Alma Mater Studiorum · Università di Bologna

DOTTORATO DI RICERCA IN MATEMATICA

XXVII Ciclo

Settore Concorsuale di afferenza: 01/A4

Settore Scientifico disciplinare: MAT/07

RIGOROUS RESULTS IN SPIN GLASSES AND MONOMER DIMER SYSTEMS

Presentata da: EMANUELE MINGIONE

Coordinatore Dottorato:
Chiar.ma Prof.ssa
GIOVANNA
CITTI

Relatore:
Chiar.mo Prof.
PIERLUIGI
CONTUCCI

Esame finale anno 2015

Contents

Acknowledgements

The amazing experience of these past three years, and in particular this thesis, are based essentially on the fact that I have been allowed to work and live in a beautiful atmosphere. I wish to express my deep gratitude to all of the people that shared and helped me during this path.

In particular, to my Parents for their constant support and for giving me the opportunity to follow this path.

Thank you Pierluigi for having shown me how to follow this path in the best way, with precious suggestions, amazing help and guidance and for all the things I learned from you in the conversations we had along the way.

I wish also to thanks Francesco Guerra for the continuous and invaluable help. He is a source of inspiration for many reasons.

Among the other scientific colleagues that I interacted with during these three years is want to thanks in particular Diego, for the precious collaboration. There are several other people I interacted with and who should be truly acknowledged: Adriano Barra, Shannon Starr, Cristian Giardinà, Claudio Giberti, Dmitry Panchenko, Wei-Kuo Chen, Daniele Tantari, Giorgio Parisi.

I would also to acknowledge the financial support of the Department of Mathematics of Università di Bologna.

Finally and most important, I wish to say thanks to Silvia , for his encouragement, patience and love. This work is dedicated to her.

Introduction

This thesis deals with the mathematical properties of a class of statistical mechanics models describing the thermodynamic behaviour of the *Disordered Spin and Monomer-Dimer Systems*. The rigorous results presented here concern the properties of the so called Boltzmann-Gibbs probability measure in the (thermodynamic) limit of infinitely many particles. While for finitely many particles all the expectations (moments) of the physical quantities display a smooth dependence with respect to the external parameters (temperature, magnetic fields, interaction strength) it is known that in the thermodynamic limit singularities may appear and are related to the phase transitions observed experimentally.

The statistical mechanics of disordered models in particular describes physical systems where the interactions between the components are inhomogeneous and can only be modelled statistically, namely considering *random interactions* with suitable distribution. The Boltzmann-Gibbs measure which describes them, the so called *quenched* state, is constructed by first fixing a realisation of the disorder, computing all the quantities in the standard setting, and then averaging in the disorder.

Spin systems and monomer-dimer systems are very different types of particles in physics. The first interact with external magnetic fields and mutual magnetic forces while the second interact mostly via the repulsive part of van der Waals force. This last in particular displays a divergence at small distances and prevent particles to get close to each other. From the mathematical point of view that is called *hard-core* interaction and is embedded in the statistical

mechanics formalism not through the classical potential term but rather by the algebraic restriction of the admissible configurations in the partition function.

The thesis, naturally split in two parts for each type of particle system, shows nevertheless that in spite of the differences among the two types of models, similar techniques can be used to study them and, most importantly, rigorous proofs and sometimes exact solutions can be obtained for both of them.

In the first part of the work we will focus the attention on a class of Disordered Spin Models, called *Spin Glasses*. They are the first and most intensively studied disordered spin models and, in the last decades, have drawn a lot of attention from the scientific community for two main reasons. First, they exhibits in the thermodynamical limit a complex behavior which is able to describe various phenomenon belonging to many areas such as condensed matter, biology, computer science, economics, etc. [11, 49, 64, 101, 15, 18]. On other hand, in the recent times, remarkable progress has been made in the rigorous mathematical of Mean Field Spin Glasses, in particular on the *Sherringhton-Kirkpatrick model*. Let us briefly recall the state-of-the-art.

In the field of Statistical Mechanics, Mean Field models are often introduced as a "solvable" version of the corresponding finite dimensional model. While the latter encodes information on the topological structure and the physical dimension of the system, the main feature of the mean field interaction is the absence of an underlying spatial structure: all the components of the system (*resp.* disordered system) interact with each other with same strength (*resp.* in distribution). One can say that the model is invariant under permutation (*resp.* in distribution) of the elementary objects on the system. However, for Mean Field Spin Glasses, something unexpected happened. Indeed, it was soon realized that the Sherringhton-Kirkpatrick model, introduced in 1975 [93] as a mean field version of its finite dimensional counterpart, the Edward-Anderson model [31], is far from being "easy" to solve.

In 1980's Parisi [110] proposed a solution which required the introduction

of completely new concepts and revealed a rich mathematical structure. Informally, the main content of the so called *Parisi theory* [15], can be split in four, strictly related points :

i) Parisi Formula: the pressure can be represented as a variational problem over the space of all the distribution functions on $[0, 1]$ whose elements are called order parameter.

ii) Replica Symmetry Breaking for low temperature the order parameter solution of the Parisi Formula is non trivial (in the sense of distributions).

iii) The order parameter which solve the Parisi formula represents the distribution of the *overlap* w.r.t. the limiting quenched measure.

iv) Ultrametricity. The support of the joint distribution of the overlaps is concentrated on an ultrametric space.

From the physical point of view parts *ii), iii), iv)* have the following interpretation: there exist a countable number of pure states organized in a hierarchical structure.

The rigorous proof of the *Parisi formula* for the pressure, was completed only thirty years later by Guerra [79] and Talagrand [119]. The result fully confirms points *i)* and *ii)* however, *iii)* and the ultrametric property *iv)* are in some sense hidden inside the proof but don't follow from it. These last two points are strictly related to the the Aizenmann-Contucci [22] and the Ghirlanda-Guerra identities [76]. Indeed, Panchenko [108] recently proved that these identities imply ultrametricity. This result combined with the Aizenmann-Simms-Starr scheme [24] provides a new and more clear proof of the Parisi Formula and as a byproduct, that *iii)* and *iv)* hold *generically* (the precise meaning of this term will be explained in chapter 4 and should be understood as a *for a slightly perturbed Hamiltonian*).

This allows us to claim that the Parisi Theory is *generically* correct but a natural question arises: is it also *universal*? Namely, can it describes, even partially and with suitable modifications, also other kind of disordered models?

The original contribution of the first part of this work is to show that the range of validity of the Parisi Theory, in particular its mathematical counterpart, is not confined to the SK model.

First, we consider a Multi-species version of the SK model called *MSK model*. This is in some sense a way to break the permutational invariance in distribution of the SK model, without losing the mean field nature, reducing it to a block permutation invariance. The model is introduced in [41] where the authors show in detail the construction of a multidimensional version of the Parisi Formula and, under some convexity assumptions, a proof of the so called *Guerra's bound* for the pressure. The proof of the reverse bound was completed by Panchenko in [109] with the same strategy used for the SK model, but revealing a new consequence of the Ghirlanda-Guerra identities and ultrametricity, the so called *sinchronization property*.

We stress the fact that this result is important for the mathematical point of view as well for the applications. Indeed, the solution is given in great generality, namely for any arbitrary integer number of blocks. Thus, this makes the model much more suitable for applications than the SK model and leaves an open door to possible extensions of the result that will be discussed in chapter 3.

The MSK model belongs to the class of Mean Field models, but what about the finite dimensional case? Is there some features of the Parisi theory that still holds for *Finite Dimensional Spin Glasses*?

The most debated point concerns the number of their equilibrium state(s). As pointed out by Newman and Stein [97], some conceptual difficulties arise even in the precise definition of a pure thermodynamic state for disordered systems. Some authors [95] believe that the main features of the mean field picture, like the existence of an infinite number of equilibrium states and their ultrametric structure, persist in the finite dimensional case, while others [71] argue, within the droplet picture, that below the critical temperature just two pure phases exist, connected by the global spin-flip symmetry. Both pictures are supported

by a mixture of numerical simulations and theoretical considerations.

Our contribution [63] is a proof of the Aizenmann-Contucci and the Ghirlanda Guerra identities for a large class of models including finite dimensional spin glasses. Thus, the limiting distribution of the *generalized overlap*, by using Panchenko's Theorem, is *generically ultrametric*, where the meaning of generically is the same as before. We want to stress the fact this result is far to being a complete answer to the original question, namely if the Parisi picture holds also in the finite dimensional case: we have only proved *iv)*, however part *ii)* is the core of the question. Indeed, it's still possible that the equilibrium state satisfies the ultrametric condition *iv)* trivially (only with two pure states) excluding the *Replica Symmetry Breaking* behavior *ii)*.

In the second part of the work we will consider two generalizations of the classical Monomer-Dimer model: the first is with random activities while in the second, in a non-random setting, we introduce an imitative interaction between the components (monomers and dimers).

The Monomer-Dimer model (*MD model*) was originally introduced [54, 113] as statistical model for the absorption of diatomic molecules on a solid, thus roughly speaking, dimers represent the molecules and monomer empty sites, while the solid can be represented by a graph. Thus, by construction, a Monomer-Dimer Model is a matching problem on graphs [78].

In general, the main feature of any monomer-dimer system is the hard-core interaction among the dimers. The physical interpretation is that two different molecules cannot deposit on the same vertex, due to the repulsivity of the van der Waals potential at short distance.

The first mathematical approach to the problem is due to Heilmann and Lieb [86, 87] where they proved, among other results, that the hard-core interaction is not enough to generate a phase transition basically for any "nice" graph.

Clearly, the hard-core interaction is not the only physical property of the molecules-solid system, thus one can think to add other structures to the MD

model that fill this gap.

In this part of the work we will address to the following question: what happens if one adds to an MD model an attractive interaction or quenched randomness in the parameters?

The first describes the attractive component of the Van der Waals potential among monomers and dimers (a different kind of interaction is also considered in [88]). The original contribution of this part of the thesis is the solution and a detailed analysis [28] of a Monomer-Dimer model with an uniform attractive interaction (*IMD model*) on the complete graph. In particular we prove that the model belongs to the *mean field universality class*.

Beside the IMD model, we will consider a Monomer-Dimer model with random activities (*RMD model*) representing the absorption of an irregular solid described in the quenched setting. Our original contribution [30] is a solution of model on the complete graph with randomness in the monomer activities and uniform dimeric one. The solution, as usual for mean field models, has to be understood as a variational representation for the pressure that turn out to be analytic in agreement with the result of Heilmann-Lieb. The case of random dimeric weights is an important open question and is also related to the following general consideration.

The spirit of these generalizations of the MD model is very close to the Spin Glass setting. Indeed, one can think the IMD and the RMD models as a preliminary step toward a *Monomer-Dimer Glassy System*, in other words a model with a combination of random hard-core and random interaction between the components. From this point of view, the RMD model play the role of a random hard-core model while the IMD model is the simpler case of a deterministic attractive interaction.

A concluding remark: besides the purely mathematical aspect, the result presented here are interesting, in our opinion, at least for two reasons. First of all, a full understanding of the mathematical structure underlying the physical

behavior of the mean field systems seems to us to be a necessary prerequisite to attack the finite dimensional counterpart. The second reason is that, as we will briefly illustrate in the following, mean field models arise naturally in many different contexts, ranging from combinatorial optimization problems and theory of neural networks to social sciences and biology.

Layout of the thesis

This thesis is split into two main parts. The first (Chapter 1, 2, 3, 4) is on Spin Glass models while the second is on Monomer Dimer models (Chapter 5, 6, 7).

• In Chapter 1, we give the mathematical background and the general formalism for Spin (Disordered) Models. Some applications to physical and mathematical problems are briefly discussed.

• In Chapter 2, we give a very short overview on some general aspects of the theory of spin glasses, illustrating its physical origin. A section is dedicated to the Sherrington-Kirkpatrick model which is of fundamental interest for the work.

• In Chapter 3, we introduce the Multi-species Sherrington-Kirkpatrick model (MSK), we prove the existence of the thermodynamical limit and the Guerra's Bound for the quenched pressure together with a detailed analysis of the annealed and the replica symmetric regime. The result is a multidimensional generalization of the Parisi's theory. Finally we briefly illustrate the strategy of the Panchenko's proof of the lower bound.

• In Chapter 4 we discuss the Aizenmann-Contucci and the Ghirlanda-Guerra identities for a wide class of Spin Glass models. As an example of application, we discuss the role of these identities in the proof of the lower

bound for the MSK model.

• In Chapter 5 we introduce the basic mathematical formalism of Monomer Dimer models. We introduce a Gaussian representation of the partition function that will be fundamental in the rest of the work.

• In Chapter 6, we introduce an interacting Monomer-Dimer model (IMD). Its exact solution is derived and a detailed study of its analytical properties and related physical quantities is performed.

• In Chapter 7, we introduce a quenched randomness in the monomer Monomer Dimer model (RMD) and show that, under suitable conditions the pressure is a self averaging quantity. The main result is that, if we consider randomness only in the monomer activity, the model is exactly solvable.

Preliminaries

Let us start giving a quick overview on the general concepts behind the statistical mechanics description of the thermal equilibrium of a system. We want to stress the fact this preliminary introduction is formal and elementary, the only purpose is to give a motivation and orientation to the non expert reader then, a reader familiar with the subject can skip this section. A detailed and advanced exposition can be found in the classical literature, see for example [5, 1] for a mathematical perspective and [4] for a discussion closest to physics.

About 1870, Ludwig Boltzmann proposed that the laws of thermodynamics should be derivable from mechanical first principles on the basis of the atomistic theory of matter. Thermodynamics describes a physical system with few macroscopic parameters (for example a gas with pressure, volume and temperature) whereas, from the mechanical point of view, usually there is an huge number of atoms (typically 10^{23}) which interacts at the microscopical level. This contrast between the microscopic and the macroscopic level is the starting point of Equilibrium Statistical Mechanics as developed by Maxwell, Boltzmann, and Gibbs. The underlying basic ideas can be illustrated in a very general setting.

The aim is to describe and explains the macroscopic behavior of large systems in thermal equilibrium in terms of the microscopic interaction between their very many constituents. Thus, we are interested to the behavior of the objects when the number of the components goes to infinity. A rigorous definition of the this limiting procedure and convergence questions depends strongly on the model consider, at this stage we suppose the system large but finite and

refer this suitable procedure as *thermodynamic limit* or *t.l.*.

Consider a large but finite set Λ which represent the components of a system. Thus, a first vague notion of *t.l.* should be understood as the $|\Lambda| \to \infty$ in a suitable way. For each $i \in \Lambda$, let σ_i be a variable taking value in a set Σ (to fix the idea suppose a finite discrete set of real numbers) which describes the possible states of each component of the system, it is called *single state space*. For example, in the case of a magnetic system, Λ consists of the sites of the crystal lattice which is formed by the positions of the atoms and Σ is the set of all possible orientations of the magnetic moments of the atoms.

Having specified the sets Λ and Σ, we can describe a particular state of the system or *microscopic configuration* by a suitable element $\sigma = (\sigma_i)_{i\in\Lambda}$ of the product set $\Sigma^{|\Lambda|}$. We will denote by $\Sigma_\Lambda \subseteq \Sigma^{|\Lambda|}$, the set of allowed configurations, called *configuration space*.

Is reasonable to suppose that the physical properties of the system are consequence of the interaction among the components of the system. In the Statistical Mechanics language this interaction is specified by the *Hamiltonian*, a function $H_\Lambda : \Sigma_\Lambda \to \mathbb{R}$, which represent the energy associated to a *microscopic configuration* $\sigma \in \Sigma_\Lambda$.

The configuration space Σ_Λ is huge, nevertheless the complexity of a microscopic description of the system, i.e. in terms of a *microscopic configuration* σ, can be overcome by a probabilistic approach. In other words, the macroscopic determinism (thermodynamic) may be regarded as a consequence of a suitable law of large numbers. According to this philosophy, it is not adequate to describe the state of the system by a particular element $\sigma \in \Sigma_\Lambda$. The system's state should rather be described considering the family $\sigma = (\sigma_i)_{i\in\Lambda}$ as a family of Σ-valued random variables, or (if we pass to the joint distribution of these random variables) by a probability measure $\mathcal{G}_\Lambda$ on Σ_Λ. In this framework, physical quantities are supposed to be represented by the expectation w.r.t. the measure $\mathcal{G}_\Lambda$ of suitable functions on the configurations space.

$[Q_1]$. *Which kind of probability measure is suitable to describe a very large physical system in thermal equilibrium*?

Let us introduce the basic quantity on which is based the answer to $[Q_1]$.

Definition 0.0.1. *For a given Hamiltonian H_Λ and for any $\beta \geq 0$, a real parameter which represent up to a constant the inverse physical temperature, the finite volume Gibbs measure $\mathcal{G}_\Lambda$ is a probability measure on Σ_Λ defined as*

$$\mathcal{G}_\Lambda(\sigma) := \frac{e^{-\beta H_\Lambda(\sigma)}}{Z_\Lambda(\beta)} \tag{1}$$

for each $\sigma \in \Sigma_\Lambda$. The normalization factor

$$Z_\Lambda := \sum_{\sigma \in \Sigma_\Lambda} e^{-\beta H_\Lambda(\sigma)} \tag{2}$$

is called partition function.

An informal answer to the question $[Q_1]$ is the following:

$[A_1]$. *When the system becomes very large, i.e. in the t.l., the probability measure which describe the thermal equilibrium of a system is given by the t.l. of Gibbs measure.*

The rigorous formulation and justification of the previous answer is a long story which is still far from being finished, we just mention the key words and relative references: *ergodic theory* [1] and *equivalence of ensembles* [3]. In this work we will focus on the mathematical analysis of the consequences of $[A_1]$ rather than on its foundation.

Let us give an equivalent formulation of $[A_1]$, closest to thermodynamics. According to the previous probabilistic description, let μ be a probability measure on Σ_Λ which encode all the thermodynamic information of the system, not necessarily at the thermal equilibrium. For example, the *internal energy density* can be computed as

$$u_\Lambda(\mu) = \frac{1}{|\Lambda|}\mathbb{E}_\mu[H_\Lambda] \tag{3}$$

The previous relation is quite intuitive. The key step and less intuitive, is the relation between the physical *entropy* and μ first proposed by Boltzmann and then explained by Gibbs. Setting the *Boltzmann constant* equal to 1 the *entropy density* is given by

$$s_\Lambda(\mu) = -\mathbb{E}_\mu[\log \mu] \tag{4}$$

The *free energy density* is given by the usual thermodynamic relation

$$f_\Lambda(\mu) = u_\Lambda(\mu) - \frac{1}{\beta} s_\Lambda(\mu) \tag{5}$$

If we impose now that the system is in thermal equilibrium, one can characterize the Gibbs measure (1), in the following way:

Gibbs Variational Principle: when the system is at thermal equilibrium, the probability that each state occurs are such as to minimize the free energy of the system.

A simple computation show that the Gibbs measure (1) is the unique minimizer of $f_\Lambda(\mu)$ on the set of all probability measure on Σ_Λ. Thus the free energy at equilibrium becomes

$$f_\Lambda := \min_\mu f_\Lambda(\mu) = -\frac{1}{\beta|\Lambda|} \log Z_\Lambda \tag{6}$$

where Z_Λ is the partition function 5.9.

Instead of the free energy, without any significant change, in this work we will consider the following quantity:

Definition 0.0.2. *The pressure density associated to an Hamiltonian is defined as*

$$p_\Lambda := \frac{1}{|\Lambda|} \log Z_\Lambda = -\beta f_\Lambda \tag{7}$$

The "pressure" is a term used in Statistical Mechanics of Lattice Gases [6], because of its analogy to the real physical concept of pressure for lattice gases.

In the rest of the work we prefer deal with the pressure because the extra factor $-1/\beta$ in the free energy becomes a nuisance when taking derivatives.

The assumption $[A_1]$ is only the beginning of the game, since after it, a natural question arises:

$[Q_2]$. *It is possible to give a rigourous formulation of the thermodynamic limit of the Gibbs measures, study the properties of its limiting object and related quantities*?

There is a beautiful and rather general theory on the *t.l.* of the Gibbs measure, know as *Gibbs formalism* or *DLR states*. A detailed exposition on this subject can be found in the standard references [5, 6, 2]. The present work is on Mean Field models for which the previous theory is not directly applicable and the questions related to the *t.l.* will be analyzed case by case.

As one expect, the limiting properties of the Gibbs measure depend strongly on the considered model. This is the most exciting feature of Statistical Mechanics, because, starting from a collection of interacting simples objects, after *t.l.*, the resulting picture can be much less simple and describe complex phenomena like *phase transition* and *symmetry breaking* [5].

In this work, for a given statistical mechanical model, we will try answer to $[Q_2]$, in particular to the following aspects:

• Existence and properties of the thermodynamic limit of the pressure density.

• Existence and properties of the thermodynamic limit of the Gibbs measure.

We will consider models belonging to two distinct classes : *Spin Models* and *Monomer-Dimer Models*. As will become clear, this preliminary division, is due to the topological properties of their configuration spaces.

Chapter 1

Spin Models formalism

The first kind of statistical mechanical models studied in this work belong to the classes of *Disordered Spin Models on Graphs*. The aim of this chapter is to introduce the necessary mathematical background.

A statistical mechanical model can specified trough three basic objects:

• A finite set Λ which represent the components of the systems.

• The *single state space* Σ and the *configuration space* Σ_Λ which represent our a priori knowledge of the properties of the components.

• The *Hamiltonian* function $H_\Lambda : \Sigma_\Lambda \to \mathbb{R}$ which represent the energy of the system trough the interaction between the components.

All the formalism developed in this chapter is for finite system (finite Λ), the definition of *thermodynamic limit* and convergence questions are analysed case by case in the next chapters. A first vague notion of *t.l.* should be understood as the $|\Lambda| \to \infty$ in a suitable way.

First at all, in this work, except for chapter 4, we will consider interactions involving only pair of components excluding the possibility of many components interactions: in other words we deal with *Models on Graphs*. Thus, let us recall the following basic definitions of graph theory.

Definition 1.0.3. *Let Λ be a finite set, $P_\Lambda^{(2)} := \{ij \equiv \{i,j\} : i \neq j \in \Lambda\}$ the set of unordered couples of different elements of Λ and consider a subset $E \subseteq P_\Lambda^{(2)}$.*

The pair $(\Lambda, E) := G$ *is called finite graph with vertex set* Λ *and edge set* E.

If $E \equiv P_\Lambda^{(2)}$ *we denote the graph by* K_Λ *and refer to it as the complete graph on* Λ.

This works focuses on models on the complete graph however, they will be introduced, without extra efforts, for a general graph G. The symbol G, unless otherwise specified, always denotes a generic graph with vertex set Λ set and edge set E which represent respectively the set of microscopic components and the set of pair interacting components of a model. Even if is not treated in detail during the work, often in the discussion we will refer also to another kind of graph called *lattice graph*.

Definition 1.0.4. *Consider an integer* d, *we call the lattice graph* $G_d := (\Lambda_d, E)$ *the finite graph with vertex set* $\Lambda \subset \mathbb{Z}^d$ *and* $E \equiv \{ij \in P_\Lambda^{(2)} : |i-j|_1 = 1\}$ *where* $\mathbb{Z}^d$ *is the d-dimensional integer lattice and* $|\cdot|_1$ *denotes the* L_1 *norm in* $\mathbb{R}^d$.

The respective terminology for statistical mechanics models is:

• *Models on the complete graph*: the choice of K_Λ can be used to describe systems with an high number of connections such as neural or social networks. If no additional structures are present in Λ, the only important thing is its cardinality, i.e. the number of components of the system that will be denoted by an integer N. Thus, we simply take $\Lambda \equiv \Lambda_N := \{1, \dots, N\}$ and $E \equiv E_N := \{ij : i < j; i, j \in \Lambda_N\}$, K_N denotes the corresponding complete graph.

• *Lattice models*: the choice of G_d and in particular $\mathbb{Z}^d$ as underlying translation invariant structure is often used to describe physical models as crystals. The condition $|i-j|_1 = 1$ identifies the nearest neighborhood vertices and we say that the resulting interaction has a *short range*. In general one can choose a real parameter $R > 0$ which tunes the range of the interaction replacing the above condition with the weaker $|i-j|_1 \leq R$.

We mention that it's also possible to consider G as a *random graph*. The

simplest example is the *Erdös - Rényi* random graph: the edges are independently present with identical probability. They will not be discussed in this work, the interested reader may consult [67, 107, 96] for spin model on random graph and [27] for monomer dimer models on random graphs.

1.1 Spin Models on Graphs

Let $G = (\Lambda, E)$ be a finite graph. Each vertex $i \in \Lambda$ is endowed with a variable $\sigma_i \in \Sigma := \{-1, 1\}$ called *spin*.

Definition 1.1.1. *A spin configuration is the family $\sigma := \{\sigma_i\}_{i\in\Lambda}$ which represent a particular microscopic configuration of the system. The configuration space is the product set $\Sigma_\Lambda := \{-1, +1\}^{|\Lambda|}$, i.e. the set of all possible spin configurations.*

The Hamiltonian which describe the interaction among the components is then a spin-spin interaction.

Definition 1.1.2. *A Spin Model on $G = (\Lambda, E)$ is defined assigning to each spin configuration $\sigma \in \Sigma_\Lambda$ an Hamiltonian function*

$$H_G(\sigma) = -\sum_{ij\in E} J_{ij}\sigma_i\sigma_j - \sum_{i\in\Lambda} h_i\sigma_i \tag{1.1}$$

where $J_E := \{J_{ij}\}_{ij\in E}$, and $h_\Lambda := \{h_i\}_{i\in\Lambda}$ are two families of real parameters.

Remark 1.1.3. *The definition 1.1.2 is slightly redundant since, strictly speaking, the choice of the graph G and the family of couplings J_E are not independent. Indeed, since the value $J_{ij} = 0$ is allowed, one can consider without loss $G \equiv K_\Lambda$. It's easy to check that a spin model on an arbitrary graph $G = (\Lambda, E)$ with family of coupling J_E, coincide with a spin model on K_Λ with family of coupling J' defined as $J'_{ij} = J_{ij}$ if $ij \in E$ and zero otherwise. Despite this fact we will use the above definition because, as we will show in the next, the graph*

G encodes in a compact way the topological properties while the couplings J_E tune the strength of the interactions.

Example 1. *The following models will be briefly discussed in chapter 2:*

i) The nearest-neighborhood ferromagnetic Ising model in $\mathbb{Z}^d$ corresponds to the choice $G = G_d$, $J_{ij} = J > 0$ for all $ij \in E$ and $h_i = h \in \mathbb{R}$ for all $i \in \Lambda$

ii) The Curie-Weiss model corresponds to the choice $G = K_N$, $J_{ij} = J > 0$ for all $ij \in E$ and $h_i = h \in \mathbb{R}$ for all $i \in \Lambda$

All the properties of the Hamiltonian are then encoded in the choice of the graph G and the families J_E and h_Λ that are called respectively *couplings* and *external fields*. For any fixed such as choice, let us recall the basic objects of interest.

Definition 1.1.4. *Let $\beta \geq 0$ be a real parameter which represent up to a constant the inverse physical temperature, the finite volume Gibbs measure $\mathcal{G}_G$ associated to the Hamiltonian (1.1), is a probability measure on Σ_Λ defined as*

$$\mathcal{G}_G(\sigma) := \frac{e^{-\beta H_G(\sigma)}}{Z_G} \tag{1.2}$$

for each $\sigma \in \Sigma_\Lambda$. The normalization factor

$$Z_G := \sum_{\sigma \in \Sigma_\Lambda} e^{-\beta H_G(\sigma)} \tag{1.3}$$

is called partition function.

NOTATION WARNING: *Clearly the Hamiltonian, the Gibbs measure all the related quantities that we are going to introduce, depend on the fixed choice of the parameters J_E and h_Λ and the inverse temperature β. In order to lighten the notation, we keep this dependance implicit in almost all situation. However, the reader should always keep in mind this dependance.*

One of the most important quantity beside the Gibbs measure is

Definition 1.1.5. *The pressure density associated to an Hamiltonian (1.1) is defined as*

$$p_G := \frac{1}{|\Lambda|} \log Z_G \tag{1.4}$$

In probabilistic terms, the pressure density, up to a non relevant normalization, is the cumulant generating function of the variable $-H_G$ w.r.t. the counting measure on Σ_Λ, then its β-derivatives are the cumulants of $-H_G$ w.r.t. the Gibbs measure.

Definition 1.1.6. *We call the finite volume Gibbs state ω_Λ the expectation w.r.t. the Gibbs measure 1.2, i.e. for any bounded function $f : \Sigma_\Lambda \to \mathbb{R}$,*

$$\omega_G(f) := \frac{1}{Z_G} \sum_{\sigma \in \Sigma_\Lambda} e^{-\beta H_G(\sigma)} f(\sigma) \tag{1.5}$$

Thus, for example

Definition 1.1.7. *The internal energy density is defined as*

$$e_G := \frac{\omega_\Lambda(H_G)}{|\Lambda|} = -\frac{\partial}{\partial \beta} p_G \tag{1.6}$$

The ground state of the system is defined as

$$e_G^{(0)} := \frac{1}{|\Lambda|} \min_{\sigma \in \Sigma_\Lambda} H_G(\sigma) \tag{1.7}$$

The origin of the term *ground state* is due to the fact that $\beta \to \infty$ correspond to the zero temperature limit. Indeed one can prove that $e_G^{(0)} = \lim_{\beta \to \infty} e_G$. This relation furnishes a direct link with statistical mechanics and optimization problems.

1.2 Disordered Spin Models

At this stage, the definition (1.1.2) is rather general. However, depending on the phenomenon to describe, we can make additional assumption on the graph G and the parameters J_E and h_Λ to construct particular spin models that

capture the relevant properties of the phenomenon and allows a satisfactory mathematical treatment. We have already anticipated that the main topic of the work will be models on the complete graph, so let us introduce the basic assumption on the parameters.

Let us suppose we don't know exactly the values of the parameters J_E and h_Λ but only their statistical properties, in other words their distribution. It would be interesting to try to understand what happen in the *thermodynamic limit* for a "typical" choice of the parameters. One natural way to give this question some meaning is to model the parameters as random variables and translate the word *typical* into *almost all realization*. The physical motivations behind the introduction of randomness on the parameters will be explained in section 2.1. Here we introduce the mathematical framework.

Let $J_E := \{J_{ij}\}_{ij\in E}$ and $h_\Lambda := \{h_i\}_{i\in\Lambda}$ be two families of random variables defined in some *standard probability space* $(\Xi, \mathcal{B}, \mathbb{P})$ with expectation denoted by $\mathbb{E}$, $\eta \in \Xi$ denotes a particular realization of the disorder.

Definition 1.2.1. *A Disordered Spin Model on G is defined, for any realization η, assigning to each spin configuration $\sigma \in \Sigma_\Lambda$ an Hamiltonian function:*

$$H_{G,\eta}(\sigma) = -\sum_{ij\in E} J_{ij}(\eta)\sigma_i\sigma_j - \sum_{i\in\Lambda} h_i(\eta)\sigma_i \tag{1.8}$$

NOTATION WARNING: *In the next chapters, we will consider particular Disordered Spin Models specifying the distribution of the random parameters, then we will drop the explicit dependence on η.*

Since the Hamiltonian (1.8) is a random variable, all the quantities related to it include randomness. Thus, the description of the equilibrium properties of a disordered system with a random Hamiltonian requires the introduction of a new notion of equilibrium state. For a general discussion on this topic see [12] and references therein. For our purpose, we can use the notion of finite volume *quenched state*.

Recall that, from the dynamical point of view [1], the use of the Gibbs measure as equilibrium measure is based on the assumption that time averages will converge to averages with respect to the Gibbs measures. In this case, we deal two with kind of variables, the spins $\sigma \in \Sigma_\Lambda$ and the random variables in $(\Xi, \mathcal{B}, \mathbb{P})$. Thus a proper notion of equilibrium measure must be a measure on the product space $\Sigma_\Lambda \times \Xi$ taking into account the fact the equilibrium time-scales of the spins and the random couplings can be different.

There are two notions conventionally used in the physics, which correspond to opposite cases:

• *Annealed state*. This procedure corresponds to treating the random variables as dynamical variables on an equal footing with the spin variables, namely they equilibrate on the same time-scale. More precisely, we consider the pair (σ, η) as a "generalized spin configuration" on $\Sigma_\Lambda \times \Xi$ and apply the standard notion of Gibbs measure 1.2 to the a priori product measure $\mathbb{P}(d\eta)\mu_\Lambda(d\sigma)$, where μ_Λ is the discrete measure on Σ_Λ, i.e. the sum over all spin configurations.

Much more interesting from the physical point of view is the opposite case, which is the case considered from now on:

• *Quenched state*. This procedures correspond to consider the time scale of the spin variables relaxation much shorter than the time scale of the random variables. Thus, we split the average into two steps, first one computes for a fixed realization of the disorder η the average of the spins w.r.t. a Gibbs measure $\mathcal{G}_{G,\eta}$ on Σ_Λ, and then performs an average over the disorder η with $\mathbb{P}$. More precisely,

Definition 1.2.2. *For a given realization η and for any $\beta \geq 0$, the finite volume random Gibbs measure $\mathcal{G}_{G,\eta}$ associated to the random Hamiltonian $H_{G,\eta}$ (1.8) is a random probability measure on Σ_Λ defined as*

$$\mathcal{G}_{G,\eta}(\sigma) := \frac{e^{-\beta H_{G,\eta}(\sigma)}}{Z_{G,\eta}} \tag{1.9}$$

for each $\sigma \in \Sigma_\Lambda$. *The normalization factor*

$$Z_{G,\eta} := \sum_{\sigma \in \Sigma_\Lambda} e^{-\beta H_{G,\eta}(\sigma)} \tag{1.10}$$

is a random variable called random partition function.

Definition 1.2.3. *We call the finite volume random Gibbs state* $\omega_{\Lambda,\eta}$ *the expectation w.r.t. the random Gibbs measure 5.8. In particular, for any a bounded function* $f : \Sigma_\Lambda \to \mathbb{R}$

$$\omega_{G,\eta}(f) := \sum_{\sigma \in \Sigma_\Lambda} \mathcal{G}_{G,\eta}(\sigma) f(\sigma) \tag{1.11}$$

Notice that in general one can consider the average (1.11) for measurable functions $f(\sigma, \eta)$ on $\Sigma_\Lambda \times \Xi$. However, functions which depends only on the spin configurations will play an important role in the next chapters, hence we restrict our attention to them introducing the following important generalization of definition 1.2.2.

Definition 1.2.4. *For any of* $n \in \mathbb{N}$ *let us consider* Σ_Λ^n, *the cartesian product of* n *copies of* Σ_Λ *called replicas. We call the finite volume replica-Gibbs state* $\Omega_{G,\eta}$ *the expectation w.r.t. to* $\mathcal{G}_{G,\eta}^{\otimes n}$ *on* Σ_Λ^n, *i.e. for any bounded function* $f : \Sigma_\Lambda^n \to \mathbb{R}$

$$\Omega_{G,\eta}(f) := \sum_{(\sigma^1, \dots, \sigma^n \in \Sigma_\Lambda^n)} \mathcal{G}_{\Lambda,\eta}(\sigma^1) \dots \mathcal{G}_{G,\eta}(\sigma^n) f(\sigma^1, \dots, \sigma^n) \tag{1.12}$$

Clearly the specification of the number n *is redundant since one can always consider the previous product of an infinite number of copies without affect the average of a function on* Σ_Λ^n.

Finally, averaging over the disorder, one obtains the *quenched Gibbs state*, denoted

$$\langle f \rangle_G := \mathbb{E}\, \Omega_{G,\eta}(f) \tag{1.13}$$

Definition 1.2.5. *The random pressure density associated to an random Hamiltonian (1.8) is the random variable* $|\Lambda|^{-1} \log Z_{G,\eta}$. *Averaging over the disorder we obtain the quenched pressure density is*

$$p_G := \frac{1}{|\Lambda|} \mathbb{E} \log Z_{G,\eta} \tag{1.14}$$

The annealed pressure density

$$p_G^A := \frac{1}{|\Lambda|} \log \mathbb{E}\, Z_{G,\eta} \tag{1.15}$$

For random quantities as the random pressure defined 1.2.5, an important property is the following.

Definition 1.2.6. *Let* $(X_N)_{n\in\mathbb{N}}$ *be a sequence of random variables. We say that it is self-averaging iff*

$$\lim_{N\to\infty} \mathbb{E}\,(X_N - \mathbb{E}\, X_N)^2 = 0 \tag{1.16}$$

Clearly the limit $N \to \infty$ should be to be understood as the *t.l.*, to fix the idea simply think $N \equiv |\Lambda|$. For all the random models considered in this work we are able to prove that the random pressure density is self averaging, and then converge in probability to its expectation, namely the *t.l.* of quenched pressure density. Actually we can prove more, namely that the convergence is $\mathbb{P}$-almost surely.

1.3 Some examples of applications

The class of spin models on graphs defined in (1.1.2) and its random version 1.8, can be tough as a preliminary setting for a statistical mechanic description of various systems. Some classical examples are listed below.

• *Magnetic systems*: Suppose you want to describe the thermal equilibrium properties of magnetic solids. In principle, there are a lot of degrees of freedom on the interactions, however as first approximation one can describe the physical

properties of this system using a spin model on a graph defined in (1.1.2). Keep in mind the general framework introduced above, let us take magnetic solid: the components of the system are then the atoms. The finite set Λ labels the atoms of solid and, for each $i \in \Lambda$, the spin σ_i represent the magnetic moment of the atom indexed by i that is quantized and can take only the two values $+1$ and 1. The graph G represent the molecular structure of the solid, namely the atoms and their connections. If the solid has a crystal structure one can choose $G \equiv G_d$, the lattice graph in dimension d. The Hamiltonian (1.1) represent the magnetic energy of a configuration. In particular, the parameter J_{ij} tunes the magnetic interaction between the atoms i and j and the parameter h_i is an external magnetic field acting on the atom i. In the random framework, the physical meaning of a choice of a random Hamiltonian (1.8) will be discussed in detail in chapter 2.

• *Absorption of monoatomic gas molecules*: Consider a monoatomic gas adsorbed on a solid material. A simple statistical mechanic model for this system can be the following. The finite set Λ labels the allowed sites of the solid, and, for each $i \in \Lambda$, let us consider an occupancy number $\alpha_i \in \{0, 1\}$ representing respectively an empty or an occupied site of the surface. This is a rough approximation for hard-core repulsion between the absorbed gas molecules. A microscopic configuration, i.e. a possible configuration of occupied or empty sites, is then represented by an element $\alpha \in \{0, 1\}^{|\Lambda|}$. As before, the graph G represents the molecular structure of the solid, and in order to taking account the crystal structure of the solid one can assume $G \equiv G_d$. The energy associated to a possible configuration of the system gas-surface is represented by the following Hamiltonian

$$H_G(\alpha) := -\sum_{ij \in E} J_{ij}\alpha_i\alpha_j - \sum_{h_i \in \Lambda} \alpha_i \tag{1.17}$$

In particular, the parameter J_{ij} tune the interaction between gas atoms on the site i and j and the parameter h_i is the chemical potential acting on the site i.

This model can be easy mapped into a spin model on G defined in 1.1.2. Indeed, there exists a bijection between the two configurations spaces, namely

$$\{0,1\}^{|\Lambda|} \ni \alpha \leftrightarrow \sigma \in \{-1,1\}^{|\Lambda|}, \quad \sigma_i = 2\alpha_i - 1 \ \forall\, i \in \Lambda$$

By consequence, its easy to show that the two Hamiltonian differ by a non relevant constant.

• *Dean's problem*: This is a classical optimization problem. Because to the lack of physical dimension it's a good example of system that can be represented as a model on the complete graph K_N. Anyway, in general, let us suppose we have a network of people represented by a finite graph G and a collection of real parameters J_{ij} for $ij \in E$, which describe how much people i and j like or dislike each other. Naturally, a positive parameter means that they like each other and a negative parameter means that they dislike each other. However, unrealistic it may seem, we will assume that the feeling is mutual. We will consider different ways to divide the initial group into two subgroups, say A and B: it will be convenient to describe them using vectors of ± 1 labels. Namely, for each $i \in \Lambda$ we assign a spin variable σ_i with the following convention: if $\sigma_i = 1$(resp. $\sigma_i = -1$) the people i belongs to the subgroup A(resp. B). Therefore a *spin configuration* $\sigma := (\sigma_i)_{i\in\Lambda}$ describe a possible partition and the *configuration space* $\Sigma_\Lambda := \{-1,+1\}^{|\Lambda|}$ correspond to the $2^{|\Lambda|}$ possible such partitions. For a given configuration σ, let us consider the following *discomfort function*:

$$H_G(\sigma) := -\sum_{ij\in E} J_{ij}\sigma_i\sigma_j \tag{1.18}$$

(1.18) is a spin Hamiltonian (1.1) at zero external field and in this context can be tough as a measure of the discomfort(=energy) generated by the partition σ.

The Deans problem is to minimize the function (1.18) over all configurations $\sigma \in \Sigma_\Lambda$, hence, in the Statistical Mechanical language, the problem is equivalent

to find the *ground state* (1.7) of the Hamiltonian.

The interpretation of this objective is clear, since minimizing the discomfort function means that we would like to keep positive interactions as much as possible within the same groups and separate negative interactions into different groups. This optimization problem is in general NP-hard [37]. One of the reasons is the presence of *frustration*, a property that will be analyzed in section 2.1. The same problem in random setting is a random combinatorial optimization problem.

Chapter 2

A short overview on Spin Glasses

Spin Glasses has been introduced in theoretical physics quite recently. In the 1970s, experiments on magnetic alloy metals like Fe, Mn and Cr weakly diluted in metals as Au, Ag and Cu, showed a thermodynamic behavior not compatible with the classical theory of ferromagnetism, such as peculiar dynamics properties reaching the equilibrium [65, 14].
In order to explain this phenomena, theoretical physicist (Edward and Anderson 1975, [31])introduced an Ising model with random interaction, know as *EA model*, which represent the archetype of Spin Glasses.

To simplify the analytical treatment, Sherrington and Kirkpatrick (1975, [93]) proposed a mean field version of the EA model, known as the *SK model*. The resulting Mean Field theory, fully developed from the work of Parisi (1985, [15]) , was based on an *ansatz* that has revealed a very rich mathematical structure. Only until recently some predictions of the above theory have received a rigorous mathematical proofs: the exact solution of the model by Guerra (2003, [79]) and Talagrand (2006, [119]) and the Parisi's ultrametric conjecture by Panchenko (2010, [108]).

These results are based on the introduction of new investigative techniques

and concepts, notably the interpolation method [79], the Aizenmann-Contucci [22] and the Ghirlanda-Guerra identities [76] that will be the essential tools used in the following chapters.

Despite these progress and the success in the applications, the underlying mathematical structure is not fully understood thus, following Talagrand [18], "*Mean Field Spin Glasses are a challenge to mathematicians*".

Spin Glasses are one example of Disorder Models on Graphs introduced in section 1.2 thus we refer to that section for the general mathematical framework. In this chapter we give a short overview on the physical motivations behind the introduction of randomness in spin models and in the last section we illustrate the basic results on the SK model.

2.1 From Ising to EA model

Before introducing Spin Glass models, let us briefly recall the *Ising model* (see [3, 6] for a complete account of the subject) emphasizing his main difference with respect to its spin glass version, the *Edwards-Anderson model*.

This model has been introduced to give an explanation of the *ferromagnetic* behavior of some kind of materials. These materials, after having been exposed to an external magnetic field develop a magnetization with the same sign of the field. When the field was then switched off, the materials showed two different behaviours depending on the temperature at which the magnetization was induced. If the temperature was below a critical value, the materials retained a degree of magnetization, called *spontaneous magnetization*, whereas they was not capable of doing this when the temperature was greater or equal to the critical value. As temperature approached the critical value from below the spontaneous magnetization vanished abruptly.

The problem is to described the magnetic properties of a crystal lattice in dimension d.

The setup of the model is the general one described in section 1.1 and in particular in example 1.

Definition 2.1.1. *The nearest-neighborhood ferromagnetic Ising model on $\mathbb{Z}^d$, for any $\Lambda \subset \mathbb{Z}^d$ is defined by the Hamiltonian function*

$$H_\Lambda(\sigma) = -J \sum_{i,j\in\Lambda:|i-j|_1=1} \sigma_i\sigma_j - h\sum_{\in\Lambda} \sigma_i \tag{2.1}$$

where $J > 0$ and $h \in \mathbb{R}$.

The choice of the integer lattice $\mathbb{Z}^d$ is motivated by translational invariance of the physical crystal, while the nearest-neighborhood condition on the edge set reflect the fact that the interactions vanish at long distances providing a spatial structure to the model. The sign of the coupling $J > 0$ entail a ferromagnetic nature of the model in the following sense: the minimization over the configuration space of the Hamiltonian (2.1) forces the neighborhood spins to be aligned.

This model represent a considerable simplification on the description of the microscopic degrees of freedom of the original system: the state of an atom is reduced to a variable taking only two values, all the complicated electromagnetic (and quantum) interactions are replaced by a simple attraction between nearest neighbours on the lattice. Despite this simplification, this model turn out to be a paradigmatic model and decisive turn in the development of Statistical Mechanics in several ways. The most important one is that furnish a mathematical explanation of the *ferromagnetism* and in particular of the *phase transition* phenomenon [3]. A mathematical discussion of this aspect is out of the scope of this work and we refer to the classical literature mentioned above. Here we give an heuristic description of the physical behaviour.

For $\beta \to 0$ ("infinite temperature") the spin variables are independent under the Gibbs measure associated to the Hamiltonian (2.1), so the model is equivalent to a fair coin tossing.

As soon as $\beta > 0$ the probability distribution starts to favour configurations with many neighbor pairs of aligned spins. This tendency becomes stronger and stronger as β increases. The limit case is $\beta \to \infty$ ("zero temperature" or ground state) were a strong order in the model appear. In the case $h = 0$, the model is symmetric under interchange of the spin values -1 and $+1$, so that there is an equal chance of having many pairs of plus spins or having many pairs of minus spins. This dichotomy gives rise to the following interesting behavior. Suppose that $d \geq 2$. If β is sufficiently small (i.e. in the high temperature regime), the interaction is not strong enough to produce any order, and the *t.l.* of the Gibbs measure is uniquely determined. In contrast, when β is sufficiently large (in the low-temperature regime), the interaction becomes so strong that a strong range order appears: the bias towards neighbor pairs of aligned spin then implies that Gibbs measures prefer configurations with either a vast majority of plus spins or a vast majority of minus spins, and this preference even survives in the infinite volume limit. We say that the system thus undergoes a *phase transition* and a spontaneous magnetization occur.

Despite his fundamental role, the ferromagnetic Ising model turn out to be inappropriate to describe the magnetic behavior of solid alloys. In order to fill this gap, theoretical physicists have constructed other kind of models taking into account the presence of *randomness* and *frustration* in the system, which, as we will see in the next, constitute the essentials ingredients of Spin Glasses.

• *Randomness*: Assume that we have a solid alloys made of a magnetic transition metal impurities in noble metal hosts, say Fe and Au. Clearly, for a given sample of the material, we don't know the exact arrangement of the impurities. However, one can hope there are physical properties that are independent of the sample. In order to formalize the previous claim, we model the spatial inhomogeneity by introducing some probability distribution on the space of possible realizations of the iron positions. As in ordinary statistical mechanics, we expect that the sample-to-sample fluctuations go to zero in the

limit of a large system. A quantity with this property is said to be *self-averaging* 1.2.6. If we know that this property holds, then not only we can expect the same results in experiments with different samples, but we can also expect that the average of the disorder give the same result of the experiments. The positions of the atoms are not the only degrees of freedom of the system, indeed iron atoms have magnetic moments (*spins*), and we are interested in the magnetic properties of the system. Thus we deal with two kind of averages, the spin and random positions averages, which play a role in the thermal equilibrium. The question is: they play the same role? The physical picture is the following. At high temperatures the atoms quickly change places, so that we guess that both the spin and the positions are free to take on values which minimize the energy of the system. This situation correspond to the *annealed average* (see section 1.2).

However, at low temperatures, the motion of atoms is strongly suppressed: one says that the positions of the atoms are *frozen*. Strictly speaking, they are not exactly frozen, but they have a dynamics which is many order of magnitudo slower than the dynamics of the spins. Indeed, unlike the positions, the spins are not *frozen*, and their behaviour could be described by a Gibbs measure. This description correspond to the *quenched average* (1.13) and is the right one for the phenomena we want to describe.

In the example we have just illustrated, randomness is in the position of the magnetic impurities. This situation does not turn out to be simple enough to allow an analytical approach. Therefore, following Edwards and Anderson [31], one usually considers models where the positions of the magnetic moments are non-random and are placed on the sites of a lattice and disorder is in the family of couplings.

Thus, keep in mind the general framework introduced in section 1.2, let us consider a *disordered spin model* the a lattice graph G_d-

Definition 2.1.2. *(EA model) The nearest-neighborhood Edward-Anderson*

model is defined, $\Lambda \subset \mathbb{Z}^d$, assigning to each spin configuration σ a random Hamiltonian function, called EA Hamiltonian, of the form

$$H_\Lambda(\sigma) = -\sum_{|i-j|_1=1} J_{ij}\sigma_i\sigma_j - h\sum_{i\in\Lambda}\sigma_i \tag{2.2}$$

where the family of couplings $J_\Lambda = \{J_{ij}\}_{i,j\in E}$, are taken to be i.i.d. standard gaussian random variables.

Thus, roughly speaking, the Hamiltonian of the EA model has the same form of the Ising Hamiltonian (2.1), apart from the fundamental difference that, the couplings J_{ij} are gaussian random variables.

• *Frustration*: The choice of a symmetric distribution for the couplings has a physical explanation: the impurity moments produce a magnetic polarization of the host metal conduction electrons, which is positive at some distances and negative at others. This simple fact has a deep consequence.

Let us start analyzing, at zero magnetic external field h, the *ground state* of the Ising model, *i.e.* the spin configuration that minimize the Hamiltonian function 2.1. Clearly, each spin-spin interaction term is minimized when the two spins are parallel, *i.e.* $\sigma_i\sigma_j = +1$ for all i, j. There are two such configurations, one with all spins equal $+1$, the other with spins -1, and they are connected by the global spin-flip symmetry $\sigma_i \to -\sigma_i$ for each $i \in \Lambda$. It is also elementary to check that any other configuration has a strictly higher energy. Therefore the ground state can be completely determined by the symmetries of the system.

Conversely in the EA model, for a given realization of the random interactions, some of them are of ferromagnetic character (*i.e.*, they favor parallel alignment of the spins) and others are anti-ferromagnetic. This is a key point: we have random, competing interactions which produce *frustration* [121]. Indeed, consider a triple $i, j, k \in \Lambda$ with $J_{ij}, J_{jk} > 0$ but $J_{ik} < 0$. The minimization of the Hamiltonian enforce the following picture: $\sigma_i\sigma_j = 1$, $\sigma_j\sigma_k = 1$, but unfortunately $\sigma_i\sigma_k = -1$. The use of the term *frustration* can be easy understood

if one think at interpretation of spin models in terms of the Dean's problem mentioned in chapter 1.

Informally, one can say that frustration is present when the Hamiltonian cannot be written as a sum of individual terms which can all be minimized by a single ground state configuration. Moreover the ground state of the system has a high degeneracy, and different ground states are not connected to one another by elementary symmetry transformations. The number of ground states in fact grows very fast as the number of spins is increased.

The EA model, and in general finite dimensional Spin Glasses, though in a sense of a simplification with respect the physical system, are still of formidable difficulty problem to attack. To simplify the analytical treatment of the EA model, Sherrington and Kirkpatrick [93] proposed a mean field version of the EA model, known as the SK model that will be discussed in the next section.

2.2 The Sherrington-Kirkpatrick model

The common feature of *Mean Field models* is that any additional structure for the graph G, as for example the spatial structure $\mathbb{Z}^d$ for lattice models, is considered in favour of a simpler setting: the graph G is now the complete graph K_N with a vertex set $\Lambda_N = \{1, \ldots, N\}$. This kind of models can be toughs in two ways.

First, they can be viewed as a *mean field approximation* of the corresponding model on the lattice because of the lacking of spatial dimensionality in the interaction: all the microscopic components of the system are supposed to interact with each other, irrespective of their physical distance. The fundamental question of the connection between the mean field picture and the corresponding short range version, at least when the space dimensionality or the range of interactions is large, is well understood for instance, for non-random ferromagnetic systems, where it is well known [4, 3] that mean field theory gives

a reasonable qualitative description of symmetry breaking and of the critical point. After almost thirty years, the question of the connection between mean field and realistic spin glass models is still to a great extent open and under discussion. We will give a partial result on this direction in chapter 4.

On other hand, one can imagine to describe a system with an high number of connections as for example a neural network, a network of people or a social system. This last point of view is also closer to optimization problems (see for example the Dean's problem (see section 1.3) where the spatial structure is irrelevant. In this case the picture where all components interact with each other is not far from the original situation and hopefully the resulting model is able to capture the relevant properties of the system.

2.2.1 A preliminary example: the Curie-Weiss model

Before introducing the SK model, we start with the prototype of mean field models: the *CurieWeiss model*. It can be viewed as the *mean field approximation* of the Ising model (2.1).

Remark 2.2.1. *Let us give an heuristic argument which gives a meaning to the term mean field approximation. Let us rewrite the Ising Hamiltonian (2.1) in the equivalent form*

$$H_\Lambda(\sigma) = -J \sum_{i\in\Lambda} \sigma_i \Big(\sum_{j\sim i} \sigma_j \Big) - h \sum_{i\in\Lambda} \sigma_i \tag{2.3}$$

where $j \sim i \leftrightarrow |i-j|_1 = 1$ denote the set of nearest-neighborhood of the site i. The term $\Big(\sum_{j\sim i} \sigma_j \Big) =: \widetilde{h}_i$ is the effective field acting on the spin σ_i. The mean field approximation is then to replace, for each $i \in \Lambda$, the effective field $\widetilde{h}_i$ with a site independent quantity, i.e. its mean value $|\Lambda|^{-1} \sum_{i\in\Lambda} \sigma_i$.

All we need to do is to replace the nearest-neighbour pair interaction of the Ising model (2.1), by another extreme choice: we assume that each spin variable interacts with each other at any site of the lattice, irrespective of their

distance. In this case, the actual structure of the lattice becomes irrelevant. The only important thing is the number of spins that will be denoted by an integer N thus, we may simply take $\Lambda \equiv \Lambda_N = \{1, \dots, N\}$. The configuration space becomes $\Sigma_\Lambda \equiv \Sigma_N = \{-1, +1\}^N$.

The resulting model is a spin model on the complete graph K_N with constant coupling $J/N > 0$ and external field $h \in \mathbb{R}$.

Definition 2.2.2. *(CW model) The Curie-Weiss model is defined assigning to each spin configuration σ the CW Hamiltonian function*

$$H_N(\sigma) = -\frac{J}{N} \sum_{1 \le i < j \le N} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i \tag{2.4}$$

The strength of the interaction should be chosen $\mathcal{O}(1/N)$, to avoid the possibility that the Hamiltonian takes on values larger than $\mathcal{O}(N)$. This request is essential to ensure the existence of the thermodynamical limit which in this case correspond to the limit of $N \to \infty$. Without loss one can considerer $J = 1$, the general case by a simple rescaling of the paramaters.

A basic quantity, the so called *order parameter* of the model, is the *empirical magnetization*.

Definition 2.2.3. *For a given configuration $\sigma \in \Sigma_N$ the empirical magnetization is defined as*

$$m_N(\sigma) := \frac{1}{N} \sum_{i=1}^{N} \sigma_i \tag{2.5}$$

This is a *macroscopic function* in the sense that it depends on all spin variables, and depends on each one of them very little. There is a rigorous mathematical definition of macroscopic variable [2] but for our purpose we can say that it is characterized by the fact in the thermodynamical limit its value is not a affected by a change of finitely many spins.

The main consequence of working on the complete graph is the following

Proposition 2.2.4. *The CW Hamiltonian (2.4) can be rewritten as a function of a single macroscopic function, i.e.* $H_N(\sigma) = N\Big(\frac{1}{2}m_N(\sigma)^2 + h\, m_N(\sigma)\Big)$.

Thanks to previous property the model turn out to be solvable in the following sense.

Theorem 2.2.5. *Let* $p_N := \frac{1}{N}\log Z_N$ *be the pressure density of the Curie-Weiss model, where* $Z_N := \sum_{\sigma\in\Sigma_N} e^{-\beta H_N(\sigma)}$ *is the partition function associated to the CW Hamiltonian (2.4). For* $m \in [-1,1]$*, let us define the function*

$$\psi(m) := \beta(hm + \frac{m^2}{2}) + H(\frac{m+1}{2}) \tag{2.6}$$

where $H(x) = -x\log x - (1-x)\log(1-x)$ *is the binary entropy function.*

Then, the thermodynamical limit of the pressure density exist and satisfy the following

$$\lim_{N\to\infty} p_N = \sup_{m\in[-1,1]} \psi(m) \tag{2.7}$$

A proof of the previous results can be obtained in various way. A good reference on the subject is [3] where, with large deviation techniques, an interested reader can also find a detailed analysis of the thermodynamical limit of the Gibbs measure. Let us briefly describe the behavior of the optimal m determined by the r.h.s. of (2.7), which represents the limiting value of the magnetization w.r.t. the Gibbs measure.

The stationary point(s) of $\psi(m)$ are solution(s) of the equation

$$m = \tanh[\beta)m + h)]. \tag{2.8}$$

If $h \neq 0$ and any β or $\beta \leq 1$ (high temperatures) and $h = 0$, $\psi(m)$ has only one global maximum, while for $\beta > 1$ has two local maxima. In particular, for $h = 0$ the function ψ is symmetric, and so takes the same value at both maxima. As a consequence, the magnetization as a function of the magnetic field, is not unique at the value $h = 0$ (and only at this value). For $h > 0$, the maximizer

is the positive solution of (2.8), while for negative h it is the negative solution. Consequently, the magnetization has a jump discontinuity at $h = 0$: we say that the CurieWeiss model exhibits a *first-order phase transition*.

2.2.2 The model

In the previous section we have seen that the Curie-Weiss model (2.4) is the mean field version of the Ising model (2.1). In the same spirit, one can think the Sherrington-Kirkpatrick model as a mean field version of the Edward-Anderson model (2.2). Despite the reasons for their original birth this model soon became of fundamental importance in the analysis of complex systems thanks to their immediate link to a class of optimization problems [106, 101]. The simplest example of such as problems is the Dean's problem mentioned section 1.3.

Applying the same considerations of the previous section, we consider a disordered spin model on the complete graph K_N with Hamiltonian:

Definition 2.2.6. *(SK model) The Sherrington-Kirkpatrick model is defined assigning a random Hamiltonian function of σ, called SK Hamiltonian:*

$$H_N(\sigma) = -\frac{1}{\sqrt{N}} \sum_{1 \le i,j \le N} J_{ij}\sigma_i\sigma_j - h\sum_{i\in\Lambda}\sigma_i \tag{2.9}$$

the family of couplings $J_N = \{J_{ij}\}_{1\le i,j\le N}$, are taken to be i.i.d. standard gaussian random variables.

Remark 2.2.7. *Strictly speaking, the edge set of the complete graph K_N is $P^{(2)}_{\Lambda_N} = \{ij \equiv \{i,j\} : 1 \le i < j \le N\}$ then the sum in (2.9) should be over $i < j$. The choice of sum over all the possible pairs is for a convenience of notation, the only difference is a factor of $\sqrt{2}$, since the contribution of the diagonal elements is negligible for large N and the sum of two i.i.d. standard gaussian Gaussian random variables $J_{ij} + J_{ji}$ is equal in distribution to $\sqrt{2}J_{ij}$.*

Notice also that, because of the normalization factor $N^{-\frac{1}{2}}$ in (2.9, unlike the CW model, the strength of the interaction is $\mathcal{O}(N^{\frac{3}{2}})$. As we will show in

the next, this is the right normalization which ensure a well defined t.l. for the pressure density and a properly scaled ground state.

It will be useful to introduce a different point of view on the SK model: one can think the SK Hamiltonian (2.9 as a *Gaussian process* indexed by the set Σ_N, i.e. by the N-dimensional hypercube. In fact, $H_N(\sigma)$ is a centered Gaussian random process which is fully characterized by its covariance function.

Definition 2.2.8. *Given two configurations $\sigma^1, \sigma^2 \in \Sigma_N$, the covariance function of the SK Hamiltonian is*

$$C_N(\sigma^1, \sigma^2) := \mathbb{E} H_N(\sigma^1) H_N(\sigma^2) = N\big(q_N(\sigma^1, \sigma^2)\big)^2 \tag{2.10}$$

where the function $q_{12} \equiv q_N(\sigma^1, \sigma^2) := \frac{1}{N}\sum_{i=1}^N \sigma_i^1 \sigma_i^2 \in [-1, 1]$ is called the overlap between the two configurations.

Notice that (2.10) is the gaussian analogous of proposition 2.2.4 for the Curie-Weiss model. Moreover, on can also consider gaussian process on Σ_N with covariance function depending on the overlap in a different way, i.e. $C_N(\sigma^1, \sigma^2) = N\xi(q_{12})$ for suitable functions ξ. A class of important examples are:

- *p-spin models*: $\xi(x) = x^p$, for any integer $p \geq 1$. These gaussian process may be represented in a form similar to the SK Hamiltonian, except that the two-spin interaction must be replaced by a p-spin interaction:

$$H_{N,p}(\sigma) := \frac{1}{\sqrt{N^{p-1}}} \sum_{i_1,\dots,i_p} J_{i_1,\dots,i_p} \sigma_{i_1} \cdots \sigma_{i_p} \tag{2.11}$$

with $J_{i_1,\dots,i_p}$ are *i.i.d.* standard gaussian random variables

- *mixed p-spin models*: $\xi(x) = \sum_{p\geq 1} \beta_p^2 x^p$, where the sequence of positive real numbers $(\beta_p)_{p\geq 1}$ is assumed to decrease fast enough to ensure the convergence of the series for $\xi(q_{12})$. The corresponding spin Hamiltonian is:

$$H_N^{mix}(\sigma) := \sum_{p\geq 1} \beta_p H_{N,p}(\sigma) \tag{2.12}$$

where $H_{N,p}$ is a sequence of independent p-spin Hamiltonian (2.11). The Hamiltonian functions of the type (2.12) plays a fundamental role in proof of the Ghirlanda-Guerra identities (see chapter 4), since their covariance contains informations about all the p-moments of the overlap.

We also notice that the overlap defined in 2.2.8 is closely related to the *Hamming distance* on Σ_N, where $d_{Hamming}(\sigma^1, \sigma_2) := \#\{i \leq N : \sigma_i^1 \neq \sigma_i^2\}$.

Further important models can be obtained by other choices of metric on Σ_N: the *Random Energy Model* (REM) [68, 18], the *Generalized Random Energy Model* (GREM) [69, 51] and the *Nonhierarchical GREM* [47]. Despite their fundamental importance in Spin Glass theory, these models will be not presented in this work, an interested reader can consult the above references.

2.2.3 The Parisi Formula and related results

The most remarkable result on the SK model is a closed expression for the *t.l.* of the pressure density expressed as variational problem know as *Parisi's Formula*. Let us specialize the general definitions 1.2.5 for the SK model.

Definition 2.2.9.

$$Z_N := \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)} \tag{2.13}$$

is the random partition function associated to the SK Hamiltonian (2.9).

The random pressure density is the random variable $N^{-1} \log Z_N$. *Averaging on the disorder we obtain quenched pressure density,*

$$p_N := \frac{1}{N} \mathbb{E} \log Z_N \tag{2.14}$$

The variational problem for the quenched pressure density is given in terms of a functional on a convex space of distributions functions.

Definition 2.2.10. *Let* $\mathcal{M}_d$ *be the set of all discrete distribution functions on* $[0, 1]$. *We notice that they are nondecreasing piecewise constant, right-continuous functions. Any* $x \in \mathcal{M}_d$, *called the functional order parameter, can*

be parameterized with an integer number $r \geq 1$ *and two sequences* $\underline{q} := (q_l)_{l=0,\dots,r}$ *and* $\underline{m} := (m_l)_{l=0,\dots,r+1}$ *such that*

$$0 = q_0 \leq q_1 \leq \cdots \leq q_{r-1} \leq q_r = 1$$
$$0 = m_0 < m_1 \ldots < m_r < m_{r+1} = 1.$$

If we denote by $\theta(\cdot)$ *the right continuous Heaviside function, then*

$$x(u) := \sum_{l=0}^{K} (m_{l+1} - m_l)\theta(u - q_l),\ u \in [0,1] \tag{2.15}$$

In other terms $x(u)$ is the cumulative distribution function of a discrete r.v. which take the value q_l with probability $m_{l+1} - m_l$. Some authors [16] use a different parametrization and m is replaced by ζ given by the relation $m_{l+1} = \zeta_l$.

Definition 2.2.11. *For a given* $x \in \mathcal{M}_d$*, any inverse temperature* β *and external field* h*, we call the Parisi functional*

$$\mathscr{P}(x) \equiv \mathscr{P}(h; x) := \log 2 + f(0, h; x) - \beta^2 \int_0^1 x(u)u\, du \tag{2.16}$$

where $f(0, h; x) \equiv f(u, y) : [0,1] \times \mathbb{R} \to \mathbb{R}$ *satisfies the Parisi's PDE:*

$$\frac{\partial f}{\partial u} + \beta^2 \Big(\frac{\partial^2 f}{\partial y^2} + x(u) \Big(\frac{\partial f}{\partial y} \Big)^2 \Big) = 0 \tag{2.17}$$

with the boundary condition

$$f(1, y) = \log \cosh(y). \tag{2.18}$$

The Parisi's PDE (2.17) is non linear, however it has nice properties. First, for any $x \in \mathcal{M}_d$, it can be solved recursively starting from the boundary condition. Indeed, suppose that $x(u)$ is constant in a interval, let say $[u_a, u_b]$.

Proposition 2.2.12. *For a given real number* $m \geq 0$ *and* $f(u, y) : [u_a, u_b] \times \mathbb{R} \to \mathbb{R}$ *let us consider the following PDE:*

$$\frac{\partial f}{\partial u} + \beta^2 \Big(\frac{\partial^2 f}{\partial u^2} + m \Big(\frac{\partial f}{\partial u} \Big)^2 \Big) = 0 \tag{2.19}$$

with final condition

$$f(u_b, y) = g(y). \tag{2.20}$$

Then the solution is:

$$f(u_b, y) = \begin{cases} \mathbb{E} g(y + z\sqrt{2\beta(u_b - u)}) & \text{if } m = 0 \\ \frac{1}{m} \log \mathbb{E} \exp\left(m g(y + z\sqrt{2\beta(u_b - u)})\right) & \text{if } m \neq 0, \end{cases} \tag{2.21}$$

where $\mathbb{E}$ denotes expectation with respect to a standard Gaussian random variable z.

The following result, due to Guerra [79], holds

Proposition 2.2.13. *The function $f(u, y; x)$ solution of 2.17 is pointwise continuous in x w.r.t. the $L_1([0,1], du)$ norm. In fact, for any $x, \tilde{x} \in \mathcal{M}_d$ we have*

$$|f(u, y; x) - f(u, y; \tilde{x})| \leq \beta^2 \int_u^1 |x(v) - \tilde{x}(v)| dv \tag{2.22}$$

Thus, proposition 2.2.13 implies that, $f(u, y; x)$ and also the Parisi functional (2.16), can be extended continuously from $\mathcal{M}_d$ to $\mathcal{M}$, where $\mathcal{M}$ is the space of all distribution functions on $[0, 1]$.

We are ready to state the celebrated *Parisi formula*.

Theorem 2.2.14. *(Parisi '85, Guerra '03, Talagrand '06)*

$$\lim_{N\to\infty} \frac{\log Z_N}{N} = \lim_{N\to\infty} \mathbb{E} \frac{\log Z_N}{N} = \inf_{x \in \mathcal{M}_d} \mathscr{P}(x), \ \mathbb{P} - a.s. \tag{2.23}$$

where $\mathscr{P}(x)$ is the Parisi functional (2.16). Moreover the minimizer is unique.

Clearly the infimum is taken over $r, \underline{m}, \underline{q}$ defined in 2.15.

The formula was discovered by G. Parisi in [110] using a non rigorous approach called *replica trick*. It was later understood in [15] that the solution

conjectured by Parisi corresponded to a number of physical properties of the Gibbs measure of the model, one of them being the *ultrametricity* of its support. The first important step toward the proof of the *Parisi formula* was the *Replica Symmetry Breaking* interpolation scheme introduced by F. Guerra in [79] which give the upper bound. The proof was completed by M. Talagrand [19] showing how to control the remainder in the *t.l.* by developing a version of Guerra's interpolation for coupled systems. Actually, Talagrand's proof applies, with little modifications of the formula, to a p-spin model (2.11) for even p. The proof all p-spin models, and then for the mixed p-spin model 2.12, was given by D. Panchenko in [16] with a different method based on the *Aizenman-Sims-Starr scheme* [24] variational principle and the Ghirlanda-Guerra identities. The uniqueness part was a longstanding open problem recently solved in [35]. In the next chapter we will consider in detail a generalization of the SK model and multidimensional analogous of Theorem 2.2.14.

A discussion about the limiting Gibbs measure of the SK model is rather delicate. The full Parisi picture contains also the following two points:

• The solution of the Parisi Formula 2.23, let say x, then represents the limiting law of the overlap w.r.t. the quenched Gibbs measure, namely for any bounded continuous function f, we should have

$$\lim_{N\to\infty} \langle f(q_{12})\rangle_N = \int f(u)x(du)$$

where $\langle\,\rangle_N$ is the quenched Gibbs measure defined in (1.13).

• The support of the joint limit law its ultrametric, namely given $\sigma^1, \sigma^2, \sigma^3 \in \Sigma_N$

$$\lim_{N\to\infty} \langle I\Big(q_{1,3} \geq \min(q_{1,2}, c_{2,3})\Big)\rangle_N = 1$$

The full picture has been rigorously proved [16] only for the mixed p-spin model (2.12). One can say that for the SK model the previous points holds *generically*: we will discuss this aspect in more detail in chapter 4.

Chapter 3

A Multi-species SK model

Multi-species spin systems at different densities are often encountered in nature. The bipartite case without disordered made its appearance since the work on meta-magnets by Cohen and Kinkaid [56]. When several types of magnetic particles like iron and manganese are diluted into a nonmagnetic metallic host the Ruderman-Kittel-Kasuya-Yosida interactions generate a multi-species spin glass phase [122]. The rich complex behaviour emerging from those physical systems revealed to be useful in a variety of applications ranging from biology to social sciences and several models were proposed and studied in the mean field approximation without disorder [58, 57, 70] and with disorder as well [43, 44, 45]. In this chapter we introduce and study a *Multi-Species SK model* (MSK model) i.e. a disordered spin model on the complete graph composed by vertex belonging to a finite number of different species. As for the SK model 2.2.6, spin couples interact through a family of independent centered gaussian random variable whose variances depends only on the species they belong to. In other words, we relax the assumption of *i.i.d.* random couplings. Under some convexity assumption (see [39, 43, 45] for a similar conditions in neural network theory) on the variances the *t.l* of the pressure density is given by a multidimensional analogous of the *Parisi formula*, namely one can prove a multidimensional analogous of Theorem 2.2.14. This is a multidimensional gen-

eralization in a proper sense, namely when there is only one species we recover the result for SK model.

This model and the related Parisi's formula were first proposed by Barra, Contucci, Mingione and Tantari in [41] where the authors, using a modification of Guerra's interpolation, showed that the formula gives an upper bound for the pressure. The matching lower bound was proved by Panchenko in [109] utilizing a new multi-species version of the Ghirlanda-Guerra identities.

The chapter is almost entirely based on [41] and is organised as follows. In the first section we introduce the MSK model and its basic properties. In section 2 we illustrate the main result, namely the related *Parisi fomula*. In section 3 we prove the existence of thermodynamic limit of the quenched pressure density 2.14 of the model. In section 4 we study the annealed region with the second moment method, while section 5 the replica symmetric solution and shows that at low temperatures it has a negative entropy. In section 6 we give a proof of the upper bound for the pressure, namely the first part of the main result, while in section 7 we explain the ideas behind the proof of the matching lower bound obtained Panchenko [109] based on a consequence of the Ghirlanda-Guerra identities, the so called *synchronization property* (see section[**?**]).

3.1 The model and the basic definitions

Let us consider a *Disordered Spin Model* on the complete graph K_N with vertex set $\Lambda_N = \{1, \dots, N\}$. Suppose that the vertex can be divided in a finite number S of different species labelled by a finite set $\mathcal{S}$, so from now the term *species* will refer to the elements of this set of cardinality S.

For each $s \in \mathcal{S}$ consider a set $\Lambda_N^{(s)} \subset \Lambda_N$ such that

$$\bigcup_{s \in \mathcal{S}} \Lambda_N^{(s)} = \Lambda_N, \tag{3.1}$$

$$\Lambda_N^{(p)} \bigcap \Lambda_N^{(s)} = \emptyset, \, \forall \, s \neq t \in \mathcal{S} \tag{3.2}$$

and define

$$N^{(s)} := |\Lambda_N^{(s)}|, \tag{3.3}$$

Clearly by definition we must have

$$\sum_{s\in\mathcal{S}} N^{(s)} = N. \tag{3.4}$$

We consider a disordered spin model defined by a collection $(\sigma^{(s)})_{s\in\mathcal{S}}$ of spin variables, meaning that $\sigma_i^{(s)} = \pm 1$ for each $\forall s \in \mathcal{S}, i \in \Lambda_N^{(s)}$.

The *configuration space* is denoted by Σ_N and is composed by the family of possible configurations $\sigma = (\sigma_i)_{i\in\Lambda_N}$ where $\sigma_i \equiv \sigma_i^{(s)}$ if $i \in \Lambda_N^{(s)}$.

Definition 3.1.1. *(MSK model) The Multi-species Sherrington-Kirkpatrick model is defined assigning to each spin configuration σ a random Hamiltonian function, called MSK Hamiltonian:*

$$H_N(\sigma) := -\frac{1}{\sqrt{N}} \sum_{i,j\in\Lambda_N} J_{ij}\sigma_i\sigma_j, \tag{3.5}$$

where the J's are independent centered gaussian r.v. such that

$$\mathbb{E}(J_{ij}) = 0, \tag{3.6}$$

and

$$\mathbb{E}(J_{ij}J_{i'j'}) = \delta_{ii'}\delta_{jj'}\Delta_{sp}^2 \mathbf{1}\{i \in \Lambda_N^{(s)}\}\mathbf{1}\{j \in \Lambda_N^{(p)}\}, \tag{3.7}$$

for some family real numbers $(\Delta_{sp}^2)_{s,p\in\mathcal{S}}$ and where $\mathbf{1}\{\cdot\}$ denotes the indicator function.

In other words the variance of the interaction between spins depends on the belonging species.

Remark 3.1.2. *The external field, for notation convenience, will be absorbed in the Gibbs measure (see eq.* (3.16) *below)*

As for the SK model, one can think the MSK Hamiltonian (3.5) as a *Gaussian process* indexed by the set Σ_N. Given two configurations $\sigma, \tau \in \Sigma_N$, by (3.6) and(3.38) the covariance matrix of the process is

$$C_N(\sigma,\tau) = \frac{1}{N}\sum_{s,p\in\mathcal{S}} \Delta_{sp}^2 \Big(\sum_{i\in\Lambda_N^{(s)}} \sigma_i^{(s)}\tau_i^{(s)} \Big)\Big(\sum_{j\in\Lambda_N^{(p)}} \sigma_j^{(p)}\tau_j^{(p)} \Big). \tag{3.8}$$

To show explicitly the dependence trough the choice of the various relative sizes $N^{(s)}$ we can define for every $s \in \mathcal{S}$ the *relative density*

$$\alpha_N^{(s)} := \frac{N^{(s)}}{N}, \tag{3.9}$$

and the relative overlap

$$q_N^{(s)}(\sigma,\tau) = \frac{1}{N^{(s)}} \sum_{i\in\Lambda_N^{(s)}} \sigma_i^{(s)}\tau_i^{(s)}, \tag{3.10}$$

then the covariance matrix can be write in the form

$$C_N(\sigma,\tau) = N \sum_{s,p\in\mathcal{S}} \Delta_{sp}^2 \alpha_N^{(s)}\alpha_N^{(p)} q_N^{(s)}(\sigma,\tau) q_N^{(p)}(\sigma,\tau). \tag{3.11}$$

It's useful to introduce the vector notation of Appendix 7.2: from now on bold letters denote vectors, and $(\,,\,)$ the usual scalar product in $\mathbb{R}^S$, where S is the number of species. Thus (3.11) rewrites as

$$C_N(\sigma,\tau) = N\Big(\mathbf{q}_N, \mathbf{\Delta}\mathbf{q}_N\Big), \tag{3.12}$$

where $\mathbf{q}_N$ is called *overlap vector* and is defined as

$$\mathbf{q}_N := \Big(q_N^{(s)}(\sigma,\tau)\Big)_{s\in\mathcal{S}} \tag{3.13}$$

while

$$\mathbf{\Delta} := (\Delta_{sp}^2 \alpha_N^{(s)}\alpha_N^{(p)})_{s,p\in\mathcal{S}} \tag{3.14}$$

is an $S \times S$ matrix. In other words the covariance matrix of the process can be tough ad a quadratic form in $\mathbb{R}^S$.

Example 2. *For example, in the case of two species, namely* $\mathcal{S} = \{a, b\}$*, the covariance matrix* $\mathbf{q}_N$ *is a 2-dimensional vector and* $\mathbf{\Delta}$ *is a* 2×2 *matrix defined by the entries*

$$\begin{bmatrix} \alpha_N^{(a)} \alpha_N^{(a)} \Delta_{aa}^2 & \alpha_N^{(a)} \alpha_N^{(b)} \Delta_{ab}^2 \\ \alpha_N^{(a)} \alpha_N^{(b)} \Delta_{ab}^2 & \alpha_N^{(b)} \alpha_N^{(b)} \Delta_{bb}^2 \end{bmatrix}$$

It's useful to define a *normalized covariance matrix* as

$$c_N(\sigma, \tau) := \frac{C_N(\sigma, \tau)}{N} = \Big(\mathbf{q}_N, \mathbf{\Delta} \mathbf{q}_N \Big) \tag{3.15}$$

The Gibbs measure extends the standard definition for disordered spin models (1.2.2),

$$\mathcal{G}_N(\sigma) := \frac{a_N(\sigma, \mathbf{h}) e^{-H_N(\sigma)}}{Z_N} \tag{3.16}$$

where

$$a_N(\sigma, \mathbf{h}) := \exp\Big(\sum_{s \in \mathcal{S}} h^{(s)} \sum_{i \in \Lambda_N^{(s)}} \sigma^{(s)} \Big), \tag{3.17}$$

and $\mathbf{h} := (h^{(s)})_{s \in \mathcal{S}}$ is a vector which represents an external magnetic field acting in each species separately. Upon reflection, the presence of a family of non negative weights a_N in the random Gibbs measure produces no change in the mathematical treatment of random Gibbs measure and related quantities. For example the random partition function is

$$Z_N := \sum_{\sigma \in \Sigma_N} a_N(\sigma, \mathbf{h}) e^{-H_N(\sigma)}, \tag{3.18}$$

Remark 3.1.3. *Notice that, to lighten the notation, in the definition of the Gibbs measure (3.16), and in the rest of this chapter, we do not write explicitly the dependence on* $\mathbf{h}$*. With the same aim, the physical inverse temperature* β*, which appears in the standard definition, in our case is set equal to* 1 *with no loss of generality as it can be recovered in every moment simply by properly rescaling of the variances* Δ_{sp}^2*. Let us briefly specialize for the MSK model, all the general definitions for disordered spin models we refer to section 1.2 to a*

more detailed explanation of them. Clearly when there is only one species, say a we recover the SK model 2.2.6 for $\beta = \Delta_{aa}$.

In this case, the general definitions given in section 1.2 become:

• *Gibbs state*

For every bounded function $f : \Sigma_N \to \mathbb{R}$ we call the Gibbs state the following r.v.

$$\omega_N(f) := \sum_{\sigma \in \Sigma_N} f(\sigma)\mathcal{G}_N(\sigma) \tag{3.19}$$

namely, the expectation w.r.t. the random Gibbs measure 3.16.

• *replica-Gibbs state*

For each $n \in \mathbb{N}$, let Σ_N^n be the cartesian product of n copies of the configuration space Σ_N, and denote its elements by $(\sigma^1, \dots, \sigma^n)$. For every bounded function $f : \Sigma_N^n \to \mathbb{R}$, we call the random replica-Gibbs state the following r.v.

$$\Omega_N(f) := \sum_{(\sigma^1,\dots,\sigma^n) \in \Sigma_N^n} f(\sigma^1, \dots, \sigma^n)\mathcal{G}_N(\sigma^1) \dots \mathcal{G}_N(\sigma^n) \tag{3.20}$$

namely the expectation w.r.t. the product measure $\mathcal{G}_N^{\otimes n}$.

• *Quenched state*

Averaging the disorder we obtain the quenched state,

$$\langle f \rangle_N := \mathbb{E}\, \Omega_N(f). \tag{3.21}$$

• *Pressure*

The *random pressure density* is the random variable $N^{-1} \log Z_N$,

averaging on the disorder we obtain the *quenched pressure density*

$$p_N := \frac{1}{N} \mathbb{E} \log Z_N. \tag{3.22}$$

3.2 The Parisi Formula for the MSK model

Let us introducing the basic object on which is based the variational representation for the *t.l.* of p_N, namely the *S-Parisi functional*.

The origin of the quantities that we are going to introduce can be easily understood recalling the Guerra's RSB interpolation [79] of the Sherringhton-Kirpatrick model. By using a clever interpolation argument, the author showed that the celebrated Parisi's solution is an upper bound for pressure. The RSB interpolation is defined through a nondecreasing, piecewise constant function (see eq. (2.2.11)) which represents the order parameter of the model. One of the key points of the proof, is that this function intrinsically defines an increasing sequence $(m_l)_{l=0,\dots,r}$, thus enabling the control of the sign of the derivative of the interpolating functional. Following the same approach, in the multi-species case we define:

Definition 3.2.1. *Let $\mathcal{M}_d^S$ be a subset of all discrete distribution functions on $[0,1]^S$, i.e. nondecreasing piecewise constant, right-continuous functions, defined as follow. For any integer $r \geq 1$, let us consider a sequence of points in $\Gamma \in [0,1]^S$, with*

$$\Gamma := (\mathbf{q}_l)_{l=0,\dots,r} = \left(q_l^{(s)}\right)_{s\in\mathcal{S}, l=0,\dots,r} \tag{3.23}$$

such that for each $s \in \mathcal{S}$, we have

$$0 = q_0^{(s)} \leq q_1^{(s)} \leq \cdots \leq q_{r-1}^{(s)} \leq q_r^{(s)} = 1.$$

Roughly speaking Γ defines a path with r steps in $[0,1]^S$ which is non decreasing in each direction.
We consider a sequence $\underline{m} = (m_l)_{l=0,\dots,r+1}$ such that $0 = m_0 < m_1 < \ldots < m_r < m_{r+1} = 1$.
If we denote by $\theta(\cdot)$ the right continuous Heaviside function, we define the $x \in \mathcal{M}_d^S$, called functional order parameter as

$$x(\mathbf{u}) := \sum_{l=0}^{r} (m_{l+1} - m_l) \prod_{s\in\mathcal{S}} \theta(u^{(s)} - q_l^{(s)}) \tag{3.24}$$

where $\mathbf{u} = (u^{(s)})_{s\in\mathcal{S}}$ is vector in $[0,1]^S$.

The function x defines an S-dimensional shape, that in the case of $S = 2$, looks like a *ziggurat*.[1]

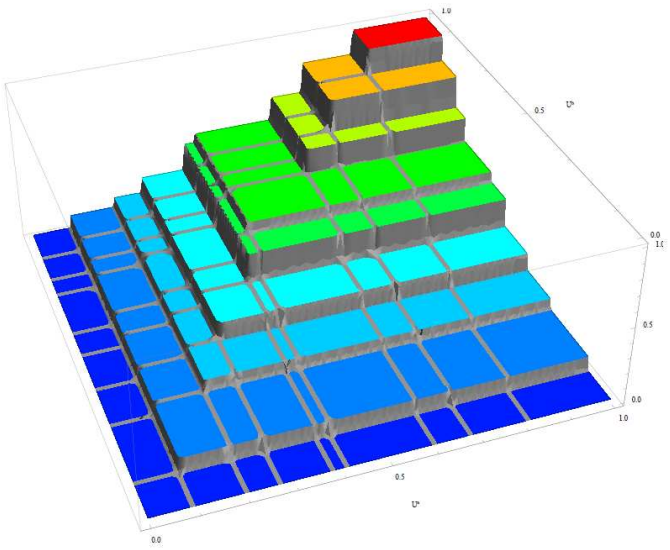

Figure 3.1: An example of the function $x(\mathbf{u})$ in the case of two species

For each $s \in \mathcal{S}$, let $\mathcal{P}_s$ be the canonical projection operator on the s direction in $\mathbb{R}^S$ and for $l = 0, \dots, r$, consider the following non decreasing sequences:

$$Q_l^{(s)} := \frac{2}{\alpha^{(s)}} \mathcal{P}_s\Big(\mathbf{\Delta q}_l\Big) \tag{3.25}$$

$$Q_l := \Big(\mathbf{q}_l, \mathbf{\Delta q}_l\Big) \tag{3.26}$$

To complete the picture we need to introduce a transformed order parameter

$$x_\Delta(\mathbf{u}) := \sum_{l=0}^{r} (m_{l+1} - m_l) \prod_{s \in \mathcal{S}} \theta(u^{(s)} - Q_l^{(s)}) \tag{3.27}$$

defined for $\mathbf{u} \in \times_{s \in \mathcal{S}} [0, Q_r^{(s)}]$.

[1] Ziggurats are pyramid-like structures found in the ancient Mesopotamian valley and western Iranian plateau.

Definition 3.2.2. *For a given $x \in \mathcal{M}_d^S$, any inverse temperature β and external field h, we call the S-Parisi functional*

$$\mathscr{P}_S(x) := \log 2 + \sum_{s\in\mathcal{S}} \alpha(s) f^{(s)}(0, h^{(s)}) - \frac{1}{2}\sum_{l=0}^{r} m_l(Q_l - Q_{l-1}) \tag{3.28}$$

where, for each $s \in \mathcal{S}$, $f^{(s)}(u^{(s)}, y)$ is the solution of the following Parisi's PDE

$$\frac{\partial f^{(s)}}{\partial u^{(s)}} + \frac{1}{2}\frac{\partial^2 f^{(s)}}{\partial y^2} + \frac{1}{2} x_{\mathbf{\Delta}}(u^{(s)})\Big(\frac{\partial f^{(s)}}{\partial y}\Big)^2 = 0 \tag{3.29}$$

where $x_{\mathbf{\Delta}}(u^{(s)})$ is the marginal value of the transformed order parameter and the boundary condition is

$$f^{(s)}(Q_r^{(s)}, y) = \log\cosh(y). \tag{3.30}$$

Clearly, for each $f^{(s)}$ solution of (3.29), the analogous of (2.2.12) and 2.2.13 hold.

Remark 3.2.3. *We mention that it's possible to rewrite the second term in r.h.s. of* (3.30) *as*

$$\frac{1}{2}\sum_{l=0}^{r} m_l(Q_l - Q_{l-1}) = \frac{1}{2}\int_{\widetilde{\Gamma}} \quad x(\mathbf{u}) \quad \nabla_{\mathbf{u}}\Big(\mathbf{u}, \mathbf{\Delta}\mathbf{u}\Big) \cdot d\mathbf{u} \tag{3.31}$$

The integral is a line integral on an arbitrary path $\widetilde{\Gamma}$ in the plan $\mathbf{u}$, starting from $\mathbf{0}$ and ending in $\mathbf{1}$, such that all the points $(\mathbf{q}_l)_{l=0,\dots,r}$ belong to Γ, in other words $\Gamma \subset \widetilde{\Gamma}$.

This representation can be useful for a continuous extension (not proved) of the space of the order parameters defined in 3.2.1.

The main result on the MSK model is the following multidimensional generalization of Theorem 2.2.14.

Theorem 3.2.4. *If the matrix $\mathbf{\Delta}$ (3.14) is positive semi-definite and for each $s \in \mathcal{S}$ the $\lim_{N\to\infty} \alpha_N^{(s)} = \alpha^{(s)}$ exist, then*

$$\lim_{N\to\infty} \frac{\log Z_N}{N} = \lim_{N\to\infty} \mathbb{E}\,\frac{\log Z_N}{N} = \inf_{x\in\mathcal{M}_d^S} \mathscr{P}_S(x), \ \mathbb{P}-a.s.$$

where $\mathscr{P}_S(x)$ is the S-dimensional Parisi functional defined in (3.2.2).

Before the proof of this theorem and the related results, let us start with some comment on the previous statement. First, the infimum is taken over $r, \underline{m}, \Gamma$ defined 3.2.1. The non-negativity assumption on the matrix $\mathbf{\Delta}$ is independent of $\alpha_N^{(s)}$ and much more important, is a fundamental point in the application of the interpolation method (see Appendix 7.2) to prove the lower bound. Basically, is the same reason that restricts the Talagrand's proof [19] to even p-spin models. The second assumption say that the *t.l.* is to be understood with constant, at least for N large enough, relatives densities between the species.

We notice that statement 3.2.4 contain two non trivial hidden results, namely the existence of the *t.l.* of the quenched pressure density and the almost surely convergence of the random pressure density $N^{-1} \log Z_N$ to this limit. The almost surely convergence of the pressure density is a classical result in Spin Glass theory. Since is based on the concentration of gaussian measure [17], it holds for a large class of spin glass models, including the MSK model and we refer to the above reference for the details. The existence of the *t.l.* of p_N will be proved in the next section. We stress the fact that the *ziggurat* ansatz forces the joint overlap distribution to have a very special structure. In [109], the author proved that the previous bound is exact, showing that this structure in fact encodes all the information of the model.

3.3 The existence of the thermodynamical limit

In this section we prove, under the assumptions of Theorem 3.2.4, the existence of the thermodynamic limit for the quenched pressure density (3.22) when, $\forall s \in \mathcal{S}$, the limit $N \to \infty$ of the relative densities $\alpha_N^{(s)}$ (3.9) exist and equal to some fixed $\alpha^{(s)} \in (0, 1)$. In order to lighten the notation from now on we denote the relative densities with their own limit $\alpha^{(s)}$.

Theorem 3.3.1. *If the matrix* $\mathbf{\Delta}$ *is positive semi-definite, then*

$$\lim_{N\to\infty} p_N = \sup_N p_N,$$

where the limit is taken at fixed densities.

The strategy of the proof follows classical Guerra-Toninelli arguments. Let us consider two non interacting and *i.i.d.* copies of the original system defined by the Hamiltonian (3.5) of sizes respectively $N_1,\ N_2$. Clearly this implies that we have to consider $\forall s \in \mathcal{S}$ the relative subsets $\Lambda^{(s)}_{N_1}, \Lambda^{(s)}_{N_2}$ defined by the equations (3.1), (3.3), (3.4) and such that

$$\begin{aligned}
\Lambda^{(s)}_{N_1} \cup \Lambda^{(s)}_{N_2} &= \Lambda^{(s)}_{N}, \\
\Lambda^{(s)}_{N_1} \cap \Lambda^{(s)}_{N_2} &= \emptyset, \\
|\Lambda^{(s)}_{N_1}| &= N^{(s)}_1, \\
|\Lambda^{(s)}_{N_2}| &= N^{(s)}_2, \\
N^{(s)}_1 + N^{(s)}_2 &= N^{(s)}.
\end{aligned}$$

More explicitly, we can define, $\forall s \in \mathcal{S}$, the following

$$\Lambda^{(s)}_{N} = \{1, \dots, N^{(s)}\}, \tag{3.32}$$

$$\Lambda^{(s)}_{N_1} = \{1, \dots, N^{(s)}_1\}, \tag{3.33}$$

$$\Lambda^{(s)}_{N_2} = \{N^{(s)}_1 + 1, \dots, N^{(s)}\}. \tag{3.34}$$

Consider the following interpolating Hamiltonian

$$H_N(\sigma, t) = \sqrt{t} H_N(\sigma) + \sqrt{1-t}\Big(H_{N_1}(\sigma) + H_{N_2}(\sigma)\Big) \tag{3.35}$$

where

$$H_{N_1}(\sigma) = -\frac{1}{\sqrt{N_1}} \sum_{i,j\in\Lambda_{N_1}} J^{(1)}_{ij}\sigma_i\sigma_j, \tag{3.36}$$

$$H_{N_2}(\sigma) = -\frac{1}{\sqrt{N_2}} \sum_{i,j\in\Lambda_{N_2}} J^{(2)}_{ij}\sigma_i\sigma_j, \tag{3.37}$$

and where $J_{ij}^{(1)}$ and $J_{ij}^{(2)}$ are two independent families of centered gaussian defined by

$$\mathbb{E}(J_{ij}^{(a)} J_{i'j'}^{(a)}) = \delta_{ii'}\delta_{jj'}\Delta_{sp}^2 \mathbf{1}\{i \in \Lambda_{N_a}^{(s)}\}\mathbf{1}\{j \in \Lambda_{N_a}^{(p)}\}, \tag{3.38}$$

for each $a \in \{1, 2\}$.

As usual we consider the interpolating pressure

$$P_N(t) = \mathbb{E}\log Z_N(t) = \mathbb{E}\log \sum_{\sigma} a_N(\sigma, \mathbf{h}) e^{-H_N(\sigma,t)}, \tag{3.39}$$

whose boundaries values are

$$P_N(1) \equiv P_N, \tag{3.40}$$

$$P_N(0) \equiv P_{N_1} + P_{N_2}, \tag{3.41}$$

since $\Sigma_N = \Sigma_{N_1} \cup \Sigma_{N_2}$ and $\Sigma_{N_1} \cap \Sigma_{N_2} = \emptyset$.

Proposition 3.3.2. *Let us denote by $\Omega_{N,t}(\,\cdot\,)$ the replica Gibbs state* (3.20) *associated to the interpolating Hamiltonian* (3.35)*. The t-derivative of the interpolating pressure is*

$$\frac{\partial}{\partial t} P_N(t) = -\frac{N}{2}\mathbb{E}\Omega_{N,t}\Big(Q_N\Big),$$

where

$$Q_N(\sigma, \tau) := \Big(\mathbf{q}_N, \mathbf{\Delta q}_N\Big) - \frac{N_1}{N}\Big(\mathbf{q}_{N_1}, \mathbf{\Delta q}_{N_1}\Big) - \frac{N_2}{N}\Big(\mathbf{q}_{N_2}, \mathbf{\Delta q}_{N_2}\Big), \tag{3.42}$$

and the vectors $\mathbf{q}_{N_1}, \mathbf{q}_{N_2}$ are defined as in (3.13)*.*

Proof. The computation of the t-derivative works essentially in the same way exploited in Proposition 7.2.2 with the following identifications:

$$i \to \sigma, \quad a_i \to a_N(\sigma, \mathbf{h}), \quad U_i \to H_N(\sigma), \quad \widetilde{U}_i \to H_{N_1}(\sigma) + H_{N_2}(\sigma)$$

The key ingredient is that the diagonal term vanishes by the condition $N = N_1 + N_2$. □

Combining the Fundamental Theorem of Calculus and the previous proposition we have that

$$P_N - P_{N_1} - P_{N_2} = -\frac{N}{2}\int_0^1 dt\mathbb{E}\Omega_{N,t}\Big(Q_N\Big). \tag{3.43}$$

To finish the proof is sufficient to show that

Proposition 3.3.3. *If the matrix* $\mathbf{\Delta}$ *is positive semi-definite, then*

$$Q_N(\sigma,\tau) \leq 0 \tag{3.44}$$

for every σ,τ *and* N*.*

Proof. First at all, we write some fundamental relations.
By definitions (3.10), (3.32), (3.33), (3.34) we have that $\forall s \in \mathcal{S}$ the following hold

$$N^{(s)}q_N^{(s)}(\sigma,\tau) = \sum_{i=1}^{N^{(s)}} \sigma_i^{(s)}\tau_i^{(s)} = \sum_{i=1}^{N_1^{(s)}} \sigma_i^{(s)}\tau_i^{(s)} + \sum_{N^{(s)}+1}^{N^{(s)}} \sigma_i^{(s)}\tau_i^{(s)}$$

then

$$q_N^{(s)}(\sigma,\tau) = \frac{N_1^{(s)}}{N^{(s)}}q_{N_1}^{(s)}(\sigma,\tau) + \frac{N_2^{(s)}}{N^{(s)}}q_{N_2}^{(s)}(\sigma,\tau).$$

Now we observe that the condition of fixed relatives densities implies that

$$\frac{N_1^{(s)}}{N^{(s)}} = \frac{N_1^{(s)}}{N_1}\frac{N}{N^{(s)}}\frac{N_1}{N} = \frac{\alpha^{(s)}}{\alpha^{(s)}}\frac{N_1}{N} = \frac{N_1}{N},$$

and in a similar fashion

$$\frac{N_2^{(s)}}{N^{(s)}} = \frac{N_2}{N},$$

then $\forall s \in \mathcal{S}$ the following holds

$$q_N^{(s)}(\sigma,\tau) = \frac{N_1}{N}q_{N_1}^{(s)}(\sigma,\tau) + \frac{N_2}{N}q_{N_2}^{(s)}(\sigma,\tau). \tag{3.45}$$

In vector notation we can write

$$\mathbf{q}_N = \frac{N_1}{N}\mathbf{q}_{N_1} + \frac{N_2}{N}\mathbf{q}_{N_2}. \tag{3.46}$$

It is easy to see that if $\mathbf{\Delta}$ is a positive semi-definite, real, symmetric matrix, hence the function

$$\mathbf{x} \to \Big(\mathbf{x}, \mathbf{\Delta x}\Big)$$

defined for $\mathbf{x} \in \mathbb{R}^S$ is convex and the conclusion follows straightforwardly from the relation (3.46). □

The last proposition, combined with equation (3.35) gives immediately the superaddivity property of the pressure. As a consequence, since the quenched pressure density is bounded from the annealed one (see the next section), then by Fekete's lemma we get the statement of the theorem.

3.4 The annealed bound

As a first analysis we can study the annealed approximation for the pressure and investigate in which case it is exact. Using Jensen inequality and the concavity of the function $x \to \log(x)$ we define the annealed approximation as a bound, i.e.

$$p_N = \frac{1}{N}\mathbb{E}\log Z_N \leq \frac{1}{N}\log \mathbb{E}Z_N = p_N^A. \tag{3.47}$$

We can easily write p_N^A as

$$\begin{aligned} p_N^A &= \frac{1}{N}\log\sum_\sigma \mathbb{E}e^{-H_N(\sigma)} = \frac{1}{N}\log\sum_\sigma e^{\frac{1}{2}C_N(\sigma,\sigma)} = \frac{1}{N}\log\sum_\sigma e^{\frac{N}{2}(\mathbf{1},\mathbf{\Delta 1})} \\ &= \log 2 + \frac{1}{2}(\mathbf{1},\mathbf{\Delta 1}). \end{aligned} \tag{3.48}$$

We define the ergodic regime as the region of the phase space in which

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}Z_N = p^A = \log 2 + \frac{1}{2}(\mathbf{1},\mathbf{\Delta 1}). \tag{3.49}$$

For this purpose a classic application of the Borel-Cantelli lemma imply that it's enough to investigate the second moment, hence checking when

$$\frac{\mathbb{E}(Z_N^2)}{\mathbb{E}^2(Z_N)} \leq C < \infty \tag{3.50}$$

for some constant $C \in \mathbb{R}$, uniformly in N. Since

$$\begin{aligned}\mathbb{E}(Z_N^2) &= \mathbb{E}\sum_{\sigma,\tau} e^{-H_N(\sigma)-H_N(\tau)} = \sum_{\sigma,\tau} e^{\frac{1}{2}\mathbb{E}(H_N(\sigma)+H_N(\tau))^2} \\ &= \sum_{\sigma,\tau} e^{N((\mathbf{1},\boldsymbol{\Delta}\mathbf{1})+(\mathbf{q}_N,\boldsymbol{\Delta}\mathbf{q}_N))} = \mathbb{E}^2(Z_N)2^{-2N}\sum_{\sigma,\tau} e^{N(\mathbf{q}_N,\boldsymbol{\Delta}\mathbf{q}_N)}\end{aligned} \tag{3.51}$$

and using the gauge transformation $\tau_i^{(s)} \to \sigma_i^{(s)}\tau_i^{(s)}$,

$$\frac{\mathbb{E}(Z_N^2)}{\mathbb{E}^2(Z_N)} = 2^{-2N}\sum_{\sigma,\tau} e^{N(\mathbf{m}_N(\tau),\boldsymbol{\Delta}\mathbf{m}_N(\tau))} = 2^{-N}\sum_{\tau} e^{N(\mathbf{m}_N(\tau),\boldsymbol{\Delta}\mathbf{m}_N(\tau))}, \tag{3.52}$$

where we define $\mathbf{m}_N(\tau) = \left(m_N^{(s)}(\tau)\right)_{s\in\mathcal{S}}$, with $m_N^{(s)}(\tau) = \frac{1}{N^{(s)}}\sum_{i=1}^{N^{(s)}} \tau_i^{(s)}$. If $\det\boldsymbol{\Delta} > 0$ we can linearize the quadratic form with a gaussian integration

$$\begin{aligned}\frac{\mathbb{E}(Z_N^2)}{\mathbb{E}^2(Z_N)} &= \frac{2^{-N}}{\sqrt{\det\boldsymbol{\Delta}}}\int\frac{d\mathbf{z}}{2\pi}e^{-\frac{1}{2}(\mathbf{z},\boldsymbol{\Delta}^{-1}\mathbf{z})}\sum_{\tau} e^{\sqrt{2N}(\mathbf{m}_N(\tau),\mathbf{z})} \\ &= \frac{1}{\sqrt{\det\boldsymbol{\Delta}}}\int\frac{d\mathbf{z}}{2\pi}e^{-\frac{1}{2}(\mathbf{z},\boldsymbol{\Delta}^{-1}\mathbf{z})}\prod_{s\in\mathcal{S}}\cosh^{N^{(s)}}\left(\frac{\sqrt{2N}}{N^{(s)}}z^{(s)}\right) \\ &= \frac{1}{\sqrt{\det(\boldsymbol{\Delta})}}\int\frac{d\mathbf{z}}{2\pi}e^{-\frac{1}{2}(\mathbf{z},\boldsymbol{\Delta}^{-1}\mathbf{z})}e^{\sum_{s\in\mathcal{S}}N^{(s)}\log\cosh\left(\frac{\sqrt{2N}}{N^{(s)}}z^{(s)}\right)}\end{aligned} \tag{3.53}$$

and, using the inequality $\log\cosh(x) \le \frac{x^2}{2}$, we obtain

$$\frac{\mathbb{E}(Z_N^2)}{\mathbb{E}^2(Z_N)} \le \frac{1}{\sqrt{\det(\boldsymbol{\Delta})}}\int\frac{d\mathbf{z}}{2\pi}e^{-\frac{1}{2}(\mathbf{z},\hat{\boldsymbol{\Delta}}\mathbf{z})}, \tag{3.54}$$

where we have defined

$$\hat{\boldsymbol{\Delta}} = \boldsymbol{\Delta}^{-1} - 2\boldsymbol{\alpha}^{-1} \tag{3.55}$$

and the diagonal matrix $\boldsymbol{\alpha} = \mathrm{diag}(\{\alpha^{(s)}\}_{s\in\mathcal{S}})$. Thus we have just proved the following

Theorem 3.4.1. *In the convex region, defined as* $\det\boldsymbol{\Delta} > 0$*, as soon as* $\hat{\boldsymbol{\Delta}}$ *is positive definite, the pressure of the model does coincide with the annealed approximation, i.e.*

$$p = \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}Z_N = p^A = \log 2 + \frac{1}{2}(\mathbf{1},\boldsymbol{\Delta}\mathbf{1}). \tag{3.56}$$

Remark 3.4.2. Note that such a region does exist and can be viewed as an high temperature region. The two regions $\det \boldsymbol{\Delta} > 0$ and $\hat{\boldsymbol{\Delta}} > 0$ have a non-zero measure intersection, because, while the first is a condition on the relative size of the covariances, the latter is related to their absolute amplitude. Indeed once fixed $\boldsymbol{\alpha}$ and $\boldsymbol{\Delta}$ satisfying $\det \boldsymbol{\Delta} > 0$, we can rescale all the covariances with a parameter β, which play the role of the inverse temperature of the system, i.e. $\Delta_{ss'} \to \beta \Delta_{ss'}$, $\forall s, s' \in \mathcal{S}$, leaving the relative sizes unaltered and the condition $\det \boldsymbol{\Delta} > 0$ is still satisfied, such that $\hat{\boldsymbol{\Delta}} \to \beta^{-S} \boldsymbol{\Delta}^{-1} - 2\boldsymbol{\alpha}^{-1}$ is positive definite for β small enough [2].

3.5 The Replica Symmetric bound

In this section we show the simplest application of the *RSB sum rule* (Theorem 3.6.1 introduced in the next section). Namely we obtain the so called RS bound.

The underlying idea is to compare the overlap vector (3.13) with a trial vector,

$$\mathbf{q}_{trial} := \Big(q^{(s)}\Big)_{s \in \mathcal{S}}. \tag{3.57}$$

We define the trial replica symmetric solution as

$$p_{RS}(\mathbf{q}_{trial}) := \log 2 + \sum_{s \in \mathcal{S}} \alpha^{(s)} p^{(s)}(\mathbf{q}_{trial}) + \frac{1}{2}\Big((\mathbf{1} - \mathbf{q}_{trial}), \boldsymbol{\Delta}(\mathbf{1} - \mathbf{q}_{trial})\Big), \tag{3.58}$$

where

$$p^{(s)}(\mathbf{q}_{trial}) := \int d\mu(z) \log\cosh\Big(\sqrt{\frac{2}{\alpha^{(s)}} \mathcal{P}_s\Big(\boldsymbol{\Delta}\mathbf{q}_{trial}\Big)}\, z + h^{(s)}\Big), \tag{3.59}$$

and

$$z \sim \mathcal{N}(0, 1).$$

Setting $r = 2$, $m_1 = 0$, $m_2 = 1$ and $\mathbf{q}_1 = \mathbf{q}_{trial}$ in Theorem 3.6.1 we obtain the following:

[2] Since $\boldsymbol{\Delta}$ is positive definite then also $\boldsymbol{\Delta}^{-1}$. Defining $a = \max_s \alpha^{(s)}$ and ρ the smallest eigenvalue of $\boldsymbol{\Delta}^{-1}$, then, for any non-null vector z, $(z, \hat{\boldsymbol{\Delta}} z) \geq (\beta^{-S}\rho - a)(z, z) > 0$ if $\beta^S < \rho/a$.

Proposition 3.5.1. *The following sum rule holds*

$$p_N = p_{RS}(\mathbf{q}_{trial}) - \frac{1}{2}\int_0^1 \mathbb{E}\Omega_{N,t}\Big((\mathbf{q}_N - \mathbf{q}_{trial}), \boldsymbol{\Delta}(\mathbf{q}_N - \mathbf{q}_{trial})\Big)dt. \tag{3.60}$$

Moreover, if the matrix $\boldsymbol{\Delta}$ *is positive semi-definite, then the following bound holds*

$$p_N \leq p_{RS}(\mathbf{q}_{trial}), \tag{3.61}$$

whose optimization gives

$$p_N \leq \inf_{\mathbf{q}_{trial}} p_{RS}(\mathbf{q}_{trial}). \tag{3.62}$$

The optimization of (3.62) on $\mathbf{q}_{trial}$, gives a system of S coupled self consistent equations, i.e. $\forall p \in \mathcal{S}$

$$\sum_{s\in\mathcal{S}} \boldsymbol{\Delta}_{ps}\left[\int d\mu(z)\tanh^2\left(\sqrt{\frac{2}{\alpha^{(s)}}\mathcal{P}_s(\boldsymbol{\Delta}\mathbf{q}_{trial})}z\right) - q^{(s)}\right] = 0, \tag{3.63}$$

This system admits a unique solution as soon as $\det(\boldsymbol{\Delta}) \neq 0$, thus whenever $\det(\boldsymbol{\Delta}) > 0$, $p_{RS}(\mathbf{q}_{trial})$ has a minimum in $\mathbf{q}_{trial} = \bar{\mathbf{q}}$ satisfying $\forall s \in \mathcal{S}$

$$\bar{q}^{(s)} = \int d\mu(z)\tanh^2\left(\sqrt{\frac{2}{\alpha^{(s)}}\mathcal{P}_s(\boldsymbol{\Delta}\bar{\mathbf{q}})}z\right) = \langle \mathcal{P}_s(\mathbf{q})\rangle_{t=0} \tag{3.64}$$

The last equalities can be easily checked thanks to the factorizability of the one-body problem at $t = 0$. In other words, the value of $\mathbf{q}_{trial}$ minimizing the overlap' s fluctuations of the original model (at $t = 1$) is just the overlap's mean of the interpolating one-body trial at $t = 0$.

Let show now how the replica symmetric bound violate the entropy positivity at low temperatures. Mirroring the scenario of the SK model, we can easily check that the replica symmetric expression for the pressure (3.62) is not the exact solution of the model in the low temperature region by studying the behavior of the entropy. We can define it as the *non-negative* quantity

$$s(\boldsymbol{\Delta}) = \lim_{N\to\infty} s_N(\boldsymbol{\Delta}) = -\frac{1}{N}\mathbb{E}\sum_{\sigma}\mathcal{G}_N(\sigma,\boldsymbol{\Delta})\log(\mathcal{G}_N(\sigma,\boldsymbol{\Delta})), \tag{3.65}$$

where $\mathcal{G}_N(\sigma, \mathbf{\Delta}) = Z_N^{-1}(\mathbf{\Delta})e^{-H_N(\sigma,\mathbf{\Delta})}$ is the Gibbs measure (3.16). Notice that, unlike before, we have write explicitly the dependance on the matrix $\mathbf{\Delta}$. Since $s_N(\mathbf{\Delta}) = p_N(\mathbf{\Delta}) - \frac{1}{N}\langle H(\sigma)\rangle_N$, we can write

$$s(\mathbf{\Delta}) = p(\mathbf{\Delta}) - \frac{d}{d\lambda}p(\lambda\mathbf{\Delta})|_{\lambda=1}. \tag{3.66}$$

Now we can define $s_{RS}(\mathbf{\Delta}) = p_{RS} - \frac{d}{d\lambda}p_{RS}(\lambda\mathbf{\Delta})|_{\lambda=1}$. We can easily show that if the amplitude of the covariances is large enough, $s_{RS}(\mathbf{\Delta})$ is strictly negative. Indeed, we have the following

Proposition 3.5.2. *In the regime of large covariances (low temperatures), the RS-entropy is strictly negative, i.e.*

$$\lim_{\beta\to+\infty} s_{RS}(\beta\mathbf{\Delta}) < 0,$$

for any choice of $\mathbf{\Delta}$ *with* $(\det(\mathbf{\Delta}) > 0)$ *and* $\boldsymbol{\alpha}$*, where* $\beta \in \mathbb{R}^+$ *plays the role of the inverse temperature.*

Proof. Using its definition

$$s_{RS}(\beta\mathbf{\Delta}) = p_{RS}(\beta\mathbf{\Delta}, \bar{\mathbf{q}}) - \frac{\partial}{\partial\lambda}p_{RS}(\lambda\beta\mathbf{\Delta}, \bar{\mathbf{q}})|_{\lambda=1}.$$

We note that, using (3.64), in the limit $\beta \to +\infty$, the optimized order parameters $\bar{\mathbf{q}} \to \mathbf{1}$. Explicating the derivative it is easy to see that

$$\lim_{\beta\to+\infty} s_{RS}(\beta\mathbf{\Delta}) = \lim_{\beta\to+\infty} -\frac{\beta^2}{2}\left((\mathbf{1}-\bar{\mathbf{q}}), \mathbf{\Delta}(\mathbf{1}-\bar{\mathbf{q}})\right) \leq 0$$

Finally we can state that the limit is strictly negative, using again (3.64) and noting that

$$\begin{aligned}\lim_{\beta\to+\infty} \beta(1-\bar{q}_{(s)}) &= \lim_{\beta\to+\infty} \beta \int d\mu(z)\Big(1-\tanh^2\left(\beta\sqrt{\frac{2}{\alpha^{(s)}}\mathcal{P}_s\Big(\mathbf{\Delta}\bar{\mathbf{q}}\Big)}z\right)\Big) \\ &= \frac{\int d\mu(z)|z|}{\sqrt{\frac{2}{\alpha^{(s)}}\mathcal{P}_s\Big(\mathbf{\Delta}\mathbf{1}\Big)}} > 0.\end{aligned}$$

□

The existence of a negative RS-entropy regime is a clear signal that the model is not always replica symmetric (certainly it is RS inside the annealed region defined in Theorem 3.4.1) but there exists a region in which the pressure $p(\mathbf{\Delta})$ is strictly lower than its RS bound $p_{RS}(\mathbf{\Delta})$. The region of validity of the replica symmetric solution in the SK model is almost fully characterized in [18].

3.6 The Broken Replica Symmetry upper bound

Keeping in mind the definition of the *S-Parisi functional* $\mathscr{P}_S(x)$ given in section 3.2, the goal of this section is to prove the following relation:

Theorem 3.6.1.

$$p_N = \mathscr{P}_S(x) - \frac{1}{2}\sum_{l=0}^{r}(m_{l+1}-m_l)\int_0^1 dt\Big\langle(\mathbf{q}_N-\mathbf{q}_l),\mathbf{\Delta}(\mathbf{q}_N-\mathbf{q}_l)\Big\rangle_{N,l,t}. \quad (3.67)$$

where $\langle\cdot\rangle_{N,l,t}$ *is a quenched state associated to a suitable interpolating Hamiltonian.*

Moreover if the matrix $\mathbf{\Delta}$ *is positive semi-definite we have the following bound*

$$p_N \leq \mathscr{P}_S(x),$$

and the optimization gives

$$p_N \leq \inf_{x\in\mathcal{M}_d^S}\mathscr{P}_S(x).$$

It is enough to show that (3.67) holds, then we have straightforward conclusions. The strategy is to apply the RSB interpolation scheme introduced in Appendix 7.2.

We define the interpolating Hamiltonian as

$$H_N(\sigma,t) := \sqrt{t}H_N(\sigma) + \sqrt{1-t}\sum_{l=1}^{r}H_N^l(\sigma,\mathbf{q}_l), \quad (3.68)$$

with

$$H_N^l(\sigma,\mathbf{q}_l) := \sum_{s\in\mathcal{S}}H_N^{l,(s)}(\sigma^{(s)},\mathbf{q}_l), \quad (3.69)$$

where $H_N(\sigma)$ is the original Hamiltonian and, for each l, $H_N^{l,(s)}(\sigma^{(s)}, \mathbf{q}_l)$ are two independent one-body interaction Hamiltonian, defined as

$$H_N^{l,(s)}(\sigma^{(s)}, \mathbf{q}_l) := -\sqrt{2}\sqrt{\mathcal{P}_s\Big(\mathbf{\Delta}(\mathbf{q}_l - \mathbf{q}_{l-1})\Big)}\frac{1}{\sqrt{\alpha^{(s)}}}\sum_{i\in\Lambda_N^{(s)}} J_i^{l,(s)}\sigma_i^{(s)} \tag{3.70}$$

where the J's are *Gaussian i.i.d.* r.v., independent of the other r.v., such that for every l, s and i we have that

$$\mathbb{E}(J_i^{l,(s)}) = 0 \tag{3.71}$$

and

$$\mathbb{E}(J_i^{l,(s)} J_{i'}^{l',(s')}) = \delta_{ll'}\delta_{ss'}\delta_{ii'}. \tag{3.72}$$

After simple computations, we get

$$\mathbb{E}\Big(H_N^{l,(s)}(\sigma^{(s)}, \mathbf{q}_l)H_N^{l',(s')}(\tau^{(s')}, \mathbf{q}_l)\Big) = \delta_{ll'}\delta_{ss'}2N\mathcal{P}_s\Big(\mathbf{\Delta}(\mathbf{q}_l - \mathbf{q}_{l-1})\Big)\mathcal{P}_s\Big(\mathbf{q}_N\Big)$$

and then by (D13) the covariance matrix of the trial Hamiltonian becomes

$$\mathbb{E}\Big(H_N^l(\sigma, \mathbf{q}_l)H_N^{l'}(\tau, \mathbf{q}_{l'})\Big) = \delta_{ll'}2N\Big((\mathbf{q}_l - \mathbf{q}_{l-1}), \mathbf{\Delta}\mathbf{q}_N\Big).$$

Keeping in mind Proposition 7.2.4, we introduce the RSB interpolation scheme with the following identifications:

$$i \to \sigma, \quad a_i \to a_N(\sigma, \mathbf{h}), \quad U_i \to H_N(\sigma), \quad B_i^l \to H_N^l(\sigma, \mathbf{q}_l)$$

and we define the interpolating pressure as

$$p_N(t) := \frac{1}{N}\mathbb{E}\log Z_{0,N}(t). \tag{3.73}$$

It is easy to check that the boundary values of $p_N(t)$ are

$$p_N(1) = p_N, \tag{3.74}$$

$$p_N(0) = \log 2 + \sum_{s\in\mathcal{S}} \alpha^{(s)} f^{(s)}(0, h^{(s)}), \tag{3.75}$$

where $f^{(s)}(u^{(s)}, h^{(s)})$ is the solution of the Parisi's PDE (3.29).
In order to apply the interpolation argument we have to compute the t-derivative of the interpolating pressure. A simple application of Proposition 7.2.4 leads to the following

$$\frac{\partial}{\partial t}p_N(t) = -\frac{1}{2}\Big(\mathbf{1}, \mathbf{\Delta 1}\Big) - \frac{1}{2}\sum_{l=0}^{r}(m_{l+1} - m_l)\Big\langle \Big(\mathbf{q}_N, \Delta\mathbf{q}_N\Big) - 2\Big(\mathbf{q}_N, \Delta\mathbf{q}_l\Big)\Big\rangle_{N,l,t}. \tag{3.76}$$

To finish the proof of the Theorem, simply add and subtract the term $\frac{1}{2}\sum_{l=0}^{r}(m_{l+1} - m_l)Q_l$ where $Q_l = \Big(\mathbf{q}_l, \mathbf{\Delta q}_l\Big)$ and observe that

$$-\frac{1}{2}\Big(\mathbf{1}, \mathbf{\Delta 1}\Big) + \frac{1}{2}\sum_{l=0}^{r}(m_{l+1} - m_l)Q_l = \frac{1}{2}\sum_{l=0}^{r}(m_{l+1}(Q_l - Q_l)$$

The integral representation is given by the following

Proposition 3.6.2. *The following representation holds*

$$-\frac{1}{2}\Big(\mathbf{1}, \mathbf{\Delta 1}\Big) + \frac{1}{2}\sum_{l=0}^{r}(m_{l+1} - m_l)\Big(\mathbf{q}_l, \mathbf{\Delta q}_l\Big) = -\frac{1}{2}\int_{\widetilde{\Gamma}} d\mathbf{u} \;\; x(\mathbf{u}) \;\; \nabla_{\mathbf{u}}\Big(\mathbf{u}, \mathbf{\Delta u}\Big)\cdot d\mathbf{u}.$$

Proof. We can use the explicit definition of $x(\mathbf{u})$ given in (3.24) to check that

$$-\frac{1}{2}\int_{\widetilde{\Gamma}} \;\; x(\mathbf{u}) \;\; \nabla_{\mathbf{u}}\Big(\mathbf{u}, \mathbf{\Delta u}\Big)\cdot d\mathbf{u} = -\frac{1}{2}\sum_{l=0}^{r}(m_{l+1} - m_l)\int_{\Gamma_l} \;\; \nabla_{\mathbf{u}}\Big(\mathbf{u}, \mathbf{\Delta u}\Big)\cdot d\mathbf{u}$$

where Γ_l is the result of the action of the θ's on the path $\widetilde{\Gamma}$, that is his component between the points $\mathbf{q}_l$ and $\mathbf{1}$. By the Gradient's Theorem, the integral is path independent and is equal to the increment of the potential function, that is the desired result. □

Finally, combining (3.74), (3.75), (3.76) and Proposition 3.6.2, the proof of (3.67) is a simple application of the fundamental theorem of calculus.

3.7 The lower bound and the synchronization

In [109], Panchenko proved that in fact the upper bound given in the previous section is exact showing that it is also a lower bound for the limiting pressure of the MSK model. The strategy of the proof is similar to the one given in [16] for SK model . In particular, it is based on the *Aizenman-Sims-Starr scheme* [24] variational principle together with the Ghirlanda-Guerra identities and related *ultrametric property*. In particular, the show that one can construct a multi-species version of the previous these identities which imply the so-called *synchronization property* 4.5.6, by consequence the overlap between different species are strongly correlated. We suppose that the reader is familiar with these concepts, otherwise a detailed description is given in Chapter 4 an references therein. In this section we only sketch the main idea behind the Panchenko's proof and we refer to the original work for the details.

First of all, the reader should keep in mind all the quantities involved in the construction of the S-Parisi functional $\mathscr{P}_S(x)$ 3.28, and the order parameter x (3.24), namely the non decreasing sequences $m_l, q_l^{(s)}, Q_l^{(s)}$ and Q_l defined in 3.2.1. It's possible to show that, for any choice of the above sequences, the corresponding S-Parisi functional can be rewritten using the Derrida-Ruelle Probability Cascades (usually denoted by RPC) [116, 16]. Here a sketch of the construction.

For a given $r \geq 1$ and $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_r) \in \mathbb{N}^r$, let $(\nu_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^r}$ be the random weights of the RPC associated to the sequence $\underline{m} = (m_l)_{l=1,\ldots,r}$ (see for example, section 2.3 in [16] for the definition). For each $1 \leq l \leq r$, let us denote by $\boldsymbol{\alpha}_l = (\alpha_1, \ldots, \alpha_l)$ the first l coordinates of $\boldsymbol{\alpha}$.

For each $s \in S$ consider two independent gaussian process $C^s(\boldsymbol{\alpha})$ and $D(\boldsymbol{\alpha})$ indexed by $\boldsymbol{\alpha} \in \mathbb{N}^r$ with covariance matrix given by

$$\mathbb{E}C^s(\boldsymbol{\alpha}^1)C^s(\boldsymbol{\alpha}^2) = Q^{(s)}_{\boldsymbol{\alpha}^1\wedge\boldsymbol{\alpha}^2} \tag{3.77}$$

$$\mathbb{E}D(\boldsymbol{\alpha}^1)D(\boldsymbol{\alpha}^2) = Q_{\boldsymbol{\alpha}^1\wedge\boldsymbol{\alpha}^2} \tag{3.78}$$

where $\boldsymbol{\alpha}^1 \wedge \boldsymbol{\alpha}^2 := \min\{0 \leq l' \leq r \,|\, \boldsymbol{\alpha}^1_l = \boldsymbol{\alpha}^2_{l'}, \ldots, \boldsymbol{\alpha}^1_l = \boldsymbol{\alpha}^2_{l'}, \boldsymbol{\alpha}^1_{l+1} \neq \boldsymbol{\alpha}^2_{l'+1}\}$.

We notice that $Q^{(s)}_{\boldsymbol{\alpha}^1 \wedge \boldsymbol{\alpha}^2}$ and $Q_{\boldsymbol{\alpha}^1 \wedge \boldsymbol{\alpha}^2}$ are random variable w.r.t. the RPC random weights, taking values respectively in the sequences $Q^{(s)}_l$ and Q_l.

Let $\alpha^{(s)}$ (don't confuse it with the RPC weights) be the relative density of the species s and suppose for the moment that is a rational number. Let us consider a finite set of natural number $I = \bigcup_{s \in \mathcal{S}} I_s$, where I_s are disjoint sets such that for each $s \in \mathcal{S}$, $\frac{|I_s|}{|I|} = \alpha^{(s)}$. For each $s \in \mathcal{S}$ and $i \in I_s$ let $C_i(\boldsymbol{\alpha})$ a copy of $C^s(\boldsymbol{\alpha}^1)$ and suppose that all these process are independent each others.

Now using the invariance properties of the RPC (Theorem 2.9 in [16]) one can prove that

$$\mathscr{P}_S(x) = \frac{1}{|I|}\Big(\mathbb{E}\log \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \nu_{\boldsymbol{\alpha}}) \prod_{i \leq N} 2\cosh(C_i(\boldsymbol{\alpha})) - \mathbb{E}\log \sum_{\boldsymbol{\alpha} \in \mathbb{N}^r} \nu_{\boldsymbol{\alpha}}) \exp(\sqrt{|I|} D(\boldsymbol{\alpha}))\Big) \tag{3.79}$$

The connection between the previous representation of the S-Parisi functional and the pressure of the MSK model is given by the so-called *Aizenmann-Simms-Starr scheme* or *cavity method*.

Let us start observing that one can consider small changes of the relative densities $(\alpha_s)_{s \in \mathcal{S}}$ since they don't affect both the pressure for large N and the S-Parisi functional (see definition 3.2.2). Thus, without loss, we can assume that all α_s are rational:

$$\alpha^{(s)} = \frac{k_s}{k} \tag{3.80}$$

for some integers $k_s \leq k$, that is we suppose that $N(s) = nk_s$ and $N = nk$ for some integer n.

Let us recall the following statement contained in the Stoltz-Cesaro Theorem [115]:

Lemma 3.7.1. *Let $(a_n)_{n \geq 1}$ and $(b_n)_{n \geq 1}$ be two sequences of real numbers. Assume that $(b_n)_{n \geq 1}$ is strictly increasing and approaches $+\infty$. Then:*

$$\liminf_{n \to \infty} \frac{a_n}{b_n} \geq \liminf_{n \to \infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} \tag{3.81}$$

Let p_N be the pressure densities of the MSK model defined in (3.22) and, for a fixed k in (3.80), we set

$$a_n \equiv \sum_{j=0}^{n} \mathbb{E}\log Z_{k(j+1)} - \mathbb{E}\log Z_{kj},\ b_n \equiv nk$$

with the convention $Z_0 = 1$. Applying the previous lemma we get the following basic relation

$$\liminf_{N\to\infty} p_N \geq \frac{1}{k} \liminf_{n\to\infty} \mathbb{E}\log Z_{k(n+1)} - \mathbb{E}\log Z_{kn} \tag{3.82}$$

The r.h.s. of (3.82) quantify the effect on the system of the addition of k spins. This kind of argument in physics is called *cavity method*. The gain is that the difference between Z_{N+k} and Z_N can be rewritten as an expectation w.r.t a suitable Gibbs measure.

First, let us consider the r.h.s. of (3.82) for a fixed $N = nk$ and consider a partition I of the k new spins, namely

$$I := \{N+1, ..., N+k\} = \bigcup_{s\in\mathcal{S}} I_s$$

into different species, so that $|I_s| = k_s$.

Using standard property of the gaussian process, one can write, up to terms of order $o(1)$

$$\mathbb{E}\log Z_{k(n+1)} - \mathbb{E}\log Z_{kn} = \mathbb{E}\log \Omega'_N\Big(\prod_{i\in I} 2\cosh\big(z_{N,i}(\sigma)+h_i\big)\Big) - \mathbb{E}\log\Omega'_N\Big(\exp\big(y_N(\sigma)\big)\Big) \tag{3.83}$$

where for each $s \in S$, $h_i \equiv h^s \mathbf{1}\{i \in \Lambda_N^{(s)}\}$ and Ω'_N is the the random Gibbs measure (3.16) associated to the a slightly modified Hamiltonian H'_N defined as

$$H'_N(\sigma) = \frac{\sqrt{N}}{\sqrt{N+k}} H_N(\sigma)$$

and $H_N(\sigma)$ is the original Hamiltonian of the MSK model defined in 3.5. The quantities $z_{N,i}(\sigma)$ and $y_N(\sigma)$ are gaussian process independent of each other and

the randomness of H'_N, thus they are determined by their respective covariance matrix. For any $s \in \mathcal{S}$ and each $i \in I_s$,

$$\mathbb{E}z_{N,i}(\sigma)z_{N,i}(\tau) = 2\sum_{p\in\mathcal{S}} \Delta_{sp}^2 \alpha^{(p)} q_N^{(p)}(\sigma,\tau) = \frac{2}{\alpha^{(s)}} \mathcal{P}_s\Big(\mathbf{\Delta q}_N\Big) + O(N^{-1}) \quad (3.84)$$

where $\mathbf{q}_N$ is the *overlap vector* defined in 3.13. The process $z_{N,i}(\sigma)$ is usually called *cavity field*, while $y_N(\sigma)$ is called *fugacity* and it is determined by

$$\mathbb{E}y_N(\sigma)y_N(\tau) = k\sum_{s,p\in\mathcal{S}} \Delta_{sp}^2 \alpha^{(s)}\alpha^{(p)}\, q^{(p)}(\sigma,\tau) q_N^{(s)}(\sigma,\tau) = \big(\mathbf{q}_N, \mathbf{\Delta q}_N\Big) + O(N^{-1}) \quad (3.85)$$

As usual, for $l, l' \geq 1$ and each $s \in S$, consider two spin configuration $\sigma^l, \sigma^{l'} \in \Sigma_N$ and let us consider the overlap matrix

$$R^{N,(s)} = (R^{N,(s)}_{l,l'})_{l,l'\geq 1} := \Big(q_N^{(s)}(\sigma^l, \sigma^{l'})\Big)_{l,l'\geq 1}. \quad (3.86)$$

One can prove (Theorem 1.3 in [16]) that

$$\mathbb{E}\log\Omega'_N\Big(\prod_{i\in I} 2\cosh\big(z_{N,i}(\sigma) + h_i\big)\Big) - \mathbb{E}\log\Omega'_N\Big(\exp\big(y_N(\sigma)\big)\Big)$$

is a continuous functional of the distribution of the random array

$$R^N = \Big(R^{N,(s)}_{l,l'}\Big)_{l,l'\geq 1, s\in\mathcal{S}} \quad (3.87)$$

under the measure $\mathbb{E}\Omega'_N$.

The basic observation is that the r.h.s. of (3.79) and (3.83) have the same structure. Indeed, the S-Parisi functional is expressed, trough (3.83), as a functional of the random array

$$Q = \big(Q^{(s)}_{\boldsymbol{\alpha}^l \wedge \boldsymbol{\alpha}^{l'}}\big)_{l,l'\geq 1, s\in\mathcal{S}} \quad (3.88)$$

$$(3.89)$$

In other words, Q is the analogous of the overlap array R^N and the RPC random weights $\nu_{\boldsymbol{\alpha}}$ play the role of the random Gibbs measure. Since the r.h.s.

of (3.79) and (3.83) are expressed by the same functional, it's "enough" to prove that the, for large N, the distribution of R_N under $\mathbb{E}\Omega'_N$ can be approximated by the distribution of Q under $\nu_{\boldsymbol{\alpha}}$ for a suitable RPC. This approximation imply that

$$\frac{1}{k}\liminf_{n\to\infty}\mathbb{E}\log Z_{k(n+1)} - \mathbb{E}\log Z_{kn} \geq \mathscr{P}_S(x) \geq \inf_x \mathscr{P}_S(x)$$

that together with (3.82), gives the desired lower bound for the pressure of the MSK model.

The main issue in the proof is that, we have to understand the (joint) distribution of the overlap array R^N. Let us start with following observation. For all $s \in \mathcal{S}$, the randomness of $Q^{(s)}_{\boldsymbol{\alpha}^l\wedge\boldsymbol{\alpha}^{l'}}$ comes from $\boldsymbol{\alpha}^l \wedge \boldsymbol{\alpha}^l$ which is the hierarchical address of the same RPC, then they are deterministic function of the same random variable. Thus, the joint distribution of the array $Q_{l,l'}$ is uniquely determined by the law of the entry:

$$\mathbb{P}\{Q^{(s)}_{1,2} = Q^{(s)}_l\} = m_{l+1} - m_l \tag{3.90}$$

for $s \in \mathcal{S}$ and $l = 0, \dots r$.

Upon reflection, one can realize that this strong correlation is the essence of the particular structure of the order parameter (3.24). The point is that, roughly speaking, since we want to use Q to approximate R_N, we have to make sure that R_N share the same property, namely that $R^{N,(s)}_{l,l'}$ are deterministic functions of a single random variable. This is precisely the content of the *synchronization property* described in Theorem 4.5.6.

The main idea follows the line of the simpler case of the SK model. Namely, we add some perturbation to the original Hamiltonian in a such a way that, without effects on the *t.l.* of the pressure, we force the limiting law of $R^{N,(s)}$, (for each subsequences without assuming of the existence of the limit) to satisfy the *Multi-species Ghirlanda-Guerra identities* 4.5.5 and allow us to use Theorem 4.5.6. Thus, let us consider the total overlap of the system

$$\widetilde{R}_{l,l'} = \sum_{s\in\mathcal{S}} R^{(s)}_{l,l'}. \tag{3.91}$$

and the array

$$\widetilde{R} = (\widetilde{R}_{l,l'})_{l,l'\geq 1} \tag{3.92}$$

then, by Theorem 4.5.6, for all $s \in \mathcal{S}$, there exist non-decreasing $(1/\alpha_s)$-Lipschitz functions L_s such that $R^{(s)} = L_s(\widetilde{R})$ almost surely for all $s \in \mathcal{S}$ and all $l, l' \geq 1$. Moreover $\widetilde{R}$ can be generated by a suitable RPC.

We would mention that there are several technical issues in the control of the sequences involved in the proof, since basically, we don't know the existence of the limit and then we have to work with subsequences and approximation arguments. An interested reader can consult the original work [109] and reference therein for the solution of this important part of the problem.

Chapter 4

Factorization properties of Spin Glasses

One the most important properties of Gaussian Spin Glasses is as set of identities which involves the *t.l.* of the quenched Gibbs measure called *Ghirlanda-Guerra identities*[76]. These follows from two general properties of Spin Glasses: the Aizenman-Contucci stochastic stability [22] and the concentration of gaussian measures [8]. In this chapter we give a proof of these identities for a large class of models. The first section shows the underlying physical ideas, section 2 contains the main results and their proof. In section 3 we illustrate the deep consequences of these identities which are essential in the proof of the lower bound of the pressure for the MSK model given in section 3.7.

4.1 The concept of stability

Let us first illustrate the statistical physics ideas that we are following. The factorization laws that we deal with can be understood as consequences of a simple stability method. Stability in Statistical Mechanics works by identifying a small (yet non-trivial) deformation of the system, prove that in the large volume limit the perturbation vanishes and, by means of the linear response

theory, compute the relations among observable quantities. This method leads to interesting consequences and applications because it reduces the *a priori* degrees of freedom of a theory. Following the ideas developed in [61] one starts considering a deterministic spin glass glass model on a graph (1.1.2) with vertex set $\Lambda = \{1 \dots, N\}$ and generic Hamiltonian H_N. For any bounded function f of spin configurations $\sigma \in \Sigma_N$ let us denote by μ_N the counting measure on Σ_N, namely

$$\mu_N(f) = \frac{1}{2^N} \sum_{\sigma} f(\sigma), \tag{4.1}$$

and defines the equilibrium state

$$\omega_{\beta,N}(f) = \frac{\mu_N(f e^{-\beta H_N})}{\mu_N(e^{-\beta H_N})}. \tag{4.2}$$

By considering the Hamiltonian per particle

$$h_N(\sigma) = \frac{H_N(\sigma)}{N} \tag{4.3}$$

the classical perturbed state is defined by

$$\omega^{(\lambda)}_{\beta,N}(f) = \frac{\omega_{\beta,N}(f e^{-\lambda h_N})}{\omega_{\beta,N}(e^{-\lambda h_N})}. \tag{4.4}$$

Since the perturbation amounts to a small change in the temperature

$$\omega^{(\lambda)}_{\beta,N}(f) = \omega_{\beta+\frac{\lambda}{N},N}(f) \tag{4.5}$$

one has that, apart from isolated singularity points, in the thermodynamic limit

$$\frac{d\omega^{(\lambda)}_{\beta,N}(f)}{d\lambda} \to 0. \tag{4.6}$$

One may appreciate the content of the previous property by showing that it implies, for the Curie-Weiss ferromagnetic model in zero magnetic field, the relation

$$\omega_\beta(\sigma_1\sigma_2\sigma_3\sigma_4) = \omega_\beta(\sigma_1\sigma_2)^2 \ . \tag{4.7}$$

Hence, although the magnetization itself may fail to concentrate due to a spin-flip symmetry breaking, the square of the magnetization does concentrate in the thermodynamic limit.

The previous approach leads to the concept of *Stochastic Stability* when applied, suitably adapted, to the spin glass phase. Consider, for smooth bounded function f of n spin configurations, the quenched equilibrium state

$$< f >_{\beta,N} = \mathbb{E}\left(\frac{\sum_\sigma f(\sigma) e^{-\beta H_N(\sigma)}}{\sum_\sigma e^{-\beta H_N(\sigma)}}\right) . \tag{4.8}$$

Define the deformation as:

$$< f >^{(\lambda)}_{\beta,N} = \frac{< f e^{-\lambda h_N} >}{< e^{-\lambda h_N} >}. \tag{4.9}$$

We observe that the previous deformation is, unlike in the classical case, not a simple temperature shift. In fact:

$$< f >^{(\lambda)}_{\beta,N} = \frac{\mathbb{E}\left(\frac{\sum_\sigma f(\sigma) e^{-(\beta+\lambda/N) H_N(\sigma)}}{\sum_\sigma e^{-\beta H_N(\sigma)}}\right)}{\mathbb{E}\left(\frac{\sum_\sigma e^{-(\beta+\lambda/N) H_N(\sigma)}}{\sum_\sigma e^{-\beta H_N(\sigma)}}\right)} ; \tag{4.10}$$

nevertheless, the system is still stable with respect to it in a sense that will be made precise in the following sections and is essentially captured by saying that apart from isolated singularity points, in the thermodynamic limit

$$\frac{d}{d\lambda} < f >^{(\lambda)}_{\beta,N} \to 0 . \tag{4.11}$$

Moreover the previous stability property implies (by use of the integration by parts techinque) that the following set of identities (Ghirlanda-Guerra), first derived in [76], holds:

$$< f c_{1,n+1} >_{\beta,N} = \frac{1}{n} < f >_{\beta,N} < c_{12} >_{\beta,N} + \tag{4.12}$$

$$+ \frac{1}{n} \sum_{j=2}^{n} < f c_{1,j} >_{\beta,N} ,$$

where the term $c_{1,n+1}$ is the overlap between a spin configuration of the set $\{1, 2, ..., n\}$ and and external one that we enumerate as the $(n+1)$-st, and $c_{1,j}$ is the overlap between two generic spin configurations among the n's.

The proof ideas can be easily summarized by the study of three quantities and their differences which encode the fluctuation properties of the spin glass system:

$$\mathbb{E}\left[\omega(H_N^2)\right] , \tag{4.13}$$

$$\mathbb{E}\left[\omega(H_N)^2\right] \ , \tag{4.14}$$

$$\mathbb{E}\left[\omega(H_N)\right]^2 \ . \tag{4.15}$$

The result is obtained by two bounds for constants $\epsilon_N^{(1)}$ and $\epsilon_N^{(2)}$ vanishing in the $N \to \infty$ limit:

- bound on averaged thermal fluctuations

$$\mathbb{E}\left[\omega(H_N^2) - \omega(H_N)^2\right] \leq \epsilon_N^{(1)} N \tag{4.16}$$

  obtained by stochastic Stochastic Stability method (see [22]) by showing that the addition of an independent term of order one to the Hamiltonian is equivalent to a small change in temperature of the entire system:

$$\beta H_N(\sigma) \to \beta H_N(\sigma) + \sqrt{\frac{\lambda}{N}} \tilde{H}_N(\sigma) \tag{4.17}$$

$$\beta \to \sqrt{\beta^2 + \frac{\lambda}{N}} \tag{4.18}$$

- bound on disorder fluctuations

$$U = \omega(H_N) \tag{4.19}$$

$$\mathbb{E}(U^2) - \mathbb{E}(U)^2 \leq \epsilon_N^{(2)} N \, , \tag{4.20}$$

  which is the self averaging of internal energy and can be proved from self averaging of pressure (with martingale methods [13] or concentration of measures [18]).

In the following sections we show how to use the previous ideas to obtain a stronger result, namely the validity of the previous properties in distribution for quenched probability measure of the Hamiltonian covariance.

4.2 Definitions and preliminary properties

The class of models we chose to work includes the most general spin glass in d-dimension in the following sense. Physical particles in fact, beside interacting in pairs have always higher order interactions, i.e. they interact in triples, quadruples etc. (see [5]). Thus, during this section we leave the framework of models o graphs in favour on a more general one.

Given a lattice in dimension d, for example $\mathbb{Z}^d$ we consider, for each finite set $\Lambda \subset \mathbb{Z}^d$, an Hamiltonian of the form

$$H_\Lambda(\sigma) \;=\; -\sum_{X\subseteq\Lambda} J_{\Lambda,X}\sigma_X\,, \tag{4.21}$$

where $\sigma_X = \prod_{x\in X}\sigma_x$ and where all the random couplings $J_{\Lambda,X}$ are independent centered Gaussian random variables with $\mathbb{E}[J^2_{\Lambda,X}] \;=\; \Delta^2_{\Lambda,X}$ for some nonnegative constants $(\Delta_{\Lambda,X})_{X\subseteq\Lambda}$.

This class of models includes also all the Spin Glass model introduced in the previous chapters.

Example 3. • *The Edwards-Anderson model [EA]. The nearest neighborhood interaction is a two-body (pair) interaction then, first at all,* $\Delta^2_{\Lambda,X} = 0$ *unless* $|X| = 2$. *Next, we take* $\Delta^2_{\Lambda,X} = 1$ *if* $X = (n, n)$ *and* $|nn| = 1$ *and zero otherwise.*

• *The Sherringhton-Kirkpatrick model [SK]. Although it is not a finite dimensional model it may still be embedded in* $\mathbb{Z}$ *setting* $|\Lambda| = N$. *Next we take* $\Delta^2_{N,X} = N^1$ *if* $|X| = 2$ *and* $\Delta^2_{N,X} = 0$ *otherwise.*

• *The Multi-species SK model [MSK]. The same as the SK model with the suitable conditions on* $\Delta^2_{N,X}$.

The thermodynamical properties of the previous models are encoded in *normalized covariance* of the gaussian process.

$$c_\Lambda(\sigma,\tau) \;=\; \frac{1}{|\Lambda|}\mathbb{E}[H_\Lambda(\sigma)H_\Lambda(\tau)] \;=\; \frac{1}{|\Lambda|}\sum_{X\subseteq\Lambda}\Delta^2_{\Lambda,X}\sigma_X\tau_X\,. \tag{4.22}$$

We will often refer to this quantity also as *generalized overlap.*

The condition for existence of the thermodynamic limit (in the sense of Fisher) called "thermodynamic stability" is:

$$\sup_{\Lambda\subseteq\mathbb{Z}^d} \frac{1}{|\Lambda|} \sum_{X\subseteq\Lambda} \Delta^2_{\Lambda,X} \;\leq\; c \;<\; \infty\,, \tag{4.23}$$

see [60, 13].

In order to introduce the necessary language to illustrate our results we start by the following:

Lemma 4.2.1. *Let c_Λ and c'_Λ be two normalised covariances of Gaussian spin glasses satisfying the condition of thermodynamic stability. Then the same condition is satisfied by the normalised covariance obtained through the operations below:*

- $c_\Lambda + c'_\Lambda$, *entry-wise addition*
- $x^2 c_\Lambda$ *for each* $x \in \mathbb{R}$, *scalar multiplication*
- $c_\Lambda c'_\Lambda$, *entry-wise multiplication.*

Proof. We first observe that the three considered operations define new covariances, in particular the last covariance is sometimes called the Schur product, Hadamard product or Schur-Hadamard product and is semidefinite positive by a lemma of Schur.

The conditions of thermodynamic stability for c_Λ and c'_Λ are

$$\sup_\Lambda c_\Lambda(\sigma,\sigma) \;<\; \infty \quad \text{and} \quad \sup_\Lambda c'_\Lambda(\sigma,\sigma) \;<\infty\,,$$

and immediately imply that

$$\sup_\Lambda [c_\Lambda(\sigma,\sigma) + c'_\Lambda(\sigma,\sigma)] \;<\; \infty\,,$$

$$\sup_\Lambda x^2 c_\Lambda(\sigma,\sigma) \;<\; \infty \quad \text{and}$$

$$\sup_\Lambda c_\Lambda(\sigma,\sigma) c'_\Lambda(\sigma,\sigma) \;<\; \infty\,.$$

The explicit inversion formula from the covariance to the Hamiltonian can be seen from Chapter 2 of [13]. □

The previous lemma says that the set of the thermodynamically stable covariances is closed under the three operations defined.

By the previous lemma, starting from the Hamiltonian (4.21) we can always construct a thermodynamically stable Hamiltonian, that we call *complete Hamiltonian*, defined by

$$\bar{H}_\Lambda(\sigma;\beta) := \sum_{p\geq 1} (\sqrt{c})^{-p} \beta_p H_\Lambda^{(p)}(\sigma), \tag{4.24}$$

where $H_\Lambda^{(1)}(\sigma) \equiv H_\Lambda(\sigma)$ is the Hamiltonian (4.21) and each p-term in the sum has a normalized covariance

$$c_\Lambda^{(p)}(\sigma,\tau) = [c_\Lambda(\sigma,\tau)]^p$$

and the family of parameters $\beta = (\beta_p)_{p\geq 1}$ is such that $\beta_p > 0$ for every p and fulfills the condition

$$\sum_{p\geq 1} \beta_p^2 = \bar{c} < \infty$$

A simple computation shows that the complete Hamiltonian has a covariance

$$\bar{c}_\Lambda(\sigma,\tau) = \sum_{p\geq 1} (c)^{-p} \beta_p^2 [c_\Lambda(\sigma,\tau)]^p$$

and is thermodynamically stable with constant $\bar{c}$.

Example 4. *If one start with a simply one body random field, namely* $\Delta_\Lambda^2 = 1$ *if* $|X| = 1$ *and zero otherwise, thus applying the previous construction one obtains the mixed p-spin model* (2.12).

In order to facilitate the reader, let us briefly recall the general framework for disorder spin systems introduced in section 1.2.

Consider n copies of the configuration space whose elements are denoted by $\sigma^1, \ldots, \sigma^n$ and, for every bounded function $f : (\sigma^1, \ldots, \sigma^n) \to \mathbb{R}$, we call the random n-Gibbs state the following r.v.

$$\Omega_{\Lambda,\beta}(f) := \sum_{\sigma^1,\ldots,\sigma^n} f(\sigma^1, \ldots, \sigma^n) \mathcal{G}_{\Lambda,\beta}(\sigma^1) \ldots \mathcal{G}_{\Lambda,\beta}(\sigma^n) \tag{4.25}$$

where

$$\mathcal{G}_{\Lambda,\beta}(\sigma) := \frac{\exp(-\bar{H}_\Lambda(\sigma;\beta))}{\sum_\sigma \exp(-\bar{H}_\Lambda(\sigma;\beta))} \tag{4.26}$$

is the random Gibbs measure. In the previous formula the dependence on the physical β is reabsorbed in the family of β_p's.

We define the quenched Gibbs state as

$$\langle f \rangle_{\Lambda,\beta} := \mathbb{E}\Omega_{\Lambda,\beta}(f) \tag{4.27}$$

4.3 Identities

Theorem 4.3.1. *The model defined by equation (4.24) satisfies with respect to the covariance (4.22) the following properties:*

(i) It is stochastically stable in the strong sense, i.e for every power $p \in \mathbb{N}$ and for almost every β_p, the following hold

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \Big\langle \sum_{\substack{j,k=1 \\ j\neq k}}^{n} f c_{j,k}^p - 2nf \sum_{k=1}^{n} c_{k,n+1}^p + n(n+1) f c_{n+1,n+2}^p \Big\rangle_{\Lambda,\beta} = 0\,, \tag{4.28}$$

where for any number of replicas $\sigma^{(1)}, \sigma^{(2)}, \ldots,$ we denote by $c_{j,k}$ the quantity $c_\Lambda(\sigma^{(j)}, \sigma^{(k)})$, and where we assume that f is a continuous function of all the variables $c_{j,k}$ for $1 \le j < k \le n$.

(ii) It fulfills the Ghirlanda-Guerra identities (GG for short) in distribution, i.e. the following identities are verified for every $n \ge 2$ and every function f of $(c_{j,k})_{j,k=1}^n$ as above, and every power $p \in \mathbb{N}$ and for almost every β_p.

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \Big[\langle f c_{n+1,n+2}^p \rangle_{\Lambda,\beta} - \frac{1}{n+1} \sum_{k=1}^{n} \langle f c_{k,n+1}^p \rangle_{\Lambda,\beta} - \frac{1}{n+1} \langle f \rangle_{\Lambda,\beta} \langle c_{1,2}^p \rangle_{\Lambda,\beta} \Big] = 0\,, \tag{4.29}$$

Moreover, let

$$\bar{p}(\beta) := \lim_{\Lambda \nearrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \mathbb{E}[\log \sum_\sigma \exp(-\bar{H}_\Lambda(\sigma,\beta))]$$

be the thermodynamical limit of the pressure.

If $\bar{p}(\beta)$ is differentiable in the β_p "direction" at the point $\beta_p = a$ then

*(iii) It is **pointwise** stochastically stable in the strong sense, i.e for every power $p \in \mathbb{N}$ and in each point $\beta_p = a$, (4.28) hold.*

*(iv) It fulfills the Ghirlanda-Guerra identities in distribution **pointwise**, i.e for every $n \geq 2$ and for every power $p \in \mathbb{N}$ and in each point $\beta_p = a$, (4.29) holds.*

Remark 4.3.2. *we notice that since the function $\overline{p}(\beta)$ is convex in each β_p then it's almost everywhere differentiable and then we have that $(iv) \Rightarrow (ii)$, but in the next sections we give an independent proof of (ii).*

Proof of Theorem 4.3.1: We fix an arbitrary $p \geq 1$ and to lighten the notation we put:

$$\beta_p \to x$$
$$\sum_{k \neq p} (\sqrt{c})^{-k} \beta_k H_\Lambda^{(k)}(\sigma) \to H_\Lambda(\sigma)$$
$$(\sqrt{c})^{-p} \beta_p H_\Lambda^{(p)}(\sigma) \to H'_\Lambda(\sigma)$$

Then the Hamiltonian defined in (4.24) becomes

$$\bar{H}_\Lambda(\sigma;\beta) \to H_\Lambda(\sigma;x) = H_\Lambda(\sigma) + xH'_\Lambda(\sigma)$$

To prove the theorem we recall a general result due to Panchenko [104]. Consider a general Hamiltonian of the type

$$H_\Lambda(\sigma;x) = H_\Lambda(\sigma) + xH'_\Lambda(\sigma)$$

where x is a real parameter and the families $(H(\sigma))_\sigma$ and $(H'(\sigma))_\sigma$ are independent jointly Gaussian families of centered r.v.s. of the type (4.21). Suppose

that the Hamiltonian are thermodynamically stable in the sense of (4.23), that is there exists a global constant c such that

$$\begin{aligned} \mathbb{E}H_\Lambda(\sigma)^2 &\leq |\Lambda|c \\ \mathbb{E}H'_\Lambda(\sigma)^2 &\leq |\Lambda|c \end{aligned} \tag{4.30}$$

Consider, the following basic quantities:

$$Z_\Lambda(x) := \sum_\sigma \exp(-H_\Lambda(\sigma; x))$$

$$\begin{aligned} p_\Lambda(x) &:= \frac{1}{|\Lambda|}\log Z_\Lambda(x) \\ p(x) &:= \lim_{\Lambda \nearrow \mathbb{Z}^d} \mathbb{E}p_\Lambda(x) \end{aligned} \tag{4.31}$$

Notice that existence of the limit in the last definition is ensured (see for example [13]) by the conditions (4.30). We define $\Omega_{\Lambda,x}(\ )$ and $\langle\ \rangle_{\Lambda,x}$ in the same way as in (4.25) and (4.27).

In the previous setting we have the following lemma:

Lemma 4.3.3. *If we denote $h'_\Lambda(\sigma) := \frac{1}{|\Lambda|}H'_\Lambda(\sigma)$ the Hamiltonian density, then we have that for every $\beta_1 < \beta_2$*

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \int_{\beta_1}^{\beta_2} \left\langle \left| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) \right| \right\rangle_{\Lambda,a} da = 0 \tag{4.32}$$

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \int_{\beta_1}^{\beta_2} \left\langle \left| h'_\Lambda(\sigma) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a} \right| \right\rangle_{\Lambda,a} da = 0 \tag{4.33}$$

On other hand, if we assume that $p(x)$ is differentiable at $x = a$, then

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \left\langle \left| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) \right| \right\rangle_{\Lambda,a} = 0 \tag{4.34}$$

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \left\langle \left| h'_\Lambda(\sigma) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a} \right| \right\rangle_{\Lambda,a} = 0 \tag{4.35}$$

It is easy to check by a simple integration-by-parts and a uniform norm bound that the relations (4.32), (4.33), (4.34), (4.35) implies the propositions $i), ii), iii), iv)$ of Theorem 4.3.1, respectively. We notice that the propositions $i), ii)$ are in almost every sense then in this case the proof of previous implication requires some elementary facts in measure theory which are explained in Remark 4.4.4.

Proof of Lemma 4.3.3: The strategy of the proof is to control all terms by the following estimation, which is essentially contained in Chapter 12 of [19].

Proposition 4.3.4. *For every $b > 0$ we have that*

$$\Big\langle \Big| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big)\Big| \Big\rangle_{\Lambda,a} \leq \sqrt{\frac{2\mathbb{E}D_\Lambda(a,b)}{b|\Lambda|}} + 8\mathbb{E}D_\Lambda(a,b) \tag{4.36}$$

$$\Big\langle \Big| \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a}\Big| \Big\rangle_{\Lambda,a} \leq \mathbb{E}D_\Lambda(a,b) + \mathbb{E}W_\Lambda(a,b) \tag{4.37}$$

where

$$D_\Lambda(x,b) := p'_\Lambda(x+b) - p'_\Lambda(x-b)$$

$$W_\Lambda(x,b) := \frac{1}{b}\Big(|p_\Lambda(x+b) - \mathbb{E}p_\Lambda(x+b)| + |p_\Lambda(x-b) - \mathbb{E}p_\Lambda(x-b)| + |p_\Lambda(x) - \mathbb{E}p_\Lambda(x)|\Big)$$

Proof of Proposition 4.3.4, equation (4.37): the function $p_\Lambda(x)$ is convex. Thus for every $b > 0$ we have that

$$p'_\Lambda(x) \leq \frac{p_\Lambda(x+b) - p_\Lambda(x)}{b} \leq W_\Lambda(x,b) + \frac{\mathbb{E}\Big(p_\Lambda(x+b) - p_\Lambda(x)\Big)}{b} \leq W_\Lambda(x,b) + \mathbb{E}p'_\Lambda(x+b)\,,$$

and then

$$p'_\Lambda(x) - \mathbb{E}p'_\Lambda(x) \leq W_\Lambda(x,b) + \mathbb{E}p'_\Lambda(x+b) - \mathbb{E}p'_\Lambda(x) \leq W_\Lambda(x,b) + \mathbb{E}D_\Lambda(x,b)\,.$$

On other hand,

$$p'_\Lambda(x) \geq -\frac{p_\Lambda(x-b) + p_\Lambda(x)}{b}$$

and then, after the same manipulations, we get

$$p'_\Lambda(x) - \mathbb{E}p'_\Lambda(x) \geq -W_\Lambda(x,b) - \mathbb{E}D_\Lambda(x,b)$$

Combining the two previous inequalities we obtain the following bound

$$|p'_\Lambda(x) - \mathbb{E}p'_\Lambda(x)| \leq W_\Lambda(x,b) + \mathbb{E}D_\Lambda(x,b) \tag{4.38}$$

This bound give immediately (4.37), since

$$\Big\langle \Big| \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a}\Big| \Big\rangle_{\Lambda,a} = \mathbb{E}\Big(\Big|p'_\Lambda(a) - \mathbb{E}p'_\Lambda(a)\Big|\Big) \leq \mathbb{E}W_\Lambda(a,b) + \mathbb{E}D_\Lambda(a,b)\,. \quad \square$$

The bound (4.36) requires an extra work.

Proposition 4.3.5. *Consider the quantity*

$$\psi_\Lambda(x) := \Omega_{\Lambda,x}\Big(\Big|h'_\Lambda(\sigma^{(1)}) - h'_\Lambda(\sigma^{(2)})\Big|\Big)$$

then we have that

$$\psi^2_\Lambda(x) \leq \frac{4}{|\Lambda|} p''_\Lambda(x)$$
$$|\psi'_\Lambda(x)| \leq 8 p''_\Lambda(x)$$

Proof of Proposition 4.3.5: During this proof we define for sake of simplicity the quantity

$$V_l := h'_\Lambda(\sigma^{(l)}) - \Omega_{\Lambda,x}\Big(h'_\Lambda(\sigma)\Big)$$

then we have that, for every l, m

$$\Omega_{\Lambda,x}\Big(V_l\Big) = \Omega_{\Lambda,x}\Big(V_m\Big) = \Omega_{\Lambda,x}\Big(V_1\Big) = 0$$
$$\Omega_{\Lambda,x}\Big(V_l V_m\Big) = \Omega_{\Lambda,x}\Big(V_l\Big)\Omega_{\Lambda,x}\Big(V_m\Big) = \Omega^2_{\Lambda,x}\Big(V_1\Big) = 0$$
$$p''_\Lambda(x) = |\Lambda|\Omega_{\Lambda,x}\Big(V_1^2\Big)$$

By Jensen's inequality and the previous equations we can obtain the first bound of the proposition, indeed

$$\psi^2_\Lambda(x) \leq \Omega_{\Lambda,x}\Big(\Big(h'_\Lambda(\sigma^{(1)}) - h'_\Lambda(\sigma^{(2)})\Big)^2\Big) = \Omega_{\Lambda,x}\Big(\Big(V_1 - V_2\Big)^2\Big) \leq \frac{4}{|\Lambda|} p''_\Lambda(x)$$

The second bound follow easily by Cauchy-Schwarz inequality, indeed

$$|\psi'_\Lambda(x)| = |\Lambda|\Omega_{\Lambda,x}\Big(\Big|h'_\Lambda(\sigma^{(1)}) - h'_\Lambda(\sigma^{(2)})\Big|\Big(h'_\Lambda(\sigma^{(1)}) + h'_\Lambda(\sigma^{(2)}) - 2h'_\Lambda(\sigma^{(3)})\Big)\Big) \leq$$

$$|\Lambda|\Omega_{\Lambda,x}\Big(\Big|V_1 - V_2\Big| \cdot \Big|V_1 - V_3 + V_2 - V_3\Big|\Big) \leq 2|\Lambda|\Omega_{\Lambda,x}\Big(\Big|V_1 - V_2\Big|\Big|V_1 - V_3\Big|\Big)$$

$$\leq 2|\Lambda|\Omega_{\Lambda,x}\Big(\Big(V_1 - V_2\Big)^2\Big) \leq 8p''_\Lambda(x)\Box$$

The last proposition can be used to obtain

Proposition 4.3.6. *Given $b > 0$ then*

$$\mathbb{E}|\psi_\Lambda(x)| \leq \sqrt{\frac{2\mathbb{E}D_\Lambda(a,b)}{b|\Lambda|}} + 8\mathbb{E}D_\Lambda(a,b)$$

Proof of Proposition 4.3.6: We observe that for $x - b \leq y \leq x + b$ we have

$$|\psi_\Lambda(x) - \psi_\Lambda(y)| \int_{x-b}^{x+b} |\psi'_\Lambda(t)|dt$$

and then

$$\Big|\int_{x-b}^{x+b}\Big(\psi_\Lambda(x) - \psi_\Lambda(y)\Big)dy\Big| \leq 2b\int_{x-b}^{x+b}|\psi'_\Lambda(t)|dt$$

The identity

$$2b\psi_\Lambda(x) = \int_{x-b}^{x+b}\psi_\Lambda(y)dy + \int_{x-b}^{x+b}\Big(\psi_\Lambda(x) - \psi_\Lambda(y)\Big)dy$$

implies that

$$|\psi_\Lambda(x)| = \frac{1}{2b}\Big|\int_{x-b}^{x+b}\psi_\Lambda(y)dy\Big| + \frac{1}{2b}\Big|\int_{x-b}^{x+b}\Big(\psi_\Lambda(x) - \psi_\Lambda(y)\Big)dy\Big|$$

and then Proposition 4.3.5 implies that

$$|\psi_\Lambda(x)| \leq \frac{1}{b|\Lambda|^{\frac{1}{2}}}\int_{x-b}^{x+b}\sqrt{p''_\Lambda(y)}dy + 8\int_{x-b}^{x+b}p''_\Lambda(y)dy$$

We can use the Jensen inequality in the first term of the r.h.s of the previous relation to get

$$|\psi_\Lambda(x)| \leq \sqrt{\frac{2}{b|\Lambda|}}\Big(\int_{x-b}^{x+b}p''_\Lambda(y)dy\Big)^{\frac{1}{2}} + 8\int_{x-b}^{x+b}p''_\Lambda(y)dy$$

To conclude, we take the expectation and using again the Jensen inequality and the obvious relation

$$\int_{x-b}^{x+b} p''_\Lambda(y)dy = p'_\Lambda(x+b) - p'_\Lambda(x-b)$$

and the proof is complete.□

Proof of Proposition 4.3.4, equation (4.36): To obtain (4.36) we simply observe that by Jensen inequality

$$\Big\langle \Big| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big)\Big| \Big\rangle_{\Lambda,a} \leq \mathbb{E}|\psi_\Lambda(a)|$$

and then by Proposition 4.3.6 we get the desired result.□

Now we are able to prove Lemma 4.3.3.

Proof of Lemma 4.3.3: We start to prove the equations (4.32) and (4.34) which give the stochastic stability.

First, using the convexity of the function $p_\Lambda(x)$ we can prove easily that

$$D_\Lambda(a,b) \leq \frac{1}{b}\Big(p_\Lambda(a+2b) - p_\Lambda(a+b) + p_\Lambda(a-2b) - p_\Lambda(a-b)\Big) \tag{4.39}$$

It is easy to check that $\mathbb{E}p_\Lambda(x)$ is bounded for every x, Λ and then for every a, b, Λ we have that

$$\mathbb{E}D_\Lambda(a,b) \leq \frac{\bar{D}}{b} < \infty\,.$$

. Then from (4.36), using Fubini's Theorem, we get the following bound: for every $b, \beta_1 < \beta_2$

$$\int_{\beta_1}^{\beta_2} \Big\langle \Big| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big)\Big| \Big\rangle_{\Lambda,a} da \leq \frac{(\beta_2-\beta_1)}{b}\sqrt{\frac{2\bar{D}}{|\Lambda|}} + 8\mathbb{E}\int_{\beta_1}^{\beta_2} D_\Lambda(a,b)da$$

and then

$$\int_{\beta_1}^{\beta_2} \Big\langle \Big| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big)\Big| \Big\rangle_{\Lambda,a} da \leq$$

$$\frac{(\beta_2-\beta_1)}{b}\sqrt{\frac{2\bar{D}}{|\Lambda|}} + 8\mathbb{E}\Big(p_\Lambda(\beta_2+b) - p_\Lambda(\beta_1+b) - p_\Lambda(\beta_2-b) + p_\Lambda(\beta_1-b)\Big)$$

Finally, we can use the $\lim_{b\to 0}\limsup_{\Lambda\nearrow\mathbb{Z}^d}$ and the continuity of the function $p(x)$ to get (4.32).

To prove (4.34), we use the hypothesis of differentiability of the function $p(x)$ at the point $x = a$ to get from (4.39) the following

$$\lim_{b\to 0}\limsup_{\Lambda\nearrow\mathbb{Z}^d} \mathbb{E}D_\Lambda(a,b) \leq \lim_{b\to 0}\frac{1}{b}\Big(p(a+2b) - p(a+b) + p(a-2b) - p(a-b)\Big) = 0 \tag{4.40}$$

and then we can bypass the intermediate integration to obtain from (4.36) the following bound: for every a, b

$$\Big\langle \Big| h'_\Lambda(\sigma) - \Omega_{\Lambda,a}\Big(h'(\sigma)\Big)\Big| \Big\rangle_{\Lambda,a} \leq \frac{1}{b}\sqrt{\frac{2\bar{D}}{|\Lambda|}} + 8\mathbb{E}D_\Lambda(a,b)$$

Finally, we can use the $\lim_{b\to 0}\limsup_{\Lambda\nearrow\mathbb{Z}^d}$ and relation (4.40) to get (4.32).

Now, we are able to prove the equations (4.33) and (4.35) which give the *GG*-identities.

We already obtained the control of the quantity D, moreover a simple inspection shows that the quantity $W_\Lambda(a,b)$ is strictly related to the self-averaging of $p_\Lambda(x)$. Then it's easy to check (see for example [60]) that the thermodynamic stability condition (4.30) ensures that there exists a finite quantity $K(a,b,c)$ which does not depend on Λ, such that:

$$\mathbb{E}W_\Lambda(a,b) \leq K(a,b,c)\frac{1}{b|\Lambda|^{\frac{1}{2}}} \tag{4.41}$$

From (4.37) and (4.41), we can obtain as before the following bound: for every $b, \beta_1 < \beta_2$

$$\int_{\beta_1}^{\beta_2} \Big\langle \Big| \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a} \Big| \Big\rangle_{\Lambda,a} da \leq$$

$$\mathbb{E}\Big(p_\Lambda(\beta_2+b) - p_\Lambda(\beta_1+b) - p_\Lambda(\beta_2-b) + p_\Lambda(\beta_1-b)\Big) + \frac{\overline{K}(\beta_1,\beta_2,b,c)}{b|\Lambda|^{\frac{1}{2}}}$$

where we have set $\overline{K}(\beta_1,\beta_2,b,c) := \int_{\beta_1}^{\beta_2} K(a,b,c)da$ to lighten the notation. Finally, we can use the $\lim_{b\to 0}\limsup_{\Lambda\nearrow\mathbb{Z}^d}$ and the continuity of the function $p(x)$ to get (4.33).

Like in the previous case, the hypothesis of differentiability allow us to bypass the intermediate integration and then, from (4.37), we have that, for every a, b

$$\Big\langle \Big| \Omega_{\Lambda,a}\Big(h'(\sigma)\Big) - \langle h'_\Lambda(\sigma)\rangle_{\Lambda,a} \Big| \Big\rangle_{\Lambda,a} \leq \mathbb{E}D_\Lambda(a,b) + K(a,b,c)\frac{1}{b|\Lambda|^{\frac{1}{2}}}$$

Finally, we can use the $\lim_{b\to 0}\limsup_{\Lambda\nearrow\mathbb{Z}^d}$ and relation (4.40) to get (4.33).□

4.4 Rate of convergence

We outline in this section a sharper version of a theorem that appears in [19] and prove it with more elementary methods for the benefit of the reader, following the approach developed in [22, 60].

As in Section 3, we consider

$$H_\Lambda(\sigma;x) \;=\; H_\Lambda(\sigma) + xH'_\Lambda(\sigma)\,,$$

where H_Λ and H'_Λ are independent, and defined as in (2). The main theorem in this section follows:

Theorem 4.4.1. *(a) Writing $h'_\Lambda = |\Lambda|^{-1} H'_\Lambda$, as before,*

$$\int_{x_1}^{x_2} \left\langle \left| h'_\Lambda(\sigma) - \Omega_{\Lambda,x}\big(h'_\Lambda(\sigma)\big) \right|^2 \right\rangle_{\Lambda,x} dx \;\le\; \frac{2(|x_1|+|x_2|)\overline{c}'_\Lambda}{|\Lambda|}\,, \tag{4.42}$$

where $c'_\Lambda(\sigma,\sigma')$ is defined to be $|\Lambda|^{-1}|\mathbb{E}[H'_\Lambda(\sigma)H'_\Lambda(\sigma')]|$, and $\overline{c}'_\Lambda \stackrel{\text{def}}{:=} \max_{\sigma,\sigma'} |c'_\Lambda(\sigma,\sigma')|$.
(b) For any $x_1 < x_2$

$$\frac{1}{x_2-x_1}\int_{x_1}^{x_2} \mathbb{E}\left[\left|\Omega_{\Lambda,x}\left(h'_\Lambda\right) - \left\langle h'_\Lambda\right\rangle_{\Lambda,x}\right|\right] dx \;\le\; 4\left(|x_1|+|x_2| + \frac{2\sqrt{x_2-x_1}}{(\overline{c}'_\Lambda|\Lambda|)^{1/4}}\right)\frac{(\overline{c}'_\Lambda)^{3/4}}{\sqrt{x_2-x_1}\,|\Lambda|^{1/4}}\,. \tag{4.43}$$

4.4.1 Application: Distributional Stochastic Stability via Perturbations

The quantitative version of the Ghirlanda-Guerra identities follows from this.

Corollary 4.4.2. *Suppose $c_\Lambda(\sigma,\sigma) = \overline{c}'_\Lambda$ for every σ. Then for every non-random function of n replicas, $\Phi^{(n)}_\Lambda(\sigma^{(1)},\dots,\sigma^{(n)})$, with maximum norm at most 1, we have the conditional expectation formula for one additional replica*

$$\begin{aligned}&\frac{1}{x_2^2-x_1^2}\int_{x_1}^{x_2}\left\langle \frac{c'_\Lambda(\sigma^{(1)},\sigma^{(n+1)})}{\overline{c}'_\Lambda}\,\Phi^{(n)}(\sigma^{(1)},\dots,\sigma^{(n)})\right\rangle_{\Lambda,x} d(x^2)\\ &\quad= \frac{1}{x_2^2-x_1^2}\int_{x_1}^{x_2}\left\langle \gamma^{(n)}_\Lambda(\sigma^{(1)},\dots,\sigma^{(n)})\Phi^{(n)}(\sigma^{(1)},\dots,\sigma^{(n)})\right\rangle_{\Lambda,x} d(x^2) + \mathrm{Rem}\end{aligned}$$

for every pair $0 \le x_1 < x_2$, where

$$\gamma^{(n)}_\Lambda(\sigma^{(1)},\dots,\sigma^{(n)}) \;=\; \frac{1}{n}\left[\sum_{k=2}^{n}\frac{c'_\Lambda(\sigma^{(1)},\sigma^{(k)})}{\overline{c}'_\Lambda} + \left\langle\frac{c'_\Lambda(\sigma,\sigma')}{\overline{c}'_\Lambda}\right\rangle_{\Lambda,x}\right],$$

and the remainder satisfies the bound

$$n\,|\mathrm{Rem}| \;\le\; \left[\frac{8}{x_1+x_2} + \frac{2^{3/2}}{\sqrt{x_2^2-x_1^2}}\right]\delta_\Lambda^{1/2} + \frac{4}{\sqrt{x_2-x_1}}\,\delta_\Lambda^{1/4}\,, \tag{4.44}$$

where δ_Λ is a small parameter $\delta_\Lambda = 1/(|\Lambda|\overline{c}_\Lambda)$.

Proof. Let us define

$$\widehat{\mathrm{Rem}} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \left(\left\langle h'_\Lambda(\sigma^{(1)}) \cdot \Phi^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right\rangle_{\Lambda,x} - \langle h'_\Lambda \rangle_{\Lambda,x} \left\langle \Phi^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right\rangle_{\Lambda,x} \right) dx \,.$$

By the triangle inequality and Cauchy-Schwarz, and the fact that $\|\Phi^{(n)}_\Lambda\|_\infty \le 1$, we know that $|\widehat{\mathrm{Rem}}|$ is bounded by

$$\left(\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \left\langle \left| h'_\Lambda(\sigma) - \Omega_{\Lambda,x}\big(h'_\Lambda(\sigma)\big) \right|^2 \right\rangle_{\Lambda,x} dx \right)^{1/2}$$

plus

$$\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} \mathbb{E}\left[\left| \Omega_{\Lambda,x}\left(h'_\Lambda\right) - \langle h'_\Lambda \rangle_{\Lambda,x} \right| \right] dx \,.$$

Using equations (4.42) and (4.43), this bound is at most $\overline{c}'_\Lambda(x_1 + x_2)/2$ times the right hand side of (4.44). In other words, we have an upper bound on $\widehat{\mathrm{Rem}}$ which is at most $\overline{c}'_\Lambda(x_1 + x_2)/(2n)$ times the bound we claimed for Rem.

But Gaussian integration by parts implies that

$$\begin{aligned} -\left\langle h'_\Lambda(\sigma^{(1)}) \Phi^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right\rangle_{\Lambda,x} &= x \sum_{k=1}^{n} \left\langle c'_\Lambda(\sigma^{(1)}, \sigma^{(k)}) \cdot \Phi^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right\rangle_{\Lambda,x} \\ &\quad - nx \left\langle c'_\Lambda(\sigma^{(1)}, \sigma^{(n+1)}) \cdot \Phi^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right\rangle_{\Lambda,x} . \end{aligned}$$

See for example Lemma 4.4.5 for a similar calculation carried out in more detail. A special case of this formula, obtained by setting $\Phi^{(n)} \equiv 1$, also gives

$$-\langle h'_\Lambda \rangle_{\Lambda,x} = x \left[\langle c'_\Lambda(\sigma, \sigma) \rangle_{\Lambda,x} - \langle c'_\Lambda(\sigma, \sigma') \rangle_{\Lambda,x} \right] .$$

If we combine these two formulas, this allows one to rewrite $\widehat{\mathrm{Rem}}$. If we assume that $c'_\Lambda(\sigma, \sigma)$ is constant as a function of σ, meaning there is a constant diagonal covariance, then the $k = 1$ term of the first formula cancels with the $c'_\Lambda(\sigma, \sigma)$ in the second formula. Therefore, we get

$$\widehat{\mathrm{Rem}} = \frac{\overline{c}'_\Lambda}{n(x_2 - x_1)} \int_{x_1}^{x_2} \left\langle \left[\frac{c'_\Lambda(\sigma^{(1)}, \sigma^{(n+1)})}{\overline{c}'_\Lambda} - \gamma^{(n)}_\Lambda(\sigma^{(1)}, \dots, \sigma^{(n)}) \right] \cdot \Phi^{(n)}(\sigma^{(1)}, \cdots, \sigma^{(n)}) \right\rangle_{\Lambda,x} x \, dx \,.$$

Note that the measure $x\,dx$ is $\frac{1}{2} \cdot d(x^2)$, writing the Riemann-Stieltjes differential form $d(x^2) = 2x\,dx$. Since

$$\int_{x_1}^{x_2} d(x^2) = x_2^2 - x_1^2 = (x_2 - x_1)(x_1 + x_2) \,,$$

we also divide by an appropriate normalization $\frac{1}{2}(x_1+x_2)$ times $\overline{c}'_\Lambda/n$ to get the bound for Rem from the bound on $\widehat{\text{Rem}}$, which gives the result. □

As an application of this result, consider the following scenario. Suppose that for each Λ, there is a given Hamiltonian H^*_Λ with covariance

$$\mathbb{E}[H^*_\Lambda(\sigma)H^*_\Lambda(\sigma')] \;=\; |\Lambda| c^*_\Lambda(\sigma,\sigma')\,,$$

where we assume that $c^*_\Lambda(\sigma,\sigma) = \overline{c}^*_\Lambda$ for all σ, and we assume that $\overline{c}^* = \sup_\Lambda \overline{c}^*_\Lambda$ is finite, in order to satisfy thermodynamic stability.

By Lemma 4.2.1 we know that we may construct *i.i.d.* Gaussian centered Hamiltonians $H^{(p)}_\Lambda(\sigma)$, for $p = 1, 2, \dots$, which are independent of H^*_Λ and such

$$\mathbb{E}[H^{(p)}_\Lambda(\sigma)H^{(p)}_\Lambda(\sigma')] \;=\; |\Lambda|\,[c^*_\Lambda(\sigma,\sigma')]^p\;.$$

For each $\epsilon > 0$ and a real sequence $\boldsymbol{x} = (x_1, x_2, \dots)$ we define the perturbed Hamiltonian

$$H_\Lambda(\sigma;\boldsymbol{x}) \;=\; H^*_\Lambda(\sigma) + \sum_{p=1}^{\infty} \frac{x_p}{p\lceil\overline{c}^*\rceil^{p/2}}\, H^{(p)}_\Lambda(\sigma)\,.$$

We denote by $\langle\cdots\rangle_{\Lambda,\boldsymbol{x}}$ the quenched multi-replica equilibrium measure with respect to $H^\epsilon_\Lambda(\sigma;\boldsymbol{x})$ and $c_{l,l'}$ is a short notation for $c^*_\Lambda(\sigma,\sigma')$. Then we may prove the following corollary.

Corollary 4.4.3. *Suppose that the sequence (ϵ_Λ) satisfies $\lim_{|\Lambda|\to\infty}|\Lambda|\epsilon^2_\Lambda = \infty$. Let $\boldsymbol{X} = (X_1, X_2, \dots)$ be an IID sequence of random variables, each uniformly distributed on $[0,1]$, all of which are independent of H^*_Λ and $H^{(p)}_\Lambda$ for $p = 1, 2, \dots$ and all Λ. Then for almost every choice of $\boldsymbol{X}$ we have stochastic stability in distribution: for each $n, p \in \{1, 2, \dots\}$,*

$$\lim_{|\Lambda|\to\infty}\max_{\Phi^{(n)}}{}' \left\langle \left([c_{1,n+1}]^p - \frac{1}{n}\sum_{k=2}^{n}[c_{1,k}]^p - \frac{1}{n}\,\langle [c_{1,2}]^p\rangle_{\Lambda,\epsilon_\Lambda\boldsymbol{X}} \right)\Phi^{(n)} \right\rangle_{\Lambda,\epsilon_\Lambda\boldsymbol{X}} \;=\; 0\,,$$

where $\max'_{\Phi^{(n)}}$ is the maximum over all non-random functions $\Phi^{(n)}(\sigma^{(1)},\dots,\sigma^{(n)})$ satisfying $\|\Phi^{(n)}\| \le 1$.

Proof. In order to prove this, for a given p, we merely split up the Hamiltonian:

$$H_\Lambda(\sigma;\epsilon\boldsymbol{x}) \;=\; H_\Lambda(\sigma) + x_p H'_\Lambda(\sigma) \;=:\; H_\Lambda(\sigma;x_p)\,,$$

where

$$H_\Lambda(\sigma) \;=\; H^*_\Lambda(\sigma) + \epsilon \sum_{\substack{k=1\\k\neq p}}^{\infty} \frac{x_k}{k[\overline{c}^*]^{k/2}}\, H^{(k)}_\Lambda(\sigma)\,,$$

and

$$H'_\Lambda(\sigma) \;=\; \frac{\epsilon}{p[\overline{c}^*]^{p/2}}\, H^{(p)}_\Lambda(\sigma)\,.$$

With this definition, we have $c'_\Lambda(\sigma,\sigma') = \frac{\epsilon^2}{p^2}\,[c^*_\Lambda(\sigma,\sigma')]^p$ but the constant prefactor has been explicitly taken into consideration in Corollary 4.4.2 by normalizing by $\overline{c}'_\Lambda$. It does enter into the definition of the remainder in (4.44) through the small parameter which is

$$\delta_\Lambda \;=\; \frac{1}{|\Lambda|\overline{c}'_\Lambda} \;\leq\; \frac{p^2}{\epsilon^2_\Lambda|\Lambda}\,.$$

That is why we required $|\Lambda|\epsilon^2_\Lambda \to \infty$, because this guarantees $\delta_\Lambda \to 0$, as is needed. □

Remark 4.4.4. *Note that technically what we proved is that for any open interval for x_p in $[0,1]$, if we average the distributional stochastic stability equation over that interval, when integrated against the measure $x\,dx$, then we obtain zero in the limit. On the other hand, the quantity in question is*

$$\left\langle \left([c^*_\Lambda(\sigma^{(1)},\sigma^{(n+1)})]^p - \frac{1}{n}\sum_{k=2}^{n} [c^*_\Lambda(\sigma^{(1)},\sigma^{(k)})]^p - \frac{1}{n}\,\langle [c^*_\Lambda(\sigma,\sigma')]^p\rangle_{\Lambda,\epsilon_\Lambda \boldsymbol{X}} \right) \Phi^{(n)} \right\rangle_{\Lambda,\epsilon_\Lambda \boldsymbol{X}}$$

and this is bounded for every x by $2[\overline{c}^*_\Lambda]^p$. Then by standard arguments from measure theory, we may conclude that for almost every choice of x with respect to the measure $d\mu(x) = 2x\,dx = d(x^2)$, the quantity is also zero. But this measure is equivalent to Lebesgue measure in the sense that they are mutually absolutely continuous with respect to each other. So the notions of measure zero sets are the same.

This means that letting X_p be random, then for almost every X_p we have the stochastic stability formula for the pth power of the overlap. But, firstly, we note that we may rigorously take an infinite number of *i.i.d.* uniform random variables $\boldsymbol{X} = (X_1, X_2, \dots)$ by Kolmogorov's principle, and secondly that the measure is precisely the product measure for all the X_p's. Therefore, knowing that for each X_p we have the stochastic stability condition for almost every X_p, by definition, this means we have the stochastic stability condition for all p for almost every $\boldsymbol{X}$.

4.4.2 Proof of Theorem 4.4.1

In the proof of Theorem 4.4.1 we will condition on $H_\Lambda(\sigma)$ in order to eliminate the need to consider it as random. But we will do this implicitly. If desired, simply interpret all expectations as conditional expectations, conditioning on $H_\Lambda(\sigma)$.

The proof will be obtained by combining several lemmas. First, we note that by usual calculations as in elementary statistical mechanics,

$$\frac{d}{dx}\, p_\Lambda(x) \;=\; -\Omega_{\Lambda,x}(h'_\Lambda)\,, \tag{4.45}$$

as has been used already in Section 3. Moreover, by performing Gaussian integration by parts, we may deduce this:

Lemma 4.4.5. *For any x*

$$\frac{d}{dx}\,\mathbb{E}[p_\Lambda(x)] \;=\; -\langle h'_\Lambda\rangle_{\Lambda,x} \;=\; \frac{x}{2}\,\left\langle c'_\Lambda\left(\sigma,\sigma\right) + c'_\Lambda\left(\sigma',\sigma'\right) - 2c'_\Lambda\left(\sigma,\sigma'\right)\right\rangle_{\Lambda,x}\,. \tag{4.46}$$

Proof. Since Λ is finite, the derivative exists and we may write

$$\frac{d}{dx}\,\mathbb{E}[p_\Lambda(x)] \;=\; \mathbb{E}[\Omega_{\Lambda,x}(h'_\Lambda)]$$

by using (4.45). Recall the definition of $\Omega_{\Lambda,x}$ as well as $\mathcal{G}_{\Lambda,x}$ from equations

(4.25), (4.26) and (4.27). Then Wick's rule gives

$$\begin{aligned}\mathbb{E}[\Omega_{\Lambda,x}(h'_\Lambda)] \;&=\; \mathbb{E}\left[\sum_\sigma h'_\Lambda(\sigma)\frac{\exp\left(-H_\Lambda(\sigma;x)\right)}{Z_\Lambda(x)}\right] \\ &=\; \sum_\sigma \mathbb{E}[h'_\Lambda(s)H'_\Lambda(\sigma)]\mathbb{E}\left[\frac{(\partial/\partial H'_\Lambda(\sigma))\exp\left(-H_\Lambda(\sigma;x)\right)}{Z_\Lambda(x)}\right] \\ &\quad - \sum_\sigma\sum_{\sigma'} \mathbb{E}[h'_\Lambda(\sigma)H'_\Lambda(\sigma')]\mathbb{E}\left[\frac{\exp(-H_\Lambda(\sigma;x))}{Z_\Lambda(x)^2}\cdot\frac{\partial}{\partial H'_\Lambda(\sigma')}\,\exp(-H_\Lambda(\sigma';x))\right],\end{aligned}$$

Since $H_\Lambda(\sigma;x) = H_\Lambda(\sigma) + xH'_\Lambda(\sigma)$, this means $(\partial/\partial H'_\Lambda(\sigma))\exp(-H_\Lambda(\sigma;x)) = -x\exp(H_\Lambda(\sigma))$. Using this and the fact that $\mathbb{E}[h'_\Lambda(\sigma)H'_\Lambda(\sigma')] = |\Lambda|^{-1}\mathbb{E}[H'_\Lambda(\sigma)H'_\Lambda(\sigma')] = c'_\Lambda(\sigma,\sigma')$, this gives the result. □

Corollary 4.4.6. *For any x*

$$|\langle h'_\Lambda\rangle_{\Lambda,x}| \;\le\; 2|x|\overline{c}'_\Lambda\,.$$

Proof. This follows from (4.46) and a uniform bound using the definition $\overline{c}'_\Lambda \stackrel{\text{def}}{:=} \max_{\sigma,\sigma'}|c_\Lambda(\sigma,\sigma')|$. □

With this, we can prove the first part of the theorem.

Proof. **Proof of Theorem 4.4.1, part (a):** Another statistical mechanics calculation following (4.45) is

$$\frac{d^2}{dx^2}\,p_\Lambda(x) \;=\; |\Lambda|(\Omega_{\Lambda,x}([h'_\Lambda]^2) - [\Omega_{\Lambda,x}(h'_\Lambda)]^2)\,.$$

So, integrating and taking expectations, we have

$$\int_{x_1}^{x_2}\mathbb{E}\left[\Omega_{\Lambda,x}\left(\left|h'_\Lambda(\sigma) - \Omega_{\Lambda,x}\big(h'_\Lambda(\sigma)\big)\right|^2\right)\right]\,dx \;=\; \frac{1}{|\Lambda|}\cdot\frac{d}{dx}\,\mathbb{E}[p_\Lambda(x)]\Big|_{x_1}^{x_2}.$$

Recall that $\langle\cdots\rangle_{\Lambda,x} = \mathbb{E}[\Omega_{\Lambda,x}(\cdots)]$. But by (4.46) and Corollary 4.4.6 this leads directly to (4.42). □

In order to obtain the proof of Theorem 4.4.1, part (b), we will use concentration of measure. Our main goal will be to obtain a bound on the random

fluctuations of the quantity $\Omega_{\Lambda,x}(h'_\Lambda) - \langle h'_\Lambda \rangle_{\Lambda,x}$. We start by quoting a result which was proved in [77]:

$$\forall t > 0\,, \qquad \mathbb{P}\left(|p_\Lambda(x) - \mathbb{E}[p_\Lambda(x)]| \geq \frac{|x|\sqrt{\overline{c}'_\Lambda}}{\sqrt{|\Lambda|}}\, t \right) \leq 2e^{-t^2/2}\,,\ a.s., \tag{4.47}$$

where recall that $\overline{c}'_\Lambda$ was defined in the statement of Theorem 4.4.1, part (b).

We now claim that the following result may be proved using this and previous results:

Lemma 4.4.7. *For any x, and for any $\epsilon > 0$, we have*

$$\mathbb{E}\left[\left| \Omega_{\Lambda,x}(h'_\Lambda) - \langle h'_\Lambda \rangle_{\Lambda,x} \right| \right] \leq \frac{4(|x|+\epsilon)\sqrt{2\overline{c}'_\Lambda}}{\epsilon\sqrt{\pi|\Lambda|}} + \epsilon \int_{-1}^{1} (1-|z|)\mathbb{E}\left[\frac{d^2 p_\Lambda}{dx^2}(x+\epsilon z) \right] dz\,. \tag{4.48}$$

Proof. By (4.45) and Taylor's theorem

$$\begin{aligned}
-\Omega_{\Lambda,x}(h'_\Lambda) &= \frac{p_\Lambda(x+\epsilon) - p_\Lambda(x-\epsilon)}{2\epsilon} + \frac{1}{2}\int_{-1}^{1} \left(\frac{dp_\Lambda}{dx}(x) - \frac{dp_\Lambda}{dx}(x+\epsilon y) \right) dy \\
&= \frac{p_\Lambda(x+\epsilon) - p_\Lambda(x-\epsilon)}{2\epsilon} \\
&\quad + \epsilon \int_{-1}^{1} \left(\int_{-1}^{1} \frac{\mathbf{1}_{\mathbb{R}_-}(y)\mathbf{1}_{(y,0)}(z) - \mathbf{1}_{\mathbb{R}_+}(y)\mathbf{1}_{(0,y)}(z)}{2} \cdot \frac{d^2 p_\Lambda}{dx^2}(x+\epsilon z)\, dz \right) dy \\
&= \frac{p_\Lambda(x+\epsilon) - p_\Lambda(x-\epsilon)}{2\epsilon} + \frac{\epsilon}{2}\int_{-1}^{1} g(z) \frac{d^2 p_\Lambda}{dx^2}(x+\epsilon z)\, dz\,,
\end{aligned}$$

where we interchanged the order of integration and

$$g(z) = \begin{cases} z-1 & \text{for } 0 < z < 1, \\ 1+z & \text{for } -1 < z < 0. \end{cases}$$

For each $z \in [-1,1]$, let us define the random variable

$$Z_\epsilon(z) = \frac{d^2 p_\Lambda}{dx^2}(x+\epsilon z)\,,$$

which is nonnegative. Let us also define

$$Y_\epsilon = \frac{\epsilon}{2}\int_{-1}^{1} g(z) Z_\epsilon(z)\, dz\,.$$

Then we obtain

$$\Omega_{\Lambda,x}(h'_\Lambda) - \mathbb{E}[\Omega_{\Lambda,x}(h'_\Lambda)] \;=\; \frac{(p_\Lambda(x+\epsilon) - \mathbb{E}[p_\Lambda(x+\epsilon)]) - (p_\Lambda(x-\epsilon) - \mathbb{E}[p_\Lambda(x-\epsilon)]}{2\epsilon} + Y_\epsilon - \mathbb{E}[Y_\epsilon]\,.$$

We may use equation (4.47) and the general subset bound, $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$, to obtain this:

$$\Omega_{\Lambda,x}\,(h'_\Lambda) - \mathbb{E}[\Omega_{\Lambda,x}(h'_\Lambda)] \;\leq\; \frac{(|x|+\epsilon)\sqrt{\overline{c}'_\Lambda}}{\epsilon\sqrt{|\Lambda|}}\,t + Y_\epsilon - \mathbb{E}[Y_\epsilon] \quad \text{ with probability } p \geq 1 - 4e^{-t^2/2}.$$

We have a similar statement for the lower bound. Therefore, again by the subset bound,

$$\mathbb{P}\left(\left|\Omega_{\Lambda,x}\,(h'_\Lambda) - \langle h'_\Lambda\rangle_{\Lambda,x}\right| \geq \frac{(|x|+\epsilon)\sqrt{\overline{c}'_\Lambda}}{\epsilon\sqrt{|\Lambda|}}\,t + |Y_\epsilon - \mathbb{E}[Y_\epsilon]|\right) \;\leq\; 8e^{-t^2/2}\,. \quad (4.49)$$

Recall that $\langle h'_\Lambda\rangle_{\Lambda,x} = \mathbb{E}[\Omega_{\Lambda,x}(h'_\Lambda)]$ by definition.

So, using the fact that for any integrable random variable X, $\mathbb{E}[X] \leq \int_0^\infty \mathbb{P}(X \geq t)\,dt$, we obtain

$$\mathbb{E}\left[\left|\Omega_{\Lambda,x}\,(h'_\Lambda) - \langle h'_\Lambda\rangle_{\Lambda,x}\right|\right] \;\leq\; \mathbb{E}\left[|Y_\epsilon - \mathbb{E}[Y_\epsilon]|\right] + \frac{8(|x|+\epsilon)\sqrt{\overline{c}'_\Lambda}}{\epsilon\sqrt{|\Lambda|}}\int_0^\infty e^{-t^2/2}\,dt\,. \quad (4.50)$$

But now note that

$$Y_\epsilon - \mathbb{E}[Y_\epsilon] \;=\; \frac{\epsilon}{2}\int_{-1}^{1} g(z)\left(Z_\epsilon(z) - \mathbb{E}\left[Z_\epsilon(z)\right]\right)\,dz\,,$$

which implies that

$$\mathbb{E}\left[|Y_\epsilon - \mathbb{E}[Y_\epsilon]|\right] \;\leq\; \frac{\epsilon}{2}\int_{-1}^{1} |g(z)|\mathbb{E}\left[|Z_\epsilon(z) - \mathbb{E}\left[Z_\epsilon(z)\right]|\right]\,dz\,.$$

But since $Z_\epsilon(z)$ is nonnegative, almost surely, this means that

$$\mathbb{E}\left[|Z_\epsilon(z) - \mathbb{E}\left[Z_\epsilon(z)\right]|\right] \;\leq\; \mathbb{E}\left[|Z_\epsilon(z)|\right] + \left|\mathbb{E}\left[Z_\epsilon(z)\right]\right| \;=\; 2\mathbb{E}\left[Z_\epsilon(z)\right]\,.$$

Combining this bound with (4.50) gives the desired inequality (4.48). □

With this, we can prove the remainder of the theorem.

Proof. **Proof of Theorem 4.4.1, part (b):** For any $x_1 < x_2$, integrating (4.48) and dividing by the length of the interval we obtain

$$\frac{1}{x_2-x_1}\int_{x_1}^{x_2}\mathbb{E}\left[\left|\Omega_{\Lambda,x}\left(h'_\Lambda\right)-\langle h'_\Lambda\rangle_{\Lambda,x}\right|\right]\,dx \;\le\; \frac{2(|x_1|+|x_2|+2\epsilon)\sqrt{2\overline{c}'_\Lambda}}{\epsilon\sqrt{\pi|\Lambda|}} + \frac{\epsilon}{x_2-x_1}\int_{-1}^{1}(1-|z|)\left(\int_{x_1}^{x_2}\mathbb{E}\left[\frac{d^2p_\Lambda}{dx^2}(x+\epsilon z)\right]\,dx\right)dz\,. \tag{4.51}$$

But, by the fundamental theorem of calculus, (4.46) and Corollary 4.4.6

$$\int_{x_1}^{x_2}\mathbb{E}\left[\frac{d^2p_\Lambda}{dx^2}(x+\epsilon z)\right]\,dx \;=\; \mathbb{E}\left[\frac{dp_\Lambda}{dx}(x+\epsilon z)\right]\Bigg|_{x_1}^{x_2},$$

so that

$$\frac{\epsilon}{x_2-x_1}\int_{-1}^{1}(1-|z|)\left(\int_{x_1}^{x_2}\mathbb{E}\left[\frac{d^2p_\Lambda}{dx^2}(x+\epsilon z)\right]\,dx\right)dz \;\le\; \frac{2\overline{c}'_\Lambda\epsilon}{x_2-x_1}\int_{-1}^{1}(1-|z|)(|x_1|+|x_2|+2\epsilon)\,dz\,.$$

Using this with (4.51) gives the result

$$\frac{1}{x_2-x_1}\int_{x_1}^{x_2}\mathbb{E}\left[\left|\Omega_{\Lambda,x}\left(h'_\Lambda\right)-\langle h'_\Lambda\rangle_{\Lambda,x}\right|\right]\,dx \;\le\; 2(|x_1|+|x_2|+2\epsilon)\left(\frac{\sqrt{2\overline{c}'_\Lambda}}{\epsilon\sqrt{\pi|\Lambda|}}+\frac{\overline{c}'_\Lambda\epsilon}{x_2-x_1}\right)\,. \tag{4.52}$$

Now we may optimize in $\epsilon>0$. We choose $\epsilon=\sqrt{x_2-x_1}/(\overline{c}'_\Lambda|\Lambda|)^{1/4}$ to get the desired result, equation (4.43). □

4.5 Consequences of the GG identities

Since the Ghirlanda-Guerra identities are limiting properties of the overlap w.r.t. the quenched Gibbs measure its convenient to introduce a mathematical framework that allow us to describe these limiting properties.

Consider a general spin glass model (4.21) with random Gibbs measure $\mathcal{G}_\Lambda$. For $l,l'\ge 1$, consider two spin configuration $\sigma^l,\sigma^{l'}\in\Sigma_\Lambda$ and let us consider the matrix of

$$R_\Lambda=\left(c_\Lambda(\sigma^l,\sigma^{l'})\right)_{l,l'\ge 1}. \tag{4.53}$$

where $c_\Lambda(\sigma^l, \sigma^{l'})$ is the generalized overlap.

Consider the law of R_Λ w.r.t the measure $\mathbb{E}\Omega_\Lambda$, where $\Omega_\Lambda := \mathcal{G}_\Lambda^{\otimes\infty}$.

Notice that each entry of R_Λ belongs to the compact space $[1, 1]$. Consider the space of positive semi-definite symmetric matrix. This is a compact separable metric space when considered as a closed subset of $[1, 1]^{\mathbb{N}\times\mathbb{N}}$ equipped with the product topology. By compactness arguments, the Borel probability measures on this space form also a compact, separable and metric space when equipped with the appropriate topology. This observation allow us to consider the limit along subsequences of the law of R_Λ.

Clearly, by construction, R^Λ is an infinite positive semi-definite symmetric matrix which is invariant in law for permutation of finitely many indices. This kind of random matrix are called *Gram-de Finetti matrix* and have a nice representation ((see [32] or [16])) in terms of Hilbert spaces. Indeed,

Theorem 4.5.1. *[(DovbyshSudakov)]*

Let $R = (R_{l,l'})_{l,l'\geq 1}$ be a Gram-de Finetti Matrix. Let $\mathcal{H}$ be an infinite dimensional separable Hilbert space. There exists a random probability measure $\mathcal{G}$ on $\mathcal{H} \times R^+$ such that the array R is equal in distribution to

$$\left(h_l \cdot h'_l + a_l\delta_{l,l'}\right)_{l,l'\geq 1}$$

where, conditionally on $\mathcal{G}$, $(h_l, a_l)_{l\geq 1}$ is a sequence of i.i.d. random variables with the distribution $\mathcal{G}$ and $h \cdot h'$ denotes the scalar product on $\mathcal{H}$.

Roughly speaking we can say that any *Gram-de Finetti matrix* can be generated by sampling a vector of an infinite dimensional separable Hilbert space $\mathcal{H}$ from a random Gibbs measure on $\mathcal{H}$ and looking at the matrix of scalar products. Basically, overlaps in spin glass models, are introduced in the same way, i.e. sampling spins from the Gibbs measure as we can regard Σ_Λ as a subset of a Hilbert space. The theorem therefore identifies the analogous for for the limit of the Gibbs measures.

Thus, for any convergent subsequence, we can identify the asymptotic law R^{Λ}, with a random probability measure in a Hilbert space using the representation theorem of Dovbysh and Sudakov.

Let $\mathcal{G}$ a random probability measure on a separable Hilbert space $\mathcal{H}$. We will denote by $(\sigma^l)_{l\geq 1}$ an i.i.d. sample from this measure, by Ω the average with respect to $\mathcal{G}^{\infty}$ and by $\mathbb{E}$ the expectation with respect to the randomness of $\mathcal{G}$. Let $c_{l,l'} := \sigma^l \cdot \sigma^{l'}$ the scalar products(overlaps) in $\mathcal{H}$ and $R = (c_{l,l'})_{l,l'\geq}$ the random matrix (Gram de Finetti).

Definition 4.5.2. *We say that $\mathcal{G}$ satisfies the Ghirlanda-Guerra identities in the distributional sense if for any $n \geq 2$, any bounded measurable function f of the overlaps $(c_{l,l'})_{l,l'\leq n}$ and any bounded measurable function ψ of one overlap, the following hold:*

$$\mathbb{E}\Omega\Big(f\psi(c_{1,n+1})\Big) = \frac{1}{n}\mathbb{E}\Omega\Big(f\Big)\mathbb{E}\Omega\Big(\psi(c_{1,2})\Big) + \frac{1}{n}\sum_{l=2}^{n}\mathbb{E}\Omega\Big(f\psi(c_{1,l})\Big) \tag{4.54}$$

4.5.1 Panchenko's Ultrametric Theorem

Let us start with the following

Definition 4.5.3. *A random measure $\mathcal{G}$ is said to be ultrametric if the distribution of $(c_{l,l'})_{l,l'\geq 1}$ satisfies*

$$\mathbb{E}\Omega\Big(I\Big(c_{1,2} \geq \min(c_{1,3}, c_{2,3})\Big)\Big) = 1 \tag{4.55}$$

where $I(A)$ is the indicator function of the event A.

A major new development in spin glass theory is Panchenko's [108] proof that the Ghirlanda-Guerra identities in the *distributional sense* imply ultrametricity. The are also other important consequences [16] summarised in the following:

Theorem 4.5.4. *Let us denote by ζ the distribution of $R_{1,2}$ under $\mathbb{E}\Omega$ and by $q^* = \max \operatorname{supp}(\zeta)$. Consider the law of $R = (c_{l,l'})_{l,l'\geq}$, that is the random matrix (Gram de Finetti) of the scalar product(overlaps), under the measure $\mathbb{E}\Omega$.*

Assume that $\mathcal{G}$ satisfies the Ghirlanda-Guerra identities in distributional sense, then

i) $R_{i,i} = q^* \ \ a.s., \forall\, i$

ii) $R_{ij} \geq 0 \ \ a.s., \forall\, i, j$ *(Talagrand's positivity principle)*

iii) $\mathcal{G}$ *is ultrametric in the sense of definition 4.5.3. (Panchenko's ultrametricity)*

We want to stress an important point now. It had been earlier known that the Ghirlanda-Guerra identities, when combined with ultrametricity, imply that the law of R is determined is uniquely by ζ ([36, 19]. As a consequence, it's also possible to show (Theorem 2.17 of [16]) that, for any given ζ, the the law of $(R_{l,l'})_{l \neq l'}$ can be represented (weak approximated) by a *Derrida-Ruelle Cascade*. The *Derrida-Ruelle Cascades* is a hierarchical point process introduced by Ruelle [116] and describes the limiting behavior of the GREM model of Derrida [69]. Their relevance in Spin Glass theory, and in particular in the SK model, was first emphasized by Bolthausen and Snzitman in [48] where the authors give a beautiful representation of these process in terms of coalescent process.

The link of Theorem 4.5.4 with the results presented in this chapter is clear. By Theorem 4.3.1, we can deduce that the asymptotic Gibbs measure of the model defined in (4.24) satisfies the condition (4.54) for every power of the covariance and, by the Stone-Weierstrass theorem, this suffices to ensure that the Ghirlanda-Guerra identities hold in the *distributional sense*. More precisely Theorem 4.3.1 has two consequences. The proposition (iv) implies that

$$\lim_{\Lambda \nearrow \mathbb{Z}^d} \mathbb{E}\Omega_{\Lambda,\beta}\Big(f\psi(c_{1,2})\Big) = \frac{1}{n}\mathbb{E}\Omega_{\Lambda,\beta}\Big(f\Big)\mathbb{E}\Omega_{\Lambda,\beta}\Big(\psi(c_{1,2})\Big) + \frac{1}{n}\sum_{l=2}^{n}\mathbb{E}\Omega_{\Lambda,\beta}\Big(f\psi(c_{1,l})\Big) \tag{4.56}$$

The proposition (ii) implies that this equation is true *for almost every* choice of parameter $\beta = (\beta_p)_{p=1}^{\infty}$. In both case, we can identify the asymptotic Gibbs measure with a random probability measure in a Hilbert-space using the repre-

sentation theorem of Dovbysh and Sudakov and the measure is *ultrametric* by the Panchenko's result.

Therefore, to summarize, for almost all the choice of the parameters:

• The complete Hamiltonian(4.24) has an ultrametric Derrida-Ruelle quenched equilibrium measure .

• By (4.4.3), a general Hamiltonian (4.21) can be suitably perturbed in order to obtain an ultrametric Derrida-Ruelle quenched equilibrium measure.

4.5.2 A comment on Finite Dimensional Spin Glasses

In this chapter we have demonstrated the following point, which we feel bears repeating. Panchenko has showed that stochastic stability implies the ultrametric scenario [108] (which does also include the possibility of trivial ultrametricity, where for any three replicas, the triangle formed by the intrinsic overlaps may be equilateral always).

Panchenko's proof uses a strong version of de Finetti's theorem. But this applies because of permutation invariance of the replicas, not the underlying model. Indeed as we have shown, stochastic stability and its implications even apply to the Edwards-Anderson model if it is perturbed by an arbitrarily small inclusion of higher order interactions. Stochastic stability is a general tool, which in principle may be applicable to any disordered model in statistical mechanics including short-ranged spin glasses. We mention that using the language of metastates [25, 97] stochastic stability has been approached in [33] where identities where proved everywhere in the parameters using periodic boundary conditions and averaging over translations. It remains open to show that their analogue of the link overlap coincides with the usual one which is what we used here.

It is finally worth to stress that the addition of the higher orders interactions

introduced in (4.24) is done in order to obtain the identities in distribution and not only in mean. The nature of those higher orders terms appear to be formally similar to the p-spin interactions of the mean field case models. In particular their interaction structure doesn't decay with distance but it's a sum normalised with the volume. At the same time we notice that the core term is still a two point interaction made of nearest neighbouring sites and as such, it retains the topological information of the lattice dimension.

4.5.3 The synchronization property

In this section we show a consequence of the ultrametric property of Theorem 4.5.4 which is the corner stone of the Panchenko's proof of the lower bound of the MSK model described in section 3.7.

As usual, let us consider an arbitrary Spin Glass model 4.21 with Hamiltonian $H_\Lambda(\sigma)$, and quenched measure $\mathbb{E}\Omega_\Lambda$ as reference measure.

Let $\mathcal{S}$ a finite set of $S = |\mathcal{S}|$ elements, which labels different independent gaussian process $H_\Lambda^{(s)}(\sigma)$ on Σ_Λ, where $H_\Lambda^{(s)}(\sigma)$ is the Hamiltonian function of an arbitrary gaussian Spin Glass 4.21 with normalized covariance matrix given by

$$\frac{1}{|\Lambda|}\mathbb{E}H_\Lambda^{(s)}(\sigma^1)H_\Lambda^{(s)}(\sigma^2) = \alpha_s c^{(s)}(\sigma^l, \sigma^{l'}) \equiv \alpha_s c_{l,l'}^{(s)}. \tag{4.57}$$

align

Let us define a new Hamiltonian $\widetilde{H}_\Lambda(\sigma) := \sum_{s\in\mathcal{S}} H_\Lambda^{(s)}(\sigma)$ which, by construction, has normalized covariance matrix given by

$$\frac{1}{|\Lambda|}\mathbb{E}\widetilde{H}_\Lambda(\sigma^1)\widetilde{H}_\Lambda(\sigma^2) = \sum_{s\in\mathcal{S}} \alpha_s c_{l,l'}^{(s)} =: \widetilde{c}_{l,l'}. \tag{4.58}$$

We refer to $c_{l,l'}^{(s)}$ as the *overlap of the species s* and to $\widetilde{c}_{l,l'}$ as the *total overlap*.

We are interested to the limiting properties of joint the distribution of the

arrays

$$R_\Lambda = \Big(c^{(s)}_{l,l'}\Big)_{l,l'\geq 1, s\in\mathcal{S}} \tag{4.59}$$

with respect to the measure $\mathbb{E}\Omega_\Lambda$. As explained in the introduction of this section, the representation Theorem 4.5.1 and compactness arguments allow us describe these limiting properties considering a random measure $\mathcal{G}$ on a separable Hilbert space.

For any bounded measurable function ϕ of the entries $c^{(s)}_{l,l'}$, consider the the array

$$Q_{l,l'} = \phi\Big(\big(c^{(s)}_{l,l'}\big)_{s\in\mathcal{S}}\Big) \tag{4.60}$$

with respect the measure $\mathbb{E}\Omega := \mathbb{E}\mathcal{G}^{\otimes\infty}$ and let us denote by $R^{(n)} = \Big(c^{(s)}_{l,l'}\Big)_{l,l'\leq n, s\in\mathcal{S}}$.

Definition 4.5.5. *We say that $\mathcal{G}$ satisfies the Multi-species Ghirlanda-Guerra identities if for any $n \geq 2$, any bounded measurable function $f = f(R^n)$, the following hold:*

$$\mathbb{E}\Omega\Big(f(R^{(n)})Q_{1,n+1)}\Big) = \frac{1}{n}\mathbb{E}\Omega\Big(f(R^{(n)})\Big)\mathbb{E}\Omega\Big(Q_{1,2}\Big) + \frac{1}{n}\sum_{l=2}^{n}\mathbb{E}\Omega\Big(f(R^{(n)})Q_{1,l}\Big) \tag{4.61}$$

These identities have a deep consequence in the joint distribution of the array (4.59). In fact one can prove [109] the following:

Theorem 4.5.6. *[Panchenko] For any array* (4.59) *that satisfies* (4.61)*, there exist non-decreasing $(1/\alpha_s)$-Lipschitz functions $L_s : [0,1] \to [0,1]$ such that $c^{(s)}_{l,l'} = L_s(\widetilde{c}_{l,l'})$ almost surely for all $s \in \mathcal{S}$ and all $l, l' \geq 1$.*

In this case we say that the array (4.59) satisfies a *synchronization property.* Indeed, the previous result implies that the joint distribution of the overlaps $c^{(s)}_{l,l'}$ for all species will be determined trivially by the array of the total overlap $\widetilde{c}_{l,l'}$, so in some sense they they are *synchronized.* Moreover, the array $\widetilde{c}_{l,l'}$ satisfies the usual Ghirlanda-Guerra identities and therbn , by 4.5.4, can be generated using the Ruelle probability cascades.

Chapter 5

Monomer-Dimer models formalism

The second kind of statistical mechanical models studied in this work belong to the class of *Monomer-Dimer models on Graphs*. The aim of this chapter, mirroring chapter 1, is to introduce necessary mathematical background. Moreover section 5.2 contain an alternative representation of the partition function that will be fundamental in chapter 6 and provide an alternative proof a classical result due to Heilmann-Lieb [86].

5.1 (Disordered) Monomer-Dimer models on Graphs

From now on the symbol G denotes a generic graph the finite vertex Λ set and edge set E which represent respectively the set of microscopic components and the set of pair interacting components of a model.

The fundamental difference between spin and monomer dimer models is the topological properties of the configuration space. Indeed, the spin configuration space Σ_Λ is a product space but, as we will see in the next this is no longer true

for *dimer configurations.*

Let us first introduce *Monomer Dimer models* on graphs in a pure probabilistic setting.

Definition 5.1.1. *Let $G = (\Lambda, E)$ be a finite simple graph. A dimer configuration (or matching) on G is a set D of pairwise non-incident edges (called dimers).*

In other terms a dimer configuration D on G is a partition of a certain set $A \subseteq \Lambda$ into pairs belonging to E :

$$\begin{aligned} &D = \big\{\{i_1, i_2\}, \dots, \{i_{|A|-1}, i_{|A|}\}\big\} \\ &\text{with } \ \{i_1, i_2, \dots, i_{|A|}\} = A \ \text{ and } \ \{i_s, i_{s+1}\} \in E \, ; \end{aligned} \tag{5.1}$$

The associated set of dimer-free vertices (called monomers) is denoted by $M(D) := \Lambda \setminus A$ Denote by $\mathscr{D}_G$ the space of all possible dimer configurations on the graph G.

The condition of non incident edges is called *hard-core interaction.*

Definition 5.1.2. *A Monomer-Dimer model on G is obtained by assigning a monomer weight $x_i > 0$ to each vertex $i \in \Lambda$, a dimer weight $w_{ij} \geq 0$ to each edge $ij \in E$ and introducing the following probability measure on $\mathscr{D}_G$:*

$$\mu_G(D) \, := \, \frac{1}{Z_\Lambda} \prod_{ij \in D} w_{ij} \prod_{i \in M(D)} x_i \quad \forall \, D \in \mathscr{D}_G \, , \tag{5.2}$$

where $Z_G := \sum_{D \in \mathscr{D}_G} \prod_{ij \in D} w_{ij} \prod_{i \in M(D)} x_i$ is the normalizing factor. The family of weights will be denoted by $x_\Lambda := (x_i)_{i \in \Lambda}$ and $w_E := (w_{ij})_{ij \in E}$

A pure Dimer model is Monomer-Dimer model with the choice $x_i = 0$ for each $i \in \Lambda$.

The *pure Dimer model* is also called perfect weighted matching on G since in this case the dimer configuration space represents the way of completely fill the graph G with dimers.

Notice that by definition the following relation hold

$$|M(D)| + 2\,|D| = |\Lambda| \;\; \forall\, D \in \mathscr{D}_G. \tag{5.3}$$

The following remark is the analogous of 1.1.3.

Remark 5.1.3. *Consider the complete graph K_N, with vertex set $\{1,\dots,N\}$ and edge set made of all possible pairs of vertices. Because of the lack of geometric structure the space of dimer configurations $\mathscr{D}_N \equiv \mathscr{D}_{K_N}$ simplifies; precisely $D \in \mathscr{D}_N$ if and only if*

$$D = \big\{\{i_1, i_2\}, \dots, \{i_{|A|-1}, i_{|A|}\}\big\} \quad \textit{with} \quad \{i_1, i_2, \dots, i_{|A|}\} = A \tag{5.4}$$

for a certain set of vertices $A \subseteq \{1,\dots,N\}$, and the monomer set associated to D is $M_(D) = \{1,\dots,N\} \setminus A$.
On the other hand any monomer-dimer model on a graph $G = (\Lambda, E)$ with N vertices can be thought as a monomer-dimer model on the complete graph K_N. Indeed the measure μ_G is equivalent to a measure $\mu_N \equiv \mu_{K_N}$ by setting $w_{ij} := 0$ for all pairs $ij \notin E$. Precisely introducing these zero dimer weights it holds $Z_N \equiv Z_{K_N} = Z_G$ and

$$\mu_N(D) = \begin{cases} \mu_G(D) & \textit{if } D \in \mathscr{D}_G \\ 0 & \textit{if } D \in \mathscr{D}_N \setminus \mathscr{D}_G \end{cases}.$$

It's possible to define a monomer dimer model with the Gibbs formalism introduced in the previous section, representing a dimer configuration $D \in \mathscr{D}_G$ and the associated $M(D)$ trough occupancy variables on G.

Definition 5.1.4. *For all $i \in \Lambda$ and $e \in E$ and $D \in \mathscr{D}_G$ let us define*

$$\alpha_i(D) := \begin{cases} 1\,, & \textit{if } i \in M(D) \\ 0\,, & \textit{otherwise} \end{cases} \qquad \textit{and} \qquad \alpha_e(D) := \begin{cases} 1\,, & \textit{if } e \in D \\ 0\,, & \textit{otherwise} \end{cases}. \tag{5.5}$$

Thus a dimer configuration D is represented by an element $\alpha(D) := (\alpha_e(D))_{e\in E} \in \{0,1\}^{|E|}$. The point is that $\{\alpha(D) : D \in \mathscr{D}_G\} \subset \{0,1\}^{|E|}$ because of the hard-core condition of non incident edges in definition 5.1.1. It's easy to see that this condition to the algebraic constraint:

$$\sum_{i\sim j} \alpha_{ij}(D) \leq 1, \ \forall i \in \Lambda \tag{5.6}$$

where $i \sim j \Leftrightarrow ij \in E$. Clearly if we define $\{0,1\}_d^{|E|} := \{\alpha \in \{0,1\}^{|E|} : (5.6) \text{ hold}\}$ there exist a bijection $\mathscr{D}_G \leftrightarrow \{0,1\}_d^{|E|}$.

Definition 5.1.5. *A Monomer-Dimer model on G is defined assigning to each dimer configuration $D \in \mathscr{D}_G$ an MD-Hamiltonian function*

$$H_G(D) = -\sum_{ij\in E} h_{ij}^{(\mathrm{d})}\alpha_{ij}(D) - \sum_{i\in\Lambda} h_i^{(\mathrm{m})}\alpha_i(D) \tag{5.7}$$

where $h_E^{(\mathrm{d})} := \{h_{ij}^{(\mathrm{d})}\}_{ij\in E}$, and $h_\Lambda^{(\mathrm{m})} := \{h_i^{(\mathrm{m})}\}_{i\in\Lambda}$ are two families of real parameters.

We want to stress an important fact. The MD-Hamiltonian is formally very similar to a spin Hamiltonian (1.8) however the meaning is totally different. The term $h_{ij}^{(\mathrm{d})}\alpha_{ij}$ is NOT an interaction between the microscopic component i and j. The interaction between the site i and j is a consequence of the hard-core constraint on the dimer configuration space. The families $h_E^{(\mathrm{d})}$ and $h_\Lambda^{(\mathrm{m})}$ are both external fields acting respectively on dimers and monomers.

Definition 5.1.6. *Let $\beta \geq 0$ be a real parameter which represent up to a constant the inverse physical temperature, the finite volume Gibbs measure $\mathcal{G}_G$ associated to the Hamiltonian (5.7), is a probability measure on $\mathscr{D}_G$ defined as*

$$\mathcal{G}_G(D) := \frac{e^{-\beta H_G(D)}}{Z_\Lambda} \tag{5.8}$$

for each $D \in \mathscr{D}_G$. The normalization factor

$$Z_G := \sum_{D\in\mathscr{D}_G} e^{-\beta H_G(D)} \tag{5.9}$$

is called partition function.

The connection with the previous definition (5.3) is straightforward. Indeed setting

$w_{ij} \equiv e^{\beta h_{ij}^{(\mathrm{d})}}$ and $x_i \equiv e^{\beta h_i^{(\mathrm{m})}}$ one obtain $\mu_G \equiv \mathcal{G}_G(D)$ for each $D \in \mathscr{D}_G$.

The extension to monomer-dimer models of the *Gibbs state* 1.1.6 and *pressure density* (1.4) definitions are obvious.

Moreover, mirroring the construction of a Disordered Spin Models (section 1.2), a *Disordered Monomer-Dimer model* is a monomer-dimer model with random parameters, namely $h_G^{(\mathrm{d})}$ and $h_\Lambda^{(\mathrm{m})}$ are two families of random variables defined in some auxiliary probability space. The definitions of *random Gibbs measure, quenched state and quenched pressure* density are analogous.

5.1.1 Some examples of applications

As for Spin Models on Graphs, the class of Monomer Dimer models on Graphs defined in (5.1.5), can be tough as a preliminary setting for a statistical mechanic description of various systems. By construction they are used to describe the systems with an hard-core constraint between pairs of components. Here the analogous of the problems listed in section 1.3.

• *Absorption of diatomic gas molecules*: Consider a diatomic gas adsorbed on a solid material. The finite set Λ labels the allowed sites of the solid, and G its molecular structure. We will assume that the gas is diatomic, namely each molecule of the gas is composed by two atoms (like the oxygen O_2). Since the chemical bonds inside the molecules do not break, we assume that two atoms belonging to the same molecule of gas can only deposit on two neighbouring sites of the graph. This is an hard core constraint between pair of sites so its natural to put the model in a monomer dimer setting. Indeed, each arrangement of the diatomic molecules on the graph is encoded in a vector $\alpha \in \{0,1\}^{|E|}$, where E is the edge set of the graph. But on the contrary not every element $\{0,1\}^{|E|}$

represents an allowed arrangement, because of the constraint given by (5.6), in other words an allowed arrangement is a dimer configuration on G and the monomers represents the empty sites.

The energy associated to a possible dimer configuration D is represented by the MD Hamiltonian (5.7). In particular, the parameter $h_i^{(\mathrm{d})}j$ can be interpreted as an *effective potential* due the solid, acting on the dimer in ij and $h_i^{(\mathrm{m})}$ the tendency of the site i to be occupied.

• *Optimization problems*: By definition the monomer dimer problem is a weighed matching problem so, for example, it can be used to describe the following situation. Let us suppose we have a network of people represented by a finite graph G. We will consider different ways to divide the group into married and single people. The monogamy constraint implies an individual can be married at most one time, in other words an allowed arrangement is a dimer configuration on G and the monomers represents single.

The energy associated to such as arrangement is represented by the MD Hamiltonian (5.7). In particular, the parameter $h_{ij}^{(\mathrm{d})}$ can be interpreted as how much people i wants be married with j and $h_i^{(\mathrm{m})}$ as the tendency of people i to be single.

5.2 The Gaussian representation

In this section we prove that the partition function of a generic MD model on a graph G admits a representation in terms of moments of a gaussian vector . Without loss of generality we work with the partition function Z_N on the complete graph K_N (see remark 5.1.3.

Proposition 5.2.1. *(Gaussian representation) The partition function of any monomer-dimer model on K_N vertices can be written as*

$$Z_N \;=\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{i=1}^{N}(\xi_i + x_i)\Big]\,, \tag{5.10}$$

where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$ is a Gaussian random vector with mean 0 and covariance matrix $W = (w_{ij})_{i,j=1,\dots,N}$. Here the diagonal entries w_{ii} are arbitrary numbers, chosen in such a way that W is a positive semi-definite matrix.

Proof. As already noticed the dimer configurations on the complete graph are the partitions into pairs of all possible $A \subseteq \{1, \dots, N\}$, hence

$$Z_N = \sum_{D \in \mathscr{D}_N} \prod_{ij \in D} w_{ij} \prod_{i \in M_N(D)} x_i = \sum_{A \subseteq \{1,\dots,N\}} \sum_{\substack{P \text{ partition} \\ \text{of } A \text{ into pairs}}} \prod_{ij \in P} w_{ij} \prod_{i \in A^c} x_i \,. \quad (5.11)$$

Now choose w_{ii} for $i = 1, \dots, N$ such that the matrix $W = (w_{ij})_{i,j=1,\dots,N}$ is positive semi-definite[1]. Then there exists an (eventually degenerate) Gaussian vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$ with mean 0 and covariance matrix W. And by the Wick-Isserlis theorem (identity (A2) in the theorem A1)

$$\mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{i \in A} \xi_i\Big] = \sum_{\substack{P \text{ partition} \\ \text{of } A \text{ into pairs}}} \prod_{ij \in P} w_{ij} \,. \quad (5.12)$$

Substituting (5.12) into (5.11) one obtains

$$Z_N = \mathbb{E}_{\boldsymbol{\xi}}\Big[\sum_{A \subseteq \{1,\dots,N\}} \prod_{i \in A} \xi_i \prod_{i \in A^c} x_i\Big] = \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{i=1}^{N} (\xi_i + x_i)\Big] \,. \quad (5.13)$$

□

Remark 5.2.2. In some sense, the Gaussian representation (5.10) "factorize" the hard-core constraints in the same way as the Hubbard-Stratonovich transform decouples the two-body interactions in spin models. Indeed, consider a generic partition function (assume without loss of generality the complete graph and the inverse temperature equals 1):

$$Z_N^{\text{Ising}} = \sum_{\sigma \in \{\pm 1\}^N} e^{\sum_{1 \le i < j \le N} J_{ij}\sigma_i\sigma_j} \, e^{\sum_{i=1}^N h_i \sigma_i} \propto \sum_{\sigma \in \{\pm 1\}^N} e^{\frac{1}{2}\sum_{1 \le i,j \le N} J_{ij}\sigma_i\sigma_j} \, e^{\sum_{i=1}^N h_i \sigma_i} \,,$$

[1]For example one can choose $w_{ii} \ge \sum_{j \ne i} w_{ij}$ for every $i = 1, \dots, N$. W can be diagonalized and has non-negative eigenvalues by the Gershgorin circle theorem, hence it is positive semi-definite.

where we set $J_{ij} = J_{ji}$ and $J_{ii} \geq \sum_{j \neq i} |J_{ij}|$. In this way $J = (J_{ij})_{i,j=1,\dots,N}$ is a real positive semi-definite matrix, by Gershgorin circle theorem. Apply the Hubbard-Stratonovich transform (namely compute the Gaussian moment generating function) to obtain

$$Z_N^{\text{Ising}} \propto \sum_{\sigma \in \{\pm 1\}^N} \mathbb{E}_{\boldsymbol{\xi}'} \left[e^{\sum_{i=1}^N \xi_i \sigma_i} \right] e^{\sum_{i=1}^N h_i \sigma_i} \propto \mathbb{E}_{\boldsymbol{\xi}'} \left[\prod_{i=1}^N \cosh(\xi_i' + h_i) \right],$$

where $\boldsymbol{\xi}' = (\xi_1', \dots, \xi_N')$ is a Gaussian random vector with mean 0 and covariance matrix J.

As an application of the Gaussian representation we show that the well-know Heilmann-Lieb recursion [86] for the partition function of monomer-dimer models can be proved by means of a Gaussian integration by parts.

Proposition 5.2.3. *(Heilmann-Lieb recursion) Let $G = (\Lambda, E)$ be a finite simple graph and consider an MD model on G. Fix $i \in \Lambda$ and look at its adjacent vertices $j \sim i$, then it holds*

$$Z_G = x_i \, Z_{G-i} + \sum_{j \sim i} w_{ij} \, Z_{G-i-j} \, . \tag{5.14}$$

Here $G - i$ is the graph obtained from G deleting the vertex i and all its incident edges.

Proof using Gaussian integration by parts. Set $N := |\Lambda|$. Introduce zero dimer weights $w_{hk} = 0$ for the pairs $hk \notin E$, so that $Z_G = Z_N$ (see remark 5.1.3). Following proposition 5.2.1, introduce an N-dimensional Gaussian vector $\boldsymbol{\xi}$ with mean 0 and covariance matrix W. Then write the identity (5.10) isolating the vertex i:

$$Z_G = \mathbb{E}_{\boldsymbol{\xi}} \left[\prod_{k=1}^N (\xi_k + x_k) \right] = x_i \, \mathbb{E}_{\boldsymbol{\xi}} \left[\prod_{k \neq i} (\xi_k + x_k) \right] + \mathbb{E}_{\boldsymbol{\xi}} \left[\xi_i \prod_{k \neq i} (\xi_k + x_k) \right] . \tag{5.15}$$

Now apply the Gaussian integration by parts (identity (A1) in the theorem A1)

to the second term on the r.h.s. of (5.15):

$$\mathbb{E}_{\boldsymbol{\xi}}\Big[\xi_i \prod_{k\neq i}(\xi_k+x_k)\Big] \;=\; \sum_{j=1}^{N} \mathbb{E}_{\boldsymbol{\xi}}[\xi_i\xi_j]\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\frac{\partial}{\partial\xi_j}\prod_{k\neq i}(\xi_k+x_k)\Big] \;=\; \sum_{j\neq i} w_{ij}\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{k\neq i,j}(\xi_k+x_k)\Big]\,. \tag{5.16}$$

Notice that summing over $j \neq i$ in the r.h.s. of (5.16) is equivalent to sum over $j \sim i$, since by definition $w_{ij} = 0$ if $ij \notin E$. Substitute (5.16) in (5.15):

$$Z_G \;=\; x_i\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{k\neq i}(\xi_k + x_k)\Big] \;+\; \sum_{j\sim i} w_{ij}\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{k\neq i,j}(\xi_k + x_k)\Big]\,. \tag{5.17}$$

To conclude observe that $(\xi_k)_{k\neq i}$ is an $(N-1)$-dimensional Gaussian vector with mean 0 and covariance $(w_{hk})_{h,k\neq i}$. Hence by proposition 5.2.1

$$Z_{G-i} \;=\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{k\neq i}(\xi_k + x_k)\Big]\,. \tag{5.18}$$

And similarly

$$Z_{G-i-j} \;=\; \mathbb{E}_{\boldsymbol{\xi}}\Big[\prod_{k\neq i,j}(\xi_k + x_k)\Big]\,. \tag{5.19}$$

Substitute the identities (5.18), (5.19) into (5.17) to obtain the identity (5.14).

□

5.3 The MD model on K_N with uniform weights

In this section we use the gaussian representation (5.2.1) to compute the pressure of the MD model of the complete graph with uniform weights. In order to lighten the notation, in this section, we assume without loss the inverse temperature $\beta = 1$. First let us illustrate some basic properties of the generic MD model. Let $G = (\Lambda, E)$ be a finite graph and $Z_G(x, w)$ the partition function with generic monomeric weights $x = (x_i)_{i\in\Lambda}$ and dimeric weights $w = (w_{ij})_{ij\in E}$

Proposition 5.3.1. *With bounded monomeric and dimeric weights $\underline{x} \leq x_i \leq \overline{x}$, $w_e \leq \overline{w}$, the following bounds for the pressure hold:*

$$\log \underline{x} \;\leq\; \frac{\log Z_G(x,w)}{|\Lambda|} \;\leq\; \log \overline{x} + \frac{|E|}{|\Lambda|}\,\log\big(1 + \frac{\overline{w}}{\overline{x}^2}\,\big)\,.$$

Moreover

Proposition 5.3.2. *If uniform dimeric (resp. monomeric) weights are considered, i.e.* $w_e \equiv w_0 \ \forall e \in E$ *(resp.* $x_i \equiv x_0 \ \forall i \in \Lambda$*), then it's possible to keep* $w = w_0$ *(resp.* $x = x_0$*) fixed and study only the dependence of the model on* x *(resp.* w*) without loss of generality. Indeed, using the relation* $|M(D)| + 2\,|D| = |\Lambda|$*, it's easy to check that*

$$Z_G(x, w_0) \;=\; (w_0)^{|\Lambda|/2}\; Z_\Lambda\big(\frac{x}{(w_0)^{1/2}}\,,\,1\big)\;; \tag{5.20}$$

$$Z_G(x_0, w) \;=\; (x_0)^{|\Lambda|}\; Z\big(1\,,\,\frac{w}{(x_0)^2}\big)\;. \tag{5.21}$$

Let K_N be the complete graph over N vertices with vertex set E_N. Notice $|E_N| = N(N-1)/2$.

We work with uniform weights $w_e \equiv w$ and $x_i \equiv x$ and we want $\log Z_N \;=\; \mathcal{O}(N)$. For this purpose, looking to proposition 5.3.1, we have to choose x, w such that $w/x^2 = \mathcal{O}(1/N)$. By 5.3.2 we can fix without loss of generality $w = 1/N$ and study

$$Z_N^{\text{MD}}(x) \;:=\; Z_N\big(x\,,\,\frac{1}{N}\big)\,, \tag{5.22}$$

Observe that the bounds of remark 5.3.1 become

$$\log x \;\le\; \frac{\log Z_N^{\text{MD}}(x)}{N} \;\le\; \log x + \frac{N-1}{2}\,\log\big(1 + \frac{1}{Nx^2}\big) \;\le\; \log x + \frac{1}{2\,x^2}\,.$$

The Hamiltonian function corresponding to $Z_N^{\text{MD}}(x)$ is, setting $x = e^h$,

$$H_N^{\text{MD}}(D, h) \;:=\; -\sum_{i=1}^{N} h\,\alpha_i(D) \;+\; \log N \sum_{e \in E_N} \alpha_e(D)\,, \tag{5.23}$$

On the complete graph it is possible to compute explicitly the pressure density.

Proposition 5.3.1. *Setting* $x = e^h$ *for* $h \in \mathbb{R}$*. The pressure per particle of the monomer-dimer model on the complete graph defined by hamiltonian (5.23) admits finite thermodynamic limit:*

$$\exists \lim_{N\to\infty} \frac{\log Z_N^{\text{MD}}(x)}{N} \;=\; p^{\text{MD}}(x) \;\in \mathbb{R}\,. \tag{5.24}$$

Furthermore the monomer density

$$m_N(x) := \frac{1}{Z_N^{\mathrm{MD}}(x)} \sum_{D \in \mathscr{D}_{K_N}} \frac{|M(D)|}{N} \exp(-H_N(D,x))$$

admits thermodynamic limit:

$$\exists \lim_{N\to\infty} m_N(x) \;=\; g(x) \;\in]0,1[\,. \tag{5.25}$$

$$p^{\mathrm{MD}}(\xi) \;:=\; -\frac{1-g(\xi)}{2} \;-\; \frac{1}{2}\log(1-g(\xi)) \;=\; -\frac{1-g(\xi)}{2} - \log g(\xi) + \xi \quad \forall\, \xi \in \mathbb{R} \tag{5.26}$$

$$g(\xi) \;:=\; \frac{1}{2}\,(\sqrt{e^{4\xi} + 4\,e^{2\xi}} - e^{2\xi}) \quad \forall\, \xi \in \mathbb{R}\,. \tag{5.27}$$

The first proof is due to Heilmann and Lieb [86] and is based on a recurrence relation (5.2.3), a second one is based on a simple combinatorial argument [28]. We notice also that is a particular case of a more general result given in chapter 7.

Chapter 6

A Monomer-Dimer model with Imitation

Let us briefly recall what is a Monomer-Dimer model on a graph G (see section 5.1 for a rigorous definition).

Each way to fully cover the vertices of a finite graph G by non-overlapping *dimers* (molecules which occupy two adjacent vertices) and monomers (molecules which occupy a single vertex) is called a *monomer-dimer configuration*. Let us associate to each dimer and monomer a weight and consider an Hamiltonian proportional to the weighted number of monomers and dimers: this defines a *Monomer-Dimer model* on G .

Monomer dimer models (MD models) were proposed in '30 to investigate the properties of diatomic oxygen molecules deposited on tungsten [113] or to study liquid mixtures in which the molecules are unequal in size[73]. The hard-core interaction accounts for the contact repulsion generated by the Pauli principle. In order to account also for the attractive component of the Van der Waals potential among monomers and dimers, one may consider an attractive interaction[114, 52, 53] among particles occupying neighbouring sites (as it was previously done for single atoms [72, 112]). More recently monomer-dimer models on diluted networks have attracted a considerable attention [126, 50, 27] and

they have been applied, with the addition of a ferromagnetic imitative interaction, also in social sciences [42].

The first mathematical approach to the problem is due to Heilmann and Lieb [86, 87] where they proved the celebrated recursion relation for the partition function 5.2.3 and used it to locate its the complex zeros. As a byproduct they show that, under general assumptions, the hard-core interaction is not enough to generate a phase transition. Exact solutions are know for the model on the complete graph, the 1-dimensional case [86] and on locally three like graphs [27]. In two dimension, i.e. for planar graph, a solution for pure dimer model was discovered independently by [92, 120].

In order to account also for the attractive component of the Van der Waals potential among monomers and dimers, one may consider an attractive interaction [114, 52, 53] among particles occupying neighbouring sites. If one includes attractive interactions among dimers with the same orientation Heilmann-Lieb proved [88] that on certain 2-dimensional regular lattices there is a phase transition: at low temperatures there is an orientational ordering, which is absent at high temperatures.

Here, we consider a Monomer-Dimer model with an attractive interaction called *IMD model* on the complete graph.

This chapter is entirely based on [28] and is organized as follow: in section 6.1 we introduce the model and show that the monomer density, i.e. the expectation value w.r.t. the Gibbs measure of the fraction of sites occupied by monomers, is the order parameter of the model in the usual sense: the *t.l.* of the pressure is expressed a one-dimensional variational principle in the monomer density.

In section 6.2 we prove that there is a phase transition between a high monomer density phase and a high dimer density (i.e. low monomer density) phase and we characterize this transition in details.

6.1 The model and the main result

Let $G = (\Lambda, E)$ be a finite simple graph and let us denote by $\mathscr{D}_G$ the dimer configuration space on G defined in 5.1.1. Keep in mind its representation 5.1.4 in terms of occupancy variables α. We want modify the classical MD model defined in the previous section introducing an imitative interaction among the particles (monomers and dimers). As in the spin models this interaction can be represented , in the Hamiltonian formulation with the occupancy variables, by a term tuned by a positive parameter J and connecting two $\alpha's$.

Definition 6.1.1. The *Imitative Monomer-Dimer model* (IMD model) on G is obtained by assigning an external field $h \in \mathbb{R}$ and an imitation coefficient $J \geq 0$ and then considering the *hamiltonian* $H_G^{\text{IMD}} \colon \mathscr{D}_G \to \mathbb{R}$,

$$H_G^{\text{IMD}} := -\sum_{v \in \Lambda} h\, \alpha_v \; - \sum_{uv \in E} J \left(\alpha_u\, \alpha_v + (1 - \alpha_u)\, (1 - \alpha_v) \right) . \tag{6.1}$$

The choice of the hamiltonian naturally induces a Gibbs probability measure on the space of configurations $\mathscr{D}_G$:

$$\mu_G^{\text{IMD}}(D) := \frac{1}{Z_G^{\text{IMD}}} \exp(-H_G^{\text{IMD}}(D)) \quad \forall\, D \in \mathscr{D}_G \, , \tag{6.2}$$

where $Z_G^{\text{IMD}} := \sum_{D \in \mathscr{D}_G} \exp(-H_G^{\text{IMD}}(D))$ is the normalizing factor, called *partition function*. Its natural logarithm $\log Z_G^{\text{IMD}}$ is called *pressure*.

The *monomer density*, that is the expected fraction of monomers on the graph, can be obtained computing the derivative of the pressure per particle with respect to h :

$$m_G^{\text{IMD}} := \sum_{D \in \mathscr{D}_G} \frac{|M(D)|}{|V|}\, \mu_G^{\text{MD}}(D) \, = \, \frac{\partial}{\partial h} \frac{\log Z_G^{\text{IMD}}}{|V|} \, . \tag{6.3}$$

We explicitly observe that taking $J = 0$ the model reduces to the usual MD model 5.3, studied by Heilmann and Lieb [86, 87]. At $J = 0$ the model is characterised only by a topological interaction, that is the hard-core constraint which defines the space of states $\mathscr{D}_G$. As proved by Heilmann and Lieb [86, 87] this interaction is not sufficient to originate a phase transition: when the

thermodynamic limit of the normalized pressure exists, it has to be an analytic function of the parameter h. Introducing the parameter $J \geq 0$ we consider add another type of interaction: the state of a vertex conditions the state of its neighbours, pushing each other to behave in the same way (*imitative interaction* among sites, *attractive interaction* among similar particles).

Remark 6.1.2. At first sight it seems that a more general hamiltonian than (6.1) could be considered. Fix a monomer external field $h^{(\mathrm{m})} \in \mathbb{R}$, a dimer eternal field $h^{(\mathrm{d})} \in \mathbb{R}$, a monomer imitation coefficient $J^{(\mathrm{m})} \in \mathbb{R}$, a dimer imitation coefficient $J^{(\mathrm{d})} \in \mathbb{R}$ and a counter-imitation coefficient $J^{(\mathrm{md})} \in \mathbb{R}$ and set:

$$\begin{aligned}\widetilde{H}_G^{\mathrm{IMD}} := &- \sum_{v\in\Lambda} h^{(\mathrm{m})}\,\alpha_v \;- \sum_{e\in E} h^{(\mathrm{d})}\,\alpha_e \;- \sum_{uv\in E} J^{(\mathrm{m})}\,\alpha_u\,\alpha_v \;- \sum_{uv\in E} J^{(\mathrm{d})}\,(1-\alpha_u)\,(1-\alpha_v) \;+ \\ &- \sum_{uv\in E} J^{(\mathrm{md})}\left(\alpha_u\,(1-\alpha_v) + (1-\alpha_u)\,\alpha_v\right).\end{aligned} \tag{6.4}$$

Now it easy to check that the following two relations hold

$$\alpha_u\,(1-\alpha_v) + (1-\alpha_u)\,\alpha_v = -\alpha_u\,\alpha_v - (1-\alpha_u)\,(1-\alpha_v) + 1 \tag{6.5}$$

and as a consequence the hamiltonian (6.4) rewrites as

$$\widetilde{H}_G^{\mathrm{IMD}} \;=\; -\,C' \;- \sum_{v\in\Lambda} h'\,\alpha_v \;- \sum_{uv\in E} J'\,\alpha_u\,\alpha_v \;- \sum_{uv\in E} J''\,(1-\alpha_u)\,(1-\alpha_v) \tag{6.6}$$

with $h' = h^{(\mathrm{m})} - h^{(\mathrm{d})}/2$, $J' = J^{(\mathrm{m})} - J^{(\mathrm{md})}$, $J'' = J^{(\mathrm{d})} - J^{(\mathrm{md})}$ and $C' = |V|/2$.
Now if the graph G is regular of degree r, also the following relation holds

$$\sum_{uv\in E} (\alpha_u + \alpha_v) = r\; Card\{M(D), D \in \mathscr{D}_G\} \tag{6.7}$$

and as a consequence the hamiltonian (6.6) rewrites as

$$\widetilde{H}_G^{\mathrm{IMD}} \;=\; -\,C \;- \sum_{v\in\Lambda} h\,\alpha_v \;- \sum_{uv\in E} J\left(\alpha_u\,\alpha_v + (1-\alpha_u)\,(1-\alpha_v)\right) \tag{6.8}$$

with $h = h' + (J' - J'')\,r$, $J = (J' + J'')/2$ and $C = C' + (J'' - J')\,|E|$.
In conclusion this last identity shows that on any regular graph the general hamiltonian (6.4) is in fact equivalent to the hamiltonian (6.1); indeed they

differ only by a constant which does not modify the probability measure (6.2). Moreover observe that the imitation condition $J \geq 0$ corresponds to $J^{(\mathrm{m})}+J^{(\mathrm{d})} \geq 2J^{(\mathrm{md})}$.

In this chapter we study the IMD model on the complete graph, that is we take $G = K_N = (\Lambda_N, E_N)$ with $\Lambda_N = \{1, \dots, N\}$ and $E_N = \{\{u, v\} \,|\, u, v \in \Lambda_N,\, u < v\}$.

Since the number of edges is of order N^2, in order to keep the pressure of order N we need to normalize the external field and the imitation coefficient. Therefore we will consider the hamiltonian $H_N^{\mathrm{IMD}} \colon \mathscr{D}_{K_N} \to \mathbb{R}$,

$$H_N^{\mathrm{IMD}} := -\sum_{v \in V_N} h\,\alpha_v + \log N \sum_{e \in E_N} \alpha_e - \sum_{uv \in E_N} \frac{J}{N} \big(\alpha_u\,\alpha_v + (1-\alpha_u)\,(1-\alpha_v)\big)\,, \tag{6.9}$$

the partition function $Z_N^{\mathrm{IMD}} := \sum_{D \in \mathscr{D}_{K_N}} \exp(-H_N^{\mathrm{IMD}}(D))$ and the monomer density $m_N^{\mathrm{IMD}} := \frac{1}{Z_N^{\mathrm{IMD}}} \sum_{D \in \mathscr{D}_{K_N}} \frac{|\mathscr{M}(D)|}{N} \exp(-H_N^{\mathrm{IMD}}(D))$.

The main result of this section is the following theorem, where in the limit $N \to \infty$ the model is solved in terms of a one-dimensional variational principle.

Theorem 6.1.3. *Let $h \in \mathbb{R}$, $J \geq 0$. The pressure per particle of the imitative monomer-dimer model on the complete graph defined by hamiltonian (6.9) admits finite thermodynamic limit:*

$$\exists \lim_{N \to \infty} \frac{\log Z_N^{\mathrm{IMD}}}{N} =: p^{\mathrm{IMD}} \in \mathbb{R}\,. \tag{6.10}$$

This limit satisfies a variational principle:

$$p^{\mathrm{IMD}} = \sup_m \widetilde{p}\,(m) \tag{6.11}$$

where the sup *can be taken indifferently over $m \in [0,1]$ or $m \in \mathbb{R}$, and*

$$\widetilde{p}\,(m) := -J\,m^2 + \frac{1}{2}J + p^{\mathrm{MD}}\big((2m-1)\,J + h\big) \quad \forall\, m \in \mathbb{R} \tag{6.12}$$

$$p^{\mathrm{MD}}(\xi) := -\frac{1-g(\xi)}{2} - \frac{1}{2}\log(1-g(\xi)) = -\frac{1-g(\xi)}{2} - \log g(\xi) + \xi \quad \forall\, \xi \in \mathbb{R} \tag{6.13}$$

$$g(\xi) := \frac{1}{2}\,(\sqrt{e^{4\xi} + 4\,e^{2\xi}} - e^{2\xi}) \quad \forall\, \xi \in \mathbb{R}\,. \tag{6.14}$$

Furthermore the function $m \mapsto \widetilde{p}(m)$ *attains its maximum in (at least) one point* $m^* \in]0,1[$ *, which is a solution of the the consistency equation*

$$m \;=\; g\big((2m-1)\,J + h\big)\;. \tag{6.15}$$

At each value of the parameters (h, J) *such that* $h \mapsto m^*(h, J)$ *is differentiable, the monomer density admits thermodynamic limit and precisely:*

$$\exists \lim_{N\to\infty} m_N^{\text{IMD}} \;=\; m^* \;\in]0,1[\;. \tag{6.16}$$

This result relies on two main facts:

1) for $J = 0$, by proposition 5.3.1 the thermodynamic limit of the pressure per particle can be computed explicitly and turns out to be $p^{\text{MD}}(h)$;

2) for $J > 0$ the hamiltonian (6.9) can be expressed as a quadratic form in the hamiltonian with $J = 0$.

Therefore before proving the theorem we properly state and prove the second result.

Given a dimer configuration D on the complete graph K_N, denote the fraction of vertices covered by monomers by

$$m_N(D) := \frac{|M(D)|}{N} \;\in [0,1]\;.$$

On the complete graph the hamiltonian (6.1) admits a useful rewriting, which shows that it depends on a dimer configuration D only via the quantity $m_N(D)$.

Lemma 6.1.4.

$$H_N^{\text{IMD}} \;=\; -N\left(J\,m_N^2 \;+\; b_N\,m_N \;+\; c_N\right) \tag{6.17}$$

with $b_N := \frac{1}{2}\log N + h - J$ *and* $c_N := -\frac{1}{2}\log N + \frac{N-1}{2N}\,J$ *.*

Proof. Using the identity (6.7) (the complete graph is regular of degree $N-1$), the hamiltonian (6.9) rewrites as

$$H_N^{\text{IMD}} = \frac{N}{2}\log N - \frac{N(N-1)}{2}\,\frac{J}{N} - \big(h + \frac{1}{2}\log N - (N-1)\,\frac{J}{N}\big)\sum_{v\in V_N}\alpha_v - 2\,\frac{J}{N}\sum_{uv\in E_N}\alpha_u\alpha_v\;.$$

Then on the complete graph it holds

$$2 \sum_{uv \in E_N} \alpha_u \alpha_v \;=\; \Big(\sum_{v \in \Lambda_N} \alpha_v \Big)^2 \;-\; \sum_{v \in \Lambda_N} \alpha_v \;. \tag{6.18}$$

Substituting in the previous expression one obtains

$$H_N^{\mathrm{IMD}} = \frac{N}{2} \log N - \frac{N-1}{2}\, J \;-\; \big(h + \frac{1}{2} \log N - J\big) \sum_{v \in \Lambda_N} \alpha_v \;-\; \frac{J}{N} \Big(\sum_{v \in \Lambda_N} \alpha_v \Big)^2$$

and since $\sum_{v \in \Lambda_N} \alpha_v = N\, m_N$ the identity (6.17) is proved. □

Now using proposition 5.3.1 and lemma 6.1.4 we are able to prove theorem 6.1.3. Our technique is the same used by Guerra[80] to solve the ferromagnetic Ising model on the complete graph.

Proof of Theorem 6.1.3. The proof is done providing a lower and an upper bound for the pressure per particle.

[**LowerBound**] Fix $m \in \mathbb{R}$. As $\big(m_N(D) - m\big)^2 \geq 0$, clearly $m_N(D)^2 \geq 2\, m\, m_N(D) - m^2$. Hence by lemma 6.1.4, using the hypothesis $J \geq 0$,

$$\begin{aligned} -H_N^{\mathrm{IMD}}(D) \;&=\; N \left(J\, m_N(D)^2 \;+\; b_N\, m_N(D) \;+\; c_N \right) \;\geq \\ &\geq\; N \left((2J\, m + b_N)\, m_N(D) \;-\; J\, m^2 + c_N \right) \end{aligned}$$

thus

$$\begin{aligned} Z_N^{\mathrm{IMD}} \;&=\; \sum_D \exp(-H_N^{\mathrm{IMD}}(D)) \;\geq\; \sum_D \exp N \big((2J\, m + b_N)\, m_N(D) - J\, m^2 + c_N \big) \;= \\ &=\; e^{N\, \gamma_N(m)}\; Z_N^{\mathrm{MD}}\big(\xi(m)\big) \end{aligned}$$

where $\gamma_N(m) := -J\, m^2 + \frac{N-1}{2N}\, J$ and $\xi(m) := 2J\, m + h - J\,.$

[**UpperBound**] m_N takes values in the set $\mathcal{A}_N := \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, 1\}$. Clearly, writing δ for the Kronecker delta, $\sum_{m \in \mathcal{A}_N} \delta_{m, m_N(D)} = 1$ and $F(m_N(D)^2)\, \delta_{m, m_N(D)} = F(2\, m\, m_N(D) - m^2)\, \delta_{m, m_N(D)}$ for any function F. Hence by lemma 6.1.4,

$$\begin{aligned} \delta_{m, m_N(D)}\, \exp(-H_N^{\mathrm{IMD}}(D)) \;&=\; \delta_{m, m_N(D)}\, \exp N (J\, m_N(D)^2 + b_N\, m_N(D) + c_N) \;= \\ &=\; \delta_{m, m_N(D)}\, \exp N \big((2J\, m + b_N)\, m_N(D) \;- J\, m^2 + c_N \big) \end{aligned}$$

thus

$$\begin{aligned} Z_N^{\mathrm{IMD}} &= \sum_D \sum_{m\in\mathcal{A}_N} \delta_{m,m_N(D)}\ \exp(-H_N^{\mathrm{IMD}}(D)) = \\ &= \sum_D \sum_{m\in\mathcal{A}_N} \delta_{m,m_N(D)}\ \exp N\big((2J\,m + b_N)\,m_N(D)\ - J\,m^2 + c_N\big) \le \\ &\le \sum_{m\in\mathcal{A}_N} \sum_D \exp N\big((2J\,m + b_N)\,m_N(D)\ - J\,m^2 + c_N\big) = \\ &= \sum_{m\in\mathcal{A}_N} e^{N\,\gamma_N(m)}\ Z_N^{\mathrm{MD}}\big(\xi(m)\big) \le (N+1) \sup_{m\in[0,1]} \big\{e^{N\,\gamma_N(m)}\ Z_N^{\mathrm{MD}}\big(\xi(m)\big)\big\}\,. \end{aligned}$$

Therefore putting together lower and upper bound we have found:

$$\sup_{m\in[0,1]} \big\{e^{N\,\gamma_N(m)}\ Z_N^{\mathrm{MD}}\big(\xi(m)\big)\big\} \le Z_N^{\mathrm{IMD}} \le (N+1) \sup_{m\in[0,1]} \big\{e^{N\,\gamma_N(m)}\ Z_N^{\mathrm{MD}}\big(\xi(m)\big)\big\}\,.$$

Then, taking the logarithm and dividing by N,

$$0 \le \frac{\log Z_N^{\mathrm{IMD}}}{N}\ - \sup_{m\in[0,1]} \big\{\gamma_N(m)\ + \frac{\log Z_N^{\mathrm{MD}}\big(\xi(m)\big)}{N}\big\} \le \frac{\log(N+1)}{N} \xrightarrow[N\to\infty]{} 0\,.$$

Now the pressure per particle $h \mapsto \frac{\log Z_N^{\mathrm{MD}}(h)}{N}$ is a convex function, hence as $N\to\infty$ the convergence $\frac{\log Z_N^{\mathrm{MD}}(h)}{N} \to p^{\mathrm{MD}}(h)$ of proposition 5.3.1 is uniform in h on compact sets. Moreover $\gamma_N(m) \to \gamma(m) := -J\,m^2 + \frac{1}{2}\,J$ uniformly in m as $N\to\infty$. Therefore

$$\gamma_N(m)\ + \frac{\log Z_N^{\mathrm{MD}}\big(\xi(m)\big)}{N} \xrightarrow[N\to\infty]{} \gamma(m)\ + p^{\mathrm{MD}}\big(\xi(m)\big)$$

and the convergence is uniform in m on compact sets. As a consequence also

$$\sup_{m\in[0,1]} \big\{\gamma_N(m)\ + \frac{\log Z_N^{\mathrm{MD}}\big(\xi(m)\big)}{N}\big\} \xrightarrow[N\to\infty]{} \sup_{m\in[0,1]} \big\{\gamma(m)\ + p^{\mathrm{MD}}\big(\xi(m)\big)\big\}\,.$$

This concludes the proof of (6.10) and (6.11).

It remains to prove (6.15) and (6.16). First of all observe that

$$\frac{\partial\widetilde{p}}{\partial m}(m) = -2J\,m + 2J\,g\big((2m-1)\,J + h\big)\,,$$

since $(p^{\mathrm{MD}})' = g$ (see proposition 5.3.1). It holds $\frac{\partial\widetilde{p}}{\partial m}(m) > 0$ for all $m \le 0$ and $\frac{\partial\widetilde{p}}{\partial m}(m) > 0$ for all $m \ge 1$, therefore the function $m \mapsto \widetilde{p}\,(m)$ attains its

global maximum inside the interval $]0,1[$ and any global maximum point m^* is a critical point of $\widetilde{p}$, i.e. satisfies equation (6.15).
Now $\frac{1}{N}\log Z_N^{\mathrm{IMD}}(h,J)$ is a convex function of h and, as shown before, it converges to $p^{\mathrm{IMD}}(h,J)=\widetilde{p}\,(m^*(h,J),h,J)$ as $N\to\infty$. Therefore, assuming that $m^*(h,J)$ is differentiable in h, the monomer density $m_N^{\mathrm{IMD}}=\frac{\partial}{\partial h}\frac{1}{N}\log Z_N^{\mathrm{IMD}}$ converges to $\frac{\partial}{\partial h}p^{\mathrm{IMD}}$. Thus to prove (6.16) it suffices to compute this derivative:

$$\frac{\partial p^{\mathrm{IMD}}}{\partial h} = \frac{\mathrm{d}}{\mathrm{d}h}\,\widetilde{p}\,\big(m^*(h,J),h,J\big) = \underbrace{\frac{\partial\widetilde{p}}{\partial m}(m^*)}_{=0}\,\frac{\partial m^*}{\partial h}+\underbrace{\frac{\partial\widetilde{p}}{\partial h}}_{=(p^{\mathrm{MD}})'} = g\big((2m^*-1)\,J+h\big) = m^*\,.$$

□

6.2 The properties of the solution

In this section we study the properties of the solution provided by theorem 6.1.3. We divide the analysis in three subsections. In subsection 6.2.1 we study all the *stationary points* of the function $m\mapsto\widetilde{p}\,(m,h,J)$. One of them will be the *global maximum point* m^* we are interested in since it represents the monomer density. We provide their complete classification, regularity properties and asymptotic behaviour as functions of the parameters h and J. As a byproduct in subsection 6.2.2 we are able to identify the region where there exists a unique global maximum point m^*. The resulting picture is the following: the function m^* is single-valued and continuous on the plane (h,J) with the exception of a curve Γ that is implicitly defined, and moreover m^* is smooth outside Γ union its endpoint (h_c,J_c). This curve play a crucial physical role since it represents the coexistence of two different thermodynamic phases and the point (h_c,J_c) is the *critical point* of the system. In physical jargon we say that a *phase transition* occur. In subsection 6.2.3 we compute the *critical exponents* that characterizes the behaviour of m^* near the critical point (h_c,J_c).

6.2.1 Analysis of the stationary points: classification, regularity properties, asymptotic behaviour.

Let us identify the stationary points of the function $\widetilde{p}\,(m,h,J)$ defined by (6.12). Remembering that $(p^{\mathrm{MD}})' = g$, one computes

$$\frac{\partial \widetilde{p}}{\partial m}\,(m,h,J) \;=\; -2J\,m \;+\; 2J\;g\big((2m-1)J+h\big) \tag{6.19}$$

$$\frac{\partial^2 \widetilde{p}}{\partial m^2}\,(m,h,J) \;=\; -2J \;+\; (2J)^2\;g'\big((2m-1)J+h\big) \tag{6.20}$$

Since $0 < g < 1$, it follows that for every $J > 0$, $h \in \mathbb{R}$

$$\frac{\partial \widetilde{p}}{\partial m}\,(m,h,J) > 0 \quad \forall\, m \in \,]-\infty, 0]\;, \quad \frac{\partial \widetilde{p}}{\partial m}\,(m,h,J) < 0 \quad \forall\, m \in [1,\infty[\;. \tag{6.21}$$

Therefore $\widetilde{p}\,(\cdot\,,h,J)$ attains its maximum in (at least) one point $m = m^*(h,J) \in \,]0,1[\,$, which satisfies

$$\frac{\partial \widetilde{p}}{\partial m}\,(m,h,J) = 0 \qquad \text{i.e.} \quad m = g\big((2m-1)J+h\big)\;, \tag{6.22}$$

$$\frac{\partial^2 \widetilde{p}}{\partial m^2}\,(m,h,J) \le 0 \qquad \text{i.e.} \quad g'\big((2m-1)J+h\big) \le \frac{1}{2J}\;. \tag{6.23}$$

The stationary points are characterized by equation (6.22), which can not be explicitly solved. Anyway their properties and a rough approximation of their values can be determined by studying inequality (6.23), which admits explicit solution.

The next proposition displays the intervals of concavity/convexity of the function $m \mapsto \widetilde{p}\,(m,h,J)$. Set

$$J_c := \frac{1}{4\,(3-2\sqrt{2})} \;\approx\; 1.4571\;. \tag{6.24}$$

Proposition 6.2.1. *For $0 < J < J_c$ and $h \in \mathbb{R}$*

$$\frac{\partial^2 \widetilde{p}}{\partial m^2}\,(m,h,J) \;<\; 0 \quad \forall\, m \in \mathbb{R}\;.$$

For $J \ge J_c$ and $h \in \mathbb{R}$

$$\frac{\partial^2 \widetilde{p}}{\partial m^2}\,(m,h,J) \begin{cases} < 0 & \textit{iff } m < \phi_1(h,J) \textit{ or } m > \phi_2(h,J) \\ > 0 & \textit{iff } \phi_1(h,J) < m < \phi_2(h,J) \end{cases},$$

where for $i = 1, 2$

$$\phi_i(h, J) := \frac{1}{2} - \frac{h}{2J} + \frac{1}{4J} \log a_i(J) \,, \tag{6.25}$$

$$a_i(J) := \frac{-(\frac{1}{(2J)^2} + \frac{8}{2J} - 4) + (-1)^i (2 - \frac{1}{2J}) \sqrt{\frac{1}{(2J)^2} - \frac{12}{2J} + 4}}{\frac{4}{2J}} \,. \tag{6.26}$$

Observe that $\phi_1(h, J) \leq \phi_2(h, J)$ *for all* $h \in \mathbb{R}$, $J \geq J_c$ *and equality holds iff* $J = J_c$ *(since* $a_1(J_c) = a_2(J_c)$*).*

Proof. It follows from the expression (6.20) through a direct computation done in lemma E1 of the Appendix, taking $\xi = (2m - 1)J + h$ and $c = \frac{1}{2J}$. □

Using the previous proposition we can determine *how many*, of *what kind* and *where* the stationary points of $\widetilde{p}\,(\cdot\,, h, J)$ are.

Proposition 6.2.2 (Classification)**.** *The equation (6.22) in* m *has the following properties:*

1. *If* $0 < J \leq J_c$ *and* $h \in \mathbb{R}$*, there exists only one solution* $m(h, J)$*. It is the maximum point of* $\widetilde{p}\,(\cdot\,, h, J)$*.*

2. *If* $J > J_c$ *and* $\psi_2(J) < h < \psi_1(J)$*, then there exist three solutions* $m_1(h, J)$*,* $m_0(h, J)$*,* $m_2(h, J)$*. Moreover* $m_1(h, J) < \phi_1(h, J)$ *and* $m_2(h, J) > \phi_2(h, J)$ *are two local maximum points, while* $\phi_1(h, J) < m_0(h, J) < \phi_2(h, J)$ *is a local minimum point of* $\widetilde{p}\,(\cdot\,, h, J)$*.*

3. *If* $J > J_c$ *and* $h > \psi_1(J)$*, there exists only one solution* $m_2(h, J)$*. Moreover* $m_2(h, J) > \phi_2(h, J)$ *and it is the maximum point of* $\widetilde{p}\,(\cdot\,, h, J)$*.*

4. *If* $J > J_c$ *and* $h = \psi_1(J)$*, there exist two solution* $m_1(h, J)$*,* $m_2(h, J)$ *. Moreover* $m_1(h, J) = \phi_1(h, J)$ *is a point of inflection, while* $m_2(h, J) > \phi_2(h, J)$ *is the maximum point of* $\widetilde{p}\,(\cdot\,, h, J)$*.*

5. *If* $J > J_c$ *and* $h < \psi_2(J)$*, there exists only one solution* $m_1(h, J)$*. Moreover* $m_1(h, J) < \phi_1(h, J)$ *and it is the maximum point of* $\widetilde{p}\,(\cdot\,, h, J)$*.*

6. *If $J > J_c$ and $h = \psi_2(J)$, there exist two solutions $m_1(h,J)$, $m_2(h,J)$. Moreover $m_2(h,J) = \phi_2(h,J)$ is a point of inflection, while $m_1(h,J) < \phi_1(h,J)$ is the maximum point of $\widetilde{p}\,(\cdot\,,h,J)$.*

Here ϕ_1, ϕ_2 are defined by (6.25), while for $i = 1,2$ and $J \geq J_c$

$$\psi_i(J) \;:=\; J + \frac{1}{2}\log a_i(J) - 2J\, g\big(\frac{1}{2}\log a_i(J)\big)\;, \tag{6.27}$$

where a_i and g are defined respectively by (6.26) and (6.14). Observe that $\psi_2(J) \leq \psi_1(J)$ for all $J \geq J_c$ and equality holds iff $J = J_c$.

Proof. Fix $h \in \mathbb{R}$, $J > 0$ and to shorten the notation set $G(m) := \frac{\partial \widetilde{p}}{\partial m}(m,h,J)$, observing it is a continuous (smooth) function.

• Suppose $J \leq J_c$. By proposition 6.2.1, $G'(m) \leq 0$ for all $m \in \mathbb{R}$ and equality holds iff ($J = J_c$ and $m = \phi_1(h,J_c) = \phi_2(h,J_c)$). Hence G is strictly decreasing on $\mathbb{R}$. On the other hand by (6.21), $G(m) < 0$ for all $m \leq 0$ and $G(m) > 0$ for all $m \geq 1$. Therefore there exists a unique point m ($m \in]0,1[$) such that $G(m) = 0$.

• Suppose $J > J_c$. By proposition 6.2.1, G is strictly decreasing for $m \leq \phi_1(h,J)$, strictly increasing for $\phi_1(h,J) \leq m \leq \phi_2(h,J)$ and again strictly decreasing for $m \geq \phi_2(h,J)$. On the other hand by (6.21), $G(m_+) > 0$ for some point $m_+ < \phi_1(h,J)$ and $G(m_-) > 0$ for some point $m_- > \phi_2(h,J)$. Therefore:

$$(\,\exists\,(\text{a unique})\; m_1 \in]-\infty, \phi_1(h,J)]\;\text{s.t.}\; G(m_1) = 0\,)\;\Leftrightarrow\; G(\phi_1(h,J)) \leq 0\,;$$

$$(\,\exists\,(\text{a unique})\; m_2 \in [\phi_2(h,J), \infty[\;\text{s.t.}\; G(m_2) = 0\,)\;\Leftrightarrow\; G(\phi_2(h,J)) \geq 0\,;$$

$$(\,\exists\,(\text{a unique})\; m_0 \in [\phi_1(h,J), \phi_2(h,J)]\;\text{s.t.}\; G(m_0) = 0\,)\;\Leftrightarrow\; G(\phi_1(h,J)) \leq 0\,,\; G(\phi_2(h,J)) \geq 0\,.$$

And now, using identity (6.19) and definitions (6.25), (6.27)

$$G(\phi_1(h,J)) \underset{(=)}{\leq} 0 \;\Leftrightarrow\; g\big((2\phi_1(h,J)-1)J + h\big) \underset{(=)}{\leq} \phi_1(h,J) \;\Leftrightarrow\; h \underset{(=)}{\leq} \psi_1(J)$$

and similarly $G(\phi_2(h,J)) \underset{(=)}{\geq} 0 \;\Leftrightarrow\; h \underset{(=)}{\geq} \psi_2(J)$.

The first • allows to conclude in case *1.*, while the second • allows to conclude in all the other cases. Notice that the nature of the stationary points of $\widetilde{p}\,(\cdot\,, h, J)$ is determined by the sign of the second derivative $\frac{\partial^2 \widetilde{p}}{\partial m^2}$ studied in proposition 6.2.1. □

A special role is played by the point (h_c, J_c), where we set

$$h_c := \psi_1(J_c) = \psi_2(J_c) \,=\, \frac{1}{2}\,\log(2\sqrt{2}-2) - \frac{1}{4} \,\approx\, -0.3441\,, \tag{6.28}$$

indeed in the next sub-sections it will turn out to be the *critical point* of the system. It is also useful to define

$$m_c := \phi_1(h_c, J_c) = \phi_2(h_c, J_c) \,=\, 2-\sqrt{2} \,\approx\, 0.5857\,, \tag{6.29}$$

$$\xi_c := (2m_c - 1)J_c + h_c \,=\, \frac{1}{2}\,\log(2\sqrt{2}-2) \,\approx\, -0.0941\,. \tag{6.30}$$

The computations are done observing that $a_1(J_c) \,=\, a_2(J_c) \,=\, 2\sqrt{2}-2$ and $g(\frac{1}{2}\log(2\sqrt{2}-2)) = 2-\sqrt{2}$.

Remark 6.2.3. We notice that m_c is the (unique) solution of equation (6.22) for $h = h_c$ and $J = J_c$, that is $m(h_c, J_c) = m_c$. Indeed a direct computation using (6.14) shows

$$g\big((2m_c-1)J_c + h_c\big) \,=\, g\big(\xi_c\big) \,= m_c.$$

Observe that as a consequence m_c is a solution of equation (6.22) for all (h, J) such that $h - h_c = (1-2m_c)(J - J_c)$.

In the next proposition we analyse the regularity of the solutions of equation (6.22).

Proposition 6.2.4 (Regularity properties)**.** *Consider the stationary points of* $\widetilde{p}\,(\cdot\,, h, J)$ *defined in proposition 6.2.2:* $m(h,J)$, $m_1(h,J)$, $m_0(h,J)$, $m_2(h,J)$

for suitable values of h, J. The functions

$$\mu_1(h,J) := \begin{cases} m(h,J) & \text{if } 0 < J \le J_c\,,\ h \in \mathbb{R} \\ m_1(h,J) & \text{if } J > J_c\,,\ h \le \psi_1(J) \end{cases}, \tag{6.31}$$

$$\mu_2(h,J) := \begin{cases} m(h,J) & \text{if } 0 < J \le J_c\,,\ h \in \mathbb{R} \\ m_2(h,J) & \text{if } J > J_c\,,\ h \ge \psi_2(J) \end{cases}, \tag{6.32}$$

$$\mu_0(h,J) := \begin{cases} m(h,J) & \text{if } 0 < J \le J_c\,,\ h \in \mathbb{R} \\ m_0(h,J) & \text{if } J > J_c\,,\ \psi_2(J) \le h \le \psi_1(J) \end{cases} \tag{6.33}$$

have the following properties:

i) are continuous on the respective domains;

ii) are C^∞ in the interior of the respective domains;

iii) for $i = 0, 1, 2$ and (h, J) in the interior of the domain of μ_i

$$\frac{\partial}{\partial h}\,\widetilde{p}\,(\mu_i(h,J),h,J) \;=\; \mu_i\;, \quad \frac{\partial}{\partial J}\,\widetilde{p}\,(\mu_i(h,J),h,J) \;=\; -\,\mu_i\,(1-\mu_i)\;; \tag{6.34}$$

$$\frac{\partial \mu_i}{\partial h} \;=\; \frac{2\,\mu_i\,(1-\mu_i)}{2-\mu_i-4J\,\mu_i\,(1-\mu_i)}\;, \qquad \frac{\partial \mu_i}{\partial J} \;=\; (2\mu_i-1)\,\frac{\partial \mu_i}{\partial h}\;. \tag{6.35}$$

Proof. i) First prove the continuity of μ_1. Observe that by propositions 6.2.2, 6.2.1:

- for (h, J) in $D_1 := \{(h,J)\,|\,(0 < J \le J_c,\, h \in \mathbb{R}) \text{ or } (J > J_c,\, h \le \psi_2(J))\}$, $\mu_1(h, J)$ is the *only* maximum point of $\widetilde{p}\,(\cdot\,, h, J)$ on the interval $[0, 1]$;
- for (h, J) in $D_2 := \{(h,J)\,|\,J \ge J_c,\, h \le \psi_1(J)\}$, $\mu_1(h, J)$ is the *only* maximum point of $\widetilde{p}\,(\cdot\,, h, J)$ on the interval $[0, \phi_1(h, J)]$.

Hence by the Berge's maximum theorem [102], continuity of the functions $\widetilde{p}$ and ϕ_1 implies continuity of the function μ_1 on the sets D_1 and D_2. As D_1 and

D_2 are both closed subsets of $\mathbb{R} \times \mathbb{R}_+$, by the pasting lemma μ_1 is continuous on their union

$$D_1 \cup D_2 \,=\, \{(h,J) \mid (0 < J \leq J_c,\, h \in \mathbb{R}) \text{ or } (J > J_c,\, h \leq \psi_1(J))\}\,.$$

A similar argument proves the continuity of μ_2 and μ_0.

ii) Now prove the smoothness of μ_1, μ_2, μ_0 in the *interior* of their domains. Set $G(m,h,J) := \frac{\partial \widetilde{p}}{\partial m}(m,h,J)$. As just seen $m = \mu_1(h,J)$, $\mu_2(h,J)$, $\mu_0(h,J)$ are *continuous* solutions of

$$G(m,h,J) = 0\,,$$

for values of h, J in the respective domains. Observe that $G \in C^\infty(\mathbb{R}\times\mathbb{R}\times\mathbb{R}_+)$ and by propositions 6.2.1, 6.2.2 it can happen

$$\begin{cases} \dfrac{\partial G}{\partial m}\,(m,h,J) = 0 \\ G(m,h,J) = 0 \end{cases} \Leftrightarrow \begin{cases} J \geq J_c\,,\ (m = \phi_1(h,J) \text{ or } m = \phi_2(h,J)) \\ G(m,h,J) = 0 \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} J \geq J_c,\, m = \phi_1(h,J) \\ h = \psi_1(J) \end{cases} \text{ or } \begin{cases} J \geq J_c,\, m = \phi_2(h,J) \\ h = \psi_2(J) \end{cases}\,.$$

$m = \mu_1(h,J)$ can fall only within the first case, while $m = \mu_2(h,J)$ can fall only within the second case. Therefore by the implicit function theorem [115] μ_1, μ_2, μ_0 are C^∞ on the interior of the respective domains.

iii) Let $i = 0,1,2$ and (h,J) in the interior of the domain of μ_i. Using (6.12), $(p^{\text{MD}})' = g$ and the fact that $\mu_i(h,J)$ satisfies equation (6.22), compute

$$\begin{aligned} \frac{\partial}{\partial h}\,\widetilde{p}\,(\mu_i,h,J) \,&=\, -2J\,\frac{\partial \mu_i}{\partial h} + (p^{\text{MD}})'\big((2\mu_i - 1)J + h\big)\,(2J\,\frac{\partial \mu_i}{\partial h} + 1) \\ &= -2J\,\frac{\partial \mu_i}{\partial h} + \mu_i\,(2J\,\frac{\partial \mu_i}{\partial h} + 1) \,=\, \mu_i\,; \end{aligned}$$

and similarly $\frac{\partial}{\partial J}\,\widetilde{p}\,(\mu_i,h,J) = \mu_i^2 - \mu_i\,.$

Using the fact that $\mu_i(h,J)$ satisfies equation (6.22) compute

$$\begin{aligned} &\frac{\partial \mu_i}{\partial h} \,=\, \frac{\partial}{\partial h}\,g\big((2\mu_i - 1)J + h\big) \,=\, g'\big((2\mu_i - 1)J + h)\big)\,(1 + 2J\,\frac{\partial \mu_i}{\partial h}) \\ &\Rightarrow\ \frac{\partial \mu_i}{\partial h} \,=\, \frac{g'\big((2\mu_i - 1)J + h\big)}{1 - 2J\,g'\big((2\mu_i - 1)J + h\big)}\,; \end{aligned}$$

and similarly $\frac{\partial \mu_i}{\partial J} = \frac{(2\mu_i-1)\, g'\big((2\mu_i-1)J+h\big)}{1-2J\, g'\big((2\mu_i-1)J+h\big)}$. Then observe that $g' = 2\, g\,(1-g)/(2-g)$ (identity (E2) in the Appendix), hence since $\mu_i(h,J)$ satisfies equation (6.22)

$$g'\big((2\mu_i-1)J+h\big) \;=\; \frac{2\,\mu_i\,(1-\mu_i)}{2-\mu_i}\;;$$

substituting this in the previous identities concludes the proof. □

To end this subsection we study the asymptotic behaviour of the stationary points of $\widetilde{p}\,(\cdot\,,h,J)$ for large J.

Proposition 6.2.5 (Asymptotic behaviour)**.** *Consider the stationary points $m_1(h,J)$, $m_0(h,J)$, $m_2(h,J)$ defined in proposition 6.2.2 for suitable values of h, J.*

i) For all fixed $h \in \mathbb{R}$

$$m_1(h,J) \xrightarrow[J\to\infty]{} 0\;, \quad m_2(h,J) \xrightarrow[J\to\infty]{} 1\;, \quad m_0(h,J) \xrightarrow[J\to\infty]{} \frac{1}{2}\;.$$

ii) Moreover for all fixed $h \in \mathbb{R}$

$$J\,m_1(h,J) \xrightarrow[J\to\infty]{} 0\;, \quad J\,(1-m_2(h,J)) \xrightarrow[J\to\infty]{} 0\;.$$

iii) And taking the sup *and* inf *over $h \in [\psi_2(J), \psi_1(J)]$*

$$\sup_h m_1(h,J) \xrightarrow[J\to\infty]{} 0\;, \quad \inf_h m_2(h,J) \xrightarrow[J\to\infty]{} 1\;.$$

Proof. i) First observe from the definition (6.27) that $\psi_2(J) \to -\infty$, $\psi_1(J) \to \infty$ as $J \to \infty$. Hence for any fixed $h \in \mathbb{R}$ there exists $\bar{J} > 0$ such that $\psi_2(J) < h < \psi_1(J)$ for all $J > \bar{J}$. This means that the limits in the statement make sense.

Now remind that by proposition 6.2.2, for $J > \bar{J}$

$$m_1(h,J) < \phi_1(h,J) < m_0(h,J) < \phi_2(h,J) < m_2(h,J)\;.$$

Observe from the definition (6.25) that $\phi_1(h,J) \to \frac{1}{2}$, $\phi_2(h,J) \to \frac{1}{2}$ as $J \to \infty$. It follows immediately that also $m_0(h,J) \to \frac{1}{2}$ as $J \to \infty$.

Moreover definition (6.25) entails that $J\big(\frac{1}{2}-\phi_1(h,J)\big)\to\infty$, $J\big(\phi_2(h,J)-\frac{1}{2}\big)\to\infty$ as $J\to\infty$. Exploit the fact that $m_1(h,J)$ is a solution of equation (6.22):

$$\begin{aligned} m_1(h,J) \;&=\; g\big((2m_1(h,J)-1)\,J+h\big) \;\le\; g\big((2\phi_1(h,J)-1)\,J+h\big) \;= \\ &=\; g\big(-2J\,(\frac{1}{2}-\phi_1(h,J))+h\big) \;\xrightarrow[J\to\infty]{}\; 0\,, \end{aligned}$$

where also the facts that the function g is increasing and $g(\xi)\to 0$ as $\xi\to-\infty$ are used. Since m_1 takes values in $]0,1[$, conclude that $m_1(h,J)\longrightarrow 0$ as $J\to\infty$. Similarly it can be shown that $m_2(h,J)\longrightarrow 1$ as $J\to\infty$.

ii) Start observing that, by a standard computation from the definition (6.14), $\xi\,g(-\xi)\longrightarrow 0$ and $\xi\,\big(1-g(\xi)\big)\longrightarrow 0$ as $\xi\to+\infty$. Then exploit the fact that, for fixed h and J sufficiently large, $m_1=m_1(h,J)$ is a solution of equation (6.22):

$$\begin{aligned} &J\,m_1 \;=\; J\,g\big((2m_1-1)J+h\big) \;= \\ &=\frac{\big((1-2m_1)J-h\big)\;g\big(-(1-2m_1)J+h\big)}{1-2m_1}+\frac{h\;g\big(-(1-2m_1)J+h\big)}{1-2m_1} \;\xrightarrow[J\to\infty]{}\; \frac{0}{1}+\frac{h\,0}{1}=0\,, \end{aligned}$$

using also that $m_1\to 0$ as $J\to\infty$ by *i)*. Similarly it can be shown that $J\,(1-m_2)\longrightarrow 0$ as $J\to\infty$.

iii) Start observing that, by a standard computation from the definition (6.27), $-J+\psi_1(J)\longrightarrow-\infty$ and $J+\psi_2(J)\longrightarrow\infty$ as $J\to\infty$. Then exploit the fact that, for $J>J_c$ and $h\in[\psi_2(J),\psi_1(J)]$, $m_1=m_1(h,J)$ is a solution of equation (6.22):

$$\begin{aligned} \sup_{h\in[\psi_2,\psi_1]} m_1 \;&=\; \sup_{h\in[\psi_2,\psi_1]} g\big((2m_1-1)J+h\big) \;\le\; g\big((2m_1-1)J+\psi_1(J)\big) \;= \\ &=\; g\big(2J\,m_1-J+\psi_1(J)\big) \\ &\xrightarrow[J\to\infty]{}\; 0\,, \end{aligned}$$

using also the facts that g is an increasing function, $g(\xi)\to 0$ as $\xi\to-\infty$, and $J\,m_1\to 0$ as $J\to\infty$ by *ii)*. Similarly it can be shown that $\inf_{h\in[\psi_2,\psi_1]} m_2\longrightarrow 1$ as $J\to\infty$. □

6.2.2 The "wall": existence and uniqueness, regularity and asymptot

In the previous subsection we have studied all the solutions of equation (6.22), that is all the stationary points of $m \mapsto \widetilde{p}\,(m,h,J)$. One of them is the point where the global maximum is attained and, because of theorem 6.1.3, we are interested in this one.

Consider the points m, m_1, m_0, m_2 defined in proposition 6.2.2 and look for the global maximum point of $m \mapsto \widetilde{p}\,(m,h,J)$:

- for $0 < J < J_c$ and $h \in \mathbb{R}$, $m(h,J)$ is the only local maximum point, hence it is the global maximum point;
- for $J > J_c$ and $h \leq \psi_2(J)$, $m_1(h,J)$ is the only local maximum point, hence it is the global maximum point;
- for $J > J_c$ and $h \geq \psi_1(J)$, $m_2(h,J)$ is the only local maximum point, hence it is the global maximum point;
- for $J > J_c$ and $\psi_2(J) < h < \psi_1(J)$, there are two local maximum points $m_1(h,J) < m_2(h,J)$, hence at least one of them is the global maximum point.

To answer which one is the global maximum point in the last case, we have to investigate the sign of the following function

$$\Delta(h,J) := \widetilde{p}\,\big(m_2(h,J),h,J\big) - \widetilde{p}\,\big(m_1(h,J),h,J\big) \tag{6.36}$$

for $J > J_c$ and $\psi_2(J) \leq h \leq \psi_1(J)$.

Proposition 6.2.6 (Existence and Uniqueness)**.** *For all $J > J_c$ there exists a unique $h = \gamma(J) \in]\psi_2(J), \psi_1(J)[$ such that $\Delta(h,J) = 0$. Moreover*

$$\Delta(h,J) \begin{cases} < 0 & \text{if } \; J > J_c, \; \psi_2(J) \leq h < \gamma(J) \\ > 0 & \text{if } \; J > J_c, \; \gamma(J) < h \leq \psi_1(J) \end{cases} .$$

Proof. It is an application of the intermediate value theorem. Fix $J > J_c$. It suffices to observe that

i. $\Delta\big(\psi_2(J), J\big) < 0$, because for $h = \psi_2(J)$ the only maximum point of the function $\widetilde{p}\,(\cdot\,, h, J)$ is $m_1(h, J)$;

ii. $\Delta\big(\psi_1(J), J\big) > 0$, because for $h = \psi_1(J)$ the only maximum point of the function $\widetilde{p}\,(\cdot\,, h, J)$ is $m_2(h, J)$;

iii. $h \mapsto \Delta(h, J)$ is a continuous function, by continuity of $\widetilde{p}$, m_1, m_2 (see proposition 6.2.4);

iv. $h \mapsto \Delta(h, J)$ is strictly increasing; indeed it is C^∞ on $]\psi_2(J), \psi_1(J)[$ by smoothness of $\widetilde{p}$, m_1, m_2 (see proposition 6.2.4) and, by formula (6.34),

$$\begin{aligned}\frac{\partial \Delta}{\partial h}\,(h, J) \;&=\; \frac{\partial}{\partial h}\,\widetilde{p}\,\big(m_2(h, J), h, J\big) - \frac{\partial}{\partial h}\,\widetilde{p}\,\big(m_1(h, J), h, J\big) \;= \\ &=\; m_2(h, J) - m_1(h, J) \;>\; \phi_2(h, J) - \phi_1(h, J) \;>\; 0\end{aligned}$$

for all $h \in]\psi_2(J), \psi_1(J)[$. □

Remark 6.2.7. By the previous results the global maximum point of $m \mapsto \widetilde{p}\,(m, h, J)$ is

$$m^*(h, J) := \begin{cases} m(h, J) & \text{if } \; 0 < J \le J_c\,, \; h \in \mathbb{R} \\ m_1(h, J) & \text{if } \; J > J_c\,, \; h < \gamma(J) \\ m_2(h, J) & \text{if } \; J > J_c\,, \; h > \gamma(J) \end{cases} \tag{6.37}$$

where the function γ is defined by proposition 6.2.6. Set also

$$\Gamma := \{(h, J) \mid J > J_c,\; h = \gamma(J)\}\,, \qquad \overline{\Gamma} := \Gamma \cup \{(h_c, J_c)\}\,. \tag{6.38}$$

Notice that proposition 6.2.6 guarantees that there is only a curve Γ in the plane (h, J) where the global maximum point of $m \mapsto \widetilde{p}\,(m, h, J)$ is not unique.

Remark 6.2.8. The techniques developed in this work do not allow us to conclude the existence of the monomer density on the wall. Nevertheless it

is easy to show that, using Theorem 6.1.3, its limsup and liminf are included between m_1 and m_2. In the standard mean-field ferromagnetic model (Curie-Weiss) the existence of the magnetization on the wall ($h = 0$) is achieved by symmetry, a property that we do not have in the present case.

By proposition 6.2.4 it follows that the function m^* is continuous on its domain $(\mathbb{R} \times \mathbb{R}_+) \setminus \Gamma$ and it is C^∞ on $(\mathbb{R} \times \mathbb{R}_+) \setminus \overline{\Gamma}$. The behaviour of m^* at the critical point (h_c, J_c) will be investigated in the next subsection.

Now we investigate the main properties of the curve $\overline{\Gamma}$, which we call "*the wall*". Extend the function γ defined by proposition 6.2.6 by

$$\overline{\gamma}(J) := \begin{cases} \gamma(J) & \text{if } J > J_c \\ h_c & \text{if } J = J_c \end{cases} . \tag{6.39}$$

Proposition 6.2.9 (Regularity properties)**.** *The function $\overline{\gamma}$ is C^∞ on $]J_c, \infty[$ and (at least) C^1 on $[J_c, \infty[$. In particular*

$$\gamma'(J) \,=\, 1 - m_1\big(\gamma(J), J\big) - m_2\big(\gamma(J), J\big) \quad \forall\, J > J_c \,,$$

and

$$\overline{\gamma}\,'(J_c) \,=\, 1 - 2\, m_c \,=\, -(3 - 2\sqrt{2}) \,.$$

Proof. I. First prove that the function $\gamma \in C^\infty(\,]J_c, \infty[\,)$.
By proposition 6.2.6 for all $J > J_c$, $h = \gamma(J)$ is the *unique* solution of equation

$$\Delta(h, J) = 0$$

where Δ is defined by (6.36). Moreover $\psi_2(J) < \gamma(J) < \psi_1(J)$. Observe that Δ is C^∞ on $\{(h, J) \,|\, J > J_c,\ \psi_2(J) < h < \psi_1(J)\}$ by smoothness of $\widetilde{p}$ and m_1, m_2 on this region (see proposition 6.2.4). And furthermore, as shown in the proof of proposition 6.2.6,

$$\frac{\partial \Delta}{\partial h}\,(h, J) \neq 0 \quad \forall\, (h, J) \text{ s.t. } h = \gamma(J) \,.$$

Therefore by the implicit function theorem [115] $\gamma \in C^\infty(\,]J_c, \infty[\,)$. Now

$$\Delta(\gamma(J), J)) \equiv 0 \;\Rightarrow\; 0 = \frac{\mathrm{d}}{\mathrm{d}J}\,\Delta(\gamma(J), J) = \frac{\partial \Delta}{\partial h}(\gamma(J), J)\,\gamma'(J) + \frac{\partial \Delta}{\partial J}(\gamma(J), J)$$
$$\Rightarrow\; \gamma'(J) \,= -\,\frac{\partial \Delta}{\partial J} \Big/ \frac{\partial \Delta}{\partial h}\,(\gamma(J), J)\;;$$

by formulae (6.34) $\frac{\partial \Delta}{\partial h} = m_2 - m_1$ and $\frac{\partial \Delta}{\partial J} = (m_2^2 - m_2) - (m_1^2 - m_1)$; therefore

$$\gamma'(J) \;=\; 1 - (m_2 + m_1)\,(\gamma(J), J)\;.$$

II. Now prove that the extended function $\overline{\gamma} \in C^1([J_c, \infty[\,)$.
First observe that $\overline{\gamma}$ is continuous also in J_c, indeed:

$$\psi_2(J) < \gamma(J) < \psi_1(J) \;\; \forall\, J > J_c \quad \Rightarrow \quad \lim_{J \to J_c +} \gamma(J) = h_c$$

by definition of h_c (6.28) and continuity of ψ_1, ψ_2. Then observe that

$$\gamma'(J) \;=\; 1 - (m_2 + m_1)\,(\gamma(J), J) \;\xrightarrow[J \to J_c +]{}\; 1 - 2m_c$$

because $m(h_c, J_c) \,=\, m_c$ (remark 6.2.3) and the functions μ_1, μ_2 defined in proposition 6.2.4 are continuous. By an immediate application of the mean value theorem, this proves that there exists $\overline{\gamma}\,'(J_c) = 1 - 2m_c$. □

Proposition 6.2.10 (Asymptote)**.** *The function $\overline{\gamma}$ has an asymptote, precisely*

$$\gamma(J) \;\xrightarrow[J \to \infty]{}\; -\frac{1}{2}\;.$$

Proof. I. Consider the function Δ defined by (6.36). The first step is to prove that $\Delta(h, J) \longrightarrow 0$ as $J \to \infty$, $h = -\frac{1}{2}$. Use definitions (6.12), (6.13) and the fact that for fixed h and J sufficiently large $m_1 \,=\, m_1(h, J)$, $m_2 \,=\, m_2(h, J)$ satisfy equation (6.22), in two different ways:

$$\widetilde{p}\,(m_1, h, J) \;=\; -J\,m_1^2 + \frac{J}{2} - \frac{1 - m_1}{2} - \log g\big((2m_1 - 1)J + h\big) + (2m_1 - 1)J + h\;,$$
$$\widetilde{p}\,(m_2, h, J) \;=\; -J\,m_2^2 + \frac{J}{2} - \frac{1 - m_2}{2} - \log m_2 + (2m_2 - 1)J + h\;.$$

Hence, reminding that $m_1 \to 0$ and $m_2 \to 1$ as $J \to \infty$ by proposition 6.2.5 part *i)*,

$$\begin{aligned}\Delta(h,J) \;&=\; \widetilde{p}\,(m_2,h,J) - \widetilde{p}\,(m_1,h,J) \;= \\ &=\; J\,(-m_2^2+2m_2+m_1^2-2m_1) + \log g\big((2m_1-1)J+h\big) + \frac{1}{2} + o(1)\;,\end{aligned}$$

Set $\delta := -m_2^2+2m_2+m_1^2-2m_1$ and $\xi := (2m_1-1)J+h$ and prove that in general

$$J\,\delta + \log g(\xi) \;\xrightarrow[J\to\infty]{}\; h\;; \tag{6.40}$$

in particular it will follow that for $h = -\frac{1}{2}$

$$\Delta\big(-\frac{1}{2},\,J\big) \;\xrightarrow[J\to\infty]{}\; 0\;. \tag{6.41}$$

Now proving (6.40) is equivalent to prove $\exp(J\delta)\,g(\xi) \longrightarrow \exp(h)$ as $J \to \infty$; and using definition (6.14)

$$e^{J\delta}\,g(\xi) \;=\; e^{J\delta}\,\frac{\sqrt{e^{4\xi}+4e^{2\xi}}-e^{2\xi}}{2} \;=\; \frac{\sqrt{e^{2(J\delta+2\xi)}+4e^{2(J\delta+\xi)}}-e^{J\delta+2\xi}}{2} \;\xrightarrow[J\to\infty]{}\; e^h\;,$$

because, since $J\,m_1 \to 0$ and $J\,(1-m_2) \to 0$ as $J \to \infty$ by proposition 6.2.5 part *ii)*,

$$\begin{aligned}J\delta + 2\xi \;&=\; J\,\big(-\,(1-m_2)^2 + m_1^2 - 2m_1 - 1\,\big) + 2h \;\xrightarrow[J\to\infty]{}\; -\infty\;, \\ J\delta + \xi \;&=\; J\,\big(-\,(1-m_2)^2 + m_1^2\,\big) + h \;\xrightarrow[J\to\infty]{}\; h\;.\end{aligned}$$

II. Remember that by definition of γ in proposition 6.2.6

$$\Delta\big(\gamma(J),J\big) \;=\; 0 \quad \forall\, J > J_c\;; \tag{6.42}$$

hence using (6.41) will not be hard to prove that $\gamma(J) \longrightarrow -\frac{1}{2}$ as $J \to \infty$. Let $\epsilon > 0$. By (6.41) there exists $\bar{J}_\epsilon > J_c$ such that

$$\big|\Delta\big(-\frac{1}{2},\,J\big)\big| < \epsilon \quad \forall\, J > \bar{J}_\epsilon\;. \tag{6.43}$$

Now by the mean value theorem for all $J > J_c$ and $h \in [\psi_2(J), \psi_1(J)]$,

$$\big|\Delta(h,J) - \Delta\big(-\frac{1}{2},\,J\big)\big| \;\geq\; \inf_{[\psi_2(J),\psi_1(J)]} \big|\frac{\partial\Delta}{\partial h}\,(\,\cdot\,,J)\big|\;\big|h+\frac{1}{2}\big|\;.$$

Furthermore by identity (6.34) and proposition 6.2.5 part *iii)*

$$\inf_{[\psi_2(J),\psi_1(J)]} \big|\frac{\partial\Delta}{\partial h}(\cdot\,,J)\big| \;=\; \inf_{[\psi_2(J),\psi_1(J)]}(m_2-m_1)(\cdot\,,J) \;\geq$$

$$\geq \inf_{[\psi_2(J),\psi_1(J)]} m_2(\cdot,J) \;-\; \sup_{[\psi_2(J),\psi_1(J)]} m_1(\cdot,J) \;\xrightarrow[J\to\infty]{}\; 1\,.$$

Therefore there exist $\bar{J}$ such that

$$\big|\Delta(h,J)-\Delta\big(-\frac{1}{2},\,J\big)\big| \;\geq\; \frac{1}{2}\,\big|h+\frac{1}{2}\big| \quad \forall\, J>\bar{J},\, h\in[\psi_2(J),\psi_1(J)]\,. \tag{6.44}$$

Choosing $h=\gamma(J)$ in (6.44), by (6.42), (6.43) one obtains that for all $J>\max\{\bar{J},\,\bar{J}_\epsilon\}$

$$\big|\gamma(J)+\frac{1}{2}\big| \;\leq\; 2\,\big|\Delta(\gamma(J),J)-\Delta\big(-\frac{1}{2},J\big)\big| < 2\epsilon\,. \qquad \square$$

6.2.3 Critical exponents

As observed in remark 6.2.7 the global maximum point $m^*(h,J)$ is a continuous function on $(\mathbb{R}\times\mathbb{R}^+)\setminus\Gamma$, but it is smooth only outside the critical point (h_c,J_c). In this section we study the behaviour of the solutions of equation (6.22) near the critical point, with particular interest in the function m^*.

As usual the notation $f=\mathcal{O}(g)$ as $x\to x_0$ means that there exists a neighbourhood U of x_0 and a constant $C\in\mathbb{R}$ such that $|f(x)|\leq C\,|g(x)|$ for all $x\in U$. The notation $f\sim g$ as $x\to x_0$ means that $f(x)/g(x)\longrightarrow 1$ as $x\to x_0$. Finally $f=o(g)$ as $x\to x_0$ means that $f(x)/g(x)\longrightarrow 0$ as $x\to x_0$.

We call *critical exponent* of a function f at a point x_0 the following limit

$$\lim_{x\to x_0}\frac{\log|f(x)-f(x_0)|}{\log|x-x_0|}\,.$$

The main result of this section is the following:

Theorem 6.2.11. *Consider the global maximum point $m^*(h,J)$ of the function $m\mapsto\widetilde{p}\,(m,h,J)$ defined by (6.12).*

i) m^* is continuous on $(\mathbb{R}\times\mathbb{R}_+)\setminus\Gamma$ and smooth on $(\mathbb{R}\times\mathbb{R}_+)\setminus\overline{\Gamma}$, where $\overline{\Gamma}=\Gamma\cup\{(h_c,J_c)\}$ and the "wall" curve Γ is the graph of the function γ defined by proposition 6.2.6.

ii) The critical exponents of m^* at the critical point (h_c, J_c) are:

$$\boldsymbol{\beta} = \lim_{J \to J_c+} \frac{\log |m^*(\delta(J), J) - m_c|}{\log(J - J_c)} = \frac{1}{2}$$

along any curve $h = \delta(J)$ with $\delta \in C^2([J_c, \infty[)$, $\delta(J_c) = h_c$, $\delta'(J_c) = 1 - 2m_c$ (i.e. if the curve is tangent to the "wall" in the critical point);

$$\frac{1}{\boldsymbol{\delta}} = \lim_{J \to J_c} \frac{\log |m^*(\delta(J), J) - m_c|}{\log |J - J_c|} = \frac{1}{3}$$
$$\frac{1}{\boldsymbol{\delta}} = \lim_{h \to h_c} \frac{\log |m^*(h, \delta(h)) - m_c|}{\log |h - h_c|} = \frac{1}{3}$$

along any curve $h = \delta(J)$ with $\delta \in C^2(\mathbb{R}_+)$, $\delta(J_c) = h_c$, $\delta'(J_c) \neq 1 - 2m_c$ or along a curve $J = \delta(h)$ with $\delta \in C^2(\mathbb{R})$, $\delta(h_c) = J_c$, $\delta'(h_c) = 0$ (i.e. if the curve is not tangent to the "wall" in the critical point).

iii) Denote by $m^*(h^\pm, J) := \lim_{h' \to h\pm} m^*(h', J)$. The critical exponent of $m^*(h^+, J)$ and $m^*(h^-, J)$ at the critical point (h_c, J_c) along the "wall" $h = \gamma(J)$ is still

$$\boldsymbol{\beta} = \lim_{J \to J_c+} \frac{\log |m^*(\gamma(J)^+, J) - m_c|}{\log(J - J_c)} = \frac{1}{2}$$
$$\boldsymbol{\beta} = \lim_{J \to J_c+} \frac{\log |m^*(\gamma(J)^-, J) - m_c|}{\log(J - J_c)} = \frac{1}{2}$$

Proof. As observed in remark 6.2.7, the global maximum point m^* is expressed piecewise using the two local maximum points μ_1, μ_2 and inherits their continuity properties outside Γ and their regularity properties outside $\overline{\Gamma}$. Thus part *i)* of the theorem is already proved by proposition 6.2.4.
The proof of the other parts of the theorem, regarding the behaviour of m^* at the critical point (h_c, J_c), is long and rather technical, then we sketch only the major points. For the benefit to the reader, the remaining parts of the proof are given in Appendix B.

In the following proposition we find the fundamental equation characterizing the behaviour of the solutions of equation (6.22) near the critical point (h_c, J_c).

Proposition 6.2.12. *Here for $h \in \mathbb{R}$, $J > 0$ let $m = m(h, J)$ be any solution of the consistency equation (6.22):*

$$m = g\big((2m-1)J + h\big) \ .$$

Then m is continuous at (h_c, J_c) and furthermore, setting $\xi := (2m-1)J + h$, it satisfies

$$(\xi - \xi_c)^3 - \kappa_1 \, (J - J_c) \, (\xi - \xi_c) - \kappa_2 \, \rho(h, J) + \mathcal{O}\big((\xi - \xi_c)^4\big) = 0 \qquad (6.45)$$

as $(h, J) \to (h_c, J_c)$, where we set $\kappa_1 := 3 \, \frac{J_c}{J} \, (2 - m_c)$, $\kappa_2 := 3 \, \frac{J_c^2}{J} \, (2 - m_c)$ and

$$\rho(h, J) := h - h_c + (2m_c - 1)(J - J_c) \ . \qquad (6.46)$$

Proof. I. First show that m is continuous at (h_c, J_c). Exploit equation (6.22) for $m(h, J)$ and use continuity and monotonicity of g: as $(h, J) \to (h_c, J_c)$

$$\limsup m(h, J) = \limsup g\big((2\, m(h, J) - 1)\, J + h\big) = g\big((2 \limsup m(h, J) - 1)\, J_c + h_c\big) \ ,$$

$$\liminf m(h, J) = \liminf g\big((2\, m(h, J) - 1)\, J + h\big) = g\big((2 \liminf m(h, J) - 1)\, J_c + h_c\big) \ .$$

Thus $\limsup m(h, J)$ and $\liminf m(h, J)$ are both solution of equation $\mu = g\big((2\mu + 1)J_c + h_c\big)$. But this solution is unique by proposition 6.2.2, and it is m_c by remark 6.2.3. Therefore

$$\limsup_{(h,J)\to(h_c,J_c)} m(h, J) \ = \ \liminf_{(h,J)\to(h_c,J_c)} m(h, J) \ = \ m_c \ .$$

II. Make a Taylor expansion of the smooth function g at the point ξ_c (see (6.14), (6.30)). By identities (E2), (E3), (E4) and since $g(\xi_c) = m_c$ it is easy to find

$$g(\xi) \ = \ m_c + \frac{1}{2J_c} \, (\xi - \xi_c) - \frac{1}{6J_c^2(2 - m_c)} \, (\xi - \xi_c)^3 + \mathcal{O}\big((\xi - \xi_c)^4\big) \qquad (6.47)$$

as $\xi \to \xi_c$. Now choose $\xi := (2m-1)J + h$. Then $g(\xi) = m$ and

$$\xi - \xi_c \ = \ \rho(h, J) + 2J \, (m - m_c) \ , \qquad (6.48)$$

where $\rho(h, J) := h - h_c + (2m_c - 1)(J - J_c)$. Now (6.45) follows from (6.47). $\square$

Given the previous expansion, the proof of part *ii)* of the theorem 6.2.11 is rather technical and it is contained in the proposition F4 of the Appendix (and in the other results of the Appendix B).

The part *ii)* of the theorem describes the critical behaviour of the local maximum points along curves of class C^2. Notice that "the wall" $\overline{\gamma}$ belongs to $C^1([J_c,\infty[) \cap C^\infty(]J_c,\infty[)$ by proposition 6.2.9, but we did not manage to prove that it is C^2 up to J_c. Anyway we are interested in the behaviour along this coexistence curve, which separates two different phases of the system. This is provided by part *iii)* of the theorem 6.2.11. To prove it let start with the following proposition, which is bases on corollary F2 and lemma F1 in the Appendix.

Proposition 6.2.13. *Consider the "wall" curve $h = \overline{\gamma}(J)$ defined by (6.39) and proposition 6.2.6. There exist $r > 0$, $C_1 < \infty$, $C_2 > 0$ such that for all $J \in]J_c, J_c + r[$.*

$$C_2 \leq \frac{\mu_2(\overline{\gamma}(J), J) - m_c}{\sqrt{J - J_c}} \leq C_1\,, \quad C_2 \leq \frac{m_c - \mu_1(\overline{\gamma}(J), J)}{\sqrt{J - J_c}} \leq C_1$$

Proof. Observe that by definition, on the curve $h = \overline{\gamma}(J)$, $J \geq J_c$, both the local maximum points $\mu_1(h, J)$, $\mu_2(h, J)$ exist.

As $\overline{\gamma} \in C^1([J_c,\infty[)$ (see proposition 6.2.9), the existence of the lower bound $C_2 > 0$ is guaranteed by corollary F2 part *2)*.

Only the existence of an upper bound $C_1 < \infty$ has to be proven. Fix $J > J_c$ and shorten the notation by $m_i = m_i(\overline{\gamma}(J), J) = \mu_i(\overline{\gamma}(J), J)$ and $\xi_i := (2m_i - 1)\, J + \gamma(J)$ for $i = 1, 2$. By proposition 6.2.12, ξ_1, ξ_2 satisfy equation (6.45). The Taylor expansion with Lagrange remainder of $\overline{\gamma}$ is (see proposition 6.2.9)

$$\gamma(J) = h_c + (1 - 2m_c)\,(J - J_c) + \gamma''(\bar{J})\,(J - J_c)^2\,, \quad \text{with } \bar{J} \in]J_c, J[\,;$$

notice $\gamma''(\bar{J})\,(J - J_c)^2$ is not necessarily a $\mathcal{O}\big((J - J_c)^2\big)$, because we do not know the behaviour of γ'' as $J \to J_c$, but for sure it is a $o(J - J_c)$ as $J \to J_c$.

Thus (see identities (6.46), (6.48)):

$$\rho(h,J) = \gamma''(\bar{J})\,(J-J_c)^2 \quad \text{and} \quad \xi_i - \xi_c = 2J\,(m_i - m_c) + \gamma''(\bar{J})\,(J-J_c)^2$$

and equation (6.45) becomes:

$$(\xi_i-\xi_c)^3 - \kappa_1\,(J-J_c)\,(\xi_i-\xi_c) - \kappa_2\,\gamma''(\bar{J})\,(J-J_c)^2 + \mathcal{O}\big((\xi_i-\xi_c)^4\big) = 0\,,$$

which entails

$$(m_i-m_c)^3 - \frac{\kappa_1}{(2J)^2}\,(J-J_c)\,(m_i-m_c) - \frac{\kappa_2}{(2J)^3}\,\gamma''(\bar{J})\,(J-J_c)^2\,(1+o(1)) + \\ + \mathcal{O}\big((m_i-m_c)^4\big) = 0\,. \quad (6.49)$$

Distinguish two cases.

1) If $\gamma''(\bar{J})\,(J-J_c)^2 = \mathcal{O}\big((m_i-m_c)^4\big)$ (along a sequence), then (6.49) rewrites

$$(m_i-m_c)^3 - \frac{\kappa_1}{(2J)^2}\,(J-J_c)\,(m_i-m_c) + \mathcal{O}\big((m_i-m_c)^4\big) = 0\,, \quad (6.50)$$

which, dividing by $m_i - m_c$ and solving, gives

$$m_i - m_c = \pm\,\frac{\sqrt{\kappa_1}}{2J}\,(J-J_c)^{\frac{1}{2}} + \mathcal{O}\big((m_i-m_c)^{\frac{3}{2}}\big)\,;$$

hence $m_i - m_c \sim \sqrt{\kappa_1}/(2J)\,(J-J_c)^{1/2}$, proving the result (along the sequence).

2) Now suppose $(m_i-m_c)^4 = o\big(\gamma''(\bar{J})\,(J-J_c)^2\big)$ (along a sequence), then (6.49) rewrites

$$(m_i-m_c)^3 \underbrace{-\frac{\kappa_1}{(2J)^2}\,(J-J_c)}_{=:\,p}\,(m_i-m_c) \underbrace{-\frac{\kappa_2}{(2J)^3}\,\gamma''(\bar{J})\,(J-J_c)^2\,(1+o(1))}_{=:\,q} = 0\,. \quad (6.51)$$

Claim $\Delta := (\frac{q}{2})^2 + (\frac{p}{3})^3 \le 0$. Suppose by contradiction $\Delta > 0$. Then the cubic equation (6.51) has only one real solution: for $i = 1, 2$

$$m_i - m_c = u_+ + u_- \quad \text{with } u_\pm = \sqrt[3]{-\frac{q}{2} \pm \sqrt[2]{\big(\frac{q}{2}\big)^2 + \big(\frac{p}{3}\big)^3}}\,.$$

Observe that q and p are written only in terms of J, so that $u_+ + u_-$ at the main order do not depend implicitly on m_i. Therefore $m_1 - m_c$ and $m_2 - m_c$

must have the same sign for every $J > J_c$ small enough. But this contradicts proposition 6.2.2 and lemma F1, which ensures that in a right neighbourhood of J_c

$$m_2 - m_c > \phi_2 - m_c > 0 \quad \text{while} \quad m_1 - m_c < \phi_2 - m_c < 0 \, .$$

This proves $\Delta \leq 0$. And now adapting to equation (6.51) the step *ii.* of the proof of corollary F2, $\Delta \leq 0$ entails (along the sequence)

$$m - m_c \, = \, \mathcal{O}\big((J - J_c)^{\frac{1}{2}}\big) \, .$$

This completes the proof of the proposition. □

To conclude the proof of the part *iii)* of theorem 6.2.11, it suffices to have the previous proposition and observe that

$$m^*(\gamma(J)^+, J) = m_2(\gamma(J), J) \, , \quad m^*(\gamma(J)^-, J) = m_1(\gamma(J), J)$$

for all $J > J_c$, by proposition 6.2.6 and continuity of m_1, m_2. □

Chapter 7

A Monomer Dimer model with random weights

In this chapter we study an MD model on the complete graph with randomness in the monomer activities. The model describes, in the mean-field approximation, the equilibrium properties of systems of diatomic molecules (see e.g. [86, 87, 54]) depositing on a inhomogeneous lattice, where the inhomogeneity is modeled introducing a probability measure on the space of possible realizations of the site activity. From the probabilistic point of view model presented here is a pair matching of the sites with random weight. The latter point of view furnishes a direct link with Combinatorics and Computer Science, where the study of monomer-dimer models applies to matching problems (see e.g. [78] for an overview of matching problems). For a different way of introducing randomness in monomer-dimer systems see [27], where a model on locally tree-like random graphs is solved. The combinatorial problem of perfect matchings on random graphs, already solved in [91, 50], corresponds the zero-temperature limit of the latter monomer-dimer model.

The main result is the exact solution of the model with *i.i.d.* randomness on the monomer activities x_i's. Precisely we prove that, under very general conditions on the probability distribution, the *t.l.* of the random pressure density

exists and converge almost surely to its quenched value. Moreover the *t.l.* of the quenched pressure density is given by a one dimensional variational principle. The pressure density turns out to be a smooth function of the dimer activity w. The dimer density is therefore a smooth function of w too and it will be described explicitely.

The problem, otherwise expected to be difficult due the hard-core interaction, becomes accessible with the use of a Gaussian representation 5.2.1 for the partition function . In this representation, the integrand function presents negative and singular contributions. However, a careful application of the *Uniform Law of Large Numbers* and the *Laplace method* lead to the solution of the model.

In the present chapter the Heilmann-Lieb recursion is one of the main tools, together with technical methods for martingales (like the Azuma's inequality), used to prove the self-averaging of the pressure density.

The chapter is entirely based on [30] and is organised as follows.

In the section 7.1 we the model and compute the *t.l.* pressure density in theorem 7.1.3 and the dimer density in corollary 7.1.5. In the section 7.2 we show, under suitable assumptions, that the free energy density of a monomer-dimer model with independent random activities is self-averaging. The appendix collects the main technical results used in this chapter, in order to facilitate the reader.

7.1 The model and the main result

In this section we fix a uniform dimer weight on the complete graph, while we choose *i.i.d.* random monomer weights. Under quite general integrability hypothesis, we show that this model is exactly solvable and it does not present a phase transition (in agreement with the general results by Heilmann and Lieb [86, 87]).

Let $w > 0$. Let $x_i > 0$, $i \in \mathbb{N}$, be *independent identically distributed* random variables. In order to keep the logarithm of the partition function of order N, a normalization of the dimer weight as w/N is needed (see section 5.3). Therefore during all this section we will denote

$$Z_N \;=\; \sum_{D \in \mathscr{D}_N} \big(\frac{w}{N}\big)^{|D|} \prod_{i \in M_N(D)} x_i \;. \tag{7.1}$$

μ_N will denote the corresponding Gibbs measure and $\langle \cdot \rangle_N$ will be the expected value with respect to μ_N. Notice that now the partition function is a random variable and the Gibbs measure is a random measure.

Remark 7.1.1. Since the dimer weight is uniform, the Gaussian representation of (7.1) simplifies:

$$Z_N \;=\; \mathbb{E}_\xi\Big[\prod_{i=1}^{N}(\xi + x_i)\Big] \;, \tag{7.2}$$

where ξ is a 1-dimensional Gaussian random variable with mean 0 and variance w/N.

Indeed by proposition 5.2.1, $Z_N = \mathbb{E}_{\boldsymbol{\xi}}\big[\prod_{i=1}^{N}(\xi_i + x_i)\big]$ where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N)$ is an N-dimensional Gaussian random vector with mean 0 and constant covariance matrix[1] $(w/N)_{i,j=1,\ldots,N}$. It is easy to check that $\boldsymbol{\xi}$ has the same joint distribution of the constant random vector $(\xi, \ldots, \xi)$. Therefore the identity (7.2) follows.

Remark 7.1.2. Keeping in mind the remark 5.2.2, one can observe the analogy between the formula (7.2) and the partition function of the Curie-Weiss random field model (see e.g. [118, 21, 117]), that is

$$Z_N^{\text{Curie-Weiss}} \;\propto\; \mathbb{E}_{\xi'}\Big[\prod_{i=1}^{N}\cosh(\xi' + h_i)\Big] \tag{7.3}$$

where ξ' is a 1-dimensional Gaussian random variable with mean 0 and variance $J/N \in \mathbb{R}_+$.

[1] It is important to notice that setting also the diagonal entries to w/N, the resulting matrix is positive semi-definite: $\sum_{i=1}^{N}\sum_{j=1}^{N}(w/N)\,\alpha_i\alpha_j = (w/N)\big(\sum_{i=1}^{N}\alpha_i\big)^2 \geq 0$ for every $\alpha \in \mathbb{R}^N$.

By the way, we want to stress the fact that the Laplace method applies directly to formula (7.3), while the presence of negative and singular contributions in (7.2) requires a supplementary work.

Let us rewrite (7.2) as an explicit integral in $\mathrm{d}\xi$:

$$Z_N \;=\; \frac{\sqrt{N}}{\sqrt{2\pi w}} \int_{\mathbb{R}} e^{-\frac{N}{2w}\,\xi^2} \prod_{i=1}^{N} (\xi + x_i)\; \mathrm{d}\xi \;. \tag{7.4}$$

Theorem 7.1.3. *Let $w > 0$. Let $x_i > 0$, $i \in \mathbb{N}$ be i.i.d. random variables. Denote by x a random variable distributed like x_i; suppose that $\mathbb{E}_x[x] < \infty$ and $\mathbb{E}_x[(\log x)^2] < \infty$. Then:*

$$\exists \lim_{N\to\infty} \frac{1}{N}\, \mathbb{E}_{\boldsymbol{x}}[\log Z_N] \;=\; \sup_{\xi \geq 0} \Phi(\xi) \tag{7.5}$$

where

$$\Phi(\xi) \;:=\; -\frac{\xi^2}{2w} + \mathbb{E}_x[\log(\xi + x)] \quad \forall\, \xi \geq 0 \;. \tag{7.6}$$

Furthermore the function Φ attains its maximum at a unique point ξ^*. ξ^* is the only solution in $[0, \infty[$ of the fixed point equation

$$\xi^* \;=\; \mathbb{E}_x\left[\frac{w}{\xi^* + x}\right] . \tag{7.7}$$

Thus the following bounds hold:

$$\frac{-\mathbb{E}_x[x] + \sqrt{\mathbb{E}_x[x]^2 + 4w}}{2} \vee \sup_{t>0} \frac{-t + \sqrt{t^2 + 4w\, \mathbb{P}_x(x \leq t)}}{2} \;\leq\; \xi^* \;\leq\; \sqrt{w} \wedge \mathbb{E}_x\left[\frac{w}{x}\right] . \tag{7.8}$$

In consequence of the theorem 7.1.3 it is not hard to prove that the system does not present a phase transition in the parameter $w > 0$. It is also easy to compute the main macroscopic quantity of physical interest, that is the *dimer density*, in terms of the positive solution ξ^* of the fixed point equation (7.7). Therefore we state the following two corollaries before starting to prove the theorem.

Corollary 7.1.4. *In the hypothesis of the theorem 7.1.3, consider the limiting pressure density function* $p(w) := \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{x}}\big[\log Z_N(w)\big]$ *for all* $w > 0$. *Then* $p \in C^\infty\big(\,]0,\infty[\,\big)$.

Proof. By the theorem 7.1.3 $p(w) = \Phi(w, \xi^*)$, where $\Phi(w,\xi) = -\xi^2/(2w) + \mathbb{E}_x[\log(\xi + x)]$ and $\xi^* = \xi^*(w)$ is *the only* positive solution of the equation $F(w,\xi) = 0$ with $F := \frac{\partial \Phi}{\partial \xi}$.

F is a smooth function on $]0,\infty[\times]0,\infty[$, because Φ is smooth as it will be proven in the lemma 7.1.6. In addition $\frac{\partial F}{\partial \xi}(w,\xi^*) \neq 0$ for all $w > 0$, by the lemma 7.1.6 equation (7.11).

As a consequence, by the implicit function theorem (see e.g. [115]), ξ^* is a smooth function of $w \in]0,\infty[$. Hence, by composition, also $p(w) = \Phi\big(w, \xi^*(w)\big)$ is a smooth function of $w \in]0,\infty[$. □

Corollary 7.1.5. *In the hypothesis of the theorem 7.1.3, the limiting dimer density*

$$d := \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{\boldsymbol{x}}\big[\big\langle\, |D| \,\big\rangle_N\big]$$

can be computed as

$$d \;=\; w\,\frac{\mathrm{d}\,p}{\mathrm{d}w} \;=\; \frac{(\xi^*)^2}{2w}\;. \tag{7.9}$$

Proof. Set $p_N := \frac{1}{N}\log Z_N$ and perform the change of parameter $w =: e^h$. Clearly $\frac{\mathrm{d}}{\mathrm{d}h} = w\,\frac{\mathrm{d}}{\mathrm{d}w}$ and it is easy to check that

$$\frac{\mathrm{d}\,\mathbb{E}_{\boldsymbol{x}}[p_N]}{\mathrm{d}h} \;=\; \mathbb{E}_{\boldsymbol{x}}\big[\big\langle\, |D| \,\big\rangle_N\big]\;.$$

By the theorem 7.1.3 and its corollary 7.1.4, $\mathbb{E}_{\boldsymbol{x}}[p_N]$ converges pointwise to a smooth function p as $N \to \infty$ for all values of $h \in \mathbb{R}$. A standard computation shows that $\mathbb{E}_{\boldsymbol{x}}[p_N]$ is a convex function of h. Therefore

$$\frac{\mathrm{d}\,\mathbb{E}_{\boldsymbol{x}}[p_N]}{\mathrm{d}h} \;\xrightarrow[N\to\infty]{}\; \frac{\mathrm{d}\,p}{\mathrm{d}h}\;.$$

Since $p(h) = \Phi\big(h, \xi^*(h)\big)$, where ξ^* is the critical point of Φ and is a smooth function of h, it is easy to compute

$$\frac{\mathrm{d}\, p}{\mathrm{d}h}(h) \;=\; \frac{\partial \Phi}{\partial h}(h, \xi^*) \;+\; \underbrace{\frac{\partial \Phi}{\partial \xi}(h, \xi^*)}_{=0} \; \frac{\mathrm{d}\xi^*}{\mathrm{d}h}(h) \;=\; \frac{(\xi^*)^2}{2\, e^h} \;. \qquad \square$$

Now let us start to prove the theorem 7.1.3. The logic structure of the proof is divided in three main parts. First we study the basic properties of the function Φ. Then we use the uniform law of large numbers and other observations to show that for large N the integrated function in (7.4) can be well approximated by $e^{N\Phi}$. Finally we will be able to exploit the Laplace's method in order to compute a lower and an upper bound for $\frac{1}{N}\, \mathbb{E}_{\boldsymbol{x}}[\, \log Z_N]$.

Lemma 7.1.6. *Φ is continuous on $[0, \infty[$, it is smooth on $]0, \infty[$ and the derivatives can be taken inside the expectation. In particular for all $\xi > 0$ it holds*

$$\Phi'(\xi) \;=\; -\frac{\xi}{w} + \mathbb{E}_x\left[\frac{1}{\xi + x}\right] ; \tag{7.10}$$

$$\Phi''(\xi) \;=\; -\frac{1}{w} - \mathbb{E}_x\left[\frac{1}{(\xi + x)^2}\right] < 0 \;. \tag{7.11}$$

As a consequence Φ has exactly one critical point ξ^* in $]0, \infty[$, that is the equation (7.7) has exactly one solution in $]0, \infty[$. ξ^* is the only global maximum point *of Φ on $[0, \infty[$.*

Proof. **I.** First of all $\Phi(\xi)$ is well-defined for all $\xi \geq 0$. Indeed for $\xi > 0$

$$\log(\xi + x) \begin{cases} \leq\; \xi + x - 1 \;\in L^1(\mathbb{P}_x) \\ \geq\; 1 - \frac{1}{\xi + x} \;\geq\; 1 - \frac{1}{\xi} \;\in L^1(\mathbb{P}_x) \end{cases} ;$$

while for $\xi = 0$, $\mathbb{E}_x[|\log x|] \leq \mathbb{E}_x[(\log x)^2]^{1/2} < \infty$ by the Hölder inequality.
Φ is continuous at $\xi = 0$ by monotone convergence: $\log(\xi + x)$ decreases to $\log x$ as $\xi \searrow 0$ and $\mathbb{E}_x[\, \log(\xi + x)] < \infty$.
Let now $\xi > 0$ and let $\delta > 0$ such that $\xi - \delta > 0$. The first derivative of Φ at ξ can be computed inside the expectation, obtaining (7.10), since the difference

quotient of $\xi \mapsto \log(\xi + x)$ satisfies the dominated convergence hypothesis. Indeed for all $\xi' \in]\xi - \delta, \xi + \delta[$

$$\left|\frac{\log(\xi' + x) - \log(\xi + x)}{\xi' - \xi}\right| \leq \sup_{\widetilde{\xi} \in [\xi, \xi']} \frac{1}{\widetilde{\xi} + x} \leq \sup_{\widetilde{\xi} \in [\xi, \xi']} \frac{1}{\widetilde{\xi}} \leq \frac{1}{\xi - \delta} \in L^1(\mathbb{P}_x) .$$

Now the second derivative of Φ at ξ can be computed inside the expectation, obtaining (7.11), since the difference quotient of $\xi \mapsto \frac{1}{\xi + x}$ satisfies the dominated convergence hypothesis. Indeed for all $\xi' \in]\xi - \delta, \xi + \delta[$

$$\left|\frac{\frac{1}{\xi' + x} - \frac{1}{\xi + x}}{\xi' - \xi}\right| \leq \sup_{\widetilde{\xi} \in [\xi, \xi']} \frac{1}{(\widetilde{\xi} + x)^2} \leq \sup_{\widetilde{\xi} \in [\xi, \xi']} \frac{1}{\left(\widetilde{\xi}\right)^2} \leq \frac{1}{(\xi - \delta)^2} \in L^1(\mathbb{P}_x) .$$

This reasoning can be iterated up to the derivative of any order, since $1/\left(\widetilde{\xi} + x\right)^k \leq 1/\left(\widetilde{\xi}\right)^k \leq 1/(\xi - \delta)^k \in L^1(\mathbb{P}_x)$ for all $\widetilde{\xi} \in]\xi - \delta, \xi + \delta[$ and all $k \geq 1$.

II. In virtue of (7.11) Φ is a strictly convex function on $]0, \infty[$. At the boundaries of this domain $\lim_{\xi \to 0+} \Phi'(\xi) = \mathbb{E}_x[x^{-1}] > 0$ and $\lim_{\xi \to \infty} \Phi'(\xi) = -\infty < 0$ by (7.10) and monotone converge. Therefore Φ has exactly one critical point ξ^* in $]0, \infty[$ and it is the only global maximum point of Φ. □

Remark 7.1.7. Since ξ^* satisfies the fixed point equation (7.7), it is easy to obtain the bounds (7.8) for ξ^*. Since $\xi^* > 0$ and $x > 0$,

$$\xi^* = \mathbb{E}_x\left[\frac{w}{\xi^* + x}\right] \leq \frac{w}{\xi^*} \Rightarrow \xi^* \leq \sqrt{w} ; \quad \xi^* = \mathbb{E}_x\left[\frac{w}{\xi^* + x}\right] \leq \mathbb{E}_x\left[\frac{w}{x}\right] .$$

Using the Jensen inequality,

$$\xi^* = \mathbb{E}_x\left[\frac{w}{\xi^* + x}\right] \geq \frac{w}{\xi^* + \mathbb{E}_x[x]} \Rightarrow (\xi^*)^2 + \xi^* \mathbb{E}_x[x] - w \geq 0 \Rightarrow \xi^* \geq \frac{-\mathbb{E}_x[x] + \sqrt{\mathbb{E}_x[x]^2 + 4w}}{2} .$$

Finally, since $\xi^* + x > 0$, it holds for all $t > 0$

$$\xi^* = \mathbb{E}_x\left[\frac{w}{\xi^* + x}\right] \geq \frac{w}{\xi^* + t}\,\mathbb{P}_x(x \leq t) \Rightarrow (\xi^*)^2 + \xi^* t - w\,\mathbb{P}_x(x \leq t) \geq 0 \Rightarrow$$
$$\Rightarrow \xi^* \geq \frac{-t + \sqrt{t^2 + 4w\,\mathbb{P}_x(x \leq t)}}{2} .$$

Lemma 7.1.8. *Define the random function*

$$\Phi_N(\xi) := -\frac{\xi^2}{2w} + \frac{1}{N}\sum_{i=1}^{N} \log|\xi + x_i| \quad \forall\, \xi \in \mathbb{R} . \tag{7.12}$$

This function is defined also for negative values of ξ and it takes the value $-\infty$ at the random points $-x_1, \dots, -x_N$. It is important to observe that

$$\Phi_N(-\xi) \,<\, \Phi_N(\xi) \quad \forall\, \xi > 0\,. \tag{7.13}$$

i. *Let $0 < M < \infty$. Then for all $\varepsilon > 0$*

$$\mathbb{P}_{\boldsymbol{x}}\Big(\,\forall\,\xi\in[0,M]\;\;|\Phi_N(\xi)-\Phi(\xi)|<\varepsilon\,\Big)\;\xrightarrow[N\to\infty]{}\;1\,. \tag{7.14}$$

ii. *Let $0 < m < M < \infty$. Then there exists $\lambda_{m,M} > 0$ such that*

$$\mathbb{P}_{\boldsymbol{x}}\Big(\,\forall\,\xi\in[m,M]\;\;\Phi_N(-\xi)<\Phi_N(\xi)-\lambda_{m,M}\,\Big)\;\xrightarrow[N\to\infty]{}\;1\,. \tag{7.15}$$

iii. *Let $C \in \mathbb{R}$. Then there exists $M_C > 0$ such that*

$$\mathbb{P}_{\boldsymbol{x}}\Big(\,\forall\,\xi\in[M_C,\infty[\;\;\Phi_N(\xi)<C\;\;\text{and}\;\;\Phi_N(\xi)<\varphi(\xi)\,\Big)\;\xrightarrow[N\to\infty]{}\;1\,; \tag{7.16}$$

where φ is the following deterministic function

$$\varphi(\xi)\;:=\;-\frac{\xi^2}{2w}+\log\xi+\frac{1}{\xi}\,(\mathbb{E}_x[x]+1)\quad\forall\,\xi>0\,. \tag{7.17}$$

Notice that $\Phi_N(\xi) - \Phi(\xi) = \frac{1}{N}\sum_{i=1}^N \log(\xi + x_i) - \mathbb{E}_x[\log(\xi + x)]$ for all $\xi > 0$. Since the x_i, $i \in \mathbb{N}$ are i.i.d., the basic idea behind the lemma 7.1.8 is to approximate Φ_N with Φ by the law of large numbers. But this approximation is needed to hold at every ξ at the same time, hence a *uniform* law of large numbers is required.

To prove the theorem 7.1.3 it will be important to have found a good uniform approximation near the global maximum point ξ^* of Φ. Far from ξ^* instead such a uniform approximation cannot hold: for example Φ_N diverges to $-\infty$ at certain negative points, while, if the distribution of x is absolutely continuous and satisfies some integrability hypothesis, it is possible to show that $\Phi(\xi) = -\frac{\xi^2}{2w} + \mathbb{E}_x[\log|\xi + x|]$ is continuous on $\mathbb{R}$. But fortunately, far from ξ^*, it will be sufficient for our purposes to bound suitably Φ_N from above.

Proof. **i.** For every $x > 0$ the function $\xi \mapsto \log(\xi + x)$ is continuous on $[0, M]$ compact. Moreover there is domination:

$$\log(\xi + x) \begin{cases} \leq \log(M + x) \ \in L^1(\mathbb{P}_x) \\ \geq \log x \ \in L^1(\mathbb{P}_x) \end{cases} \qquad \forall\, \xi \in [0, M] \ .$$

Therefore (7.14) holds by the uniform weak law of large numbers (theorem C1).

ii. Clearly $\log(\xi + x) > \log|-\xi + x|$ for all $\xi, x > 0$. Furthermore an elementary computation shows that for all $\xi, x, \tau > 0$

$$\log(\xi + x) - \log|-\xi + x| \ \geq \ \tau \quad \Leftrightarrow \quad \frac{e^\tau - 1}{e^\tau + 1}\,\xi \ \leq \ x \ \leq \ \frac{e^\tau + 1}{e^\tau - 1}\,\xi \ .$$

Therefore for all $\xi \in [m, M]$ and all $\tau > 0$,

$$\begin{aligned} \Phi_N(\xi) - \Phi_N(-\xi) \ &= \ \frac{1}{N} \sum_{i=1}^{N} \big(\log(\xi + x_i) - \log|-\xi + x_i| \big) \ \geq \\ &\geq \ \frac{1}{N} \sum_{i=1}^{N} \tau \ \mathbb{1}\!\left(\frac{e^\tau - 1}{e^\tau + 1}\,\xi \ \leq \ x_i \ \leq \ \frac{e^\tau + 1}{e^\tau - 1}\,\xi \right) \geq \\ &\geq \ \tau \ \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\!\left(\frac{e^\tau - 1}{e^\tau + 1}\,M \ \leq \ x_i \ \leq \ \frac{e^\tau + 1}{e^\tau - 1}\,m \right) . \end{aligned} \tag{7.18}$$

Set $I^\tau_{m,M} := \left[\frac{e^\tau - 1}{e^\tau + 1}\,M \, , \, \frac{e^\tau + 1}{e^\tau - 1}\,m \right]$. Now by the weak law of large numbers, for all $\varepsilon > 0$

$$\mathbb{P}_{\boldsymbol{x}}\!\left(\frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\big(x_i \in I^\tau_{m,M}\big) \ > \ \mathbb{P}_x\big(x \in I^\tau_{m,M}\big) - \varepsilon \right) \ \xrightarrow[N \to \infty]{} \ 1 \ . \tag{7.19}$$

Hence, using (7.18) and (7.19), for all $\tau, \varepsilon > 0$

$$\mathbb{P}_{\boldsymbol{x}}\!\left(\Phi_N(\xi) - \Phi_N(-\xi) \ > \ \tau \left(\mathbb{P}_x(x \in I^\tau_{m,M}) - \varepsilon \right) \right) \ \xrightarrow[N \to \infty]{} \ 1 \ . \tag{7.20}$$

To conclude observe that $I^\tau_{m,M} \nearrow \,]0, \infty[$ (which is the support of the distribution of x) as $\tau \searrow 0$. Hence there exists $\tau_0 > 0$ such that $\mathbb{P}_x(x \in I^{\tau_0}_{m,M}) > 0$. Choose $0 < \varepsilon_0 < \mathbb{P}_x(x \in I^{\tau_0}_{m,M})$ and set

$$\lambda_{m,M} \ := \ \tau_0 \left(\mathbb{P}_x(x \in I^{\tau_0}_{m,M}) - \varepsilon_0 \right) \ > 0 \ .$$

Then (7.15) follows from (7.20).

iii. For all $\xi > 0$ the following bound holds:

$$\Phi_N(\xi) \,=\, -\frac{\xi^2}{2w} + \frac{1}{N}\sum_{i=1}^{N}\log(\xi + x_i) \,=\, -\frac{\xi^2}{2w} + \log\xi + \frac{1}{N}\sum_{i=1}^{N}\log\big(1+\frac{x_i}{\xi}\big) \,\leq$$
$$\leq\, -\frac{\xi^2}{2w} \,+\, \log\xi \,+\, \frac{1}{\xi}\,\frac{1}{N}\sum_{i=1}^{N} x_i\,. \tag{7.21}$$

Now by the weak law of large numbers (no uniformity in ξ is needed here), for all $\varepsilon > 0$

$$\mathbb{P}_{\boldsymbol{x}}\bigg(\frac{1}{N}\sum_{i=1}^{N} x_i \,<\, \mathbb{E}_x[x] + \varepsilon\bigg) \,\xrightarrow[N\to\infty]{}\, 1\,. \tag{7.22}$$

Hence, using (7.21) and (7.22), for all $0 < \varepsilon < 1$

$$\mathbb{P}_{\boldsymbol{x}}\Big(\forall\,\xi{>}0 \;\;\Phi_N(\xi) < \varphi(\xi)\Big) \,\xrightarrow[N\to\infty]{}\, 1\,. \tag{7.23}$$

Furthermore it holds $\varphi(\xi) \to -\infty$ as $\xi \to \infty$. Hence for all $C \in \mathbb{R}$ there exists $M_C > 0$ such that

$$\varphi(\xi) < C \quad \forall\,\xi > M_C\,. \tag{7.24}$$

In conclusion (7.16) follows from (7.23) and (7.24). □

Lemma 7.1.9. *There exists a constant $C_0 < \infty$ such that*

$$\mathbb{E}_{\boldsymbol{x}}\bigg[\bigg(\frac{\log Z_N}{N}\bigg)^2\bigg] \,\leq\, C_0 \quad \forall\, N \in \mathbb{N}\,. \tag{7.25}$$

Proof. Since $x \mapsto (\log x)^2$ is concave for $x \geq e$, the Jensen inequality can be used as follows:

$$\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(Z_N \geq e)\big] \,=\, \mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2 \,\big|\, Z_N \geq e\big]\;\mathbb{P}_{\boldsymbol{x}}(Z_N \geq e) \,\leq$$
$$\leq\, \big(\log \mathbb{E}_{\boldsymbol{x}}\big[Z_N \,\big|\, Z_N \geq e\big]\,\big)^2\;\mathbb{P}_{\boldsymbol{x}}(Z_N \geq e) \,=$$
$$=\, \bigg(\log\frac{\mathbb{E}_{\boldsymbol{x}}\big[Z_N\,\mathbb{1}(Z_N \geq e)\big]}{\mathbb{P}_{\boldsymbol{x}}(Z_N \geq e)}\bigg)^2\;\mathbb{P}_{\boldsymbol{x}}(Z_N \geq e) \,\leq$$
$$\leq\, 2\,\big(\log \mathbb{E}_{\boldsymbol{x}}\big[Z_N\big]\big)^2 \,+\, 2\max_{p\in[0,1]}(\log p)^2\,p\,. \tag{7.26}$$

Since the x_i, $i \in \mathbb{N}$ are i.i.d. $\mathbb{E}_{\boldsymbol{x}}[Z_N]$ equals a deterministic partition function with uniform weights. Hence it is easy to bound it as follows:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}}\big[Z_N\big] \;&=\; \sum_{D\in\mathscr{D}_N}\left(\frac{w}{N}\right)^{|D|}\mathbb{E}_x[x]^{|M(D)|} \;\le\; \sum_{d=0}^{|E_N|}\binom{|E_N|}{d}\left(\frac{w}{N}\right)^{d}\mathbb{E}_x[x]^{N-2d} \;= \\ &=\; \mathbb{E}_x[x]^N\left(1+\frac{w}{N}\,\mathbb{E}_x[x]^{-2}\right)^{|E_N|} \;\le\; \mathbb{E}_x[x]^N\,\exp\left(\frac{N-1}{2}\,\frac{w}{\mathbb{E}_x[x]^2}\right)\end{aligned} \tag{7.27}$$

(here $|E_N| = \frac{N(N-1)}{2}$ denotes the number of edges in the complete graph over N vertices). Therefore, substituting (7.27) into (7.26),

$$\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(Z_N \ge e)\big] \;\le\; 2\,N^2\left(\log\mathbb{E}_x[x] + \frac{w}{2\,\mathbb{E}_x[x]^2}\right)^2 + \,2\max_{p\in[0,1]}(\log p)^2\,p\;. \tag{7.28}$$

It remains to deal with the case $Z_N < e$. When $1 < Z_N < e$, it holds $0 < \log Z_N < 1$ hence trivially

$$\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(1 < Z_N < e)\big] \;\le\; \mathbb{E}_{\boldsymbol{x}}\big[(\log e)^2\,\mathbb{1}(1 < Z_N < e)\big] \;\le\; 1\;. \tag{7.29}$$

When instead $Z_N \le 1$, it holds $\log Z_N \le 0$ hence we need a lower bound for Z_N. For example, considering only the configuration with no dimers, $Z_N \ge \prod_{i=1}^N x_i$. Therefore:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(Z_N \le 1)\big] \;&\le\; \mathbb{E}_{\boldsymbol{x}}\left[\left(\log\prod_{i=1}^N x_i\right)^2\mathbb{1}(Z_N\le 1)\right] \;\le\; \mathbb{E}_{\boldsymbol{x}}\left[\left(\sum_{i=1}^N \log x_i\right)^2\right] \le \\ &\le\; N^2\,\mathbb{E}_x\big[\log x\big]^2 \;+\; N\,\mathbb{E}_x\big[(\log x)^2\big]\;.\end{aligned} \tag{7.30}$$

In conclusion the lemma is proved splitting $\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\big]$ as $\mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(Z_N \ge e)\big] + \mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(1 < Z_N < e)\big] + \mathbb{E}_{\boldsymbol{x}}\big[(\log Z_N)^2\,\mathbb{1}(Z_N \le 1)\big]$ and applying the bounds (7.28), (7.29), (7.30). □

Proof of the theorem 7.1.3. It remains to prove only the convergence (7.5). Fix $C < \Phi(\xi^*)$. Fix $0 < m < M_C =: M < \infty$ such that (7.16) holds and $m < \xi^* < M$: it is possible to make such a choice thanks to the bounds (7.8) for

ξ^* proven in the remark 7.1.7. Fix $\lambda_{m,M} =: \lambda > 0$ such that (7.15) holds. Let $\varepsilon > 0$. Then consider the following random events depending on $x_1, \dots, x_N$

$$E^1_{N,\varepsilon} := \{\, \forall\, \xi \in [0, M] \;\; |\Phi_N(\xi) - \Phi(\xi)| < \varepsilon \,\}$$
$$E^2_N := \{\, \forall\, \xi \in [m, M] \;\; \Phi_N(-\xi) < \Phi_N(\xi) - \lambda \,\}$$
$$E^3_N := \{\, \forall\, \xi \in [M, \infty[\;\; \Phi_N(\xi) < C\,, \;\; \Phi_N(\xi) < \varphi(\xi) \,\}$$

and set $E_{N,\varepsilon} := E^1_{N,\varepsilon} \cap E^2_N \cap E^3_N$. It is convenient to split the expectation of $\log Z_N$ as follows:

$$\mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N\right] \,=\, \mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N \;\mathbb{1}\big(E_{N,\varepsilon}\big)\right] \,+\, \mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N \;\mathbb{1}\big((E_{N,\varepsilon})^c\big)\right]. \tag{7.31}$$

In the following we are going to see that in the limit $N \to \infty$ the second term on the r.h.s. of (7.31) is negligible, while the first term can be computed using the Laplace's method.

By the lemma 7.1.8, using the Hölder inequality and the lemma 7.1.9,

$$\left|\, \mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N \;\mathbb{1}\big((E_{N,\varepsilon})^c\big)\right] \,\right| \,\leq\, \mathbb{E}_{\boldsymbol{x}}\left[\left(\frac{1}{N}\log Z_N\right)^2\right]^{1/2} \mathbb{P}_{\boldsymbol{x}}\big((E_{N,\varepsilon})^c\big)^{1/2} \,\xrightarrow[N\to\infty]{}\, 0\,. \tag{7.32}$$

[*Upper bound*] Using the Gaussian representation (7.4), a simple upper bound for Z_N is

$$Z_N \,\leq\, \frac{\sqrt{N}}{\sqrt{2\pi w}} \int_{\mathbb{R}} e^{-\frac{N}{2w}\,\xi^2} \prod_{i=1}^{N} |\xi + x_i| \,\mathrm{d}\xi \,=\, \frac{\sqrt{N}}{\sqrt{2\pi w}} \int_{\mathbb{R}} e^{N\,\Phi_N(\xi)} \,\mathrm{d}\xi\,. \tag{7.33}$$

If the event $E_{N,\varepsilon}$ holds true, remembering also the inequality (7.13), then the

following upper bound holds:

$$\begin{aligned}
&\int_{\mathbb{R}} e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi \;\le \\
&\le\; 2\int_0^m e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi \;+\; \int_m^M e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi \;+\; \int_m^M e^{N\,(\Phi_N(\xi)-\lambda)}\,\mathrm{d}\xi \;+\; 2\int_M^\infty e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi \;\le \\
&\le\; 2\int_0^m e^{N\,(\Phi(\xi)+\varepsilon)}\,\mathrm{d}\xi \;+\; \int_m^M e^{N\,(\Phi(\xi)+\varepsilon)}\,\mathrm{d}\xi \;+\; \int_m^M e^{N\,(\Phi(\xi)+\varepsilon-\lambda)}\,\mathrm{d}\xi \;+\; 2\,e^{(N-1)\,C}\int_M^\infty e^{\varphi(\xi)}\,\mathrm{d}\xi \;= \\
&\underset{N\to\infty}{=}\; O\big(e^{N\,(\max_{[0,m]}\Phi+\varepsilon)}\big) \;+\; e^{N\,(\Phi(\xi^*)+\varepsilon)}\,\frac{\sqrt{2\pi}\,(1+o(1))}{\sqrt{-N\,\Phi''(\xi^*)}} \;+\; O\big(e^{N\,(\Phi(\xi^*)+\varepsilon-\lambda)}\big) \;+\; O\big(e^{N\,C}\big)\;;
\end{aligned} \tag{7.34}$$

the last step is obtained by applying the Laplace's method (theorem B1) to the function Φ, which by lemma 7.1.6 satisfies all the necessary hypothesis. Now since $\max_{[0,m]}\Phi$, $\Phi(\xi^*)-\lambda$ and C are strictly smaller than $\Phi(\xi^*)$, it holds

$$\text{r.h.s. of (7.34)} \;\underset{N\to\infty}{\sim}\; e^{N\,(\Phi(\xi^*)+\varepsilon)}\,\frac{\sqrt{2\pi}}{\sqrt{-N\,\Phi''(\xi^*)}}\;. \tag{7.35}$$

As a consequence of (7.33), (7.34), (7.35),

$$\frac{1}{N}\log Z_N\;\mathbb{1}(E_{N,\varepsilon}) \;\le\; \Phi(\xi^*)+\varepsilon+O\!\left(\frac{\log N}{N}\right),$$

where the $O(\frac{\log N}{N})$ is deterministic. Therefore for all $\varepsilon>0$

$$\limsup_{N\to\infty}\mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N\;\mathbb{1}(E_{N,\varepsilon})\right] \;\le\; \Phi(\xi^*)+\varepsilon\;. \tag{7.36}$$

[*Lower bound*] Observe that the product $\prod_{i=1}^N(\xi+x_i)$ is always positive for $\xi\ge 0$, while it is negative for some $\xi<0$. Hence using the Gaussian representation (7.4), a lower bound for Z_N is

$$\begin{aligned}
Z_N \;&\ge\; \frac{\sqrt{N}}{\sqrt{2\pi w}}\left(\int_0^\infty e^{-\frac{N}{2w}\,\xi^2}\,\prod_{i=1}^N|\xi+x_i|\,\mathrm{d}\xi \;-\; \int_{-\infty}^0 e^{-\frac{N}{2w}\,\xi^2}\,\prod_{i=1}^N|\xi+x_i|\,\mathrm{d}\xi\right) = \\
&=\; \frac{\sqrt{N}}{\sqrt{2\pi w}}\left(\int_0^\infty e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi \;-\; \int_{-\infty}^0 e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi\right).
\end{aligned} \tag{7.37}$$

If the event $E_{N,\varepsilon}$ holds true, remembering also the inequality (7.13), then the following lower bound holds:

$$\begin{aligned}
&\int_0^\infty e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi\,-\int_{-\infty}^0 e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi\,\geq\\
&\geq\int_m^M e^{N\,\Phi_N(\xi)}\,\mathrm{d}\xi\,-\int_m^M e^{N\,(\Phi_N(\xi)-\lambda)}\,\mathrm{d}\xi\,\geq\\
&\geq\int_m^M e^{N\,(\Phi(\xi)-\varepsilon)}\,\mathrm{d}\xi\,-\int_m^M e^{N\,(\Phi(\xi)+\varepsilon-\lambda)}\,\mathrm{d}\xi\,=\\
&\underset{N\to\infty}{=}\,e^{N\,(\Phi(\xi^*)-\varepsilon)}\,\frac{\sqrt{2\pi}\,(1+o(1))}{\sqrt{-N\,\Phi''(\xi^*)}}\,-\,e^{N\,(\Phi(\xi^*)+\varepsilon-\lambda)}\,\frac{\sqrt{2\pi}\,(1+o(1))}{\sqrt{-N\,\Phi''(\xi^*)}}\,;
\end{aligned}\tag{7.38}$$

the last step is obtained by applying the Laplace's method (theorem B1) to the function Φ, which by lemma 7.1.6 satisfies all the necessary hypothesis. Now since $\Phi(\xi^*)+\varepsilon-\lambda<\Phi(\xi^*)-\varepsilon$ for all $0<\varepsilon<\frac{1}{2}\lambda$, for such a choice of ε it holds

$$\text{r.h.s. of (7.38)}\;\underset{N\to\infty}{\sim}\;e^{N\,(\Phi(\xi^*)-\varepsilon)}\,\frac{\sqrt{2\pi}}{\sqrt{-N\,\Phi''(\xi^*)}}\,.\tag{7.39}$$

As a consequence of (7.37), (7.38), (7.39), for all $0<\varepsilon<\frac{1}{2}\lambda$

$$\frac{1}{N}\log Z_N\;\mathbb{1}(E_{N,\varepsilon})\;\geq\;\left(\Phi(\xi^*)-\varepsilon+O\!\left(\frac{\log N}{N}\right)\right)\mathbb{1}(E_{N,\varepsilon})\,,$$

where the $O(\frac{\log N}{N})$ is deterministic. Therefore, using also the lemma 7.1.8, for all $0<\varepsilon<\frac{1}{2}\lambda$

$$\liminf_{N\to\infty}\mathbb{E}_{\boldsymbol{x}}\left[\frac{1}{N}\log Z_N\;\mathbb{1}(E_{N,\varepsilon})\right]\;\geq\;\liminf_{N\to\infty}\left(\Phi(\xi^*)-\varepsilon+O\!\left(\frac{\log N}{N}\right)\right)\mathbb{P}_{\boldsymbol{x}}(E_{N,\varepsilon})\;=\;\Phi(\xi^*)-\varepsilon\,.\tag{7.40}$$

In conclusion the convergence $\mathbb{E}_{\boldsymbol{x}}[\frac{1}{N}\log Z_N]\to\Phi(\xi^*)$ as $N\to\infty$ is proven by considering (7.31) for $0<\varepsilon<\frac{1}{2}\lambda$, then letting $N\to\infty$ exploiting (7.32), (7.36), (7.40), and finally letting $\varepsilon\to 0+$. □

Remark 7.1.10. In the deterministic case, namely when the distribution of the x_i's is a Dirac delta centred at a point x, the theorem 7.1.3 and its corollary 7.1.5 reproduce the results obtained in the Proposition 6 of [28] by a combinatorial

computation. Indeed the fixed point equation (7.7) reduces to $\xi^* = \frac{w}{\xi^*+x}$, whose positive solution is

$$\xi^* = \frac{-x + \sqrt{x^2+4w}}{2} \, .$$

As a consequence, by (7.9) the limiting dimer and monomer density are respectively

$$d = \frac{(\xi^*)^2}{2w} = \frac{x^2 - x\sqrt{x^2+4w} + 2w}{2w} \, , \quad m = 1 - 2\,d = \frac{-x^2 + x\sqrt{x^2+4w}}{2w} \, .$$

Moreover by (7.5) and (7.9) the limiting pressure can be written as

$$p = \Phi(\xi^*) = -\frac{(\xi^*)^2}{2w} + \log(\xi^* + x) = -d - \frac{1}{2}\log\frac{2\,d}{w} \, .$$

7.2 Concentration inequality for random monomer-dimer models

In this section we prove that under quite general hypothesis an the pressure of a random MD model with independent random weights satisfy an exponential concentration inequality. In particular it will follows, by classical spin glass arguments [13], that the pressure is self averaging and the convergence (7.5) of the theorem 7.1.3 can be strengthen as

$$\mathbb{P}_{\boldsymbol{x}} \text{ - almost surely } \; \exists \lim_{N\to\infty} \frac{1}{N}\log Z_N \; = \; \sup_{\xi\geq 0} \Phi(\xi) \, , \tag{7.41}$$

when in the hypothesis of the theorem 7.1.3 one substitutes $\mathbb{E}_x[x] < \infty$, $\mathbb{E}_x[(\log x)^2] < \infty$ with the stronger $\mathbb{E}_x[x] < \infty$, $\mathbb{E}_x[x^{-1}] < \infty$.

In general let $w_{ij}^{(N)} \geq 0$, $1 \leq i < j \leq N$, $N \in \mathbb{N}$, and $x_i > 0$, $i \in \mathbb{N}$, be *independent* random variables. Since the dimer weights may be allowed to take the value 0 (or to be identically 0), we do not really know on which kind of graph the model lives, on the contrary the framework is very general (for example the complete graph is included, but also finite-dimensional lattices or diluted random graphs are). This is why we allow a generic dependence of the

dimer weights on N, in case a normalisation is needed. During all this section we will denote

$$Z_N := \sum_{D\in\mathscr{D}_N} \prod_{ij\in D} w_{ij}^{(N)} \prod_{i\in M_N(D)} x_i \,. \tag{7.42}$$

Denote simply by $\mathbb{E}[\,\cdot\,]$ the expectation with respect to all the weights and assume that

$$\sup_N \sup_{1\le i<j\le N} \mathbb{E}[w_{ij}^{(N)}] =: C_1 < \infty\,, \quad \sup_{i\in\mathbb{N}} \mathbb{E}[x_i] =: C_2 < \infty\,, \quad \sup_{i\in\mathbb{N}} \mathbb{E}[x_i^{-1}] =: C_3 < \infty\,. \tag{7.43}$$

Clearly the pressure $p_N := \frac{1}{N}\log Z_N$ is a random variable and it has finite expectation, indeed

$$N\,p_N \begin{cases} \geq \log\prod_{i=1}^N x_i \,=\, \sum_{i=1}^N \log x_i \,\geq\, \sum_{i=1}^N (1+x_i^{-1}) \;\in L^1(\mathbb{P}) \\ \leq\, Z_N - 1 \;\in L^1(\mathbb{P}) \end{cases}.$$

The following theorem shows that in the limit $N\to\infty$ the pressure p_N concentrates around its expectation, or in other terms it tends to become a deterministic quantity.

Theorem 7.2.1. *Let $w_{ij}^{(N)} \geq 0$, $1\le i<j\le N$, $N\in\mathbb{N}$, and $x_i>0$, $i\in\mathbb{N}$, be independent random variables that satisfy* (7.43)*. Then for all $t>0$, $N\in\mathbb{N}$, $q\geq 1$*

$$\mathbb{P}\big(\,|p_N - \mathbb{E}[p_N]|\geq t\big) \,\leq\, 2\,\exp\left(-\frac{t^2\,N}{4\,q^2\,\log^2 N}\right) \,+\, (a+b\,N)\,N^{1-q}\,, \tag{7.44}$$

where $a := 4 + 2C_2C_3$, $b := 2C_1C_3^2$. As a consequence, choosing $q>3$,

$$|p_N - \mathbb{E}[p_N]| \xrightarrow[N\to\infty]{} 0 \;\; \mathbb{P}\text{-almost surely}\,. \tag{7.45}$$

If the random variables $w_{ij}^{(N)}$, x_i, x_i^{-1} are bounded, then one could obtain an exponential rate of convergence instead of (7.44), but here we prefer to obtain the result (7.45) with minimal assumptions.

Proof. Fix $N\in\mathbb{N}$. Set $w_i := \big(w_{i(i+1)}^{(N)},\dots,w_{iN}^{(N)}\big)$ for all $i=1,\dots,N-1$. We consider the filtration of length $2N-1$ such that in the first N steps the

monomer weights x_i are exposed, while in the last $N-1$ steps the vectors w_i of dimer weights are exposed. Since p_N is a function of $x_1, \dots, x_N$, $w_1, \dots, w_{N-1}$ and $\mathbb{E}[|p_N|] < \infty$, we may define the Doob martingale of p_N with respect to this filtration

$$M_i := \mathbb{E}\big[\, p_N \,\big|\, x_1, \dots, x_i\big] \quad \forall\, i = 0, \dots, N \;,$$

$$M_{N+i} := \mathbb{E}\big[\, p_N \,\big|\, x_1, \dots, x_N, w_1, \dots, w_i\big] \quad \forall\, i = 1, \dots, N-1 \;;$$

in particular it holds $M_0 = \mathbb{E}[\, p_N]$ and $M_{2N-1} = p_N$.
Now we want to bound the increments $|M_i - M_{i-1}|$ for every $i = 1, \dots, 2N-1$, in order to apply the Azuma's inequality. By hypothesis $x_1, \dots, x_N$, $w_1, \dots, w_{N-1}$ are stochastically independent, hence the conditional expectations are simply $M_i = \mathbb{E}_{\mathbf{x}^{i+1},\, \mathbf{w}}[p_N]$ for $i = 0, \dots, N$ and $M_{N+i} = \mathbb{E}_{\mathbf{w}^{i+1}}[p_N]$ for $i = 1, \dots, N-1$.
As a consequence it is easy to check that for $i = 1, \dots, N$ it holds

$$|M_i - M_{i-1}| \;\le \sup_{\tilde{\mathbf{x}}_{i-1},\, \tilde{\mathbf{x}}^{i+1},\, \tilde{\mathbf{w}}} \Big|\, p_N\big(\tilde{\mathbf{x}}_{i-1},\, x_i,\, \tilde{\mathbf{x}}^{i+1},\, \tilde{\mathbf{w}}\big) - \mathbb{E}_{x_i}\big[\, p_N\big(\tilde{\mathbf{x}}_{i-1},\, x_i,\, \tilde{\mathbf{x}}^{i+1},\, \tilde{\mathbf{w}}\big)\big]\Big| \tag{7.46}$$

and for $i = 1, \dots, N-1$ it holds

$$|M_{N+i} - M_{N+i-1}| \;\le \sup_{\tilde{\mathbf{w}}_{i-1},\, \tilde{\mathbf{w}}^{i+1}} \Big|\, p_N\big(\mathbf{x},\, \tilde{\mathbf{w}}_{i-1},\, w_i,\, \tilde{\mathbf{w}}^{i+1}\big) - \mathbb{E}_{w_i}\big[\, p_N\big(\mathbf{x},\, \tilde{\mathbf{w}}_{i-1},\, w_i,\, \tilde{\mathbf{w}}^{i+1}\big)\big]\Big| \;. \tag{7.47}$$

Here we have adopted the following notation $\mathbf{x} := (x_1, \dots, x_N)$, $\mathbf{x}_k := (x_1, \dots, x_k)$, $\mathbf{x}^k := (x_k, \dots, x_N)$ and similarly $\mathbf{w} := (w_1, \dots, w_{N-1})$, $\mathbf{w}_k := (w_1, \dots, w_k)$, $\mathbf{w}^k := (w_k, \dots, w_N)$; the symbols with a tilde denote a deterministic value taken by the corresponding random quantity.

First fix $i = 1, \dots, N$, fix the deterministic vectors $\tilde{\mathbf{x}}_{i-1}$, $\tilde{\mathbf{x}}^{i+1}$, $\tilde{\mathbf{w}}$ and let x_i', x_i'' be two independent random variables distributed as x_i. Set

$$p_N' := p_N\big(\tilde{\mathbf{x}}_{i-1}, x_i',\, \tilde{\mathbf{x}}^{i+1},\, \tilde{\mathbf{w}}\big) \;, \quad p_N'' := p_N\big(\tilde{\mathbf{x}}_{i-1},\, x_i'',\, \tilde{\mathbf{x}}^{i+1},\, \tilde{\mathbf{w}}\big) \;.$$

To estimate the difference between p_N', p_N'' we use the Heilmann-Lieb recursion

for the partition function of a monomer-dimer model (see [86] and the proposition 5.2.3):

$$p'_N - p''_N \;=\; \frac{1}{N}\log\frac{Z'_N}{Z''_N} \;=\; \frac{1}{N}\log\frac{x'_i\,Z_{-i} \;+\; \sum_{j=1}^{i-1}\tilde{w}_{ji}\,Z_{-j-i} \;+\; \sum_{j=i+1}^{N}\tilde{w}_{ij}\,Z_{-i-j}}{x''_i\,Z_{-i} \;+\; \sum_{j=1}^{i-1}\tilde{w}_{ji}\,Z_{-j-i} \;+\; \sum_{j=i+1}^{N}\tilde{w}_{ij}\,Z_{-i-j}} \;\leq$$
$$\leq\; \frac{1}{N}\log\left(\frac{x'_i}{x''_i} \;+\; 1\right); \tag{7.48}$$

here we denote by Z_{-i}, Z_{-i-j} the partitions function of the model over the vertices $\{1,\ldots,N\}\smallsetminus\{i\}$, $\{1,\ldots,N\}\smallsetminus\{i,j\}$ respectively, with weights $\tilde{\mathbf{x}}_{i-1}$, $\tilde{\mathbf{x}}^{i+1}$, $\tilde{\mathbf{w}}_{i-1}$, $\tilde{\mathbf{w}}^{i+1}$. It is important (for the inequality in (7.48)) to notice that these partition functions do not depend on the weights x'_i, x''_i. In the same way one finds

$$p''_N - p'_N \;\leq\; \frac{1}{N}\log\left(\frac{x''_i}{x'_i} \;+\; 1\right). \tag{7.49}$$

Denote by $\mathbb{E}''$ the expectation with respect to the random variable x''_i only. Then the inequalities (7.48), (7.49) provide respectively the following random bounds

$$p'_N - \mathbb{E}[p''_N] \;=\; \mathbb{E}''[p'_N - p''_N] \overset{(7.48)}{\leq}\; \mathbb{E}''\left[\frac{1}{N}\log\left(\frac{x'_i}{x''_i} \;+\; 1\right)\right] \;\leq\; \frac{1}{N}\log\left(x'_i\,\mathbb{E}[x_i^{-1}] + 1\right); \tag{7.50}$$

$$\mathbb{E}[p''_N] - p'_N \;=\; \mathbb{E}''[p''_N - p'_N] \overset{(7.49)}{\leq}\; \mathbb{E}''\left[\frac{1}{N}\log\left(\frac{x''_i}{x'_i} \;+\; 1\right)\right] \;\leq\; \frac{1}{N}\log\left(\mathbb{E}[x_i]\,(x'_i)^{-1} + 1\right). \tag{7.51}$$

Choose $q > 0$ and the previous inequalities provide a bound for $|M_i - M_{i-1}|$

that holds true "with high probability":

$$\begin{aligned}
&\mathbb{P}\Big(|M_i - M_{i-1}| > \frac{q}{N}\log N\Big) \overset{(7.46)}{\leq} \mathbb{P}\Big(\sup_{\tilde{\mathbf{x}}_{i-1},\,\tilde{\mathbf{x}}^{i+1},\,\tilde{\mathbf{w}}} \big|p'_N - \mathbb{E}[p''_N]\big| > \frac{q}{N}\log N\Big) \leq \\
&\leq \mathbb{P}\Big(\sup_{\tilde{\mathbf{x}}_{i-1},\,\tilde{\mathbf{x}}^{i+1},\,\tilde{\mathbf{w}}} \big(p'_N - \mathbb{E}[p''_N]\big) > \frac{q}{N}\log N\Big) + \mathbb{P}\Big(\sup_{\tilde{\mathbf{x}}_{i-1},\,\tilde{\mathbf{x}}^{i+1},\,\tilde{\mathbf{w}}} \big(\mathbb{E}[p''_N] - p'_N\big) > \frac{q}{N}\log N\Big) \overset{(7.50),(7.51)}{\leq} \\
&\leq \mathbb{P}\Big(\frac{1}{N}\log\big(x_i\,\mathbb{E}[x_i^{-1}] + 1\big) > \frac{q}{N}\log N\Big) + \mathbb{P}\Big(\frac{1}{N}\log\big(\mathbb{E}[x_i]\,x_i^{-1} + 1\big) > \frac{q}{N}\log N\Big) = \\
&= \mathbb{P}\Big(1 + x_i\,\mathbb{E}[x_i^{-1}] > N^q\Big) + \mathbb{P}\Big(1 + \mathbb{E}[x_i]\,x_i^{-1} > N^q\Big) \leq \qquad (7.52)\\
&\leq \mathbb{E}\Big[1 + x_i\,\mathbb{E}[x_i^{-1}]\Big]\,N^{-q} + \mathbb{E}\Big[1 + \mathbb{E}[x_i]\,x_i^{-1}\Big]\,N^{-q} \leq \\
&\leq 2\,(1 + C_2C_3)\,N^{-q}\,;
\end{aligned}$$

here at the penultimate step we have used the Markov inequality.

Now instead fix $i = 1, \ldots, N-1$, fix the deterministic vectors $\tilde{\mathbf{w}}_{i-1}$, $\tilde{\mathbf{w}}^{i+1}$, let w'_i, w''_i be two independent random vectors distributed as w_i and leave the vector of monomer weights $\mathbf{x}$ random (choose w'_i, w''_i independent of $\mathbf{x}$ too). Reassign the notation previously used, setting now:

$$p'_N := p_N\big(\mathbf{x},\,\tilde{\mathbf{w}}_{i-1},\,w'_i,\,\tilde{\mathbf{w}}^{i+1}\big)\,,\quad p''_N := p_N\big(\mathbf{x},\,\tilde{\mathbf{w}}_{i-1},\,w''_i,\,\tilde{\mathbf{w}}^{i+1}\big)\,.$$

To estimate the difference between p'_N, p''_N we use again the Heilmann-Lieb recursion for the partition function (see [86] and the proposition 5.2.3):

$$\begin{aligned}
p'_N - p''_N \;&=\; \frac{1}{N}\log\frac{Z'_N}{Z''_N} \;=\; \frac{1}{N}\log\frac{x_i\,Z_{-i} + \sum_{j=1}^{i-1}\tilde{w}_{ji}\,Z_{-j-i} + \sum_{j=i+1}^{N} w'_{ij}\,Z_{-i-j}}{x_i\,Z_{-i} + \sum_{j=1}^{i-1}\tilde{w}_{ji}\,Z_{-j-i} + \sum_{j=i+1}^{N} w''_{ij}\,Z_{-i-j}} \leq \\
&\leq \frac{1}{N}\log\left(1 + \frac{\sum_{j=i+1}^{N} w'_{ij}\,Z_{-i-j}}{x_i\,Z_{-i}}\right) = \frac{1}{N}\log\left(1 + \sum_{j=i+1}^{N}\frac{w'_{ij}}{x_i\,x_j}\,\langle \mathbb{1}_{j\in M}\rangle_{-i}\right) \leq \\
&\leq \frac{1}{N}\log\left(1 + \sum_{j=i+1}^{N}\frac{w'_{ij}}{x_i\,x_j}\right);
\end{aligned} \qquad (7.53)$$

we have denoted by Z_{-i}, Z_{-i-j} the partitions function of the model over the vertices $\{1,\ldots,N\}\smallsetminus\{i\}$, $\{1,\ldots,N\}\smallsetminus\{i,j\}$ respectively, with weights $\mathbf{x}_{i-1}$, $\mathbf{x}^{i+1}$, $\tilde{\mathbf{w}}_{i-1}$, $\tilde{\mathbf{w}}^{i+1}$.

It is important (for the first inequality in (7.53)) to notice that these partition functions do not depend on the weights w'_i, w''_i. In the same way one finds

$$p''_N - p'_N \,\le\, \frac{1}{N}\log\left(1 \,+\, \sum_{j=i+1}^{N} \frac{w''_{ij}}{x_i\, x_j}\right). \tag{7.54}$$

Denote by $\mathbb{E}''$ the expectation with respect to the random vector w''_i only. Then the inequalities (7.53), (7.54) provide respectively the following random bounds

$$p'_N - \mathbb{E}''[p''_N] \,=\, \mathbb{E}''[p'_N - p''_N] \overset{(7.54)}{\le} \frac{1}{N}\log\left(1 \,+\, \sum_{j=i+1}^{N} \frac{w'_{ij}}{x_i\, x_j}\right); \tag{7.55}$$

$$\begin{aligned} \mathbb{E}''[p''_N] - p'_N \,&=\, \mathbb{E}''[p''_N - p'_N] \overset{(7.55)}{\le} \mathbb{E}''\left[\frac{1}{N}\log\left(1 \,+\, \sum_{j=i+1}^{N} \frac{w''_{ij}}{x_i\, x_j}\right)\right] \le \\ &\le\, \frac{1}{N}\log\left(1 \,+\, \sum_{j=i+1}^{N} \frac{\mathbb{E}[w_{ij}]}{x_i\, x_j}\right). \end{aligned} \tag{7.56}$$

Choose $q > 0$ and the previous inequalities provide a bound for $|M_{N+i} - M_{N+i-1}|$ that holds true "with high probability":

$$\begin{aligned} &\mathbb{P}\left(|M_{N+i} - M_{N+i-1}| > \frac{q}{N}\log N\right) \overset{(7.47)}{\le} \mathbb{P}\left(\sup_{\tilde{\mathbf{w}}_{i-1},\, \tilde{\mathbf{w}}^{i+1}} \left|p'_N - \mathbb{E}''[p''_N]\right| > \frac{q}{N}\log N\right) \le \\ &\le\, \mathbb{P}\left(\sup_{\tilde{\mathbf{w}}_{i-1},\, \tilde{\mathbf{w}}^{i+1}} \left(p'_N - \mathbb{E}''[p''_N]\right) > \frac{q}{N}\log N\right) \,+\, \mathbb{P}\left(\sup_{\tilde{\mathbf{w}}_{i-1},\, \tilde{\mathbf{w}}^{i+1}} \left(\mathbb{E}''[p''_N] - p'_N\right) > \frac{q}{N}\log N\right) \overset{(7.54),(7.55)}{\le} \\ &\le\, \mathbb{P}\left(\frac{1}{N}\log\left(1 + \sum_{j=i+1}^{N} \frac{w_{ij}}{x_i\, x_j}\right) > \frac{q}{N}\log N\right) \,+\, \mathbb{P}\left(\frac{1}{N}\log\left(1 + \sum_{j=i+1}^{N} \frac{\mathbb{E}[w_{ij}]}{x_i\, x_j}\right) > \frac{q}{N}\log N\right) \,= \\ &\le\, \mathbb{P}\left(1 + \sum_{j=i+1}^{N} \frac{w_{ij}}{x_i\, x_j} > N^q\right) \,+\, \mathbb{P}\left(1 + \sum_{j=i+1}^{N} \frac{\mathbb{E}[w_{ij}]}{x_i\, x_j} > N^q\right) \le \\ &\le\, \mathbb{E}\left[\,1 + \sum_{j=i+1}^{N} \frac{w_{ij}}{x_i\, x_j}\,\right] N^{-q} \,+\, \mathbb{E}\left[\,1 + \sum_{j=i+1}^{N} \frac{\mathbb{E}[w_{ij}]}{x_i\, x_j}\,\right] N^{-q} \,\le \\ &\le\, 2\,(1 + N\, C_1 C_3^2)\, N^{-q}\,; \end{aligned} \tag{7.57}$$

here at the penultimate step we have applied the Markov inequality.

As an immediate consequence of (7.52) and (7.57),

$$\begin{aligned}
&\mathbb{P}\Big(\exists\, i = 1, \dots, 2N-1 \text{ s.t. } |M_i - M_{i-1}| > \frac{q}{N} \log N\Big) \;\leq \\
&\leq\; N\left(2\,(1 + C_2C_3)\,N^{-q}\right) \;+\; (N-1)\left(2\,(1 + N\,C_1C_3^2)\,N^{-q}\right) \\
&\leq\; 2\left(2 + C_2C_3 + C_1C_3^2\,N\right)N^{1-q}\,.
\end{aligned} \tag{7.58}$$

Therefore by the extended Azuma's inequality (theorem C2), for all $t > 0$ it holds

$$\mathbb{P}\big(|M_{N-1} - M_0| \geq t\big) \;\leq\; 2\,\exp\left(-\frac{t^2}{2}\,\frac{N}{2\,q^2\,\log^2 N}\right) + 2\left(2 + C_2C_3 + C_1C_3^2\,N\right)N^{1-q} \tag{7.59}$$

and the proof of (7.44) is concluded. Choosing $q > 3$ the r.h.s. of (7.44) is summable with respect to $N \in \mathbb{N}$, hence (7.45) follows by a standard application of the Borel-Cantelli lemma. □

Outlooks

In the present work I reported new results that I have obtained, trough various collaborations, during my PhD course. The main field is the Statistical Mechanics of Equilibrium, and they cover two main topics: Spin Glass models (Chapters 1-4) and Monomer Dimer models (Chapters 1-7). Despite we gave a complete answer to all questions raised, that is we founded a more or less explicit representation of the limiting pressure and the Gibbs measure, several related and important problems are still open. I would like to make a little bit more explicit this aspect.

The main contribution of Chapter 3 is the development of a multidimensional version of the Parisi Theory which describes the thermodynamical limit of the Multi-species SK model. However, our results hold in a subset of the domain of the parameters which satisfies a convexity condition: the mutual interaction does not exceed a threshold where the strength of the off-diagonal terms prevails on the inter-party interactions. I would be very interesting to extend the results to the opposite case. The latter is intrinsically different because the model approaches the Hopfield model for neural network, a well know longstanding open problem in spin glass theory.

In Chapter 4 we developed a general framework for the study of the stability, under suitable perturbations, of the limiting Gibbs measure. The framework is general in the sense that covers a large class of spin glass models, including short range interactions. This allow us to deduce, let say for the Edward Anderson model, the ultrametric property of the limiting Gibbs measure under small

perturbations. However, is still an open question if or not this ultrametric structure gives non trivial information about the physical properties of the system.

In Chapters 6 and 7 we introduce two modifications of the classical Monomer Dimer model. The first one is obtained adding an imitative attraction among the particles (monomer and dimer) while the second is characterized by the quenched randomness on the activities. In both cases, we have shown that the models are solvable in the complete graph. The natural next step would be to study the previous models on a lattice, the difficulty of the mathematical treatment presumably increases by several orders of magnitudo as well the interest of possible results in this direction.

One can also think to mix the two previous features introducing a random interaction among particles. In this way one can hope to obtain a glassy behavior, namely a kind of *Monomer Dimer Glassy system*.

Appendix

A. Wick's Theorem

In this appendix we state the main technical results used in the paper. We omit their proofs, that can be found in the literature.

Theorem A1 (Gaussian integration by parts; Wick-Isserlis formula)**.** *Let* $(\xi_1, \dots, \xi_n)$ *be a Gaussian random vector with mean* 0 *and positive semi-definite covariance matrix* $C = (c_{ij})_{i,j=1,\dots,n}$*. Let* $f : \mathbb{R}^{n-1} \to \mathbb{R}$ *be a differentiable function such that* $\mathbb{E}\big[\big|\xi_1 f(\xi_2, \dots, \xi_n)\big|\big] < \infty$ *and* $\mathbb{E}\big[\big|\frac{\partial f}{\partial \xi_j}(\xi_2, \dots, \xi_n)\big|\big] < \infty$ *for all* $j = 2, \dots, n$*. Then:*

$$\mathbb{E}\big[\xi_1 f(\xi_2, \dots, \xi_n)\big] = \sum_{j=2}^{n} c_{1j}\, \mathbb{E}\left[\frac{\partial f}{\partial \xi_j}(\xi_2, \dots, \xi_n)\right] . \tag{A1}$$

As a consequence one can prove the following:

$$\mathbb{E}\left[\prod_{i=1}^{n} \xi_i\right] = \sum_{\substack{P \text{ partition of} \\ \{1,\dots,n\} \text{ into pairs}}} \prod_{\{i,j\} \in P} c_{ij} \ . \tag{A2}$$

The Gaussian integration by parts (A1) can be found in [17]. The Wick-Isserlis formula (A2) follows by (A1) using an induction argument; but it appeared for the first time in [89].

B. Laplace method

Theorem B1 (Laplace's method)**.** *Let* $\phi\colon [a, b] \to \mathbb{R}$ *be a function of class* C^2*. Suppose that there exists* $x_0 \in]a, b[$ *such that*

i. $\phi(x_0) > \phi(x)$ *for all* $x \in [a,b]$ *(i.e.* x_0 *is the only global maximum point of* ϕ*);*

ii. $\phi''(x_0) < 0$.

Then as $n \to \infty$

$$\int_a^b e^{n\,\phi(x)}\,\mathrm{d}x \;=\; e^{n\,\phi(x_0)}\;\frac{\sqrt{2\pi}}{\sqrt{-n\,\phi''(x_0)}}\,\big(1+o(1)\big)\;. \tag{B1}$$

A formal proof of the Laplace's method can be found in [66].

C. Probability estimation

Theorem C1 (uniform weak law of large numbers)**.** *Let* $\mathcal{X}$, Θ *be metric spaces. Let* X_i, $i \in \mathbb{N}$ *be i.i.d. random variables taking values in* $\mathcal{X}$*. Let* $f\colon \mathcal{X} \times \Theta \to \mathbb{R}$ *be a function such that* $f(\cdot,\theta)$ *is measurable for all* $\theta \in \Theta$*. Suppose that:*

i. Θ *is compact;*

ii. $\mathbb{P}\big(f(X_1,\cdot)$ *is continuous at* $\theta\big) = 1$ *for all* $\theta \in \Theta$ *;*

iii. $\exists\ F\colon \mathcal{X} \to [0,\infty]$ *such that* $\mathbb{P}\big(|f(X_1,\theta)| \le F(X_1)\big) = 1$ *for all* $\theta \in \Theta$ *and* $\mathbb{E}[F(X_1)] < \infty$.

Then for all $\varepsilon > 0$

$$\mathbb{P}\bigg(\sup_{\theta\in\Theta}\,\bigg|\,\frac{1}{n}\sum_{i=1}^{n} f(X_i,\theta) - \mathbb{E}[f(X,\theta)]\,\bigg|\;\ge\varepsilon\bigg)\;\xrightarrow[n\to\infty]{}\;0\;. \tag{C1}$$

The uniform law of large number appeared in [90]. It is based on the (standard) law of large numbers and on a compactness argument.

Theorem C2 (extension of the Azuma's inequality)**.** *Let* $M = (M_i)_{i=0,\dots,n}$ *be a real martingale with respect to a filter. Suppose that there exist constants* $\varepsilon > 0$ *and* $c_1,\dots,c_n < \infty$ *such that*

$$\mathbb{P}\big(\exists\, i = 1,\dots,n \text{ s.t. } |M_i - M_{i-1}| > c_i\big)\;\le\;\varepsilon\;.$$

Then for all $t > 0$

$$\mathbb{P}\big(|M_n - M_0| > t\big) \;\leq\; 2\,\exp\left(-\,\frac{t^2}{2\,\sum_{i=1}^{n} c_i^2}\right) + \,\varepsilon\,. \tag{C2}$$

The Azuma's inequality is a useful tool in the martingale theory that allows to obtain concentration results. Its usual formulation is given with $\varepsilon = 0$. The extension with $\varepsilon > 0$ can be found in [55]; but it can be proven also starting from the usual formulation and introducing a suitable stopping time, following the ideas in [123].

D. Interpolation method

Much of the recent progresses in the study of mean field spin glass models is based on methods and arguments introduced by Guerra in a series of works (see e.g. [79, 85, 84]), constituting the so called *interpolation method*. Beyond the original works, the interested reader can found a detailed and complete exposition with several applications of this method in [17]. In order to present a self-consistent exposition, hereafter we outline briefly the basic ideas.

Let N be an integer, and for $i \in I = \{1, ..., N\}$, let U_i and $\widetilde{U}_i$ be two families of centered Gaussian random variables, independent each other, uniquely determined by the respective covariance matrices $\mathbb{E}(U_i U_j) = C_{ij}$ and $\mathbb{E}(\widetilde{U}_i \widetilde{U}_j) = \widetilde{C}_{ij}$. We treat the set of indices i as configuration space for some statistical mechanics system. Let $a_i \in \mathbb{R}^+$ for each $i \in I$ be an arbitrary (finite) weight. We define the Hamiltonian interpolating function as the following random variable

$$H_i(t) := \sqrt{t} U_i + \sqrt{1-t}\tilde{U}_i,$$

where $t \in [0, 1]$ is the real parameter used for interpolation.

Let us introduce also the so-called *quenched measures*. First, we define the random partition function of the system as

$$Z(t) := \sum_i a_i e^{H_i(t)},$$

and the random Gibbs measure as

$$G_i(t) := \frac{a_i e^{H_i(t)}}{Z(t)}.$$

Then let $F : (I \times I) \longrightarrow \mathbb{R}$ be an observable defined in the duplicated configuration space, we define the *quenched measure* as

$$\langle F \rangle_t := \mathbb{E}\Big(\Omega_t(F)\Big), \tag{D1}$$

where

$$\Omega_t(F) := \sum_{i,j} G_i(t) G_j(t) F_{ij}. \tag{D2}$$

The measure Ω_t is called the duplicated random Gibbs measure.
Keeping in mind definition (D1), it is possible to prove (see [17]) the following

Proposition 7.2.1. *Consider the function*

$$\varphi(t) := \mathbb{E} \log Z(t) \tag{D3}$$

then for its t-derivative the following expression holds

$$\varphi'(t) = \frac{1}{2}\langle C_{ii} - \tilde{C}_{ii} \rangle_t - \frac{1}{2}\langle C_{ij} - \tilde{C}_{ij} \rangle_t. \tag{D4}$$

The generalization to multi-partite systems requires only minor modifications. Suppose that the system is composed by a finite number S of species indexed by $s \in \mathcal{S}$, then $|\mathcal{S}| = S$. Consider a generic statistical mechanic system as before and assume that:

- the configuration space is decomposed in a disjoint union $I = \bigcup_{s \in \mathcal{S}} I^{(s)}$,

- the U's are also decomposed in the following way

$$U_i = \sum_{s,p \in \mathcal{S}} U_i^{(sp)}, \tag{D5}$$

where $U_i^{(sp)}$ is a family of gaussian r.v. such that the covariance matrix is of the form

$$\mathbb{E}(U_i^{(sp)}U_j^{(s'p')}) = \Delta_{sp}^2\delta_{ss'}\delta_{pp'}C_{ij}^{(s)}C_{ij}^{(p)} \tag{D6}$$

where $C_{ij}^{(s)}$ is a covariance matrix defined on $I^{(s)} \times I^{(s)}$.

Notice that the covariance matrix defined in (D6) is the Schur-Hadamard product of the $C_{ij}^{(s)}$ and then is positive definite. The family of positive parameters $(\Delta_{sp}^2)_{s,p\in\mathcal{S}}$ tunes the interactions between the various species.
For a fixed couple (i,j) we can think at each $C_{ij}^{(s)}$ as a component of a vector in the space $\mathbb{R}^S$ and then, thanks to (D5) and (D6), the covariance matrix of the entire system can be rewritten, with a little abuse of notation, as a quadratic form in $\mathbb{R}^S$, namely as

$$C_{ij} = \mathbb{E}(U_iU_j) = \sum_{s,p\in\mathcal{S}} C_{ij}^{(s)}\Delta_{sp}^2 C_{ij}^{(p)} = \Big(\mathbf{C}, \mathbf{\Delta C}\Big), \tag{D7}$$

where $\mathbf{C} := (C_{ij}^{(s)})_{s\in\mathcal{S}}$ is a vector in $\mathbb{R}^S$ and $\mathbf{\Delta}$ is the real symmetric matrix defined by the entries

$$\mathbf{\Delta} := (\Delta_{sp}^2)_{s,p\in\mathcal{S}}.$$

Suppose for simplicity that $C_{ii}^{(s)} = \sqrt{c}$ for some $c \in \mathbb{R}^+$ for each $i \in I, s \in \mathcal{S}$, that is

$$C_{ii} = c(\mathbf{1}, \mathbf{\Delta 1}), \tag{D8}$$

where

$$\mathbf{1} := (1)_{s\in\mathcal{S}}.$$

Under the assumption that an analogous decomposition holds for the $\widetilde{U}$'s too, then

$$\widetilde{C}_{ij} = \mathbb{E}(\widetilde{U}_i\widetilde{U}_j) = \sum_{s,p\in\mathcal{S}} \widetilde{C}_{ij}^{(s)}\Delta_{sp}^2 \widetilde{C}_{ij}^{(p)} = \Big(\widetilde{\mathbf{C}}, \mathbf{\Delta}\widetilde{\mathbf{C}}\Big), \tag{D9}$$

and

$$\widetilde{C}_{ii} = \widetilde{c}(\mathbf{1}, \mathbf{\Delta 1}). \tag{D10}$$

In the multipartite framework, by (D7,D8,D9,D10), Proposition 7.2.1 becomes

Proposition 7.2.2. *Consider the functional defined in (D3), then for its t-derivative the following holds*

$$\varphi'(t) = \frac{1}{2}(c - \widetilde{c})(\mathbf{1}, \mathbf{\Delta 1}) - \frac{1}{2}\langle(\mathbf{C}, \mathbf{\Delta C}) - (\widetilde{\mathbf{C}}, \mathbf{\Delta}\widetilde{\mathbf{C}})\rangle_t. \tag{D11}$$

In order to separate the contribution of the various species, let us introduce the operator $\mathcal{P}_s$ as the canonical projector in $\mathbb{R}^S$.
For any $s \in \mathcal{S}$ and for any vector $\mathbf{u} = (u^{(s)})_{s\in\mathcal{S}}$ in $\mathbb{R}^S$, we have that

$$\mathcal{P}_s\Big(\mathbf{u}\Big) := u^{(s)}. \tag{D12}$$

Clearly, for two vectors $\mathbf{u}, \mathbf{v}$, the following relation holds

$$\Big(\mathbf{u}, \mathbf{\Delta v}\Big) = \sum_{s\in\mathcal{S}} \mathcal{P}_s\Big(\mathbf{u}\Big)\mathcal{P}_s\Big(\mathbf{\Delta v}\Big) = \sum_{s\in\mathcal{S}} \mathcal{P}_s\Big(\mathbf{\Delta u}\Big)\mathcal{P}_s\Big(\mathbf{v}\Big). \tag{D13}$$

If we denote by $(\mathbf{e}_s)_{s\in\mathcal{S}}$ the canonical basis of $\mathbb{R}^S$, the canonical projection can be expressed as a scalar product, that is

$$\mathcal{P}_s\Big(\mathbf{u}\Big) = (\mathbf{e}_s, \mathbf{u}). \tag{D14}$$

Let us recall briefly the Guerra's RSB scheme. Let U_i be a family of centered Gaussian random variables uniquely determined by the covariance matrix $\mathbb{E}(U_iU_j) = C_{ij}$ and let us introduce the integer K, associated to the number of levels of Replica Symmetry Breaking (RSB in the following). For each couple $(l, i) \in \{1, 2, ..., K\} \times I$, let us introduce further the family of centered Gaussian random variables B_i^l independent from the U_i and uniquely defined through the covariances

$$\mathbb{E}(B_i^l B_j^{l'}) = \delta_{ll'}\widetilde{B}_{ij}^l,$$

and point out that there is independence between different l, l' levels of symmetry breaking.
Further, we need some preliminary definitions:

For the average with respect to B_i^l and U_i we use the following notation

$$\mathbb{E}_l(\cdot) = \int \prod_i d\mu(B_i^l)(\cdot), \qquad \forall l = 1, ..., K, \tag{D15}$$

$$\mathbb{E}_0(\cdot) = \int \prod_i d\mu(U_i)(\cdot), \tag{D16}$$

$$\mathbb{E}(\cdot) = \mathbb{E}_0\mathbb{E}_1...\mathbb{E}_K(\cdot). \tag{D17}$$

We need also a sequence of non negative real numbers $(m_0, m_1, ..., m_K, m_{K+1})$ with $m_0 = 0,\ m_{K+1} = 1$ and we define recursively the following the random variables

$$Z_K(t) := \sum_i \omega_i \exp\left(\sqrt{t}U_i + \sqrt{1-t}\sum_{l=1}^K B_i^l\right), \tag{D18}$$

$$Z_{l-1}^{m_l} := \mathbb{E}_l(Z_l^{m_l}), \tag{D19}$$

$$f_l := \frac{Z_l^{m_l}}{\mathbb{E}_l(Z_l^{m_l})}, \tag{D20}$$

and the following modified Gibbs states,

$$\widetilde{\omega}_{K,t}(\cdot) := \omega_t(\cdot) \qquad \widetilde{\Omega}_{K,t} = \Omega_t(\cdot), \tag{D21}$$

$$\widetilde{\omega}_{l,t}(\cdot) := \mathbb{E}_{l+1}...\mathbb{E}_K(f_{l+1}...f_K\omega_t(\cdot)) \quad \forall l = 0, ..., K, \tag{D22}$$

$$\widetilde{\Omega}_{l,t}(\cdot) := \mathbb{E}_{l+1}...\mathbb{E}_K(f_{l+1}...f_K\Omega_t(\cdot)) \quad \forall l = 0, ..., K, \tag{D23}$$

$$\langle\cdot\rangle_{l,t} := \mathbb{E}(f_1...f_K\widetilde{\Omega}_{l,t}(\cdot)) \quad \forall l = 0, ..., K. \tag{D24}$$

Bearing in mind the previous definitions, it is possible to prove (see [79]) the following

Proposition 7.2.3. *Consider the function*

$$\varphi(t) = \mathbb{E}_0 \log(Z_0(t)), \tag{D25}$$

then for its t-derivative the following relation holds

$$\varphi'(t) = \frac{1}{2}\langle C_{ii} - \widehat{B}_{ii}^K\rangle_{K,t} - \frac{1}{2}\sum_{l=0}^K (m_{l+1} - m_l)\langle C_{ij} - \widehat{B}_{ij}^l\rangle_{l,t} \tag{D26}$$

where $\widehat{B}_{ij}^0 = 0$ *and* $\widehat{B}_{ij}^l = \sum_{l'=1}^l \widetilde{B}_{ij}^{l'}$.

We discuss now a generalization of the previous scheme for multipartite systems.

Let K be an integer and consider an arbitrary sequence of points $\Gamma := (\mathbf{q}_l)_{l=1,...,K} \in [0,1]^S$. For each triple (l,i,s) with $l = 1,2,...,K$, $i \in I$, $s \in \mathcal{S}$, let us introduce the family of centered Gaussian random variables $B_i^{l,(s)}$ independent from the U_i and uniquely defined through the covariances

$$\mathbb{E}(B_i^{l,(s)} B_j^{l',(s')}) = \delta_{ss'}\delta_{ll'}\mathcal{P}_s\Big(\mathbf{\Delta}\mathbf{u}_l(\Gamma)\Big)\mathcal{P}_s\Big(\widetilde{\mathbf{C}}_l\Big) \tag{D27}$$

where, for each value of l, the component of the vector $\widetilde{\mathbf{C}}_l = (\widetilde{C}_{l,ij}^{(s)})_{s\in\mathcal{S}}$, are covariance matrix defined on $I^{(s)} \times I^{(s)}$ and $\mathbf{u}_l(\Gamma)$ is an arbitrary vector in $\mathbb{R}^S$ which depends on the choice of the sequence Γ .

Notice that (D27) implies independence between two different $l^{(s)}, l'^{(s)}$ levels of symmetry breaking of each s-species. For each $l = 1,2,...,K$ and $i \in I$, we can define the following family of random variables

$$B_i^l := \sum_{s\in\mathcal{S}} B_i^{l,(s)}$$

then by (D13) we have that

$$\mathbb{E}(B_i^l B_j^{l'}) = \delta_{ll'}\Big(\mathbf{u}_l(\Gamma), \mathbf{\Delta}\widetilde{\mathbf{C}}_l\Big). \tag{D28}$$

Suppose for simplicity that $\widetilde{C}_{l,ii}^{(s)} = 1$ for each l,i,s, that is

$$\mathbb{E}(B_i^l B_i^{l'}) = \delta_{ll'}\Big(\mathbf{u}_l(\Gamma), \mathbf{\Delta}\mathbf{1}\Big).$$

Let us introduce the following notations for the average with respect to B_i^l, U_i,

$$\mathbb{E}_l(\cdot) = \int \prod_i d\mu(B_i^l)(\cdot) \qquad \forall l = 1,...,K, \tag{D29}$$

$$d\mu(B_i^l) = \prod_{s\in\mathcal{A}} d\mu(B_i^{l,(s)}), \tag{D30}$$

$$\mathbb{E}_0(\cdot) = \int \prod_i d\mu(U_i)(\cdot), \tag{D31}$$

$$\mathbb{E}(\cdot) = \mathbb{E}_0\mathbb{E}_1...\mathbb{E}_K(\cdot). \tag{D32}$$

Hence, the multi-species analogous of the Proposition 7.2.3, is the following

Proposition 7.2.4. *Consider the functional*

$$\varphi(t) = \mathbb{E}_0 \log(Z_0(t)), \tag{D33}$$

then for its t-derivative the following relation holds

$$\varphi'(t) = \frac{1}{2}(\mathbf{1}, \mathbf{\Delta 1}) - \sum_{l=1}^{K} \Big(\mathbf{u}_l(\Gamma), \mathbf{\Delta 1}\Big) - \frac{1}{2}\sum_{l=0}^{K}(m_{l+1} - m_l)\langle\Big(\mathbf{C}, \mathbf{\Delta C}\Big) - \widehat{B}^l\rangle_{l,t} \tag{D34}$$

where $\widehat{B}^0 = 0$ *and* $\widehat{B}^l = \sum_{l'=1}^{l} \Big(\mathbf{u}_l(\Gamma), \mathbf{\Delta}\widetilde{\mathbf{C}}_{l'}\Big)$.

E. Properties of the function g

We study the main properties of the function g defined by (6.14), which are often used in the following. Remind

$$g(\xi) \;=\; \frac{1}{2}\,(\sqrt{e^{4h} + 4\,e^{2h}} - e^{2h}) \quad \forall\, h \in \mathbb{R}\;.$$

Standard computations show that g is analytic on $\mathbb{R}$, $0 < g < 1$, $\lim_{h\to-\infty} g(\xi) = 0$, $\lim_{h\to\infty} g(\xi) = 1$, g is strictly increasing, g is strictly convex on $]-\infty, \frac{\log(2\sqrt{2}-2)}{2}]$ and strictly concave on $[\frac{\log(2\sqrt{2}-2)}{2}, \infty[$, $g(\frac{\log(2\sqrt{2}-2)}{2}) = 2 - \sqrt{2}$.

Solving in h the equation $g(\xi) = k$ for any fixed $k \in]0, 1[$, one finds the inverse function:

$$g^{-1}(k) \;=\; \frac{1}{2}\log\frac{k^2}{1-k} \quad \forall k \in]0, 1[\;. \tag{E1}$$

It is useful to write the derivatives of g in terms of lower order derivatives of g itself. For the first derivative, think g as $(g^{-1})^{-1}$ and exploit (E1):

$$g'(h) \;=\; \frac{1}{(g^{-1})'(k)}\Big|_{k=g(\xi)} \;=\; \frac{2\,k\,(1-k)}{2-k}\Big|_{k=g(\xi)} \;=\; \frac{2\,g(\xi)\,(1-g(\xi))}{2-g(\xi)} \tag{E2}$$

Then for the second derivative, differentiate the rhs of (E2) and substitute (E2) itself in the expression:

$$g'' = \frac{2\,g'}{2-g}\big(1 - 2\,g + \frac{g\,(1-g)}{2-g}\big) \;=\; \frac{2\,g'\,(1-2\,g) + (g')^2}{2-g}\;. \tag{E3}$$

The same for the third derivative: differentiate the rhs of (E3) and substitute (E3) itself in the expression:

$$\begin{aligned} g''' &= \frac{1}{2-g}\left(2\,g''(1-2\,g+g') - 4\,(g')^2 + g'\,\frac{2\,g'\,(1-2\,g)+(g')^2}{2-g}\right) = \\ &= \frac{g''\,(2-4\,g+3\,g') - 4\,(g')^2}{2-g}\,. \end{aligned} \tag{E4}$$

Lemma E1. *For* $c > 6-4\sqrt{2}$,

$$g'(\xi) < c \quad \forall\, \xi \in \mathbb{R}\,.$$

For $0 < c \le 6-4\sqrt{2}$,

$$g'(\xi) \begin{cases} < c & \textit{iff } \ \xi < \frac{1}{2}\log\alpha_-(c) \ \textit{ or } \ \xi > \frac{1}{2}\log\alpha_+(c) \\ > c & \textit{iff } \ \frac{1}{2}\log\alpha_-(c) < \xi < \frac{1}{2}\log\alpha_+(c) \end{cases},$$

where

$$\alpha_\pm(c) := \frac{-(c^2+8c-4) \,\pm\, (2-c)\sqrt{c^2-12c+4}}{4\,c}\,.$$

Proof. Investigate for example the inequality $g'(\xi) < c$. By (E2) clearly $0 < g' < 2$, hence the inequality is trivially true for $c \ge 2$ and false for $c \le 0$.

Using identity (E2) one finds

$$g' < c \quad \Leftrightarrow \quad 2\,g^2 - (2+c)\,g + 2c > 0\,;$$

this is a second degree inequality in g with $\Delta = c^2 - 12c + 4$.

If $6-4\sqrt{2} < c < 6+4\sqrt{2}$, it is verified for any value of g.

If instead $c \le 6-4\sqrt{2}$ or $c \ge 6+4\sqrt{2}$, it is verified if and only if

$$g(\xi) < \frac{2+c-\sqrt{c^2-12c+4}}{4} =: s_-(c) \quad \text{or} \quad g(\xi) > \frac{2+c+\sqrt{c^2-12c+4}}{4} =: s_+(c)\,.$$

For $0 < c < 2$, $s_\pm(c) \in \,]0,1[$ hence one can apply g^{-1}, which is strictly increasing:

$$\xi < g^{-1}(s_-(c)) \quad \text{or} \quad \xi > g^{-1}(s_+(c))\,.$$

This concludes the proof because identity (E1) and standard computations show that

$$g^{-1}(s_\pm(c)) \,=\, \frac{1}{2}\log\alpha_\pm(c)\,. \qquad \square$$

F. Critical exponents: technical proofs

Let us prove the results used in subsection 6.2.3 to compute the critical exponents.

Lemma F1. *Consider the inflection points ϕ_1, ϕ_2 of $\widetilde{p}$ defined by (6.25). Their behaviour at the critical point (h_c, J_c) along any curve $\delta \in C^1([J_c, \infty[)$, with $\delta(J_c) = h_c$, is*

$$\frac{\phi_1(\delta(J), J) - m_c}{\sqrt{J - Jc}} \xrightarrow[J \to J_c+]{} -C\,, \quad \frac{\phi_2(\delta(J), J) - m_c}{\sqrt{J - Jc}} \xrightarrow[J \to J_c+]{} C\,.$$

where $C = \sqrt[4]{2}/(2J_c) > 0$.

Proof. For $i = 1, 2$ and $J \geq J_c$ definition (6.25), observing that $(2m_c - 1)J = -h_c + (2m_c - 1)(J - J_c) + \xi_c$, gives

$$2J\left(\phi_i(\delta(J), J) - m_c\right) = \frac{1}{2}\log a_i(J) - \xi_c - (\delta(J) - h_c) - (2m_c - 1)(J - J_c)\,.$$

Now the definition (6.26) may be rewritten as

$$a_i(J) = \underbrace{(2J - 2 - \frac{1}{8J})}_{=: b(J)} \mp \underbrace{4\,(\frac{1}{2} - \frac{1}{8J})\sqrt{J - \tfrac{3-2\sqrt{2}}{4}}}_{=: c(J)} \sqrt{J - Jc}\,.$$

Thus, exploiting $\log(x + y) = \log x + \log(1 + y/x) = \log x + y/x + \mathcal{O}((y/x)^2)$ as $y/x \to 0$, $\frac{1}{2}\log b(J_c) = \xi_c$ and $\log b(J)$ differentiable at $J = J_c$,

$$\begin{aligned}\frac{1}{2}\log a_i(J) - \xi_c &= \frac{1}{2}\,\frac{\log b(J) - \log b(J_c)}{(J - J_c)}\,(J - J_c) \mp \frac{1}{2}\,\frac{c(J)}{b(J)}\,\sqrt{J - Jc} + \mathcal{O}(J - J_c) \\ &= \mp\frac{1}{2}\,\frac{c(J)}{b(J)}\,\sqrt{J - J_c} + \mathcal{O}(J - J_c)\,.\end{aligned}$$

To conclude put things together and use also δ differentiable at J_c:

$$\begin{aligned}2J\,\frac{\phi_i(\delta(J), J) - m_c}{\sqrt{J - Jc}} &= \frac{\frac{1}{2}\log a_i(J) - \xi_c}{\sqrt{J - J_c}} - \frac{\delta(J) - h_c}{\sqrt{J - J_c}} - (2m_c - 1)\sqrt{J - J_c} \\ &= \mp\frac{1}{2}\,\frac{c(J)}{b(J)} + \mathcal{O}(\sqrt{J - J_c}\,) \xrightarrow[J \to J_c+]{} \pm\sqrt[4]{2}\,.\end{aligned}$$

□

Next corollary gives a first bound for the critical exponents.

Corollary F2. *Here for $h \in \mathbb{R}$, $J > J_c$ let $m = m(h, J)$ be any solution of the consistency equation (6.22).*

1) There exist $r_1 > 0$, $C_1 < \infty$ such that for all $(h, J) \in B\big((h_c, J_c), r_1\big)$ with $J > J_c$

$$|m - m_c| \,\leq\, C_1\big(\,|h - h_c|^{\frac{1}{3}} + |J - J_c|^{\frac{1}{3}}\big)\,.$$

2) Assume that m pointwise coincides with one of the local maximum points m_1, m_2 (see proposition 6.2.2). There exist $r_2 > 0$, $C_2 > 0$ such that for all $(h, J) \in B\big((h_c, J_c), r_2\big)$ with $J > J_c$ and $h = \delta(J)$ for some $\delta \in C^1([J_c, \infty[)$, $\delta(J_c) = h_c$

$$|m - m_c| \,\geq\, C_2\,|J - J_c|^{\frac{1}{2}}\,.$$

Proof. 1) Set $\xi := (2m - 1)J + h$. By proposition 6.2.12, ξ satisfies equation (6.45), which can be treated as a third degree algebraic equation in $\xi - \xi_c$:

$$(\xi - \xi_c)^3 \underbrace{-\kappa_1\,(J - J_c)}_{=:\,p}\,(\xi - \xi_c)\,\underbrace{-\kappa_2\,\rho(h, J) + \mathcal{O}\big((\xi - \xi_c)^4\big)}_{=:\,q} = 0\,.$$

Analyse the real solutions of this equation. Set $\Delta := (\frac{q}{2})^2 + (\frac{p}{3})^3$ and observe that $(\frac{q}{2})^2 > 0$ while $(\frac{p}{3})^3 < 0$ as we are assuming $J > J_c$.

i. If $\Delta > 0$, the only real solution of (6.45) is

$$\xi - \xi_c = u_+ + u_- \quad \text{with } u_\pm = \sqrt[3]{-\frac{q}{2} \pm \sqrt[2]{\Delta}}\,.$$

On the other hand

$$\Delta > 0 \;\Rightarrow\; \big(\frac{p}{3}\big)^3 = \mathcal{O}\big(\big(\frac{q}{2}\big)^2\big) \;\Rightarrow\; \Delta = \mathcal{O}\big(\big(\frac{q}{2}\big)^2\big)\,.$$

Therefore, reminding also definition (6.46),

$$\xi - \xi_c \,=\, \mathcal{O}\big(\big(\frac{q}{2}\big)^{\frac{1}{3}}\big) \,=\, \mathcal{O}\big((h - h_c)^{\frac{1}{3}}\big) + \mathcal{O}\big((J - J_c)^{\frac{1}{3}}\big) + \mathcal{O}\big((\xi - \xi_c)^{\frac{4}{3}}\big)\,,$$

hence $\xi - \xi_c \,=\, \mathcal{O}\big((h - h_c)^{\frac{1}{3}}\big) + \mathcal{O}\big((J - J_c)^{\frac{1}{3}}\big)$ because $(\xi - \xi_c)^{\frac{4}{3}-1} \,\to\, 0$ as $(h, J) \to (h_c, J_c)$.

ii. If $\Delta = 0$ or $\Delta < 0$ there are respectively two or three distinct real solutions of (6.45) and, from their explicit form, it is immediate to see that they all satisfy

$$\xi - \xi_c \,=\, \mathcal{O}\big(\sqrt[2]{-\frac{p}{3}}\,\big) \,=\, \mathcal{O}\big((J-J_c)^{\frac{1}{2}}\big)\,.$$

Conclude that for any possible value of Δ,

$$\xi - \xi_c = \mathcal{O}\big((h-h_c)^{\frac{1}{3}}\big) + \mathcal{O}\big((J-J_c)^{\frac{1}{3}}\big)\,.$$

Now, as observed in (6.48), $\xi - \xi_c = h - h_c + (2m_c - 1)(J - J_c) + 2J\,(m - m_c)$. Therefore also $m - m_c = \mathcal{O}\big((h-h_c)^{\frac{1}{3}}\big) + \mathcal{O}\big((J-J_c)^{\frac{1}{3}}\big)$, and this concludes the proof of the first statement.

2) Now consider the two maximum points m_1, m_2. By proposition 6.2.2

$$m_1 < \phi_1 < \phi_2 < m_2$$

where ϕ_1, ϕ_2 are the inflection points defined by (6.25). Hence applying lemma F1 one finds:

$$\frac{m_2 - m_c}{\sqrt{J-J_c}} > \frac{\phi_2 - m_c}{\sqrt{J-J_c}} \longrightarrow C\,, \quad \frac{m_c - m_1}{\sqrt{J-J_c}} > \frac{m_c - \phi_1}{\sqrt{J-J_c}} \longrightarrow C\,,$$

as $J \to J_c+$ and $h = \delta(J)$ with $\delta(J_c) = h_c$ and δ differentiable in J_c. And this proves the second statement. □

The next lemma tells us in which region of the plane (h, J) a curve passing through the point (h_c, J_c) lies.

Lemma F3. *Let $\delta \in C^2([J_c, \infty[)$ such that $\delta(J_c) = h_c$, $\delta'(J_c) =: \alpha$. There exists $r > 0$ such that for all $J \in\,]J_c, J_c + r[$*

- *if $\alpha = 1 - 2m_c$, $\psi_2(J) < \delta(J) < \psi_1(J)$;*
- *if $\alpha < 1 - 2m_c$, $\delta(J) < \psi_2(J)$;*
- *if $\alpha > 1 - 2m_c$, $\delta(J) > \psi_1(J)$.*

Proof. I. Observe that $a_i(J)$ is continuous for $J \geq J_c$ and smooth for $J > J_c$. Moreover $g'(\frac{1}{2}\log a_i(J)) = \frac{1}{2J}$ by definition (6.26) and lemma E1, and $g(\frac{1}{2}\log a_i(J_c)) = g(\xi_c) = m_c$ by definition (6.30) and remark 6.2.3. Then differentiating definition (6.27) at $J > J_c$,

$$\psi_i'(J) \;=\; 1 - 2\,g(\frac{1}{2}\log a_i(J)) + \frac{1}{2}\,\frac{a_i'(J)}{a_i(J)}\,\big(\underbrace{1 - 2J\,g'(\frac{1}{2}\log a_i(J))}_{=0}\big) \;\xrightarrow[J\to J_c]{}\; 1 - 2m_c\,.$$

Hence an immediate application of the mean value theorem shows that for $i = 1, 2$ there exits $\psi_i'(J_c) = 1 - 2m_c$.

II. Differentiating definition (6.26) at $J > J_c$ shows that $a_1'(J) \to -\infty$, $a_2'(J) \to +\infty$ as $J \to J_c+$, while $a_i(J) \to 2\sqrt{2} - 2$ as $J \to J_c$. Hence

$$\psi_i''(J) \;=\; -\,g'(\frac{1}{2}\log a_i(J))\,\frac{a_i'(J)}{a_i(J)} \;=\; -\frac{1}{2J}\,\frac{a_i'(J)}{a_i(J)} \;\xrightarrow[J\to J_c+]{}\; \begin{cases} +\infty & \text{for } i = 1 \\ -\infty & \text{for } i = 2 \end{cases}.$$

The result is provided comparing the first order Taylor expansions at J_c with Lagrange remainder of ψ_1, ψ_2 and δ. □

The following proposition essentially contain the proof of part *ii)* of theorem 6.2.11.

Proposition F4. *Let $(h, J) \to (h_c, J_c)$ along a curve $h = \delta(J)$ with $\delta \in C^2(\mathbb{R}_+)$, $\delta(J_c) = h_c$, $\delta'(J_c) =: \alpha$ or along a curve $J = \delta(h)$ with $\delta \in C^2(\mathbb{R})$, $\delta(h_c) = J_c$, $\delta'(h_c) = 0$, then*

$$\mu_1(h, J) - m_c \;\sim\; \begin{cases} -\,C\,(J - J_c)^{\frac{1}{2}} & \text{if } h = \delta(J),\ \alpha = 1 - 2m_c \text{ and } J > J_c \\ C_\alpha\,(J - J_c)^{\frac{1}{3}} & \text{if } h = \delta(J),\ \alpha < 1 - 2m_c \\ C_\infty\,(h - h_c)^{\frac{1}{3}} & \text{if } J = \delta(h) \end{cases}$$

$$\mu_2(h, J) - m_c \;\sim\; \begin{cases} C\,(J - J_c)^{\frac{1}{2}} & \text{if } h = \delta(J),\ \alpha = 1 - 2m_c \text{ and } J > J_c \\ C_\alpha\,(J - J_c)^{\frac{1}{3}} & \text{if } h = \delta(J),\ \alpha > 1 - 2m_c \\ C_\infty\,(h - h_c)^{\frac{1}{3}} & \text{if } J = \delta(h) \end{cases}$$

where $C = \frac{1}{2J_c}\sqrt{3(2-m_c)}$, $C_\alpha = \frac{1}{2J_c}\sqrt[3]{\frac{3}{2}J_c(2-m_c)(2m_c-1+\alpha)}$, $C_\infty = \frac{1}{2J_c}\sqrt[3]{3J_c(2-m_c)}$. *To complete the cases, along the line* $h = h_c+(1-2m_c)(J-J_c)$, *when* $J \leq J_c$

$$\mu_1(h,J) = \mu_2(h,J) = m_c.$$

Proof. Fix (h,J) on the curve given by the graph of δ and in the rest of the proof denote by m a solution of the consistency equation (6.22), i.e. $m = g\big((2m-1)J+h\big)$. Furthermore when necessary m is assumed to be a local maximum point of $\widetilde{p}$. Set $\xi := (2m-1)J+h$. By proposition 6.2.12, $\xi-\xi_c \to 0$ as $(h,J) \to (h_c,J_c)$ and it satisfies (6.45). Solve this equation in the different cases.

i) Suppose $h = \delta(J)$ with $\alpha = 1-2m_c$. Hence $h-h_c = (1-2m_c)(J-J_c) + \mathcal{O}\big((J-J_c)^2\big)$. Observe that by (6.46), (6.48)

$$\rho(h,J) = \mathcal{O}\big((J-J_c)^2\big) \quad \text{and} \quad \xi-\xi_c = 2J\,(m-m_c) + \mathcal{O}\big((J-J_c)^2\big)\,.$$

Hence equation (6.45) becomes

$$(\xi-\xi_c)^3 - \kappa_1\,(J-J_c)\,(\xi-\xi_c) + \mathcal{O}\big((J-J_c)^2\big) + \mathcal{O}\big((\xi-\xi_c)^4\big) = 0\,.$$

Observe that if $J > J_c$ by corollary F2 part *2)*,

$$(J-J_c)^{\frac{1}{2}} = \mathcal{O}(\xi-\xi_c)\,;$$

therefore when $J > J_c$ the previous equation rewrites

$$(\xi-\xi_c)^3 - \kappa_1\,(J-J_c)\,(\xi-\xi_c) + \mathcal{O}\big((\xi-\xi_c)^4\big) = 0\,.$$

This one simplifies in

$$\xi = \xi_c \quad \text{or} \quad (\xi-\xi_c)^2 - \kappa_1\,(J-J_c) + \mathcal{O}\big((\xi-\xi_c)^3\big) = 0\,,$$

giving $\xi = \xi_c$ or, as we are assuming $J > J_c$,

$$\xi-\xi_c = \pm\sqrt{\kappa_1}\,(J-J_c)^{\frac{1}{2}} + \mathcal{O}\big((\xi-\xi_c)^{\frac{3}{2}}\big)\,.$$

This entails

$$m - m_c \,=\, \pm \frac{\sqrt{\kappa_1}}{2J}\,(J-J_c)^{\frac{1}{2}} + \mathcal{O}\big((J-J_c)^2\big) + \mathcal{O}\big((m-m_c)^{\frac{3}{2}}\big)$$

and dividing both sides by $m - m_c$, since $(m-m_c)^{\frac{1}{2}} \to 0$, one finds

$$m - m_c \,\sim\, \pm \frac{\sqrt{\kappa_1}}{2J}\,(J-J_c)^{\frac{1}{2}}\,. \tag{F1}$$

ii) Suppose $J = \delta(h)$ with $\delta'(h_c) = 0$. Hence $J - J_c = \mathcal{O}\big((h-h_c)^2\big)$. (6.46) and (6.48) give

$$\rho(h,J) = h - h_c + \mathcal{O}\big((h-h_c)^2\big) \quad \text{and} \quad \xi - \xi_c = 2J\,(m-m_c) + h - h_c + \mathcal{O}\big((h-h_c)^2\big)\,.$$

Hence equation (6.45) becomes

$$(\xi-\xi_c)^3 - \kappa_2\,(h-h_c) + \mathcal{O}\big((h-h_c)^2\big) + \mathcal{O}\big((\xi-\xi_c)^4\big) = 0\,.$$

giving

$$\xi - \xi_c \,=\, \sqrt[3]{\kappa_2}\,(h-h_c)^{\frac{1}{3}} + \mathcal{O}\big((h-h_c)^{\frac{2}{3}}\big) + \mathcal{O}\big((\xi-\xi_c)^{\frac{4}{3}}\big)\,.$$

This entails

$$m - m_c \,=\, \frac{\sqrt[3]{\kappa_2}}{2J}\,(h-h_c)^{\frac{1}{3}} + \mathcal{O}\big((h-h_c)^{\frac{2}{3}}\big) + \mathcal{O}\big((m-m_c)^{\frac{4}{3}}\big)$$

and dividing both sides by $m - m_c$, since $(m-m_c)^{\frac{1}{3}} \to 0$, one finds

$$m - m_c \,\sim\, \frac{\sqrt[3]{\kappa_2}}{2J}\,(h-h_c)^{\frac{1}{3}}\,. \tag{F2}$$

iii) Suppose $h = \delta(J)$ with $\alpha \neq 1 - 2m_c$. Hence $h - h_c = \alpha\,(J-J_c) + \mathcal{O}\big((J-J_c)^2\big)$. Observe that by (6.46), (6.48)

$$\rho(h,J) = (\alpha + 2m_c - 1)(J-J_c) + \mathcal{O}\big((J-J_c)^2\big)\,,$$
$$\xi - \xi_c = 2J\,(m-m_c) + (\alpha + 2m_c - 1)(J-J_c) + \mathcal{O}\big((J-J_c)^2\big)\,.$$

Hence equation (6.45) becomes

$$(\xi-\xi_c)^3 \underbrace{-\kappa_1\,(J-J_c)}_{=:\,p}\,(\xi-\xi_c) \underbrace{-\kappa_2\,(\alpha+2m_c-1)\,(J-J_c) + \mathcal{O}\big((J-J_c)^2\big) + \mathcal{O}\big((\xi-\xi_c)^4\big)}_{=:\,q} = 0\,.$$

This third order equation has $\Delta := (\frac{q}{2})^2 + (\frac{p}{3})^3 > 0$ for $|J - J_c|$ small enough, indeed if $J < J_c$ then $p > 0$, while if $J > J_c$ then by corollary F2 part *1)* $(\xi - \xi_c)^4 = \mathcal{O}\big((J - J_c)^{\frac{4}{3}}\big) = o(J - J_c)$ hence

$$q = -\kappa_2 \, (\alpha + 2m_c - 1) \, (J - J_c) + o\big(J - J_c\big) \quad \Rightarrow$$
$$\big(\frac{q}{2}\big)^2 + \big(\frac{p}{3}\big)^3 \; = \; \frac{\kappa_2^2}{4} \, \underbrace{(\alpha + 2m_c - 1)}_{\neq 0}{}^2 \, (J - J_c)^2 \, (1 + o(1)) - \frac{\kappa_1^3}{27} \, (J - J_c)^3 \; > \; 0 \; .$$

Then, using Cardano's formula for cubic equations: $\xi - \xi_c = u_+ + u_-$ with

$$u_\pm \; = \; \sqrt[3]{-\frac{q}{2} \pm \sqrt[2]{\big(\frac{q}{2}\big)^2 + \big(\frac{p}{3}\big)^3}} \; = \; \sqrt[3]{-\frac{q}{2} \pm \big|\frac{q}{2}\big|} \; + \mathcal{O}\big(\big|\frac{p}{3}\big|^{\frac{1}{2}}\big) \; ;$$

hence

$$\begin{aligned} \xi - \xi_c \; &= \; \sqrt[3]{-q} \; + \mathcal{O}\big(\big|\frac{p}{3}\big|^{\frac{1}{2}}\big) \; = \\ &= \; \sqrt[3]{\kappa_2 \, (\alpha + 2m_c - 1)} \, (J - J_c)^{\frac{1}{3}} + \mathcal{O}\big((J - J_c)^{\frac{2}{3}}\big) + \mathcal{O}\big((\xi - \xi_c)^{\frac{4}{3}}\big) + \mathcal{O}\big((J - J_c)^{\frac{1}{2}}\big) \; . \end{aligned}$$

This entails

$$m - m_c \; = \; \frac{\sqrt[3]{\kappa_2 \, (\alpha + 2m_c - 1)}}{2J} \, (J - J_c)^{\frac{1}{3}} + \mathcal{O}\big((J - J_c)^{\frac{1}{2}}\big) + \mathcal{O}\big((m - m_c)^{\frac{4}{3}}\big)$$

and dividing both sides by $m - m_c$, since $(m - m_c)^{\frac{1}{2}} \to 0$, one finds

$$m - m_c \; \sim \; \frac{\sqrt[3]{\kappa_2 \, (\alpha + 2m_c - 1)}}{2J} \, (J - J_c)^{\frac{1}{3}} \; . \tag{F3}$$

Now by propositions 6.2.2, 6.2.4 and lemma F3, μ_1 and μ_2 are solutions of the consistency equation (6.22) defined near (h_c, J_c) along the curves $h = \delta(J)$ respectively with $\alpha \leq 1 - 2m_c$ and $\alpha \geq 1 - 2m_c$. Moreover for $\alpha = 1 - 2m_c$ and $J > J_c$ sufficiently small, by lemma F1,

$$\mu_2 - m_c > \phi_2 - m_c > 0 \quad \text{while} \quad \mu_1 - m_c < \phi_1 - m_c < 0 \; .$$

These facts together with (F1), (F2), (F3) allow to conclude the proof. □

Bibliography

Statistical Mechanics

[1] Gallavotti, G.: *Statistical mechanics.* Texts and Monographs in Physics. Springer-Verlag, 1999.

[2] Georgii, H.-O.: *Gibbs Measures and Phase Transitions.* vol. 9 of de Gruyter Studies in Mathematics. Berlin: Walter de Gruyter & Co., 1988.

[3] Ellis, R. S.: *Entropy, Large Deviations, and Statistical Mechanics.* vol. 271 of Fundamental Principles of Mathematical Sciences. New York: Springer-Verlag, 1985.

[4] Landau, L. D., Lifshitz, E. M.: *Statistical physics* Pergamon Press, London, 1958.

[5] Ruelle, D.: *Statistical mechanics: Rigorous results.* World Scientific, 1999.

[6] Simon, B.: *The statistical mechanics of lattice gases. Vol. I.* Princeton Series in Physics. Princeton, NJ: Princeton University Press, 1993.

Probability Theory

[7] Feller, W.: *An introduction to probability theory and its applications*, vol. I and II, J. Wiley ed., New York (1971).

[8] Ledoux, M.: *The concentration of measure phenomenon*, Mathematical Surveys and Monographs **89**, Am. Math. Soc., Providence, RI (2001).

[9] Ledoux, M., Talagrand, M.: *Probability in Banach spaces. Isoperimetry and processes*, Springer, Berlin (1991).

[10] Kallenberg, O.: *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York (1997).

Spin Glass Theory

[11] Amit, D.J.: *Modeling brain function* , Cambridge University Press, 1989.

[12] Bovier, A.: *Statistical Mechanics of Disordered Systems .A Mathematical Perspective*, Cambridge University Press, (2006)

[13] Contucci, P., Giardina, C.: *Perspective on Spin Glasses* , Cambridge University Press, 2012.

[14] Fischer, K.H. Hertz, J.H. *Spin glasses* Cambridge University Press, Cambridge 1991.

[15] Mezard, M., Parisi, G., Virasoro, M. A.: *Spin glass theory and beyond*, World Scientific, Singapore, 1987.

[16] Panchenko, D.: *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer-Verlag, (2013).

[17] Talagrand, M.: *Spin glasses: A challenge for mathematicians.* Springer, Berlin (2003).

[18] Talagrand, M.: *Mean Field Models for Spin Glasses Volume I: Basic Examples.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge A Series of Modern Surveys in Mathematics, Volume 54. Springer-Verlag, 2011.

[19] Talagrand, M.: *Mean Field Models for Spin Glasses Volume II: Advanced Replica-Symmetry and Low Temperature.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge A Series of Modern Surveys in Mathematics, Volume 55. Springer-Verlag, 2011.

Specific references

[20] Abramowitz M., Stegun I., *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (U.S. Government Printing Office, Washington, 1964) pp.771-791

[21] Aharony, A.: Tricritical points in systems with random fields. *Physical Review B* **18**(7), 3318-3327 (1978)

[22] Aizenman, M., Contucci, P.: On the Stability of the Quenched state in Mean Field Spin Glass Models. *Journal of Statistical Physics* **92**(5/6), 765-783, 1998.

[23] Aizenman M., Lebowitz, J.L., Ruelle, D. : Some rigorous results on the Sherrington-Kirkpatrick model. *Comm. Math. Phys.* **112**, 320, 1987.

[24] Aizenman, M., Sims, R., Starr, S.L.: An extended variational principle for the SK spin-glass model. *Phys. Rev. B* **68**, 214403, (2003)

[25] Aizenman, M., Wehr, J.: Rounding effects of quenched randomness on first-order phase transitions. *Comm. Math. Phys.*, **130**(3), 489528, 1990.

[26] Aizenman, M. , Arguin, L. P.: On the structure of quasi-stationary competing particle systems. *Ann. Probab.* **37**, no. 3, 1080-1113, (2009).

[27] Alberici, D., Contucci, P.: Solution of the monomer-dimer model on locally tree-like graphs. Rigorous results, *Communications in Mathematical Physics* **331**, 975-1003 (2014)

[28] Alberici, D., Contucci, D., Mingione, E.: A mean-field monomer-dimer model with attractive interaction. Exact solution and rigorous results, *Journal of Mathematical Physics* **55**, 063301 (2014)

[29] Alberici, D., Contucci, D., Mingione, E.: The exact solution of a mean-field monomer-dimer model with attractive potential, *Europhysics Letters* **106**, 10001-10005 (2014)

[30] Alberici, D., Contucci, D., Mingione, E.: A mean-field monomer-dimer model with random monomer activities. Exact solution and rigorous results. Preprint arXiv:1409.6192.

[31] Anderson, P.W., Edwards, S.F.: Theory of spin glasses, *J. Phys. F* **5**, 965974 1975.

[32] Arguin, L.P., Chatterjee, S.: Random Overlap Structures: Properties and Applications to Spin Glasses. *Probab. Theory Relat. Fields* DOI 10.1007/s00440-012-0431-6

[33] Arguin, L.P., Damron, M.: Short-range spin glasses and Random Overlap Structures. *J. Statist. Phys.* **143**, no. 2, 226–250 (2011).

[34] Arguin, L.P., Newman, C. M., Stein, D. L., Wehr, J.: Fluctuation Bounds For Interface Free Energies in Spin Glasses *Journal of Statistical Physics*, **156**(2), 221-238, 2014.

[35] Auffinger, A., Chen, W.: The Parisi Formula has a Unique Minimizer. *Communications in Mathematical Physics*, DOI 10.1007/s00220-014-2254-z, (2014) Date: 27 Nov 2014

[36] Baffioni, F. Rosati, F.: On the Ultrametric Overlap Distribution for Mean Field Spin Glass Models (I). *The European Physical Journal B* **17**, 439-447, (2000).

[37] Barahona, F.: On the computational complexity of Ising spin glass models. *Journal of Physics A*, **15**, 3241, 1982.

[38] Barra, A. *Irreducible free energy expansion and overlap locking in mean field spin glasses*, J. Stat. Phys. **123**, 3, 601-614, 2006.

[39] Barra, A., Bernacchia, A., Santucci, E., Contucci, P.: On the equivalence of Hopfield Networks and Boltzmann Machines. *Neural Networks* **34**, 1-9, 2012.

[40] Barra, A., P. Contucci, P.: Toward a quantitative appproach to migrant's integration. *Eur. Phys. Lett.* **89**, 6, 68001, 2010.

[41] Barra, A., Contucci, P., Mingione, E., Tantari, D.: Multi-species mean-field Spin Glasses. Rigorous results. *Annales Henri Poincaré*, DOI: 10.1007/s00023-014-0341-5, 2014.

[42] Barra, A., Contucci, P., Sandell, R., Vernia, C.: An analysis of a large dataset on immigrant integration in Spain. The Statistical Mechanics perspective on Social Action. *Sci. Rep.* **4**, doi:10.1038/srep04174 (2014)

[43] Barra, A., Genovese, G., Guerra, F.: The replica symmetric behavior of the analogical neural network. *J. Stat. Phys.* **140**, 784-796, 2010.

[44] Barra, A., Genovese, G., Guerra, F.: Equilibrium statistical mechanics of bipartite spin systems. *J.Phys. A* **44**, 24, 245002, 2011.

[45] Barra, A., Genovese, G., Guerra, F., Tantari, D.: How glassy are neural networks?, *J. Stat. Mech.,* 07009, 2012.

[46] Billoire, A., *et al.*: Comment on "Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model". http://arxiv.org/abs/1211.0843

[47] Bolthausen, E., Kistler, N.: On a nonhierarchical version of the Generalized Random Energy Model, II: Ultrametricity. *Stochastic Processes and their Applications* **119**, Issue 7, 23572386, (2009).

[48] Bolthausen, E., Sznitman, A.-S.: On Ruelles probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**(2), 247276, (1998).

[49] Bouchaud, J. P., Potters, M.: *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, Cambridge University Press, 2004.

[50] Bordenave, C., Lelarge, M., Salez, J.: Matchings on infinite graphs. *Probab. Theory Relat. Fields* **157**, 183-208 (2013)

[51] Bovier, A., Kurkova, I.: Derrida's generalized random energy models. 1. Models with finitely many hierarchies". *Ann. Inst. H. Poincare. Prob. et Statistiques (B) Prob. Stat.* **40**, 439-480 (2004).
Bovier A. and Kurkova I., Derrida's generalized random energy models. 2. Models with continuous hierarchies". *Ann. Inst. H. Poincare. Prob. et Statistiques (B) Prob. Stat.* **40**, 481-495 (2004).

[52] Chang, T.S.: Statistical theory of the adsorption of double molecules. *Proc. Roy. Soc. London A* **169**, 512-531 (1939)

[53] Chang, T.S.: The number of configurations in an assembly and cooperative phenomena. *Proc. Cambridge Phil. Soc.* **38**, 256-292 (1939)

[54] Chang, T.S.: Statistical theory of the adsorption of double molecules. *Proceedings of the Royal Society of London A* **169**, 512-531 (1939)

[55] Chung, F., Lu, L.: Concentration inequalities and martingale inequalities - a survey. *Internet Mathematics* **3**(1) 79-127 (2006)

[56] Cohen, E.G.D., Kincaid, J.M.: Phase diagrams of liquid helium mixtures and metamagnets: experiment and mean field theory, *Physics Letters C*, **22**, 58-142, 1975.

[57] Contucci, P., Fedele, M.: Scaling Limits for Multispecies Statistical Mechanics Mean-Field Models *Journal of Statistical Physics,* **144**, N. 6, 1186-1205, 2011.

[58] Contucci, P., I. Gallo, I. Bipartite mean field spin systems. Existence and solution *Math. Phys. E. J.* **14**, 463, 2008.

[59] Contucci, P., Giardina, G., Giberti, C., Vernia, C.: Overlap Equivalence in the Edwards-Anderson Model. *Physical Review Letters*, **96**, 217204, 2006.

[60] Contucci, P., Giardina, C.: The Ghirlanda-Guerra identities. *J. Statist. Phys.* **126**, no. 4, 917-931, (2007).

[61] Contucci, P., Giardina, C., Giberti., C.: Stability of the Spin Glass Phase under Perturbations. *Europhysics Letters* **96**, no. 1, 17003-17006, (2011).

[62] Contucci, P., Giardina, C., Giberti., C.: Overlap Equivalence in the Edwards-Anderson Model. *Physical Review Letters*, **96**, 217204 (2006).

[63] Contucci, P., Mingione, E., Starr.,S.L.: Factorization properties in d-dimensional spin glasses. Rigorous results and some perspectives. *J. Stat. Phys.* **151**, Issue 5, pages 809-829, 2013.

[64] Coolen, A.C.C.: The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents, Oxford University Press, 2007.

[65] Cugliandolo, L.F., Kurchan, J.: The out-of-equilibrium dynamics of the Sherrington-Kirkpatrick model. *Journal of Physics A: Mathematical and Theoretical*, **41**, 324018, 2008.

[66] De Bruijn, N.G: *Asymptotic methods in Analysis* 2nd ed., pp.63-65, North-Holland, 1961

[67] Dembo, A., Montanari, A.: Ising models on locally tree-like graphs. *Ann. Appl. Probab.*,**20**, 565-592, 2010.

[68] Derrida, B.: Random energy model: An exactly solvable model of disordered systems, *Phys. Rev. B* **24**, 2613 (1981).

[69] Derrida, B. Gardner, E.: Solution of the generalized random energy model, *J. Phys. C* **19**, 2253 (1986).

[70] Fedele, M., Unguendoli, F. Rigorous results on the bipartite mean-field model *Jou. Phys. A, Math. and Theo.*, **45**, 3850-3860, 2012.

[71] see, for instance:

Fisher, D. S.D., Huse, D. A.: em Ordered Phase of Short-Range Ising Spin-Glasses. *Phys. Rev. Lett.* **56**, 1601 (1986);

Fisher, D. S.D., Huse, D. A.: em Equilibrium behavior of the spin-glass ordered phase. *Phys. Rev. B* **38**, 386 (1988).

[72] Fowler, R.H.: Adsorption isotherms. Critical conditions. *Math. Proc. Cambridge Phil. Soc.* **32**, 144-151 (1936)

[73] Fowler, R.H., Rushbrooke, G.S.: An attempt to extend the statistical theory of perfect solutions. *Trans. Faraday Soc.* **33**, 1272-1294 (1937)

[74] Franz, S., Toninelli, F. L.: *The Kac limit for finite-range spin glasses*, Phys. Rev. Lett. **92** 030602 (2004).

[75] Fröhlich, J., Zegarlinski, B.: Some comments on the Sherrington- Kirkpatrick model of spin glasses. *Comm. Math. Phys.* 112, 553566, 1987.

[76] Ghirlanda, S., Guerra, F.: General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A: Math. Gen.* **31**, 9149-9155, 1998.

[77] Giardinà, S., Starr., S.: Variational bounds for the Generalized Random Energy Model. *J. Statist. Phys.* **127**, no. 1, 1-20, (2007).

[78] Gibbons, A.: *Algorithmic graph theory*, pp.125-152, Cambridge University Press, 1985

[79] Guerra, F.: Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model. *Comm. Math. Phys.*, **233**, 1-12, 2003.

[80] Guerra, F.: Mathematical aspects of mean field spin glass theory. *4th European Congress of Mathematics, Stockholm June 27-July 2 2004*, Laptev A. editor, European Mathematical Society (2005)

[81] Guerra, F.: Sum rules for the free energy in the mean field spin glass model, in *Mathematical Physics in Mathematics and Physics: Quantum and Operator Algebraic Aspects*, Fields Institute Communications **30**, American Mathematical Society (2001).

[82] Guerra, F.: *Proceeding of the conference "Equilibrium and Dynamics of Spin Glasses"*, Ascona, Switzerland, A.Bovier, E. Bolthausen Ed.s (2004).

[83] Guerra, F.: The replica symmetric region in the Sherrington-Kirkpatrick mean field spin glass model: The Almeida-Thouless line. arXiv:cond-mat/0604674

[84] Guerra, F., Toninelli, F.L.: The Thermodynamic Limit in Mean Field Spin Glass Models, *Commun. Math. Phys.* **230:1**, 71-79 (2002).

[85] Guerra, F., Toninelli, F.L.: The high temperature region of the Viana-Bray diluted spin glass model, *J. Stat. Phys.* **115** (2004).

[86] Heilmann, O.J., Lieb, E.H.: Theory of monomer-dimer systems. *Communications in Mathematical Physics* **25**, 190-232 (1972)

[87] Heilmann, O.J., Lieb, E.H.: Monomers and dimers. *Physical Review Letters* **24**(25), 1412-1414 (1970)

[88] Heilmann, O.J., Lieb, E.H.: Lattice models for liquid crystals. *J. Stat. Phys.* **20**, 680-693 (1979)

[89] Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **12**, 134-139 (1918)

[90] Jennrich, J.R.: Asymptotic properties of non-linear least squares estimators. *The Annals of Mathematical Statistics* **40**(2), 633-643 (1969)

[91] Karp, J., Sipser, M.: Maximimum matchings in sparse random graphs, *Proceedings of the 22nd Annual Symposium on Foundations of Computer Science*, IEEE Computer Society Press, 364-375 (1981)

[92] Kasteleyn, P. W.: The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice. *Physica* **27** (12): 12091225 (1961)

[93] Kirkpatrick, S., Sherrington, D.:"Solvable model of spin-glass", *Physics Review Letters* **35**(26), 1792-1796 (1975)

[94] Krzakala, F., Martin, O.C.: Spin and Link Overlaps in Three-Dimensional Spin Glasses. *Phys. Rev. Lett.* **85**, 3013, (2000)

[95] see, for instance:

Marinari, E., Parisi,G., Ruiz-Lorenzo, J. J., Ritort, F.: Numerical Evidence for Spontaneously Broken Replica Symmetry in 3D Spin Glasses, *Phys. Rev. Lett.* **76**, 843 (1996).

[96] Mezard, M., Parisi, G.: The Bethe lattice spin glass revisited. *Eur. Phys. J. B Condens. Matter Phys.* **20**, no. 2, 217233, (2001)

[97] see, for instance:

Newman, C. M., Stein, D. L.: *Simplicity of State and Overlap Structure in Finite-Volume Realistic Spin Glasses*, Phys. Rev. E, **57**, 1356 (1998);

Newman, C. M., Stein, D. L.: *The Metastate approach to thermodynamic chaos*, Phys. Rev. E **55**, 5194 (1997).

[98] Newman, C. M., Stein, D. L.: Short-range spin glasses: a status report. In *XIIth International Congress of Mathematical Physics (ICMP 97) (Brisbane)*, pp. 167172. Cambridge, MA: International Press, 1999.

[99] Newman, C. M., Stein, D. L.: Thermodynamic chaos and the structure of short-range spin glasses. In *Mathematical Aspects of Spin Glasses and Neural Networks*, vol. **156** of Progr. Probab., 243287, 1998.

[100] Newman, C. M., Stein, D. L.: *The State(s) of Replica Symmetry Breaking: Mean Field Theories vs. Short-Ranged Spin Glasses*, J. Stat. Phys. **106**, 213-244, (2004).

[101] Nishimori, H.: *Statistical physics of spin glasses and information process ing: an introduction* Oxford University Press, 2001.

[102] Ok E. A.: *Real analysis with economics applications* (Princeton University Press, Princeton, 2007) pp.306-311

[103] Palassini, M., Young, A.P.: Nature of the Spin Glass State. *Phys, Rev. Lett.* **85**, 3017, 2000.

[104] Panchenko, D.: The Ghirlanda-Guerra identities for the mixed p-spin model. *Comptes Rendus Acad. Sci.* **348**, no. 34, 189192, (2010).

[105] Panchenko, D.: The Sherrington-Kirkpatrick model: an overview. em J. Statist. Phys. **149**, no. 2, pp 362-383, (2012).

[106] Panchenko, D.: *Introduction to the SK model*, Preprint arXiv:1412.0170.

[107] Panchenko, D.: Structure of finite-RSB asymptotic Gibbs measures in the diluted spin glass models, Preprint arXiv:1406.4702, (2014)

[108] Panchenko, D.: The Parisi ultrametricity conjecture. *Annals of Mathematics* **177**, Issue 1, 383393, 2013.

[109] Panchenko, D.: The free energy in a multi-species Sherrington-Kirkpatrick model. Preprint on arXiv: 1310.6679, 2013.

[110] Parisi, G.: The order parameter for spin glasses: a function on the interval $0-1$. *J. Phys. A* **13**, 1101, (1980)

[111] Parisi, G., Talagrand, M.: On the distribution of the overlaps at given disorder. *Comptes Rendus Math.* **339**, no. 4, 303-306, (2004).

[112] Peierls, R.: Statistical theory of adsorption with interaction between the adsorbed atoms. *Math. Proc. Cambridge Phil. Soc.* **32**, 471-476 (1936)

[113] Roberts, J.K.: Some properties of mobile and immobile adsorbed films. *Proc. Cambridge Phil. Soc.* **34**, 399-411 (1938)

[114] Roberts, J.K., Miller, R.: The application of statistical methods to immobile adsorbed films. *Proc. Cambridge Phil. Soc.* **35**, 293-297 (1939)

[115] Rudin, W.: *Principles of Mathematical Analysis*, McGraw-Hill International Editions, United States, 3rd ed., 1976, pp.223-228

[116] Ruelle, D.:A mathematical reformulation of Derrida's REM and GREM. *Commun. Math. Phys.* **108**, 225 (1987).

[117] Salinas, S.R., Wreszinski, W.F.: On the mean-field Ising model in a random external field. *Journal of Statistical Physics* **41**(l/2), 299-313 (1985)

[118] Schneider, T., Pytte, E.: Random-field instability of the ferromagnetic state. *Physical Review B* **15**(3), 1519-1522 (1977)

[119] Talagrand, M. The Parisi formula. *Annals of Mathematics*, **163**, N. 1, 221-263, 2006.

[120] Temperley, H. N. V., Fisher, M. E.:. Dimer problem in statistical mechanics-an exact result. *Philosophical Magazine* **6** (68): 10611063. (1961)

[121] Toulouse, G.: Theory of Frustration Effect in Spin Glasses. *Commun. Phys.* **2**, 115, (1977)

[122] Wijn H.P.J. ed., *Magnetics Properties of Metals.: D-Elements, Alloys and Compounds* Springer, Berlin, 1991.

[123] Wormald, N.C.: The differential equation method for random processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, pp. 73-155, M. Karonski and H.J. Proemel editors, PWN, Warsaw, 1999.

[124] Vladimirov, I.G.: The monomer-dimer model and Lyapunov exponents of homogeneous gaussian random fields. *Discrete and Continuous Dynamical Systems B* **18**(2), 575-600 (2013)

[125] Yucesoy, B., Katzgraber, H.G., Machta, J.: Evidence of Non-Mean-Field-Like Low-Temperature Behavior in the Edwards-Anderson Spin-Glass Model. *Physical Review Letters* **109**, (17), pp177204. , (2012)

[126] Zdeborová, L., Mézard, M.: The number of matchings in random graphs, *J. Stat. Mech.* **5**, P05003 (2006)