT J. 1990. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 12. Hayward, CA: Institute of Mathematical Statistics.

ADLER, ROBERT J. & TAYLOR, JONATHAN E. 2007. *Random fields and geometry*. Springer Monographs in Mathematics. New York: Springer.

AIZENMAN, MICHAEL; SIMS, ROBERT & STARR, SHANNON L. 2003. An Extended Variational Principle for the SK Spin-Glass Model. *Phys. Rev. B*, **68**, 214403.

AIZENMAN, MICHAEL; SIMS, ROBERT & STARR, SHANNON L. 2007. Mean-field spin glass models from the cavity-ROSt perspective. *Pages 1–30 of: Prospects in mathematical physics*. Contemp. Math., vol. 437. Providence, RI: Amer. Math. Soc. arXiv:math-ph/0607060.

AMIT, DANIEL J. 1989. *Modeling brain function*. Cambridge: Cambridge University Press. The world of attractor neural networks.

ARBIB, MICHAEL A. (ed). 2003. *The handbook of brain theory and neural networks*. Second edn. A Bradford Book. Cambridge, MA: MIT Press.

ARGUIN, LOUIS-PIERRE. 2007. Spin glass computations and Ruelle's probability cascades. *J. Stat. Phys.*, **126**(4-5), 951–976. arXiv:math-ph/0608045v1.

ARRATIA, RICHARD; BARBOUR, ANDREW D. & TAVARÉ, SIMON. 2003. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich.

BASDEVANT, ANNE-LAURE. 2006. Ruelle's probability cascades seen as a fragmentation process. *Markov Process. Related Fields*, **12**(3), 447–474.

BAUKE, HEIKO & MERTENS, STEPHAN. 2004. Universality in the level statistics of disordered systems. *Phys. Rev. E*, **70**, 025102(R).

BEN AROUS, GÉRARD; DEMBO, AMIR & GUIONNET, ALICE. 2001. Aging of spherical spin glasses. *Probab. Theory Related Fields*, **120**(1), 1–67.

BEN AROUS, GÉRARD; BOGACHEV, LEONID V. & MOLCHANOV, STANISLAV A. 2005. Limit theorems for sums of random exponentials. *Probab. Theory Related Fields*, **132**(4), 579–612.

BEN AROUS, GÉRARD; GAYRARD, VERONIQUE & KUPTSOV, ALEXEY. 2006. A new REM conjecture. *Preprint*. arXiv:0706.2135v1 [math.PR].

BERTOIN, JEAN. 2006. *Random fragmentation and coagulation processes*. Cambridge Studies in Advanced Mathematics, vol. 102. Cambridge: Cambridge University Press.

BERTOIN, JEAN & LE GALL, JEAN-FRANÇOIS. 2000. The Bolthausen-Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields*, **117**(2), 249–266.

BOGACHEV, VLADIMIR I. 1998. *Gaussian measures*. Mathematical Surveys and Monographs, vol. 62. Providence, RI: American Mathematical Society.

BOLTHAUSEN, ERWIN. 2007. Random media and spin glasses: an introduction into some mathematical results and problems. *Pages 1–44 of: Spin glasses*. Lecture Notes in Math., vol. 1900. Berlin: Springer.

BOLTHAUSEN, ERWIN & KISTLER, NICOLA. 2006. On a nonhierarchical version of the generalized random energy model. *Ann. of Appl. Prob.*, **16**, 1.

BOLTHAUSEN, ERWIN & KISTLER, NICOLA. 2008. On a nonhierarchical version of the Generalized Random Energy Model. II. Ultrametricity. *Preprint*.

BOLTHAUSEN, ERWIN & SZNITMAN, ALAIN-SOL. 1998. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, **197**(2), 247–276.

BOLTHAUSEN, ERWIN & SZNITMAN, ALAIN-SOL. 2002. *Ten Lectures on Random Media*. Birkhäuser Verlag.

BOLTHAUSEN, ERWIN (ED.) & BOVIER, ANTON (ED.). 2007. *Spin glasses*. Lecture Notes in Mathematics 1900. Berlin: Springer. ix, 179 p. .

BORGS, CHRISTIAN; CHAYES, JENNIFER & PITTEL, BORIS. 2001. Phase transition and finite-size scaling for the integer partitioning problem. *Random Structures Algorithms*, **19**(3-4), 247–288. Analysis of algorithms (Krynica Morska, 2000).

BORGS, CHRISTIAN; CHAYES, JENNIFER; MERTENS, STEPHAN & NAIR, CHANDRA. 2005a. Proof of the local REM conjecture for number partitioning I: Constant energy scales. *Preprint*. arXiv:cond-mat/0501760v2 [cond-mat.dis-nn].

BORGS, CHRISTIAN; CHAYES, JENNIFER; MERTENS, STEPHAN & NAIR, CHANDRA. 2005b. Proof of the local REM conjecture for number partitioning II: growing energy scales. *Preprint*. arXiv:cond-mat/0508600v1 [cond-mat.dis-nn].

BOUCHAUD, JEAN-PHILIPPE & POTTERS, MARC. 2003. *Theory of Financial Risk and Derivative Pricing*. Cambridge: Cambridge University Press. From statistical physics to risk management.

BOVIER, ANTON. 2001. *Statistical mechanics of disordered systems*. MaPhySto Lecture Notes, no. 10. Aarhus University.

BOVIER, ANTON. 2006. *Statistical mechanics of disordered systems*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press. A mathematical perspective.

BOVIER, ANTON & KURKOVA, IRINA. 2003a. Gibbs measures of Derrida's generalized random energy models and the genealogy of Neveu's continuous state branchng process. *Preprint*.

BOVIER, ANTON & KURKOVA, IRINA. 2003b. Rigorous results on some simple spin glass models. *Markov Process. Related Fields*, **9**(2), 209–242. Inhomogeneous random systems (Cergy-Pontoise, 2002).

BOVIER, ANTON & KURKOVA, IRINA. 2004a. Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, **40**(4), 439–480.

BOVIER, ANTON & KURKOVA, IRINA. 2004b. Derrida's generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.*, **40**(4), 481–495.

BOVIER, ANTON & KURKOVA, IRINA. 2004c. Poisson convergence in the restricted k-partitioning problem. *WIAS preprint 964*. arXiv:cond-mat/0504366v1 [cond-mat.dis-nn], to appear in Random Structures & Algorithms.

BOVIER, ANTON & KURKOVA, IRINA. 2006a. Local energy statistics in disordered systems: a proof of the local REM conjecture. *Comm. Math. Phys.*, **263**(2), 513–533.

BOVIER, ANTON & KURKOVA, IRINA. 2006b. A tomography of the GREM: beyond the REM conjecture. *Comm. Math. Phys.*, **263**(2), 535–552.

BOVIER, ANTON & KURKOVA, IRINA. 2007. Much ado about Derrida's GREM. *Pages 81–115 of: Spin glasses*. Lecture Notes in Math., vol. 1900. Berlin: Springer.

BOVIER, ANTON; KURKOVA, IRINA & LÖWE, MATTHIAS. 2002. Fluctuations of the free energy in the REM and the p-spin SK models. *Ann. Probab.*, **30**(2), 605–651.

BRIAND, PHILIPPE & HU, YING. 2007. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields*, –. arXiv:math/0703423.

CARMONA, PHILIPPE & HU, YUEYUN. 2006. Universality in Sherrington-Kirkpatrick's spin glass model. *Ann. Inst. H. Poincaré Probab. Statist.*, **42**(2), 215–222.

CHALLET, DAMIEN; MARSILI, MATTEO & ZHANG, YI-CHENG. 2005. *Minority Games*. Oxford University Press.

CHATTERJEE, SOURAV & PAL, SOUMIK. 2007. A phase transition behavior for Brownian motions interacting through their ranks. *Preprint*. arXiv:0706.3558v1 [math.PR].

COMETS, FRANCIS. 1989. Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probab. Theory Related Fields*, **80**(3), 407–432.

CONTUCCI, PIERLUIGI & GRAFFI, SANDRO (eds). 2007. *How Can Mathematics Contribute to Social Sciences*. Quality and Quantity, vol. 41, no. 4.

CONTUCCI, PIERLUIGI; DEGLI ESPOSTI, MIRKO; GIARDINÀ, CRISTIAN & GRAFFI, SANDRO. 2003. Thermodynamical limit for correlated Gaussian random energy models. *Comm. Math. Phys.*, **236**(1), 55–63.

COOLEN, A. C. C. 2005. *The mathematical theory of minority games*. Oxford: Oxford University Press. Statistical mechanics of interacting agents.

CRISANTI, ANDREA & SOMMERS, HANS-JÜRGEN. 1992. The spherical p-spin interaction spin glass model: the statics. *Zeitschrift für Physik B Condensed Matter*, **87**(3), 341–354.

DA LIO, FRANCESCA & LEY, OLIVIER. 2006. Uniqueness Results for Second-Order Bellman–Isaacs Equations under Quadratic Growth Assumptions and Applications. *SIAM J. Control Optim.*, **45**(1), 74–106.

DALEY, DARYL J. & VERE-JONES, DAVID. 2003. *An introduction to the theory of point processes. Vol. I*. Second edn. Probability and its Applications (New York). New York: Springer-Verlag. Elementary theory and methods.

DEN HOLLANDER, FRANK. 2000. *Large deviations*. Fields Institute Monographs, vol. 14. Providence, RI: American Mathematical Society.

DERRIDA, BERNARD. 1980. Random-energy model: limit of a family of disordered models. *Phys. Rev. Lett.*, **45**(2), 79–82.

DERRIDA, BERNARD. 1981. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, **24**(5), 2613–2626.

DERRIDA, BERNARD. 1985. A generalization of the Random Energy Model which includes correlations between energies. *Journal de physique. Lettres*, **46**(9), 401–407.

DERRIDA, BERNARD & GARDNER, ELISABETH. 1986. Magnetic properties and the function $q(x)$ of the generalised random-energy model. *Journal of Physics C: Solid State Physics*, **19**(29), 5783–5798.

DORLAS, TEUNIS C. & WEDAGEDERA, JANAK R. 2001. Large deviations and the random energy model. *Int. J. Mod. Phys. B*, **15**, 1–15.

DOTSENKO, VIKTOR. 2001. *Introduction to the Replica Theory of Disordered Statistical Systems*. Cambridge University Press.

DUDLEY, RICHARD M. 1999. *Uniform central limit theorems*. Cambridge Studies in Advanced Mathematics, vol. 63. Cambridge: Cambridge University Press.

DURLAUF, STEVEN N. 1999. How can statistical mechanics contribute to social science? *Proc. Natl. Acad. Sci. USA*, **96**(19), 10582–10584 (electronic).

EDWARDS, S.F. & ANDERSON, P.W. 1975. Theory of spin glasses. *Journal of Physics F: Metal Physics*, **5**(5), 965–974.

EISELE, THEODOR. 1983. On a third-order phase transition. *Comm. Math. Phys.*, **90**(1), 125–159.

ELLIS, RICHARD S. 2006. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Berlin: Springer-Verlag. Reprint of the 1985 original.

EVANS, LAWRENCE C. 1998. *Partial differential equations*. Graduate Studies in Mathematics, vol. 19. Providence, RI: American Mathematical Society.

FERNIQUE, XAVIER. 1997. *Fonctions aléatoires gaussiennes, vecteurs aléatoires gaussiens*. Montreal, QC: Université de Montréal Centre de Recherches Mathématiques.

FISCHER, K.H. & HERTZ, J.A. 1991. *Spin Glasses*. Cambridge University Press.

FRÖHLICH, JÜRG & ZEGARLIŃSKI, BOGUSŁAW. 1987. Some comments on the Sherrington-Kirkpatrick model of spin glasses. *Comm. Math. Phys.*, **112**(4), 553–566.

GRIFFITHS, ROBERT B. 1964. A proof that the free energy of a spin system is extensive. *J. Mathematical Phys.*, **5**, 1215–1222.

GUERRA, FRANCESCO. 2003. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, **233**(1), 1–12.

GUERRA, FRANCESCO. 2005. *Spin Glasses*. Preprint. arXiv:cond-mat/0507581v1.

GUERRA, FRANCESCO & TONINELLI, FABIO LUCIO. 2002. The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.*, **230**(1), 71–79.

GUERRA, FRANCESCO & TONINELLI, FABIO LUCIO. 2003. The infinite volume limit in generalized mean field disordered models. *Markov Process. Related Fields*, **9**(2), 195–207.

HARTMANN, ALEXANDER K. & WEIGT, MARTIN. 2005. *Phase transitions in combinatorial optimization problems*. Wiley-VCH Verlag, Weinheim. Basics, algorithms and statistical mechanics.

JANA, NABIN KUMAR. 2007. *Contributions to Random Energy Models*. Ph.D. thesis, Indian Statistical Institute.

JANA, NABIN KUMAR & RAO, B. V. 2006. Large Deviations and Random Energy Models. *Preprint*. arXiv:math/0612836v1 [math.PR].

JOAG-DEV, KUMAR; PERLMAN, MICHAEL D. & PITT, LOREN D. 1983. Association of normal random variables and Slepian's inequality. *Ann. Probab.*, **11**(2), 451–455.

KAHANE, JEAN-PIERRE. 1986. Une inégalité du type de Slepian et Gordon sur les processus gaussiens. *Israel J. Math.*, **55**(1), 109–110.

KALLENBERG, OLAV. 1983. *Random measures*. Third edn. Berlin: Akademie-Verlag.

KLIMOVSKY, ANTON. 2005. Learning and generalization errors for the 2D binary perceptron. *Math. Comput. Modelling*, **42**(11-12), 1339–1358.

LEADBETTER, M. R.; LINDGREN, GEORG & ROOTZÉN, HOLGER. 1983. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. New York: Springer-Verlag.

LEDOUX, MICHEL. 2001. *The concentration of measure phenomenon*. Mathematical Surveys and Monographs, vol. 89. Providence, RI: American Mathematical Society.

LEDOUX, MICHEL & TALAGRAND, MICHEL. 1991. *Probability in Banach spaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23. Berlin: Springer-Verlag. Isoperimetry and processes.

LIFSHITS, M. A. 1995. *Gaussian random functions*. Mathematics and its Applications, vol. 322. Dordrecht: Kluwer Academic Publishers.

MÉZARD, MARC & MONTANARI, ANDREA. 2007. *Constraint satisfaction networks in Physics and Computation*. Oxford University Press. To appear.

MÉZARD, MARC; PARISI, GIORGIO & VIRASORO, MIGUEL ANGEL. 1987. *Spin glass theory and beyond*. World Scientific Lecture Notes in Physics, vol. 9. Teaneck, NJ: World Scientific Publishing Co. Inc.

NEWMAN, CHARLES M. & STEIN, DANIEL L. 2007. Short-range spin glasses: results and speculations. *Pages 159–178 of: Spin glasses*. Lecture Notes in Math., vol. 1900. Berlin: Springer.

NISHIMORI, HIDETOSHI. 2001. *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford University Press, USA.

OLIVIERI, ENZO & PICCO, PIERRE. 1984. On the existence of thermodynamics for the random energy model. *Comm. Math. Phys.*, **96**(1), 125–144.

OPPER, MANFRED & SAAD, DAVID (eds). 2001. *Advanced mean field methods*. Neural Information Processing Series. Cambridge, MA: MIT Press. Theory and practice, Papers from the workshop held at Aston University, Birmingham, 1999, A Bradford Book.

PANCHENKO, DMITRY. 2005a. A question about the Parisi functional. *Electron. Commun. Probab.*, **10**, 155–166.

PANCHENKO, DMITRY. 2005b. Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.*, **17**(7), 793–857.

PANCHENKO, DMITRY & TALAGRAND, MICHEL. 2007a. Guerra's interpolation using Derrida-Ruelle cascades. *Preprint*. arXiv:0708.3641v2 [math.PR].

PANCHENKO, DMITRY & TALAGRAND, MICHEL. 2007b. On the overlap in the multiple spherical SK models. *Ann. Probab.*, **35**(6), 2321–2355.

PARISI, GIORGIO. 2007. Mean field theory of spin glasses: statics and dynamics. *In: Proceedings of the 2006 Les Houches Summer School*.

PARISI, GIORGIO & TALAGRAND, MICHEL. 2004. On the distribution of the overlaps at given disorder. *C. R. Math. Acad. Sci. Paris*, **339**(4), 303–306.

PASTUR, LEONID & SHCHERBINA, MARIYA. 1991. Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model. *J. Statist. Phys.*, **62**(1-2), 1–19.

PITMAN, JIM. 2006. *Combinatorial stochastic processes*. Lecture Notes in Mathematics, vol. 1875. Berlin: Springer-Verlag. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.

PITMAN, JIM & YOR, MARC. 1997. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.*, **25**(2), 855–900.

PRESTON, CHRIS. 1976. *Random fields*. Berlin: Springer-Verlag. Lecture Notes in Mathematics, Vol. 534.

RESNICK, SIDNEY I. 1987. *Extreme values, regular variation, and point processes*. Applied Probability. A Series of the Applied Probability Trust, vol. 4. New York: Springer-Verlag.

RUELLE, DAVID. 1987. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, **108**(2), 225–239.

RUZMAIKINA, A. & AIZENMAN, M. 2005. Characterization of invariant measures at the leading edge for competing particle systems. *Ann. Probab.*, **33**(1), 82–113.

SHCHERBINA, MARIYA. 2005. Some mathematical problems of neural networks theory. *Pages 425–444 of: European Congress of Mathematics*. Eur. Math. Soc., Zürich.

SHCHERBINA, MARIYA & TIROZZI, BRUNELLO. 2002. Generalization and Learning Error for Nonlinear Perceptron. *Mathematical and Computer Modelling*, **35**, 259–271.

SHERRINGTON, DAVID. 2007. Spin glasses: a perspective. *Pages 45–62 of: Spin glasses*. Lecture Notes in Math., vol. 1900. Berlin: Springer.

SHERRINGTON, DAVID & KIRKPATRICK, SCOTT. 1975. Solvable Model of a Spin-Glass. *Physical Review Letters*, **35**(26), 1792–1796.

SIMON, BARRY. 1993. *The statistical mechanics of lattice gases. Vol. I*. Princeton Series in Physics. Princeton, NJ: Princeton University Press.

SLEPIAN, DAVID. 1962. The one-sided barrier problem for Gaussian noise. *Bell System Tech. J.*, **41**, 463–501.

TALAGRAND, MICHEL. 2000a. Large deviation principles and generalized Sherrington-Kirkpatrick models. *Ann. Fac. Sci. Toulouse Math. (6)*, **9**(2), 203–244.

TALAGRAND, MICHEL. 2000b. Rigorous low-temperature results for the mean field p-spins interaction model. *Probab. Theory Related Fields*, **117**(3), 303–360.

TALAGRAND, MICHEL. 2003. *Spin glasses: a challenge for mathematicians*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 46. Berlin: Springer-Verlag. Cavity and mean field models.

TALAGRAND, MICHEL. 2006a. Free energy of the spherical mean field model. *Probab. Theory Related Fields*, **134**(3), 339–382.

TALAGRAND, MICHEL. 2006b. The Parisi formula. *Ann. of Math. (2)*, **163**(1), 221–263.

TALAGRAND, MICHEL. 2006c. Parisi measures. *J. Funct. Anal.*, **231**(2), 269–286.

TALAGRAND, MICHEL. 2007a. Large Deviations, Guerra's and A.S.S. Schemes, and the Parisi Hypothesis. *Journal of Statistical Physics*, **126**(4), 837–894.

TALAGRAND, MICHEL. 2007b. Mean field models for spin glasses: some obnoxious problems. *Pages 63–80 of: Spin glasses*. Lecture Notes in Math., vol. 1900. Berlin: Springer.

TEH, YEE WHYE; JORDAN, MICHAEL I.; BEAL, MATTHEW J. & BLEI, DAVID M. 2006. Hierarchical Dirichlet processes. *J. Amer. Statist. Assoc.*, **101**(476), 1566–1581.

Toubol, Alain. 1998. High temperature regime for a multidimensional Sherrington-Kirkpatrick model of spin glass. *Probab. Theory Related Fields*, **110**(4), 497–534.

Winkler, Gerhard. 2003. *Image analysis, random fields and Markov chain Monte Carlo methods*. Second edn. Applications of Mathematics (New York), vol. 27. Berlin: Springer-Verlag. A mathematical introduction, Stochastic Modelling and Applied Probability.

Zegarliński, Bogusław. 1998. Random spin systems with long-range interactions. *Pages 289–320 of: Mathematical aspects of spin glasses and neural networks*. Progr. Probab., vol. 41. Boston, MA: Birkhäuser Boston.