Electronic Journal of Probability

Vol. 14 (2009), Paper no. 8, pages 161–241.

Journal URL
http://www.math.washington.edu/~ejpecp/

The Aizenman-Sims-Starr and Guerra's schemes for the SK model with multidimensional spins

Anton Bovier
Institut für Angewandte Mathematik
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstrasse 6
53115 Bonn
Germany
e-mail: bovier@uni-bonn.de

Anton Klimovsky*
Department Mathematik
Friedrich-Alexander-Universität Erlangen-Nürnberg
Bismarckstrasse $1\frac{1}{2}$
91054 Erlangen
Germany
e-mail: klimovsk@math.tu-berlin.de

Abstract

We prove upper and lower bounds on the free energy of the Sherrington-Kirkpatrick model with multidimensional spins in terms of variational inequalities. The bounds are based on a multidimensional extension of the Parisi functional. We generalise and unify the comparison scheme of Aizenman, Sims and Starr and the one of Guerra involving the GREM-inspired processes and Ruelle's probability cascades. For this purpose, an abstract quenched large deviations principle

*Supported in part by the DFG Research Training Group "Stochastic Processes and Probabilistic Analysis" and by the Helmholz-Gemeinschaft.

of the Gärtner-Ellis type is obtained. We derive Talagrand's representation of Guerra's remainder term for the Sherrington-Kirkpatrick model with multidimensional spins. The derivation is based on well-known properties of Ruelle's probability cascades and the Bolthausen-Sznitman coalescent. We study the properties of the multidimensional Parisi functional by establishing a link with a certain class of semi-linear partial differential equations. We embed the problem of strict convexity of the Parisi functional in a more general setting and prove the convexity in some particular cases which shed some light on the original convexity problem of Talagrand. Finally, we prove the Parisi formula for the local free energy in the case of multidimensional Gaussian a priori distribution of spins using Talagrand's methodology of a priori estimates.

Key words: Sherrington-Kirkpatrick model, multidimensional spins, quenched large deviations, concentration of measure, Gaussian spins, convexity, Parisi functional, Parisi formula.

AMS 2000 Subject Classification: Primary 60K35, 82B44, 60F10.

Submitted to EJP on February 23, 2008, final version accepted December 18, 2008.

Contents

1 Introduction

The Sherrington-Kirkpatrick (SK) model of a mean-field spin-glass has long been one of the most enigmatic models of statistical mechanics. The recent rigorous proof of the celebrated *Parisi formula* for its free energy, due to Talagrand [30], based on the ingenious interpolation schemes of Guerra [17] and Aizenman, Sims, and Starr [1] constitutes one of the major recent achievements of probability theory. Recently, these results have been generalised to spherical SK-models [29] and to models with spins taking values in a bounded subset of $\mathbb{R}$ [21].

In this paper, we are mainly concerned with the question of the validity of the Parisi formula in the case where spins take values in a d-dimensional Riemannian manifold. We address the issue of extending the approach of Aizenman, Sims and Starr, and the one of Guerra to the multidimensional spins. We study the properties of the multidimensional Parisi functional. Motivated by a problem posed by [31], we show the strict convexity of the local Parisi functional in some cases.

We partially extend Talagrand's methodology of estimating the remainder term to the multidimensional setting. In the case of the multidimensional Gaussian a priori distribution of spins we prove the validity of the Parisi formula in the low temperature regime.

Definition of the model

Let $\Sigma \subset \mathbb{R}^d$ and denote $\Sigma_N \equiv \Sigma^N$. We define a family of Gaussian processes $X \equiv \{X(\sigma)\}_{\sigma \in \Sigma_N}$ as follows

$$X(\sigma) = X_N(\sigma) \equiv \frac{1}{N} \sum_{i,j=1}^{N} g_{i,j} \langle \sigma_i, \sigma_j \rangle, \tag{1.1}$$

where the *interaction matrix* $G \equiv \{g_{i,j}\}_{i,j=1}^N$ consists of i.i.d. standard normal random variables and, for $x, y \in \mathbb{R}^d$, $\langle x, y \rangle \equiv \sum_{u=1}^d x_u y_u$ is the standard Euclidean scalar product. In what follows all random variables and processes are assumed to be centred. We shall call $H_N(\sigma) \equiv -\sqrt{N} X_N(\sigma)$ a *random Hamiltonian* of our model.

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

We shall be interested mainly in the *free energy*

$$p_N(\beta) \equiv \frac{1}{N} \log \int_{\Sigma_N} \exp\left(\beta\sqrt{N} X(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma), \tag{1.2}$$

where $\beta \geq 0$ is the *inverse temperature* and $\mu \in \mathscr{M}_{\mathrm{f}}(\Sigma)$ is some arbitrary (not necessarily uniform or discrete) finite *a priori measure*. We assume that the a priori measure μ is such that (1.2) is finite. We shall be interested in proving bounds on the thermodynamic limits of these quantities, e.g., on

$$p(\beta) \equiv \lim_{N\uparrow+\infty} p_N(\beta). \tag{1.3}$$

Remark 1.1. *Note that there is no need to include the additional external field terms into the Hamiltonian* (1.1)*, since they could be absorbed into the a priori measure μ.*

Mean-field spin-glass models (see, e.g., [7]) with multidimensional *(Heisenberg)* spins were considered in the theoretical physics literature, see, e.g., [25] and references therein. Rigorous results are, however, rather scarce. An early example is [15], where the authors get bounds on the free energy in the high temperature regime. Methods of stochastic analysis and large deviations are used in [34] to identify the limiting distribution of the partition function and also to obtain some information about the geometry of the Gibbs measure for small β. More recent treatments of the high temperature regime using the very different methods are due to Talagrand [27], see also [28, Subsection 2.13]. The importance of the SK model with multidimensional spins for understanding the ultrametricity of the original SK model [26] (which corresponds to $d = 1$ and μ being the Rademacher measure in the above notations) was emphasised in [33].

For the SK model, Guerra's scheme gave historically the first way to obtain the variational upper bound on the free energy in terms of the Parisi functional. The scheme is based on the comparison between two Gaussian processes: the first one being the original SK Hamiltonian (1.1) and the other one being a carefully chosen GREM inspired process indexed by $\sigma \in \Sigma_N$. The second important ingredient is a recursively defined non-linear comparison functional acting on the Gaussian processes indexed by $\sigma \in \Sigma_N$.

The Aizenman-Sims-Starr (AS^2) scheme [1; 2] gives an intrinsic way to obtain variational upper bounds on the free energy in the SK model. The scheme is also based on a comparison between two Gaussian processes. The first process is the sum of the original SK Hamiltonian X and a GREM-inspired process indexed by additional index space $\mathscr{A} \equiv \mathbb{N}^n$. The second one is another GREM-inspired process indexed by the extended configuration space $\Sigma_N \times \mathscr{A}$. The scheme uses a comparison functional defined on Gaussian processes indexed by the extended configuration space equipped with the product measure between the original a priori measure and Ruelle's probability cascade (RPC) [24]. The role of the comparison functional in the AS^2 scheme is played by a free energy functional acting on the Gaussian processes indexed by the extended configuration space. In [22] Panchenko and Talagrand have reexpressed Guerra's scheme for the SK model using the RPC.

Talagrand [30] using Guerra's scheme and the wealth of other ingenious analytical insights showed that the variational upper bound is also the lower bound for the free energy in the SK model. This established, hence, the remaining half of the Parisi formula.

A particular case ($d = 1$, μ with bounded support) of the model we are considering here was treated by Panchenko in [21]. He used the techniques of [30] to prove that in the case $d = 1$ upper and lower bounds on the free energy coincide (cf. (1.14) and (1.22) in this chapter). However, the results of [21, Section 5 and the proofs of Theorems 2, 5 and 9] are based on relatively detailed differential properties of the optimal Lagrange multipliers in the saddle point optimisation problem of interest. These properties are harder to obtain in multidimensional situations such as that we are dealing with here. In fact, as we show in Theorems 1.1 and 1.2, one can obtain the same saddle point variational principles without invoking the detailed properties of the optimal Lagrange multipliers. This is achieved using a quenched large deviations principle (LDP) of the GÃď'rtner-Ellis type.

The most advanced recent study of spin-glass models with multidimensional spins was attempted by Panchenko and Talagrand in [23], where the multidimensional spherical spin-glass model was considered. The authors combined the techniques of [30; 21] to obtain partial results on the ultrametricity and also get some information on the local free energy for their model.

Main results

In this paper, we prove upper and lower bounds on the free energy in the SK model with multidimensional spins in terms of variational inequalities involving the corresponding multidimensional generalisation of the Parisi functional (Theorems 1.1, 1.2, 5.1, 5.18). For this purpose, we generalise and unify the AS^2 and Guerra's schemes for the case of multidimensional spins, and employ a quenched LDP which may be of independent interest (Theorems 3.1 and 3.2). Both schemes are formulated in a unifying framework based on the same comparison functional. The functional acts on Gaussian processes indexed by an extended configuration space as in the original AS^2 scheme. As a by-product, we provide also a short derivation of the remainder term in multidimensional Guerra's scheme (Theorem 5.4) using well-known properties of the RPC and the Bolthausen-Sznitman coalescent. This gives a clear meaning to the remainder in terms of averages with respect to a measure changed disorder. The change of measure is induced by a reweighting of the RPC using the exponentials of the GREM-inspired process[1]. See [22] for another approach in the case of the SK model ($d = 1$).

We study the properties of the multidimensional Parisi functional by establishing a link between the functional and a certain class of non-linear partial differential equations (PDEs), see Propositions 6.1, 6.2 and Theorem 6.2. We extend the Parisi functional to a continuous functional on a compact space (Theorems 6.1, 6.2). We show that the class of PDEs corresponds to the Hamilton-Jacobi-Bellman (HJB) equations induced by a linear problem of diffusion control (Proposition 6.4). Motivated by a problem posed by [31], we show the strict convexity of the local Parisi functional in some cases (Theorem 6.4).

We partially extend Talagrand's methodology of estimating the remainder term to the multidimensional setting (Theorem 5.4, Proposition 7.1, Theorem 7.1). In the case of multidimensional Gaussian a priori distribution of spins we prove the validity of the Parisi formula (Theorem 1.3).

We partially extend Talagrand's methodology of estimating the remainder term to the multidimensional setting (Theorem 5.4, Proposition 7.1, Theorem 7.1). Though the main technical problem of the methodology in the general multidimensional setting remains (Remark 7.5). In the case of the

[1]In $d = 1$ the latter fact was also known to the author of [3], private communication.

multidimensional Gaussian a priori distribution of spins we prove the validity of the Parisi formula (Theorem 1.3).

Below we introduce the notations, assumptions and formulate our main results. The other results (mentioned above) are formulated and proved in the subsequent sections.

Assumption 1.1. *Suppose that the configuration space* Σ *is bounded and such that* $0 \in \operatorname{int}\operatorname{conv}\Sigma$, *where* $\operatorname{conv}\Sigma$ *denotes the convex hull of* Σ.

The examples listed below verify this assumption:

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

Remark 1.2. *The boundedness assumption can be relaxed and replaced by concentration properties of the a priori measure. In Section 8 we will exemplify this in the case of a Gaussian a priori distribution. In general a subgaussian distribution will suffice.*

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

$$\operatorname{Sym}^+(d) \equiv \left\{\Lambda \in \operatorname{Sym}(d) \mid \Lambda \succeq 0\right\},$$

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

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

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

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

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

$$R_N(\sigma^{(1)}, \sigma^{(2)})_{u,v} \equiv \frac{1}{N}\sum_{i=1}^{N} \sigma_{i,u}^{(1)}\sigma_{i,v}^{(2)}, \quad u,v \in [1;d] \cap \mathbb{N}. \tag{1.4}$$

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

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

Next, define the *local free energy*

$$p_N(\mathscr{V}) \equiv \frac{1}{N} \log \int_{\Sigma_N(\mathscr{V})} \mathrm{e}^{\beta\sqrt{N}X(\sigma)} \mathrm{d}\mu^{\otimes N}(\sigma). \tag{1.5}$$

We also define

$$p(\mathscr{V}) \equiv p(\beta, \mathscr{V}) \equiv \lim_{N\uparrow+\infty} p_N(\mathscr{V}), \tag{1.6}$$

where the existence of the limit follows from a result of Guerra and Toninelli [19, Theorem 1]. Consider a sequence of matrices $\mathscr{Q} \equiv \{Q^{(k)} \in \mathrm{Sym}(d)\}_{k=0}^{n+1}$ such that

$$0 \equiv Q^{(0)} \prec Q^{(1)} \prec \ldots \prec Q^{(n+1)} \equiv U, \tag{1.7}$$

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

$$0 \equiv x_0 < x_1 < \ldots < x_{n+1} \equiv 1. \tag{1.8}$$

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

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

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

$$X_{n+1}(x, \mathscr{Q}, U, \Lambda) \equiv \log \int_{\Sigma} \exp\left(\sqrt{2}\beta \left\langle \sum_{k=0}^{n} z_k, \sigma \right\rangle + \langle \Lambda\sigma, \sigma\rangle\right) \mathrm{d}\mu(\sigma). \tag{1.9}$$

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

$$X_k(x, \mathscr{Q}, U, \Lambda) \equiv \frac{1}{x_k} \log \mathbb{E}_{z^{(k)}}\left[\exp\left(x_k X_{k+1}(x, \mathscr{Q}, U, \Lambda)\right)\right] \tag{1.10}$$

with

$$X_0(x, \mathscr{Q}, U, \Lambda) \equiv \mathbb{E}_{z^{(0)}}\left[X_1(x, \mathscr{Q}, U, \Lambda)\right], \tag{1.11}$$

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

Remark 1.3. *Section 5.4 contains the more general framework of dealing with the recursive quantities* (1.11) *which in particular brings to light the links with certain non-linear parabolic PDEs. For these PDEs the recursion* (1.2) *is closely related to an iterative application of the well-known Hopf-Cole transformation, see, e.g., [14].*

Define the *local Parisi functional* as

$$f(x, \mathscr{Q}, U, \Lambda) \equiv -\langle \Lambda, U\rangle - \frac{\beta^2}{2}\sum_{k=1}^{n} x_k \left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2\right) + X_0(x, \mathscr{Q}, U, \Lambda). \tag{1.12}$$

Assumption 1.2 (Hadamard squares). *We shall say that a sequence, $\{Q^{(i)}\}_{i=1}^{n}$, of matrices satisfies Assumption 1.2, if*

$$\left(Q^{(1)}\right)^{\odot 2} \prec \ldots \prec \left(Q^{(n)}\right)^{\odot 2} \prec \left(Q^{(n+1)}\right)^{\odot 2}. \tag{1.13}$$

Remark 1.4. *The above assumption on the matrix order parameters $\mathscr{Q}$ is necessary only to employ the AS^2 scheme. In contrast, Guerra's scheme (Theorems 5.1 and 5.18) does not require the above assumption.*

One may verify that the matrices q and ρ in [28, Theorems 2.13.1 and 2.13.2] correspond to the matrices $Q^{(1)}$ and $Q^{(2)}$ of this paper ($n=1$). (See also (1.25) below.) Furthermore, a straightforward application of the Cauchy-Schwarz inequality shows that the matrices q and ρ actually satisfy Assumption 1.2. We also note that in the simultaneous diagonalisation scenario in which the matrices in (1.7) are diagonalisable in the same orthogonal basis (see Sections 6.3 and 8.2) this assumption is also satisfied.

The first main result of the present paper uses the AS^2 scheme to establish the upper bound on the limiting free energy $p(\beta)$ in terms of the saddle point problem for the local Parisi functional (1.12).

Theorem 1.1. *For any closed set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \leq \sup_{U \in \mathscr{V} \cap \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{1.14}$$

where the infimum runs over all x satisfying (1.8)*, all $\mathscr{Q}$ satisfying both* (1.7) *and Assumption 1.2, and all $\Lambda \in \mathrm{Sym}(d)$.*

We were not able to prove in general that the r.h.s. of (1.14) gives also the lower bound to the thermodynamic free energy. See, however, Theorem 1.3 for a positive example.

To formulate the lower bound on (1.3) we need some additional definitions.

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

$$Q(\alpha^{(1)}, \alpha^{(2)}) \equiv Q^{(q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)}))}, \tag{1.15}$$

where $q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)})$ is the (normalised) *lexicographic overlap* defined as follows

$$q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)}) \equiv 1 + \begin{cases} 0, & \alpha_1^{(1)} \neq \alpha_1^{(2)} \\ \max\left\{k \in [1;n] \cap \mathbb{N} : [\alpha^{(1)}]_k = [\alpha^{(2)}]_k\right\}, & \text{otherwise.} \end{cases} \tag{1.16}$$

Given a $d \times d$-matrix M and $p \in \mathbb{R}$, we denote by $M^{\odot p}$ the $d \times d$-matrix with entries

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

The matrix valued lexicographic overlap (1.15) can be used to construct the multidimensional ($d \geq 1$) versions of the GREM (see, e.g., [8] and references therein for a review of the results on the one-dimensional case of the model). Here we shall need the following two GREM-inspired real-valued Gaussian processes: $A \equiv \{A(\sigma,\alpha)\}_{\sigma \in \Sigma_N, \alpha \in \mathscr{A}}$ and $B \equiv \{B(\alpha)\}_{\alpha \in \mathscr{A}}$ with covariance structures

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

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

$$A(\sigma,\alpha) = \left(\frac{2}{N}\right)^{1/2} \sum_{i=1}^{N} \langle A_i(\alpha), \sigma_i \rangle, \tag{1.17}$$

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

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

Given $t \in [0;1]$, we define the *interpolating AS² Hamiltonian*

$$H_t(\sigma,\alpha) \equiv \sqrt{t}\,(X(\sigma)+B(\alpha)) + \sqrt{1-t}A(\sigma,\alpha). \tag{1.18}$$

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

$$\pi_N \equiv \mu^{\otimes N} \otimes \xi,$$

where $\xi = \xi(x)$ is the RPC [24]. We denote by $\{\xi(\alpha)\}_{\alpha\in\mathscr{A}}$ the enumeration of the atom locations of the RPC and consider the enumeration as a random measure on $\mathscr{A}$ (independent of all other random variables around). Define the *local AS² Gibbs measure* $\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})$ by

$$\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\,[f] \equiv \frac{1}{Z_N(t,\mathscr{V})}\int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} f(\sigma,\alpha)\mathrm{e}^{\sqrt{N}\beta H_t(\sigma,\alpha)}\mathrm{d}\pi_N(\sigma,\alpha), \tag{1.19}$$

where $f : \Sigma_N \times \mathscr{A} \to \mathbb{R}$ is an arbitrary measurable function for which the right-hand side of (1.19) is finite. For $\mathscr{V} \subset \mathscr{U}$, define the *AS² remainder term* as

$$\begin{aligned}&\mathscr{R}_N(x,\mathscr{Q},U,\mathscr{V})\\ &\equiv -\frac{1}{2}\int_0^1 \mathbb{E}\left[\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\otimes\mathscr{G}_N(t,x,\mathscr{Q},U,\mathscr{V})\left[\|R_N(\sigma^{(1)},\sigma^{(2)})-Q(\alpha^{(1)},\alpha^{(2)})\|_{\mathrm{F}}^2\right]\right]\mathrm{d}t.\end{aligned} \tag{1.20}$$

We define also the *limiting AS² remainder term*

$$\mathscr{R}(x,\mathscr{Q},U) \equiv \lim_{\varepsilon\downarrow+0}\lim_{N\uparrow\infty}\mathscr{R}_N(x,\mathscr{Q},B(U,\varepsilon)) \leq 0, \tag{1.21}$$

where $B(U,\varepsilon)$ is the ball with centre U and radius ε. (The existence of the limiting remainder term is proved in Theorem 1.2.)

The second main result of this paper uses the AS² scheme to establish a lower bound on (1.3) in terms of the same saddle point Parisi-type functional as in the upper bound which includes, however, the non-positive remainder term (1.21). In one-dimensional situations Talagrand [30] and Panchenko [21], respectively, have shown that the corresponding error term vanishes on the optimiser of the Parisi functional.

Theorem 1.2. *For any open set* $\mathscr{V} \subset \mathrm{Sym}(d)$*, we have*

$$p(\mathscr{V}) \geq \sup_{U\in\mathscr{V}\cap\mathscr{U}}\inf_{(x,\mathscr{Q},\Lambda)}\left[f(x,\mathscr{Q},\Lambda,U)+\mathscr{R}(x,\mathscr{Q},U)\right], \tag{1.22}$$

where the infimum runs over all x satisfying (1.8)*, all* $\Lambda \in \mathrm{Sym}(d)$*, and all* $\mathscr{Q}$ *satisfying both* (1.7) *and Assumption 1.2.*

Remark 1.5. *The comparison scheme of Guerra [17] (see also more recent accounts [32], [18] and [2]) is also applicable to our model and is covered by our quenched LDP approach, see Theorems 5.1 and 5.18 for the formal statements. Guerra's scheme seems to be more amenable (compared to the Aizenman-Sims-Starr one) for Talagrand's remainder estimates [30], see Section 7. The scheme is based on the following interpolation*

$$\widetilde{H}_t(\sigma,\alpha) \equiv \sqrt{t}X(\sigma) + \sqrt{1-t}A(\sigma,\alpha) \tag{1.23}$$

which induces the corresponding local Gibbs measure (1.19) *and remainder term* (1.20) *by substituting* (1.18) *with* (1.23)*. Guerra's scheme does not include the process B and, hence, does not require Assumption 1.2. Recovering the terms corresponding to $\Phi_N(x,\mathscr{U})[B]$ (see,* (4.23)*) in the Parisi functional requires then a short additional calculation (Lemma 5.1).*

Note that the results of Talagrand [28, Theorems 2.13.2 and 2.13.3] imply that at least in the high temperature region (i.e., for small enough β) the Parisi formula for the SK model with multidimensional spins is valid with $n=1$

$$p(\beta) = f(x,\mathscr{Q}^*,0,U^*) = \sup_{U\in\mathscr{U}} \inf_{(\mathscr{Q},\Lambda)} f(x,\mathscr{Q},\Lambda,U), \tag{1.24}$$

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

$$\begin{cases} \partial_{Q^{(2)}_{u,v}} f(x,\mathscr{Q}^*,0,U^*) = 0, & u,v\in[1;d]\cap\mathbb{N}, \\ \partial_{Q^{(1)}_{u,v}} f(x,\mathscr{Q}^*,0,U^*) = 0, & u,v\in[1;d]\cap\mathbb{N}. \end{cases} \tag{1.25}$$

Note that the system (1.25) coincides with the mean-field equations obtained in [28, see (2.469) and (2.470)].

Let $\Sigma \equiv \mathbb{R}^d$ and fix some vector $h\in\mathbb{R}^d$. Let $\mu\in\mathscr{M}_{\mathrm{f}}(\Sigma)$ be the finite measure with the following density (with respect to the Lebesgue measure λ on Σ)

$$\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}(\sigma) = \left(\frac{\det C}{(2\pi)^d}\right)^{1/2} \exp\left(-\frac{1}{2}\langle C\sigma,\sigma\rangle + \langle h,\sigma\rangle\right), \tag{1.26}$$

where $C\in\mathrm{Sym}^+(d)$. Note that, given $m\in\mathbb{R}^d$ and $C\in\mathrm{Sym}^+(d)$ such that $\det C\neq 0$, the density (1.26) with $h\equiv Cm$ coincides (up to the constant factor $\exp\left(-\frac{1}{2}\langle Cm,m\rangle\right)$) with the Gaussian density with covariance matrix C^{-1} and mean m.

Remark 1.6. *It turns out that only matrices C with sufficiently large eigenvalues will result in finite global free energy, cf. Lemma 8.8. The local free energy is, in contrast, always finite, see Lemma 8.7 and Theorem 1.3.*

Consider the function $f:(0:+\infty)^2\to\mathbb{R}$ given by

$$f(c,u) = \begin{cases} \beta^2u^2 + \log cu - cu + 1, & u\in(0;\frac{\sqrt{2}}{2\beta}], \\ (2\sqrt{2}\beta - c)u + \log\frac{c}{\beta} - \frac{1}{2}\left(1+\log 2\right), & u\in(\frac{\sqrt{2}}{2\beta};+\infty]. \end{cases} \tag{1.27}$$

The following result shows that, at least, in the highly symmetric situation (1.26) with $h=0$ the multidimensional Parisi formula indeed holds true (see Lemma 8.7 for an explanation why the result is indeed a Parisi formula).

Theorem 1.3. *Let μ satisfy* (1.26) *with* $h = 0$*. Assume that the matrices U and C are simultaneously diagonalisable in the same basis. Denote by* $\{c_v \in \mathbb{R}_+\}_{v=1}^d$ *and* $\{u_v \in \mathbb{R}_+\}_{v=1}^d$ *the eigenvalues of the matrices C and U, respectively. Moreover, assume that* $\min_v u_v > 0$ *and* $\min_v c_v > 0$.

Then we have

$$\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow+\infty} p_N(\Sigma_N(B(U,\varepsilon))) = \sum_{v=1}^{d} f(c_v, u_v).$$

Remark 1.7. *Close results have previously been obtained in the case of the spherical model in [23], from where we borrow the general methodology of the proof of the Theorem 1.3. As noted in [23], another more straightforward way to obtain the Theorem 1.3 is to diagonalise the interaction matrix G and use the properties of the corresponding random matrix ensemble.*

Organisation of the paper

The rest of the present paper is organised as follows. In Section 2 we record some basic properties of the covariance structure of the process X and establish the relevant concentration of measure results. The section contains also the tools allowing to compare and interpolate between the free energy-like functionals of different Gaussian processes. In Section 3 we derive a quenched LDP of the GÃď'rtner-Ellis type under measure concentration assumptions. Section 4 contains the derivation (based on the AS^2 scheme) of the upper and lower bounds on the free energy of the SK model with multidimensional spins in terms of the saddle point of the Parisi-like functional. In Section 5 we employ the ideas of Guerra's comparison scheme in order to obtain the upper and lower bounds on the free energy and we also get a useful analytic representation of the remainder term. In Section 6 we study the properties of the multidimensional Parisi functional. Section 7 contains the partial extension of Talagrand's remainder term estimates to the case of multidimensional spins. In Section 8 a case of Gaussian a priori distribution of spins is considered and the corresponding local Parisi formula is proved. In the appendix we prove the almost super-additivity of the local free energy, as an application of the Gaussian comparison results of Subsection 2.3.

2 Some preliminary results

2.1 Covariance structure

Our definition of the overlap matrix in (1.4) is motivated by the fact that, as can be seen from a straightforward computation

$$\mathbb{E}\left[X_N(\sigma^{(1)})X_N(\sigma^{(2)})\right] = \sum_{u,v=1}^{d}\left(R_N(\sigma^{(1)},\sigma^{(2)})_{u,v}\right)^2 = \|R_N(\sigma^{(1)},\sigma^{(2)})\|_2^2, \tag{2.1}$$

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

Proposition 2.1. *We have, for all* $\sigma^{(1)},\sigma^{(2)},\sigma \in \Sigma_N$, *1.* Matrix representation. $R_N(\sigma^{(1)}, \sigma^{(2)}) = \frac{1}{N}\left(\sigma^{(1)}\right)^* \sigma^{(2)}$.

2. Symmetry #1. $R_N^{u,v}(\sigma^{(1)}, \sigma^{(2)}) = R_N^{v,u}(\sigma^{(2)}, \sigma^{(1)})$.

3. Symmetry #2. $R_N^{u,v}(\sigma, \sigma) = R_N^{v,u}(\sigma, \sigma)$.

4. Non-negative definiteness #1. $R_N(\sigma, \sigma) \succeq 0$.

5. Non-negative definiteness #2.

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

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

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

Proof. The proof is straightforward. □

2.2 Concentration of measure

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

Proposition 2.2. *Let $(\Sigma, \mathfrak{S})$ be a Polish space. Suppose μ is a random finite measure on Σ. Suppose, moreover, that $X(\sigma)$, $\sigma \in \Sigma$ is the family of Gaussian random variables independent of μ which possesses a bounded covariance, i.e.,*

$$\text{there exists } K > 0 \text{ such that } \sup_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma} |\operatorname{Cov}(X(\sigma^{(1)}), X(\sigma^{(2)}))| \leq K. \tag{2.2}$$

Assume that

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

Then

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

Remark 2.1. *An analogous result was given in a somewhat more specialised case in [21].*

Proof. This is an adaptation of the proof of [28, Theorem 2.2.4]. We can not apply the comparison Theorem 2.5 directly, so we resort to the basic interpolation argument as stated in Proposition 2.1. For $j = 1, 2$, let the processes $X_j(\cdot)$ be the two independent copies of the process $X(\cdot)$. For $t \in [0; 1]$, let

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

and

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

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

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

Hence, differentiation gives

$$\dot{\varphi}_s(t) = s\mathbb{E}\left[\exp\left(s(F_1 - F_2)\right)(\dot{F}_1 - \dot{F}_2)\right] \tag{2.3}$$

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

$$\begin{aligned}\dot{F}_j(t) =& \frac{1}{2}\left(\int_\Sigma \exp\left(X_{j,t}(\sigma)\right) \mathrm{d}\mu(\sigma)\right)^{-1} \\ &\times \int_\Sigma \left(t^{-1/2}X_j(\sigma) - (1-t)^{-1/2}X(\sigma)\right) \exp\left(X_{j,t}(\sigma)\right) \mathrm{d}\mu(\sigma).\end{aligned} \tag{2.4}$$

Now, we substitute (2.4) back to (2.3) and apply Corollary 2.1 to the result. After some tedious but elementary calculations we get

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

Thus, thanks to (2.2), we obtain

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

The conclusion of the theorem follows now exactly as in the proof of [28, Theorem 2.2.4].

□

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

Proposition 2.3. *Suppose $\Sigma \subset B(0,r)$, for $r > 0$. For $\Omega \subset \Sigma_N$, denote*

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

and

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

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

1. For any $t > 0$,

$$\mathbb{P}\left\{\left|P_N^{SK}(\beta,\Omega) - \mathbb{E}\left[P_N^{SK}(\beta,\Omega)\right]\right| > t\right\} \leq 2\exp\left(-\frac{t^2}{4\beta^2 r^4 N}\right). \tag{2.5}$$

2. For any $t > 0$,

$$\mathbb{P}\left\{\left|P_N^{GREM}(\beta,\Omega) - \mathbb{E}\left[P_N^{GREM}(\beta,\Omega)\right]\right| > t\right\} \leq 2\exp\left(-\frac{t^2}{8\beta^2 r^4 N}\right). \tag{2.6}$$

Proof. 1. We would like to use Proposition 2.2. By (2.1) and the Cauchy-Bouniakovsky-Schwarz inequality, we have, for all $N \in \mathbb{N}$, $\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N$, that

$$\mathrm{Cov}(X_N(\sigma^{(1)},\sigma^{(2)})) = \|R_N(\sigma^{(1)},\sigma^{(2)})\|_{\mathrm{F}}^2 = \frac{1}{N^2}\sum_{i,j=1}^{N}\langle\sigma_i^{(1)},\sigma_j^{(1)}\rangle\langle\sigma_i^{(2)},\sigma_j^{(2)}\rangle \leq r^4. \tag{2.7}$$

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

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

Thus (2.5) is proved.

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

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

Bound (2.7) implies that, for any $U \in \mathscr{U}$, we have $\|U\|_2 \leq r^2$. Since $Q(\alpha^{(1)},\alpha^{(2)}) \in \mathscr{U}$, we obtain

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

Therefore, using Proposition 2.2, we obtain (2.6).

□

2.3 Gaussian comparison inequalities for free energy-like functionals

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

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

$$\partial_{x\rightsquigarrow e}F(x) \equiv \partial_t F(x+te)\Big|_{t=0}. \tag{2.8}$$

With this notation the following lemma holds.

Lemma 2.1. *Let $\{g(i)\}_{i\in I}$ be a real-valued Gaussian process (the set I is an arbitrary index set), and h be some Gaussian random variable. Define the vector $e \in \mathbb{R}^I$ as $e(i) \equiv \mathbb{E}\left[hg(i)\right]$, $i \in I$. Let $F : \mathbb{R}^I \to \mathbb{R}$ such that, for all $f \in \mathbb{R}^I$, the function*

$$\mathbb{R} \ni t \mapsto F(f + te) \in \mathbb{R} \tag{2.9}$$

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

Then

$$\mathbb{E}[hF(g)] = \mathbb{E}\left[\partial_{g\rightsquigarrow e}F(g)\right]. \tag{2.10}$$

The previous proposition coincides with [21, Lemma 4] (modulo the differentiability condition on (2.9) and the integrability assumptions which are needed, e.g., for [5, Theorem 5.1.2]).

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

Proposition 2.4. *Consider a Polish measure space $(\Sigma, \mathfrak{S})$ and a random measure μ on it. Let $X = \{X(\sigma)\}_{\sigma\in\Sigma}$ and $Y \equiv \{Y(\sigma)\}_{\sigma\in\Sigma}$ be two independent Gaussian real-valued processes. For $u \in \mathbb{R}$, we define*

$$H_u(\sigma) \equiv uX(\sigma) + Y(\sigma).$$

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

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

almost surely, and also that

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

Then we have

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

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

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

Proof. We write

$$\frac{\mathrm{d}}{\mathrm{d}u} \log \int \mathrm{e}^{H_u(\sigma)} \mathrm{d}\mu(\sigma) = \int X(\sigma) \frac{\mathrm{e}^{H_u(\sigma)}}{Z_u(\beta)} \mathrm{d}\mu(\sigma), \tag{2.11}$$

where $Z_u(\beta) \equiv \int e^{\beta H_u(\sigma)} d\mu(\sigma)$. The main ingredient of the proof is the Gaussian integration by parts formula. Denote, for $\tau \in \Sigma$, $e(\tau) \equiv \mathbb{E}\left[X(\sigma)H_u(\tau)\right]$. By (2.10), we have

$$\mathbb{E}\left[X(\sigma)\frac{e^{H_u(\sigma)}}{Z_u(\beta)}\right] = \mathbb{E}\left[\partial_X\left(\frac{e^{H_u(\sigma)}}{\int e^{H_u(\tau)} d\mu(\tau)}\right)(X;e)\right]. \tag{2.12}$$

Due to the independence, we have

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

Henceforth, the computation of the directional derivative in (2.12) amounts to

$$\begin{aligned}
&\frac{\partial}{\partial t}\left[\frac{e^{H_u(\sigma)+tu\operatorname{Var}(\sigma)}}{\int e^{H_u(\tau)+tu\operatorname{Cov}(\sigma,\tau)} d\mu(\tau)}\right] \\
&\quad = \left(\int e^{H_u(\sigma)} d\mu(\sigma)\right)^{-2}\left(u\operatorname{Var}X(\sigma)e^{H_u(\sigma)}\int e^{H_u(\tau)} d\mu(\tau)\right. \\
&\qquad \left. - e^{H_u(\sigma)}\int u\operatorname{Cov}\left[X(\sigma), X(\tau)\right] e^{H_u(\tau)} d\mu(\tau)\right).
\end{aligned} \tag{2.13}$$

Substituting the r.h.s. of (2.13) into (2.11), we obtain the assertion of the proposition. □

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

Proposition 2.5. *Let $(X(\sigma))_{\sigma\in\Sigma}$, $(Y(\sigma))_{\sigma\in\Sigma}$ be two independent Gaussian processes as before. Set*

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

Assume that

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

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

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

Then we have

$$\begin{aligned}
&\mathbb{E}\left[\log\int e^{X(\sigma)} d\mu(\sigma)\right] = \mathbb{E}\left[\log\int e^{Y(\sigma)} d\mu(\sigma)\right] \\
&- \frac{1}{2}\int_0^1 \mathcal{G}(t)\otimes\mathcal{G}(t)\left[\left(\operatorname{Var}X(\sigma^{(1)}) - \operatorname{Var}Y(\sigma^{(1)})\right)\right. \\
&- \left.\left(\operatorname{Cov}\left[X(\sigma^{(1)}), X(\sigma^{(2)})\right] - \operatorname{Cov}\left[Y(\sigma^{(1)}), Y(\sigma^{(2)})\right]\right)\right] dt,
\end{aligned} \tag{2.14}$$

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

$$\mathscr{G}(t)\left[f\right]=\frac{1}{Z(t)}\int_\Sigma f(\sigma)\exp\left(H_t(\sigma)\right)\mathrm{d}\mu(\sigma). \tag{2.15}$$

Proof. Let us introduce the process

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

Hence,

$$H_t(\sigma)=W_{\sqrt{t},\sqrt{1-t}}(\sigma). \tag{2.16}$$

Thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\log\int \mathrm{e}^{H_t(\sigma)}\mathrm{d}\mu(\sigma)\right]=\frac{1}{2}\Bigg(&\frac{1}{\sqrt{t}}\frac{\partial}{\partial u}\mathbb{E}\left[\log\int \mathrm{e}^{W_{u,v}(\sigma)}\mathrm{d}\mu(\sigma)\right]\\&-\frac{1}{\sqrt{1-t}}\frac{\partial}{\partial v}\mathbb{E}\left[\log\int \mathrm{e}^{W_{u,v}(\sigma)}\mathrm{d}\mu(\sigma)\right]\Bigg)\Bigg|_{u=\sqrt{t},v=\sqrt{1-t}}.\end{aligned}$$

Applying Proposition 2.4 and $\int_0^1\cdot\mathrm{d}t$ to the previous formula, we conclude the proof. □

3 Quenched Gärtner-Ellis type LDP

In this section, we derive a quenched LDP under measure concentration assumptions. Theorems 3.1 and 3.2 give the corresponding LDP upper and lower bounds, respectively. The proofs of the LDP bounds will be adapted to get the proofs of the upper and lower bounds on the free energy of the SK model with multidimensional spins. However, they may be of independent interest.

Note that the existing "level-2" quenched large deviation results of Comets [10] are applicable only to a certain class of mean-field random Hamiltonians which are required to be "macroscopic" functionals of the joint empirical distribution of the random variables representing the disorder and the independent spin variables. The SK Hamiltonian can not be represented in such form, since the interaction matrix consists of i.i.d. random variables. Moreover, it is assumed in [10] that the Hamiltonian has the form $H_N(\sigma)=NV(\sigma)$, where $\{V(\sigma)\}_{\sigma\in\Sigma_N}$ is a random process taking values in some fixed bounded subset of $\mathbb{R}$. Since the Hamiltonian of our model is a Gaussian process, this assumption is also not satisfied, due to the unboundedness of the Gaussian distribution.

3.1 Quenched LDP upper bound

The following assumption will be satisfied for the applications we have in mind. As is clear from what follows, much weaker concentration functions are also allowed.

Assumption 3.1. *Suppose $\{Q_N\}_{N=1}^\infty$ is a sequence of random measures on a Polish space $(\mathscr{X},\mathfrak{X})$. Assume that there exists some $L>0$ such that for any Q_N-measurable set $A\subset\mathscr{X}$ we have*

$$\mathbb{P}\left\{\left|\log Q_N(A)-\mathbb{E}\left[\log Q_N(A)\right]\right|>t\right\}\leq\exp\left(-\frac{t^2}{LN}\right). \tag{3.1}$$

Note that Assumption 3.1 will hold in the cases we are interested in due to Proposition 2.2.

Lemma 3.1. *Suppose $\{Q_N\}_{N=1}^{\infty}$ is a sequence of random measures on a Polish space $(\mathscr{X}, \mathfrak{X})$ and for $\{A_r \subset \mathscr{X} : r \in \{1, \dots, p\}\}$ is a sequence of Q_N-measurable sets such that, for some absolute constant $L > 0$ and some concentration function $\eta_N(t) : \mathbb{R}_+ \to \mathbb{R}_+$ with the property*

$$\int_0^{+\infty} \eta_N(tN)\mathrm{d}t \xrightarrow[N\uparrow+\infty]{} 0, \tag{3.2}$$

we have

$$\mathbb{P}\left\{\left|\log Q_N(A_r) - \mathbb{E}\left[\log Q_N(A_r)\right]\right| > t\right\} \leq \eta_N(t). \tag{3.3}$$

Then we have

$$\lim_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\left|\log Q_N\left(\bigcup_{r=1}^{p} A_r\right) - \max_{r\in\{1,\dots,p\}} \mathbb{E}\left[\log Q_N(A_r)\right]\right|\right] = 0. \tag{3.4}$$

Remark 3.1. *As is easy to extract from Assumption 3.1, we will apply this result in the very pleasant situation, where*

$$\gamma_N(t) = \exp\left(-\frac{t^2}{LN}\right).$$

However, our subsequent results hold for substantially worse concentration functions satisfying (3.2).

Proof of Lemma 3.1. First, (3.3) gives

$$\mathbb{P}\left\{\max_{r\in\{1,\dots,p\}} \left|\log Q_N(A_r) - \mathbb{E}\left[\log Q_N(A_r)\right]\right| \geq t\right\} \leq p\eta_N(t).$$

Since, for $a, b \in \mathbb{R}^p$, the following elementary inequality holds

$$\left|\max_r a_r - \max_r b_r\right| \leq \max_r |a_r - b_r|,$$

we get

$$\mathbb{P}\left\{\left|\max_{r\in\{1,\dots,p\}} \log Q_N(A_r) - \max_{r\in\{1,\dots,p\}} \mathbb{E}\left[\log Q_N(A_r)\right]\right| \geq t\right\} \leq p\eta_N(tN).$$

The last equation in turn implies that

$$\frac{1}{N}\mathbb{E}\left[\left|\max_{r\in\{1,\dots,p\}} \log Q_N(A_r) - \max_{r\in\{1,\dots,p\}} \mathbb{E}\left[\log Q_N(A_r)\right]\right|\right] \leq p\int_0^{+\infty} \eta_N(tN)\mathrm{d}t, \tag{3.5}$$

and the r.h.s. of the previous formula vanishes as $N \uparrow +\infty$ due to (3.2). $\square$

Let $Q_N \in \mathscr{M}(\mathscr{X})$, $N \in \mathbb{N}$ be a family of random measures on $(\mathscr{X}, \mathfrak{X})$. Define the Laplace transform

$$L_N(\Lambda) \equiv \int_{\mathscr{X}} \mathrm{e}^{N\langle x,\Lambda\rangle} \mathrm{d}Q_N(x).$$

Suppose that, for all $\Lambda \in \mathbb{R}^d$, we have

$$I(\Lambda) \equiv \lim_{N\uparrow\infty} \frac{1}{N} \mathbb{E}\left[\log L_N(\Lambda)\right] \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}. \tag{3.6}$$

Define the Legendre transform

$$I^*(x) \equiv \inf_{\Lambda} \left[-\langle x, \Lambda\rangle + I(\Lambda)\right]. \tag{3.7}$$

Define, for $\delta > 0$,

$$I^*_\delta(x) \equiv \max\left\{I^*(x) + \delta, -\frac{1}{\delta}\right\}. \tag{3.8}$$

Lemma 3.2. *Suppose*

$$0 \in \operatorname{int} \mathscr{D}(I) \equiv \operatorname{int}\{\Lambda : I(\Lambda) < +\infty\}. \tag{3.9}$$

Then

1. *The mapping $I^*(\cdot) : \mathscr{X} \to \mathbb{R}$ is upper semi-continuous and concave.*
2. *For all $M > 0$,*

   $$\{x \in \mathscr{X} : I^*(x) \leq M\} \text{ is a compact.}$$

Proof.
1. Since, for all $\Lambda \in \mathscr{D}(I)$, the linear mappings

   $$x \mapsto -\langle \Lambda, x\rangle + I(\Lambda)$$

   are obviously concave, the infimum of this family is upper semi-continuous and concave.
2. See, e.g., [13] for the proof.

□

Theorem 3.1. *Suppose that*

1. *The family $\{Q_N\}$ satisfies condition* (3.4).
2. *Condition* (3.6) *is satisfied.*
3. *Condition* (3.9) *is satisfied.*

Then, for any closed set $\mathscr{V} \subset \mathbb{R}^d$, we have

$$\overline{\lim_{N\uparrow\infty}} \frac{1}{N} \mathbb{E}\left[\log Q_N(\mathscr{V})\right] \leq \sup_{x\in\mathscr{V}} I^*(x). \tag{3.10}$$

Proof. 1. Suppose at first that $\mathscr{V}$ is a compact.

Thanks to (3.7), for any $x \in \mathscr{X}$, there exists $\Lambda(x) \in \mathscr{X}$ such that

$$-\langle x, \Lambda(x)\rangle + I(\Lambda(x)) \leq I_\delta^*(x). \tag{3.11}$$

For any $x \in \mathscr{X}$, there exists a neighbourhood $A(x) \subset \mathscr{X}$ of x such that

$$\sup_{y \in A(x)} \langle y - x, \Lambda(x)\rangle \leq \delta.$$

By compactness, the covering $\bigcup_{x\in\mathscr{Y}} A(x) \supset \mathscr{V}$ has the finite subcovering, say $\bigcup_{r=1}^{p} A(x_r) \supset \mathscr{V}$. Hence,

$$\frac{1}{N}\log Q_N(\mathscr{V}) \leq \frac{1}{N}\log\left(\bigcup_{r=1}^{p} Q_N(A(x_r))\right). \tag{3.12}$$

Applying condition (3.4), we get

$$\varlimsup_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\max_{r\in\{1,\dots,p\}} \log Q_N(A(x_r)) - \max_{r\in\{1,\dots,p\}} \mathbb{E}\left[\frac{1}{N}\log Q_N(A(x_r))\right]\right] \leq 0. \tag{3.13}$$

By the Chebyshev inequality,

$$\begin{aligned} Q_N(A(x)) &\leq Q_N\left\{y \in \mathscr{X} : \langle y - x, \Lambda(x)\rangle \leq \delta\right\} \\ &\leq \mathrm{e}^{-\delta N}\int_{\mathscr{X}} \mathrm{e}^{N\langle y-x,\Lambda(x)\rangle}\mathrm{d}Q_N(y) \\ &= \mathrm{e}^{-\delta N}\mathrm{e}^{-N\langle x,\Lambda(x)\rangle}L_N(\Lambda(x)). \end{aligned} \tag{3.14}$$

Hence, (3.14) together with (3.11) yields

$$\begin{aligned} \varlimsup_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\log Q_N(A(x_r))\right] &\leq \lim_{N\uparrow+\infty}\left[-\langle x_r, \Lambda(x_r)\rangle + \frac{1}{N}\log L_N(\Lambda(x_r))\right] - \delta \\ &= -\langle x_r, \Lambda(x_r)\rangle + I(\Lambda(x_r)) - \delta \\ &\leq I_\delta^*(x_r) - \delta. \end{aligned} \tag{3.15}$$

Combining (3.12), (3.13), (3.15), we obtain

$$\begin{aligned} \varlimsup_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{V})\right] &\leq \max_{r\in\{1,\dots,p\}} I_\delta^*(x_r) - \delta \\ &\leq \sup_{x\in\mathscr{V}} I_\delta^*(x) - \delta. \end{aligned}$$

Taking $\delta \downarrow +0$ limit, we get the assertion of the theorem.

2. Let us allow now the set $\mathscr{V}$ to be unbounded. We first prove that the family Q_N is quenched exponentially tight. For that purpose, let

$$R_N(M) \equiv \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{X} \setminus [-M; M]^d)\right],$$

and denote

$$R(M) \equiv \overline{\lim_{N\uparrow+\infty}} R_N(M).$$

We want to prove that

$$\lim_{M\uparrow+\infty} R(M) = -\infty. \tag{3.16}$$

Fix some $u \in \{1, \dots, d\}$. Suppose $\delta_{u,p} \in \{0, 1\}$ is the standard Kronecker symbol. Let $e_u \in \mathbb{R}^d$ be an element of the standard basis of $\mathbb{R}^d$, i.e., for all $p \in \{1, \dots, d\}$, we have

$$(e_u)_p \equiv \delta_{u,p}.$$

Thanks to the Chebyshev inequality, we have

$$Q_N\{x_u \leq -M\} \leq \mathrm{e}^{-NM} \int_{\mathbb{R}^d} \mathrm{e}^{-N\langle x,e_u\rangle} \mathrm{d}Q_N(x), \text{ a.s.} \tag{3.17}$$

Now, we get

$$\begin{aligned}\int_{\mathbb{R}^d} \mathrm{e}^{-N\langle x,e_u\rangle} \mathrm{d}Q_N(x) &= \frac{1}{L_N(\Lambda_e)} \int_{\mathbb{R}^d} \mathrm{e}^{N\langle x,\Lambda_e - e_u\rangle} \mathrm{d}Q_N(x) \\ &= \frac{L_N(\Lambda_e - e_u)}{L_N(\Lambda_e)}, \text{ a.s.}\end{aligned} \tag{3.18}$$

Hence, combining (3.17) and (3.18), we obtain

$$\frac{1}{N}\mathbb{E}\left[\log Q_N\{x_u \leq -M\}\right] \leq -M + I_N(\Lambda_e - e_u) - I_N(\Lambda_e). \tag{3.19}$$

Using the same argument, we also get

$$\frac{1}{N}\mathbb{E}\left[\log Q_N\{x_u \geq M\}\right] \leq -M + I_N(\Lambda_e + e_u) - I_N(\Lambda_e). \tag{3.20}$$

We obviously have

$$R_N(M) \leq \frac{1}{N}\mathbb{E}\left[\log Q_N\left(\bigcup_{u=1}^{d}\left(\{x_u \leq -M\} \cup \{x_u \geq M\}\right)\right)\right]. \tag{3.21}$$

Applying condition (3.4) to (3.21), we get

$$\begin{aligned}\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\Big[&\log Q_N\Big(\bigcup_{u=1}^{d}\Big(\{x_u \leq -M\} \cup \{x_u \geq M\}\Big)\Big) \\ &- \max_{u\in\{1,\dots,d\}} \max\Big\{\mathbb{E}[\log Q_N\left(\{x_u \leq -M\}\right)], \mathbb{E}[\log Q_N\left(\{x_u \geq M\}\right)]\Big\}\Big] \leq 0.\end{aligned} \tag{3.22}$$

Applying (3.19) and (3.20) in (3.22), we get

$$\begin{aligned}&\overline{\lim_{N\uparrow+\infty}} \frac{1}{N}\mathbb{E}\left[\log Q_N\left(\bigcup_{u=1}^{d}\left(\{x_u \leq -M\} \cup \{x_u \geq M\}\right)\right)\right] \\ &\leq -M - I(\Lambda_e) + \max_{u\in\{1,\dots,d\}} \max\left\{I(\Lambda_e - e_u), I(\Lambda_e + e_u)\right\}.\end{aligned} \tag{3.23}$$

The bound (3.23) assures (3.16). Now, since we have (with the help of (3.4) and (3.10))

$$\begin{aligned}\overline{\lim_{N\uparrow+\infty}}\frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{V})\right] &\le \overline{\lim_{N\uparrow+\infty}}\frac{1}{N}\mathbb{E}\left[\log Q_N((\mathscr{V}\cap[-M;M]^d)\cup(\mathscr{X}\setminus[-M;M]^d))\right] \\ &\le \max\left\{\sup_{x\in(\mathscr{V}\cap[-M;M]^d)} I^*(x), R(M)\right\}, \qquad (3.24)\end{aligned}$$

the assertion of the theorem follows from (3.16) by taking the $\overline{\lim}_{M\uparrow+\infty}$ in the bound (3.24).

□

3.2 Quenched LDP lower bound

Suppose that, for some $\Lambda\in\mathbb{R}^d$ and all $N\in\mathbb{N}$, we have

$$\int_{\mathscr{X}} e^{N\langle y,\Lambda\rangle}\mathrm{d}Q_N(y) < +\infty.$$

Let $\widetilde{Q}_{N,\Lambda}\in\mathscr{M}(\mathscr{X})$ be the random measure defined by

$$\widetilde{Q}_{N,\Lambda}(A) = \int_A e^{N\langle y,\Lambda\rangle}\mathrm{d}Q_N(y), \qquad (3.25)$$

for any Q_N measurable $A\subset\mathscr{X}$.

Lemma 3.3. *Suppose the family of random measures Q_N satisfies the following assumptions.*

1. *Measure concentration. For all $N\in\mathbb{N}$, there exists some $L>0$ and $\eta_N:\mathbb{R}_+\to\mathbb{R}_+$ such that, for any Q_N-measurable set $A\subset\mathscr{X}$, we have*

$$\mathbb{P}\left\{\left|\log Q_N(A)-\mathbb{E}\left[\log Q_N(A)\right]\right| > t\right\} \le \eta_N(t).$$

Assume, in addition, that, for some $p>0$, the concentration function satisfies

$$N^p\int_0^{+\infty}\eta_N(Nt)\mathrm{d}t \xrightarrow[N\uparrow+\infty]{} 0. \qquad (3.26)$$

2. *Tails decay condition. Let*

$$C(M)\equiv\{x\in\mathscr{X}:\|x\|<M\}.$$

There exists $p\in\mathbb{N}$ such that

$$\lim_{K\uparrow+\infty}\overline{\lim_{N\uparrow\infty}}\int_0^{+\infty}\mathbb{P}\left\{\frac{1}{N}\log\widetilde{Q}_{N,\Lambda}(\mathscr{X}\setminus C(N^p)) > -K+t\right\}\mathrm{d}t = 0. \qquad (3.27)$$

3. *Non-degeneracy. The family of the sets $\left\{B_j\subset\mathscr{X}: j\in\{1,\dots,q\}\right\}$ satisfies the following condition*

$$\text{there exists some } j_0\in\{1,\dots,q\} \text{ such that } \underline{\lim_{N\uparrow\infty}}\frac{1}{N}\mathbb{E}\left[\log\widetilde{Q}_{N,\Lambda}(B_{j_0})\right] > -\infty. \qquad (3.28)$$

Then, for any $\Lambda \in \mathbb{R}^d$, we have

$$\overline{\lim_{N\uparrow\infty}} \frac{1}{N}\mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}\left(\bigcup_{j=1}^{q} B_j\right) - \max_{j\in\{1,\ldots,q\}} \mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_j)\right]\right] \leq 0. \tag{3.29}$$

Remark 3.2. *The polynomial growth choice of $M = M_N \equiv N^p$ made in assumptions* (3.27) *and* (3.26) *is made for specificity. Inspecting the following proof, one can easily restate the conditions* (3.27) *and* (3.26) *for general M_N dependencies. Effectively, the growth rate of M_N is related to the covering dimension of the Polish space $(\mathscr{X}, \mathfrak{X})$.*

Proof of Lemma 3.3. We fix some $j \in \{1,\ldots,q\}$. Take an arbitrary $\varepsilon > 0$, $M > 0$ and denote $J_{M,\varepsilon} \equiv \mathbb{Z} \cap [-\|\Lambda\|M/\varepsilon; \|\Lambda\|M/\varepsilon]$. Consider, for $i \in J_{M,\varepsilon}$, the following closed sets

$$A_{i,j} \equiv \{x \in B_j : (j-1)\varepsilon \leq \langle \Lambda, x\rangle \leq j\varepsilon\}.$$

We get

$$\begin{aligned}
\frac{1}{N}\log \widetilde{Q}_{N,\Lambda}\left(\bigcup_{j=1}^{q} B_j\right) &\leq \frac{1}{N}\log \widetilde{Q}_{N,\Lambda}\left(\left(\bigcup_{j=1}^{q} B_j \cap C(M)\right) \cup (\mathscr{X} \setminus C(M))\right)\\
&\leq \frac{1}{N}\max\Big\{\max_{j\in\{1,\ldots,q\}} \log \widetilde{Q}_{N,\Lambda}(B_j \cap C(M)),\\
&\qquad \log \widetilde{Q}_{N,\Lambda}(\mathscr{X} \setminus C(M))\Big\} + \frac{\log(q+1)}{N}.
\end{aligned} \tag{3.30}$$

We have

$$\begin{aligned}
\frac{1}{N}\log \widetilde{Q}_{N,\Lambda}(B_j \cap C(M)) &\leq \frac{1}{N}\log\left(\sum_{i\in J_{M,\varepsilon}} \mathrm{e}^{Ni\varepsilon} Q_N(A_{i,j})\right)\\
&\leq \max_{i\in\{1,\ldots,p\}}\left[i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right] + \frac{\log(\operatorname{card} J_{M,\varepsilon})}{N}.
\end{aligned} \tag{3.31}$$

Denote

$$\alpha_N(\varepsilon) \equiv \max_{j\in\{1,\ldots,q\}} \max_{i\in J_{M,\varepsilon}} \left(i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right),$$

and

$$\beta_N \equiv \max_{j\in\{1,\ldots,q\}} \mathbb{E}\left[\log \widetilde{Q}_{N,\Lambda}(B_j)\right],$$

$$\widetilde{\beta}_N(\varepsilon) \equiv \max_{j\in\{1,\ldots,q\}} \mathbb{E}\left[\max_{i\in J_{M,\varepsilon}}\left(i\varepsilon + \frac{1}{N}\log Q_N(A_{i,j})\right)\right],$$

$$\gamma_N(M) \equiv \frac{1}{N}\log \widetilde{Q}_{N,\Lambda}(\mathscr{X} \setminus C(M)).$$

We also have

$$\begin{aligned}\frac{1}{N}\log\widetilde{Q}_{N,\Lambda}(B_j) &\geq \frac{1}{N}\log\widetilde{Q}_{N,\Lambda}(B_j\cap C(M))\\ &\geq \max_{i\in J_{M,\varepsilon}}\left[(i-1)\varepsilon+\frac{1}{N}\log Q_N(A_{i,j})\right]\\ &= \max_{i\in J_{M,\varepsilon}}\left[i\varepsilon+\frac{1}{N}\log Q_N(A_{i,j})\right]-\varepsilon. \end{aligned}\tag{3.32}$$

Due to condition (1), we have

$$\mathbb{P}\left\{\left|\alpha_N(\varepsilon)-\widetilde{\beta}_N(\varepsilon)\right|>t\right\}\leq \eta_N(tN)q\operatorname{card}J_{M,\varepsilon}.\tag{3.33}$$

We put $M\equiv M_N\equiv N^p$, and we get

$$\begin{aligned}\operatorname{card}J_{M,\varepsilon} &\leq 2\|\Lambda\|M/\varepsilon+1\\ &\leq 2\|\Lambda\|N^p/\varepsilon+1.\end{aligned}\tag{3.34}$$

Let

$$X_N(M,\varepsilon)\equiv\max\{\gamma_N(M),\alpha_N(\varepsilon)\}-\beta_N,$$

then we have

$$\mathbb{P}\{X_N(K,\varepsilon)>t\}\leq\mathbb{P}\{\gamma_N(M)>\beta_N+t\}+\mathbb{P}\{\alpha_N(\varepsilon)>\beta_N+t\}.\tag{3.35}$$

Due to property (3.28), there exists $K>0$ such that we have

$$\mathbb{P}\{\gamma_N(M)>\beta_N+t\}\leq\mathbb{P}\{\gamma_N(M)>-K+t\}.\tag{3.36}$$

Thanks to (3.32), we have

$$\mathbb{P}\{\alpha_N(\varepsilon)>\beta_N+t\}\leq\mathbb{P}\{\alpha_N(\varepsilon)>\widetilde{\beta}_N(\varepsilon)+t-\varepsilon\}.\tag{3.37}$$

For $t>\varepsilon$, we apply (3.33) and (3.34) to (3.37) to obtain

$$\mathbb{P}\{\alpha_N(\varepsilon)>\beta_N+t\}\leq(2\|\Lambda\|N^p/\varepsilon+1)\,q\eta_N(tN).\tag{3.38}$$

Combining (3.30) and (3.31), we get

$$\begin{aligned}&\mathbb{E}\left[\log\widetilde{Q}_{N,\Lambda}\Big(\bigcup_{j=1}^{q}B_j\Big)-\max_{j\in\{1,\dots,q\}}\mathbb{E}\left[\log\widetilde{Q}_{N,\Lambda}(B_j)\right]\right]\\ &\leq\mathbb{E}\left[X_N(M,\varepsilon)\right]+\frac{\log(q+1)}{N}+\frac{\log(2\|\Lambda\|N^p/\varepsilon+1)}{N}.\end{aligned}\tag{3.39}$$

Now, (3.35), (3.36) and (3.38) imply

$$\begin{aligned}\mathbb{E}\left[X_N(M,\varepsilon)\right] &\leq\int_0^{+\infty}\mathbb{P}\left\{X_N(M,\varepsilon)>t\right\}\mathrm{d}t\\ &\leq\int_\varepsilon^{+\infty}\mathbb{P}\left\{X_N(M,\varepsilon)>t\right\}\mathrm{d}t+\varepsilon\\ &\leq\int_\varepsilon^{+\infty}\mathbb{P}\{\gamma_N(M)>-K+t\}\mathrm{d}t\\ &\quad+(2\|\Lambda\|N^p/\varepsilon+1)\,q\int_\varepsilon^{+\infty}\eta_N(tN)\mathrm{d}t+\varepsilon.\end{aligned}\tag{3.40}$$

Therefore, taking sequentially $\overline{\lim}_{N\uparrow+\infty}$, $\lim_{K\uparrow+\infty}$ and $\lim_{\varepsilon\uparrow+0}$ in (3.40), and using (3.26), we arrive at

$$\overline{\lim_{N\uparrow\infty}} \mathbb{E}\left[X_N(M,\varepsilon)\right] \leq 0. \tag{3.41}$$

Bound (3.41) together with (3.39) implies the assertion of the lemma.

□

Let $\hat{Q}_{N,\Lambda}$ be the (random) probability measure defined by

$$\hat{Q}_{N,\Lambda} \equiv \frac{\widetilde{Q}_N}{L_N(\Lambda)}.$$

Lemma 3.4. *Suppose that the measure Q_N satisfies the assumptions of the previous lemma. Then* (3.29) *is valid also for $\hat{Q}_{N,\Lambda}$.*

Proof. Similar to the one of the previous lemma. □

Remark 3.3. *Recall that a point $x \in \mathscr{X}$ is called an exposed point of the concave mapping I^* if there exists $\Lambda \in \mathbb{R}^d$ such that, for all $y \in \mathscr{X} \setminus \{x\}$, we have*

$$I^*(y) - I^*(x) < \langle y - x, \Lambda \rangle. \tag{3.42}$$

Theorem 3.2. *Suppose*

1. *The family $\{Q_N : N \in \mathbb{N}\} \subset \mathscr{M}(\mathbb{R}^d)$ satisfies the assumptions of Lemma 3.3.*
2. *$\mathscr{G} \subset \mathscr{X}$ is an open set.*
3. *$\emptyset \neq \mathscr{E}(I^*) \subset \mathscr{D}(I^*)$ is the set of the exposed points of the mapping I^*.*
4. *Condition* (3.9) *is satisfied.*

Then

$$\underline{\lim}_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\log Q_N(\mathscr{G}\cap\mathscr{E})\right] \geq \sup_{x\in\mathscr{G}} I^*(x). \tag{3.43}$$

Proof. Let $B(x,\varepsilon)$ be a ball of radius $\varepsilon > 0$ around some arbitrary $x \in \mathscr{X}$. It suffices to prove that

$$\lim_{\varepsilon\downarrow+0} \underline{\lim}_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] \geq I^*(x). \tag{3.44}$$

Indeed, since we have

$$Q_N(\mathscr{G}) \geq Q_N(B(x,\varepsilon)), \tag{3.45}$$

applying $\frac{1}{N}\log(\cdot)$, taking the expectation, taking $\underline{\lim}_{N\uparrow+\infty}$, $\varepsilon \downarrow +0$ and taking the supremum over $x \in \mathscr{G}$ in (3.45), we get (3.43).

Take any $x \in \mathscr{G} \cap \mathscr{E}$. Then we can find the corresponding vector $\Lambda_e = \Lambda_e(x) \in \mathbb{R}^d$ orthogonal to the exposing hyperplane at the point x, as in (3.42). Define the new ("tilted") random probability measure $\hat{Q}_N$ on $\mathbb{R}^d$ by demanding that

$$\frac{\mathrm{d}\hat{Q}_N}{\mathrm{d}Q_N}(y) = \frac{1}{L_N(\Lambda_e)} \mathrm{e}^{N\langle y, \Lambda_e\rangle}. \tag{3.46}$$

Moreover, we have

$$\begin{aligned} \frac{1}{N}\mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] &= \frac{1}{N}\mathbb{E}\left[\log \int_{B(x,\varepsilon)} \mathrm{d}Q_N(y)\right] \\ &= \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda_e)\right] + \frac{1}{N}\mathbb{E}\left[\int_{B(x,\varepsilon)} \mathrm{e}^{-N\langle y,\Lambda_e\rangle} \mathrm{d}\hat{Q}_N(y)\right] \\ &\geq \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda_e)\right] - \langle x, \Lambda_e\rangle - \varepsilon \|\Lambda_e\|_2 + \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(B(x,\varepsilon))\right]. \end{aligned}$$

Hence,

$$\lim_{\varepsilon\downarrow+0} \varliminf_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log Q_N(B(x,\varepsilon))\right] \geq \left[-\langle x,\Lambda_e\rangle + I(\Lambda_e)\right] + \lim_{\varepsilon\downarrow+0} \varliminf_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(B(x,\varepsilon))\right].$$

Since we have

$$-\langle x, \Lambda_e\rangle + I(\Lambda_e) \geq I^*(x),$$

in order to show (3.44) it remains to prove that

$$\lim_{\varepsilon\downarrow+0} \varliminf_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(B(x,\varepsilon))\right] = 0. \tag{3.47}$$

The Laplace transform of $\hat{Q}_N$ is

$$\hat{L}_N(\Lambda) = \frac{L_N(\Lambda + \Lambda_e)}{L_N(\Lambda_e)}.$$

Hence, we arrive at

$$\hat{I}(\Lambda) = I(\Lambda + \Lambda_e) - I(\Lambda_e).$$

Moreover, we have

$$\hat{I}^*(x) = I^*(x) + \langle x, \Lambda_e\rangle - I(\Lambda_e). \tag{3.48}$$

By the assumptions of the theorem, the family Q_N satisfies the assumptions of Lemma 3.3. Hence, due to Lemma 3.4, the family $\hat{Q}_N$ satisfies (3.4). Thus we can apply Theorem 3.1 to obtain

$$\varlimsup_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\log \hat{Q}_N(\mathbb{R}^d \setminus B(U,\varepsilon))\right] \leq \sup_{y\in\mathscr{U}\setminus B(x,\varepsilon)} \hat{I}^*(y). \tag{3.49}$$

Lemma 3.2 implies that there exists some $x_0 \in \mathscr{X} \setminus B(x,\varepsilon)$ (note that $x_0 \neq x$) such that

$$\sup_{y \in \mathscr{X} \setminus B(x,\varepsilon)} \hat{I}^*(y) = \hat{I}^*(x_0).$$

Since Λ_e is an exposing hyperplane, using (3.48), we get

$$\begin{aligned}\hat{I}^*(x_0) &= I^*(x_0) + \langle x_0, \Lambda_e \rangle - I(\Lambda_e) \\ &\leq [I^*(x_0) + \langle x_0, \Lambda_e \rangle] - [I^*(x) + \langle x, \Lambda_e \rangle] < 0,\end{aligned} \tag{3.50}$$

and hence, combining (3.49) and (3.50), we get

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N} \mathbb{E}\left[\log \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon))\right] < 0.$$

Therefore, due to the concentration of measure, we have almost surely

$$\overline{\lim_{N\uparrow+\infty}} \frac{1}{N} \log \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon)) < 0$$

which implies that, for all $\varepsilon > 0$, we have almost surely

$$\lim_{N\uparrow+\infty} \hat{Q}_N(\mathbb{R}^d \setminus B(x,\varepsilon)) = 0,$$

and (3.47) follows by yet another application of the concentration of measure.

□

Corollary 3.1. *Suppose that in addition to the assumptions of previous Theorem 3.2 we have*

1. *$I(\cdot)$ is differentiable on* $\operatorname{int}\mathscr{D}(I)$.
2. *Either $\mathscr{D}(I) = \mathscr{X}$ or*

$$\lim_{\Lambda \to \partial\mathscr{D}(I)} \|\nabla I(\Lambda)\| = +\infty.$$

Then $\mathscr{E}(I^*) = \mathbb{R}^d$, consequently

$$\underline{\lim_{N\uparrow+\infty}} \frac{1}{N} \mathbb{E}\left[\log Q_N(\mathscr{G})\right] \geq \sup_{x \in \mathscr{G}} I^*(x).$$

Proof. The proof is the same as in the classical GÃď rtner-Ellis theorem (see, e.g., [13]).

□

4 The Aizenman-Sims-Starr comparison scheme

In this section, we shall extend the AS^2 scheme to the case of the SK model with multidimensional spins and prove Theorems 1.1 and 1.2, as stated in the introduction. We use the Gaussian comparison results of Section 2.3 in the spirit of AS^2 scheme in order to relate the free energy of the SK model with multidimensional spins with the free energy of a certain GREM-inspired model. Comparing to [1], due to more intricate nature of spin configuration space, some new effects occur. In particular, the remainder term of the Gaussian comparison non-trivially depends on the variances and covariances of the Hamiltonians under comparison. To deal with this obstacle, we use the idea of localisation to the configurations having a given overlap (cf. (1.5)). This idea is formalised by adapting the proofs of the quenched GÃď rtner-Ellis type LDP obtained in Section 3.

4.1 Naive comparison scheme

We start by recalling the basic principles of the AS^2 comparison scheme (see, e.g., [7, Chapter 11]). It is a simple idea to get the comparison inequalities by adding some additional structure into the model. However, the way the additional structure is attached to the model might be suggested by the model itself. Later on we shall encounter a real-world use of this trick. Let $(\Sigma, \mathfrak{S})$ and $(\mathscr{A}, \mathfrak{A})$ be Polish spaces equipped with measures μ and ξ, respectively. Furthermore, let

$$X \equiv \{X(\sigma)\}_{\sigma\in\Sigma}, A \equiv \{A(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma,\\ \alpha\in\mathscr{A}}}, B \equiv \{B(\sigma)\}_{\alpha\in\mathscr{A}}$$

be independent real-valued Gaussian processes. Define the *comparison functional*

$$\Phi[C] \equiv \mathbb{E}\left[\log \int_{\Sigma\times\mathscr{A}} \mathrm{e}^{C(\sigma,\alpha)} \mathrm{d}\left(\mu\otimes\xi\right)(\sigma,\alpha)\right], \tag{4.1}$$

where $C \equiv \{C(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma\\ \alpha\in\mathscr{A}}}$ is a suitable real-valued Gaussian process. Theorem 4.1 of [2] is easily understood as an example of the following observation. Suppose $\Phi[X]$ is somehow hard to compute directly, but $\Phi[A]$ and $\Phi[B]$ are manageable. We always have the following additivity property

$$\Phi\left[X+B\right] = \Phi\left[X\right] + \Phi\left[B\right]. \tag{4.2}$$

Assume now that

$$\Phi\left[X+B\right] \leq \Phi\left[A\right] \tag{4.3}$$

which we can obtain, e.g., from Proposition 2.5. Combining (4.2) and (4.3), we get the bound

$$\Phi\left[X\right] \leq \Phi\left[A\right] - \Phi\left[B\right]. \tag{4.4}$$

4.2 Free energy upper bound

Let $\mathscr{V} \subset \mathrm{Sym}(d)$ be an arbitrary Borell set.

Remark 4.1. *Note that $\mathscr{U}$ is closed and convex.*

Let

$$\begin{aligned}\Sigma_N(\mathscr{V}) &\equiv \left\{\sigma\in\Sigma_N : R_N(\sigma,\sigma)\in\mathscr{V}\right\}\\ &= \left\{\sigma\in\Sigma_N : R_N(\sigma,\sigma)\in\mathscr{V}\cap\mathscr{U}\right\}.\end{aligned} \tag{4.5}$$

Let us define the *local comparison functional* $\Phi_N(x,\mathscr{V})$ as follows (cf. (4.1))

$$\Phi_N(x,\mathscr{V})[C] \equiv \frac{1}{N}\mathbb{E}\left[\log \pi_N\left[\mathbb{1}_{\Sigma_N(\mathscr{V})}\exp\left(\beta\sqrt{N}C\right)\right]\right], \tag{4.6}$$

where $C \equiv \{C(\sigma,\alpha)\}_{\substack{\sigma\in\Sigma\\ \alpha\in\mathscr{A}}}$ is a suitable Gaussian process. Let us consider the following family ($N \in \mathbb{N}$) of random measures on the Borell subsets of $\mathrm{Sym}(d)$ generated by the SK Hamiltonian,

$$P_N(\mathscr{V}) \equiv \int_{\Sigma_N(\mathscr{V})} \mathrm{e}^{\beta\sqrt{N}X_N(\sigma)}\mathrm{d}\mu^{\otimes N}(\sigma),$$

and consider also the following family of the random measures generated by the Hamiltonian $A(\sigma,\alpha)$

$$\widetilde{P}_N(\mathscr{V}) \equiv \widetilde{P}_N^{x,\mathscr{Q},U}(\mathscr{V}) \equiv \int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} \exp\left(\beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma,\alpha), \tag{4.7}$$

where the parameters $\mathscr{Q}$ and U are taken from the definition of the process $A(\alpha)$ (cf. (1.7)). The vector x defines the random measure $\xi \in \mathscr{M}(\mathscr{A})$ (cf. (1.8)), and, hence, also the measure $\pi_N \in \mathscr{M}(\Sigma\times\mathscr{A})$.

Remark 4.2. *To lighten the notation, most of the time we shall not indicate explicitly the dependence of the following quantities on the parameters x, $\mathscr{Q}$, U.*

Consider (if it exists) the Laplace transform of the measure (4.7)

$$\widetilde{L}_N(\Lambda) \equiv \int_{\mathscr{U}} \mathrm{e}^{N\langle U,\Lambda\rangle} \mathrm{d}\widetilde{P}_N(U). \tag{4.8}$$

Let (if it exists)

$$\widetilde{I}(\Lambda) \equiv \lim_{N\uparrow\infty} \frac{1}{N}\mathbb{E}\left[\log \widetilde{L}_N(\Lambda)\right]. \tag{4.9}$$

Define the following Legendre transform

$$\widetilde{I}^*(U) \equiv \inf_{\substack{x\in\mathscr{Q}'(1,1),\\ \mathscr{Q}\in\mathscr{Q}'(U,d),\\ \Lambda\in\mathrm{Sym}(d)}} \left[-\langle U,\Lambda\rangle - \Phi_N(x,\mathscr{V})[B] + \widetilde{I}(\Lambda)\right]. \tag{4.10}$$

Denote, for $\delta > 0$,

$$\widetilde{I}^*_\delta(U) \equiv \max\left\{\widetilde{I}^*(U) + \delta, -\frac{1}{\delta}\right\}.$$

Let

$$p(\mathscr{V}) \equiv \lim_{N\uparrow+\infty} \frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V})\right]. \tag{4.11}$$

Remark 4.3. *Note that the result of [19] assures the existence of the limit in the previous formula.*

Lemma 4.1. *We have*

1. *The Laplace transform* (4.8) *exists. Moreover, for any $\Lambda \in \mathrm{Sym}(d)$, we have*

$$\begin{aligned} &\int_{\mathscr{V}} \mathrm{e}^{N\langle U,\Lambda\rangle} \mathrm{d}P_N(U) \\ &= \int_{\Sigma_N(\mathscr{V})} \exp\left(N\langle\Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}X(\sigma)\right) \mathrm{d}\mu^{\otimes N}(\sigma), \end{aligned} \tag{4.12}$$

$$\begin{aligned} &\int_{\mathscr{V}} \mathrm{e}^{N\langle U,\Lambda\rangle} \mathrm{d}\widetilde{P}_N(U) \\ &= \int_{\Sigma_N(\mathscr{V})\times\mathscr{A}} \exp\left(N\langle\Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma,\alpha). \end{aligned} \tag{4.13}$$

2. *The quenched cumulant generating function* (4.9) *exists in the $N \uparrow \infty$ limit, for any $\Lambda \in \mathrm{Sym}(d)$. Moreover, for all $N \in \mathbb{N}$, we have*

$$I_N(\Lambda) \equiv \frac{1}{N}\mathbb{E}\left[\log L_N(\Lambda)\right] = X_0(x, \mathcal{Q}, \Lambda, U), \tag{4.14}$$

that is $I_N(\cdot)$ in fact does not depend on N.

Proof. 1. We prove (4.13), the proof of (4.12) is similar. Since $\mathcal{U}$ is a compact, it follows that, for arbitrary $\varepsilon > 0$, there exists the following ε-partition of $\mathcal{U}$

$$\mathcal{N}(\varepsilon) = \{\mathcal{V}_r \subset \mathcal{U} : r \in \{1, \dots, K\}\}$$

such that $\bigcup_r \mathcal{V}_r = \mathcal{U}$, $\mathcal{V}_r \cap \mathcal{V}_s = \emptyset$, diam $\mathcal{V}_r \leq \varepsilon$ and pick some $V_r \in \mathrm{int}\,\mathcal{V}_r$, for all $r \neq s$.
We denote

$$\widetilde{L}_N(\Lambda, \varepsilon) \equiv \sum_{r=1}^{K} \mathrm{e}^{N\langle \Lambda, V_r\rangle} \int_{\Sigma_N(\mathcal{V}_r)\times\mathcal{A}} \exp\left(\beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha), \sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma, \alpha).$$

For small enough ε, we have

$$(1 - 2N\|\Lambda\|\varepsilon)\,\mathrm{e}^{N\langle \Lambda, R_N(\sigma,\sigma)\rangle} \leq \mathrm{e}^{N\langle \Lambda, U\rangle} \leq \mathrm{e}^{N\langle \Lambda, R_N(\sigma,\sigma)\rangle}(1 + 2N\|\Lambda\|\varepsilon).$$

Therefore, if we denote

$$\widehat{L}_N(\mathcal{V}, \Lambda) \equiv \int_{\Sigma_N(\mathcal{V})\times\mathcal{A}} \exp\left(N\langle \Lambda, R_N(\sigma,\sigma)\rangle + \beta\sqrt{N}\sum_{i=1}^{N}\langle A_i(\alpha), \sigma_i\rangle\right) \mathrm{d}\pi_N(\sigma, \alpha),$$

we get

$$(1 - 2N\|\Lambda\|\varepsilon)\sum_{r=1}^{K}\widehat{L}_N(\mathcal{V}_r, \Lambda) \leq \widetilde{L}_N(\Lambda, \varepsilon) \leq (1 + 2N\|\Lambda\|\varepsilon)\sum_{r=1}^{K}\widehat{L}_N(\mathcal{V}_r, \Lambda).$$

Hence,

$$(1 - 2N\|\Lambda\|\varepsilon)\,\widehat{L}_N(\mathcal{U}, \Lambda) \leq \widetilde{L}_N(\Lambda, \varepsilon) \leq (1 + 2N\|\Lambda\|\varepsilon)\,\widehat{L}_N(\mathcal{U}, \Lambda). \tag{4.15}$$

Let $\varepsilon \downarrow +0$ in (4.15) and we arrive at

$$\widetilde{L}_N(\Lambda) = \widehat{L}_N(\mathcal{U}, \Lambda).$$

That is, the existence of $L_N(\Lambda)$ and the representation (4.13) are proved.

2. For all $N \in \mathbb{N}$, we have, by the RPC averaging property (see, e.g., [2, Theorem 5.4] or Theorem 5.3, property (4) below), that

$$\frac{1}{N}\mathbb{E}\left[\log\widetilde{L}_N(\mathcal{U}, \Lambda)\right] = \Phi_N(x, \mathcal{U})\left[A + N\langle \Lambda, R_N(\sigma,\sigma)\rangle\right] = X_0(x, \mathcal{Q}, \Lambda, U).$$

$\square$

Proof of Theorem 1.1. In essence, the proof follows almost literally the proof of Theorem 3.1. The notable difference is that we apply the Gaussian comparison inequality (Proposition 2.5) in order to "compute" the rate function in a somewhat more explicit way.

Due to (4.5), we can without loss of generality suppose that $\mathscr{V}$ is compact. For any $\delta > 0$ and $U \in \mathscr{V}$, by (4.10), there exists $\Lambda(U,\delta) \in \mathrm{Sym}(d)$, $x(U,\delta) \in \mathscr{Q}'(1,1)$ and $Q(U,\delta) \in \mathscr{Q}'(U,d)$ such that

$$-\langle U, \Lambda(U)\rangle + \widetilde{I}(\Lambda(U)) \le \widetilde{I}^*_\delta(U). \tag{4.16}$$

For any $U \in \mathscr{V}$, there exists an open neighbourhood $\mathscr{V}(U) \subset \mathrm{Sym}(d)$ of U such that

$$\sup_{V \in \mathscr{V}(U)} \langle V - U, \Lambda(U)\rangle \le \delta.$$

Fix some $\varepsilon > 0$. Without loss of generality, we can suppose that all the neighbourhoods satisfy additionally the condition $\operatorname{diam} \mathscr{V}(U) \le \varepsilon$. By compactness, the covering $\bigcup_{U\in\mathscr{V}} \mathscr{V}(U) \supset \mathscr{V}$ has a finite subcovering, say $\bigcup_{r=1}^{p} \mathscr{V}(U^{(r)}) \supset \mathscr{V}$. We denote the corresponding to this covering approximants in (4.16) by $\{x^{(r)} \in \mathscr{Q}'(1,1)\}_{r=1}^{p}$ and $\{\mathscr{Q}^{(r)} \in \mathscr{Q}'(U^{(r)},d)\}_{r=1}^{p}$. We have

$$\frac{1}{N}\log P_N(\mathscr{V}) \le \frac{1}{N}\log\Big(\bigcup_{r=1}^{p} P_N(\mathscr{V}(U^{(r)}))\Big). \tag{4.17}$$

Due to the concentration of measure Proposition 2.3, we can apply Lemma 3.1 and get

$$\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\left|\log P_N\Big(\bigcup_{r=1}^{p}\mathscr{V}(U^{(r)})\Big) - \max_{r\in\{1,\dots,p\}}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right]\right|\right] = 0. \tag{4.18}$$

In fact, since we know that (4.11) exists, (4.18) implies that

$$\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N\Big(\bigcup_{r=1}^{p}\mathscr{V}(U^{(r)})\Big)\right] = \max_{r\in\{1,\dots,p\}}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right]. \tag{4.19}$$

For $U^{(r)}$, $x = x^{(r)}$, $\mathscr{Q} = \mathscr{Q}^{(r)}$, Proposition 2.5 gives

$$\begin{aligned}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right] &= \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(\mathscr{V}(U^{(r)}))\right] - \Phi_N(x,\mathscr{U})[B]\\ &\quad + \mathscr{R}_N(x^{(r)},\mathscr{Q}^{(r)},U^{(r)},\mathscr{V}(U^{(r)})) + \mathscr{O}(\varepsilon)\\ &\le \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(\mathscr{V}(U^{(r)}))\right] - \Phi_N(x,\mathscr{U})[B] + K\varepsilon,\end{aligned} \tag{4.20}$$

where $K > 0$ is an absolute constant.

By the Chebyshev inequality and Lemma 4.1, we have

$$\begin{aligned}\widetilde{P}_N(\mathscr{V}(U)) &\le \widetilde{P}_N\{V \in \mathscr{U} : \langle V - U, \Lambda(U)\rangle \le \delta\}\\ &\le \mathrm{e}^{-\delta N}\int_{\mathscr{U}} \mathrm{e}^{N\langle V-U,\Lambda(U)\rangle}\mathrm{d}\widetilde{P}_N(V)\\ &= \mathrm{e}^{-\delta N}\mathrm{e}^{-N\langle U,\Lambda(U)\rangle}\widetilde{L}_N(\Lambda(U)).\end{aligned}$$

Thus, using (4.20) and (4.16), we get

$$\begin{aligned}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right] &\leq \lim_{N\uparrow+\infty}\left[-\langle U^{(r)},\Lambda(U^{(r)})\rangle-\Phi[B]+\frac{1}{N}\log\widetilde{L}_N(\Lambda(U^{(r)}))\right]-\delta+K\varepsilon\\ &=-\langle U_r,\Lambda(U_r)\rangle-\Phi[B]+\widetilde{I}(\Lambda(U_r))-\delta+K\varepsilon\\ &\leq \widetilde{I}^*_\delta(U_r)-\delta+K\varepsilon. \end{aligned} \tag{4.21}$$

Combining (4.17), (3.13), (4.21), we obtain

$$\begin{aligned}p(\mathscr{V})=\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V})\right] &\leq \max_{r\in\{1,\ldots,p\}}\lim_{N\uparrow+\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(\mathscr{V}(U^{(r)}))\right]\\ &\leq \max_{r\in\{1,\ldots,p\}}\widetilde{I}^*_\delta(U^{(r)})+K\varepsilon-\delta\\ &\leq \sup_{U\in\mathscr{V}}\widetilde{I}^*_\delta(V)+K\varepsilon-\delta.\end{aligned}$$

Taking $\delta\downarrow+0$ and $\varepsilon\downarrow+0$ limits, we get

$$p(\mathscr{V})\leq\sup_{V\in\mathscr{V}}\widetilde{I}^*(U). \tag{4.22}$$

The averaging property of the RPC (see, e.g., [2, Theorem 5.4] or property (4) of Theorem 5.3) gives

$$\Phi_N(x,\mathscr{U})[B]=\frac{\beta^2}{2}\sum_{k=1}^{n}x_k\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2-\|Q^{(k)}\|_{\mathrm{F}}^2\right). \tag{4.23}$$

To finish the proof it remains to show that, for any fixed $\Lambda\in\mathrm{Sym}(d)$, we have

$$\widetilde{I}(\Lambda)=X_0(x,\mathscr{Q},\Lambda,U)$$

which is assured by Lemma 4.1. □

4.3 Free energy lower bound

In this subsection, we return to the notations of Section 4.2.

Lemma 4.2. *For any $\mathscr{B}\subset\mathrm{Sym}(d)$ such that $\mathrm{int}\,\mathscr{B}\cap\mathrm{int}\,\mathscr{U}\neq\emptyset$ there exists $\Delta\subset\Sigma$ with $\mathrm{int}\,\Delta\neq\emptyset$ such that*

$$\begin{aligned}\lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\int_{\Sigma_N(\mathscr{B})\times\mathscr{A}}\exp\left(N\langle\Lambda,R_N(\sigma,\sigma)\rangle+\sum_{i=1}^{N}\langle A_i(\alpha),\sigma_i\rangle\right)\mathrm{d}\pi_N(\sigma,\alpha)\right]&\\ \geq\log\int_\Delta\exp\left(\langle(\beta^2U+\Lambda)\sigma,\sigma\rangle\right)\mathrm{d}\mu(\sigma)>-\infty.&\end{aligned} \tag{4.24}$$

Proof. In view of (1.10), iterative application of the Jensen inequality with respect to $\mathbb{E}_{z^{(k)}}$ leads to the following

$$\mathbb{E}\left[X_{n+1}(x,\mathscr{Q},\Lambda,U)\right]\leq X_0(x,\mathscr{Q},\Lambda,U).$$

Performing the Gaussian integration, we get

$$\mathbb{E}\left[X_{n+1}(x,\mathcal{Q},\Lambda,U)\right] \geq \log \int_{\Delta} \exp\left(\langle(\beta^2 U+\Lambda)\sigma,\sigma\rangle\right) \mathrm{d}\mu(\sigma),$$

where $\Delta \subset \Sigma$ is such that $\mu(\Delta) > 0$ and $\{R(\sigma,\sigma) : \sigma \in \Delta^N\} \subset \mathscr{B}$. $\square$

Define the following Legendre transform

$$\widehat{I}^*(U) \equiv \inf_{\substack{x\in\mathcal{Q}'(1,1),\\ \mathcal{Q}\in\mathcal{Q}'(U,d),\\ \Lambda\in\mathrm{Sym}(d)}} \left[-\langle U,\Lambda\rangle - \Phi[B] + \widetilde{I}(\Lambda) + \mathscr{R}(x,\mathcal{Q},U)\right]. \tag{4.25}$$

Proof of Theorem 1.2. As it is the case with the proof of Theorem 1.1, this proof also follows in essence almost literally the proof of Theorem 3.2. The notable difference is that we apply the Gaussian comparison in order to “compute” the rate function in a somewhat more explicit way.

In notations of Theorem 3.2 we are in the following situation: $\mathscr{X} \equiv \mathrm{Sym}(d)$ and $\mathfrak{X}$ is the topology induced by any norm on $\mathrm{Sym}(d)$.

Let $B(U,\varepsilon)$ be the ball (in the Hilbert-Schmidt norm) of radius $\varepsilon > 0$ around some arbitrary $U \in \mathscr{V}$. Let us prove at first that

$$\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow\infty}\frac{1}{N}\mathbb{E}\left[\log P_N(B(U,\varepsilon))\right] \geq \widehat{I}^*(U). \tag{4.26}$$

Similarly to (4.20), for any $(x,\mathcal{Q})$, we have

$$\begin{aligned}&\mathbb{E}\left[\frac{1}{N}\log P_N(B(U,\varepsilon))\right]\\ &= \frac{1}{N}\mathbb{E}\left[\log \widetilde{P}_N(B(U,\varepsilon))\right] - \Phi[B] + \mathscr{R}_N(x,\mathcal{Q},U,B(U,\varepsilon)) + \mathscr{O}(\varepsilon).\end{aligned} \tag{4.27}$$

The random measure $\widetilde{P}_N$ satisfies the assumptions of Corollary 3.1. Indeed:

1. Due to representation (4.14), mapping $I(\cdot)$ is differentiable with respect to Λ. Henceforth assumption (1) of the corollary is also fulfilled.

2. Let us note at first that, thanks to Proposition 2.3, we have $\mathscr{D}(I) = \mathbb{R}^d$. Thus, the assumption (2) of Corollary 3.1 is satisfied, as is condition (3.9).

Moreover, the assumptions of Lemma 3.3 are satisfied:

1. The concentration of measure condition is satisfied due to Proposition 2.3.

2. The tail decay is obvious since the family $\{\widetilde{P}_N : N \in \mathbb{N}\}$ has compact support. Namely, for all $N \in \mathbb{N}$, we have $\operatorname{supp}\widetilde{P}_N = \mathscr{U}$. Thus the measure $\widetilde{Q}_{N,\Lambda}$ (cf. (3.25)) generated by $\widetilde{P}_N$ has the same support. Thus, $\operatorname{supp}\widetilde{Q}_{N,\Lambda} = \mathscr{U}$.

3. The non-degeneracy is assured by Lemma 4.2.

Hence, due to (4.27), arguing in the same way as in Theorem 3.2, we arrive at (4.26). Note that the $N \uparrow +\infty$ limit of $\mathscr{R}_N(x, \mathscr{Q}, U, B(U, \varepsilon))$ exists, since in (4.27) the limits of the other two N-dependent quantities exist due to [19]. The subsequent $\varepsilon \downarrow +0$ limit of the remainder term exists due to the monotonicity.

Finally, taking the supremum over $U \in \mathscr{V}$ in (4.26), we get (1.22). $\square$

5 Guerra's comparison scheme

In this section, we shall apply Guerra's comparison scheme (see the recent accounts by [18; 32; 2]) to the SK model with multidimensional spins. However, we shall use also the ideas (and the language) of [1]. In particular, we shall use the same local comparison functional (4.6) as in the AS^2 scheme, see (5.4). The section contains the proofs of the upper (5.16) and lower (5.18) bounds on the free energy without Assumption 1.2. The proofs use the GREM-like Gaussian processes, RPCs as in the AS^2 scheme. We also obtain an analytic representation of the remainder term (which is an artifact of this scheme) using the properties of the Bolthausen-Sznitman coalescent.

5.1 Multidimensional Guerra's scheme

Let $\xi = \xi(x_1, \ldots, x_n)$ be an RPC process. Theorem 5.3 of [2] guarantees that there exists a rearrangement $\xi = \{\xi(i)\}_{i\in\mathbb{N}}$ of the ξ's atoms in a decreasing order. Recall (1.16) and define a (random) *limiting ultrametric overlap* $q_{\mathrm{L}} : \mathbb{N}^2 \to [0; n] \cap \mathbb{Z}$ as follows

$$q_{\mathrm{L}}(i, j) \equiv 1 + \max\{k \in [0; n] \cap \mathbb{Z} : [\pi(i)]_k = [\pi(j)]_k\}, \tag{5.1}$$

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

$$i \underset{k}{\sim} j \overset{\text{def}}{\Longleftrightarrow} q_{\mathrm{L}}(i, j) \geq k. \tag{5.2}$$

Given $n \in \mathbb{N}$, assume that x and $\mathscr{Q}$ satisfy (1.8) and (1.7), respectively. Recall the definitions of the Gaussian processes X and A which satisfy (1.1) and (1.17), respectively. We consider, for $t \in [0; 1]$, the following interpolating Hamiltonian on the configuration space $\Sigma_N \times \mathscr{A}$

$$H_t(\sigma, \alpha) \equiv \sqrt{t}X(\sigma) + \sqrt{1-t}A(\sigma, \alpha). \tag{5.3}$$

Given $\mathscr{U} \subset \mathrm{Sym}^+(d)$, the Hamiltonian (5.3) in the usual way induces the following *local free energy*

$$\varphi_N(t, x, Q, \mathscr{U}) \equiv \Phi_N(x, \mathscr{U})\left[H_t\right], \tag{5.4}$$

where we use the same local comparison functional (4.6) as in the AS^2 scheme. Using (1.5), we obtain then

$$\varphi(0, x, Q, \mathscr{U}) = \Phi_N(x, \mathscr{U})[A] \text{ and } \varphi(1, x, Q, \mathscr{U}) = \Phi_N(x, \mathscr{U})[X] = p_N(\mathscr{U}).$$

Now, we are going to disintegrate the Gibbs measure defined on $\mathscr{U} \times \mathscr{A}$ into two Gibbs measures acting on $\mathscr{U}$ and $\mathscr{A}$ separately. For this purpose we define the correspondent (random) *local free energy* on $\mathscr{U}$ as follows

$$\psi(t,x,Q,\alpha,\mathscr{U}) \equiv \log \int_{\Sigma_N(\mathscr{U})} \exp\left[\beta\sqrt{N}H_t(\sigma,\alpha)\right] \mathrm{d}\mu^{\otimes N}(\sigma). \tag{5.5}$$

For $\alpha \in \mathscr{A}$, we can define the (random) *local Gibbs measure* $\mathscr{G}(t,Q,\alpha,\mathscr{U}) \in \mathscr{M}_1(\Sigma_N)$ by demanding that the following holds

$$\frac{\mathrm{d}\mathscr{G}(t,x,Q,\alpha,\mathscr{U})}{\mathrm{d}\mu^{\otimes N}}(\sigma) \equiv \mathbb{1}_{\Sigma_N(\mathscr{U})}(\sigma)\exp\left[\beta\sqrt{N}H_t(\sigma,\alpha) - \psi(t,x,Q,\mathscr{U},\alpha)\right].$$

Let us define a certain reweighting of the RPC ξ with the help of (5.5). We define the random point process $\{\tilde{\xi}\}_{\alpha\in\mathscr{A}}$ in the following way

$$\tilde{\xi}(\alpha) \equiv \xi(\alpha)\exp\left(\psi(t,x,Q,\mathscr{U},\alpha)\right).$$

We also define the *normalisation operation* $\mathscr{N} : \mathscr{M}_{\mathrm{f}}(\mathscr{A}) \to \mathscr{M}_1(\mathscr{A})$ as

$$\mathscr{N}(\xi)(\alpha) \equiv \frac{\xi(\alpha)}{\sum_{\alpha'\in\mathscr{A}}\xi(\alpha')}.$$

We introduce the *local Gibbs measure* $\mathscr{G}(t,x,Q,\mathscr{U}) \in \mathscr{M}_1(\mathscr{U} \times \mathscr{A})$, for any $\mathscr{V} \subset \mathscr{U} \times \mathscr{A}$, as follows

$$\mathscr{G}(t,x,Q,\mathscr{U})[\mathscr{V}] \equiv \sum_{\alpha\in\mathscr{A}_n} \mathscr{N}(\tilde{\xi})(\alpha)\mathscr{G}(t,x,Q,\alpha,\mathscr{U})[\mathscr{V}]. \tag{5.6}$$

Finally, we introduce, what shall call *Guerra's remainder term*:

$$\mathscr{R}(t,x,Q,\mathscr{U}) \equiv -\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,x,Q,\mathscr{U}) \otimes \mathscr{G}(t,x,Q,\mathscr{U})\left[\|R(\sigma^1,\sigma^2) - Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right]\right]. \tag{5.7}$$

Note that (5.7) coincides with (1.20) after substituting (1.18) with (1.23).

5.2 Local comparison

We recall for completeness the following.

Proposition 5.1 (Ruelle [24], Bolthausen and Sznitman [6]). *For any $k \in [1; n+1] \cap \mathbb{N}$, we have*

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

The results of Section 4 can be straightforwardly generalised to the comparison scheme based on (5.3). Given $\varepsilon, \delta > 0$ and $\Lambda \in \mathrm{Sym}(d)$, define

$$\mathscr{V}(\Lambda,\mathscr{U},\varepsilon,\delta) \equiv \{U' \in \mathrm{Sym}(d) : \|U' - U\|_{\mathrm{F}} < \varepsilon, \langle U' - U, \Lambda\rangle < \delta\}. \tag{5.8}$$

We now specialise to the case $\mathscr{U} = \Sigma_N(\mathscr{V}(\Lambda,U,\varepsilon,\delta))$.

Lemma 5.1. *We have*

$$\frac{\partial}{\partial t}\varphi_N(t,x,Q,\mathscr{V}(\Lambda,U,\varepsilon,\delta)) = \mathscr{R}(t,x,Q,\Sigma_N(\mathscr{A}(\Lambda,U,\varepsilon,\delta))) \\ -\frac{\beta^2}{2}\sum_{k=1}^{n} x_k\left(\|Q^{(k+1)}\|_F^2-\|Q^{(k)}\|_F^2\right)+\mathscr{O}(\varepsilon). \tag{5.9}$$

Proof. This is an immediate consequence of Proposition 2.5. Indeed, recalling that $Q(\alpha^1,\alpha^1)=U$, and setting $\mathscr{U}\equiv\Sigma(B(U,\varepsilon))$, we have

$$\begin{aligned}
&\frac{\partial}{\partial t}\varphi(t,x,Q,\mathscr{U}) \\
&=\frac{\beta^2}{2}\mathbb{E}\Big[\mathscr{G}(t,x,Q,\mathscr{U})\otimes\mathscr{G}(t,x,Q,\mathscr{U})\Big[\|R(\sigma^1,\sigma^1)-U\|_{\mathrm{F}}^2-\|R(\sigma^1,\sigma^2)-Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2 \\
&\quad-\left(\|U\|_{\mathrm{F}}^2-\|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\right)\Big]\Big] \\
&=-\frac{\beta^2}{2}\mathbb{E}\Big[\mathscr{G}(t,x,Q,\mathscr{U})\otimes\mathscr{G}(t,x,Q,\mathscr{U})\Big[\|R(\sigma^1,\sigma^2)-Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\Big]\Big] \\
&\quad-\frac{\beta^2}{2}\mathbb{E}\Big[\mathscr{G}(t,x,Q,\mathscr{U})\otimes\mathscr{G}(t,x,Q,\mathscr{U})\Big[\|U\|_{\mathrm{F}}^2-\|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\Big]\Big]+\mathscr{O}(\varepsilon).
\end{aligned}\tag{5.10}$$

Using the Proposition 5.1, we get

$$\begin{aligned}
&\frac{\beta^2}{2}\mathbb{E}\Big[\mathscr{G}(t,x,Q,\mathscr{U})\otimes\mathscr{G}(t,x,Q,\mathscr{U})\Big[\|U\|_{\mathrm{F}}^2-\|Q(\alpha^1,\alpha^2)\|_{\mathrm{F}}^2\Big]\Big] \\
&\quad=\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{N}(\xi)\otimes\mathscr{N}(\xi)\left[\sum_{k=q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})}^{n}\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2-\|Q^{(k)}\|_{\mathrm{F}}^2\right)\right]\right] \\
&\quad=\frac{\beta^2}{2}\sum_{k=1}^{n}\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2-\|Q^{(k)}\|_{\mathrm{F}}^2\right)\mathbb{E}\left[\mathscr{N}(\xi)\otimes\mathscr{N}(\xi)\{k\geq q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})\}\right] \\
&\quad=\frac{\beta^2}{2}\sum_{k=1}^{n}x_k\left(\|Q^{(k+1)}\|_{\mathrm{F}}^2-\|Q^{(k)}\|_{\mathrm{F}}^2\right).
\end{aligned}\tag{5.11}$$

Combining (5.10) and (5.11), we get (5.9)

□

Lemma 5.2. *We have*

$$p_N(\Sigma_N(B(U,\varepsilon))) = \Phi_N(x,\Sigma_N(B(U,\varepsilon)))[A]-\frac{\beta^2}{2}\sum_{k=1}^{n}x_k\left(\|Q^{(k+1)}\|_F^2-\|Q^{(k)}\|_F^2\right) \\ +\int_0^1\mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon))\mathrm{d}t+\mathscr{O}(\varepsilon). \tag{5.12}$$

Remark 5.1. *Note that the above lemma also holds if we substitute $B(U,\varepsilon)$ with the smaller set $\mathscr{V}(\Lambda,U,\varepsilon,\delta)$.*

Proof. The claim follows from (5.9) by integration. □

Proposition 5.2. *There exists $C = C(\Sigma, \mu) > 0$ such that, for all $U \in \mathrm{Sym}^+(d)$ as above, and all $\varepsilon, \delta > 0$, there exists an δ-minimal Lagrange multiplier $\Lambda = \Lambda(U, \varepsilon, \delta) \in \mathrm{Sym}(d)$ in* (1.12) *such that, for all $t \in [0;1]$, and all $(x, \mathcal{Q})$, we have*

$$p_N(\Sigma_N(\mathcal{V}(\Lambda, U, \varepsilon, \delta))) \leq \inf_{\Lambda \in \mathrm{Sym}(d)} f(x, \mathcal{Q}, U, \Lambda) + C(\varepsilon + \delta) \tag{5.13}$$

and

$$\lim_{N\uparrow+\infty} p_N(\Sigma_N(B(U, \varepsilon))) \geq \inf_{\Lambda \in \mathrm{Sym}(d)} f(x, \mathcal{Q}, U, \Lambda) + \lim_{N\uparrow+\infty} \int_0^1 \mathcal{R}(t, x, Q, \Sigma_N(B(U, \varepsilon)))\mathrm{d}t \\ - C(\varepsilon + \delta). \tag{5.14}$$

Remark 5.2. *The following upper bound also holds true. There exists $C = C(\Sigma, \mu) > 0$, such that, for any $\Lambda \in \mathrm{Sym}(d)$,*

$$p_N(\Sigma_N(B(U, \varepsilon))) \leq f(x, \mathcal{Q}, U, \Lambda) + C\|\Lambda\|_F \varepsilon. \tag{5.15}$$

Proof. The result follows from Lemma 5.2 by the same arguments as in the proofs of Theorems 1.1 and 1.2. □

5.3 Free energy upper and lower bounds

Similarly to the quenched LDP bounds for the AS^2 scheme in the SK model with multidimensional spins (see Section 3), we get the quenched LDP bounds for Guerra's scheme in the same model without Assumption 1.2 on Q.

Recall the definition of the local Parisi functional f (1.12).

Theorem 5.1. *For any closed set $\mathcal{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathcal{V}) \leq \sup_{U \in \mathcal{V} \cap \mathcal{U}} \inf_{(x, \mathcal{Q}, \Lambda)} f(x, \mathcal{Q}, \Lambda, U), \tag{5.16}$$

where the infimum runs over all x satisfying (1.8)*, all $\mathcal{Q}$ satisfying* (1.7) *and all $\Lambda \in \mathrm{Sym}(d)$.*

Proof. The proof is identical to the one of Theorem 1.1. □

Define the *local limiting Guerra remainder term $\mathcal{R}(x, \mathcal{Q}, U)$* as follows

$$\mathcal{R}(x, \mathcal{Q}, U) \equiv -\lim_{\varepsilon\downarrow+0} \lim_{N\uparrow+\infty} \int_0^1 \mathcal{R}(t, \Sigma_N(B(U, \varepsilon)))\mathrm{d}t \leq 0. \tag{5.17}$$

The existence of the limits in (5.17) is proved similar to the case of the AS^2 scheme, see the proof of Theorem 1.2.

Theorem 5.2. *For any open set $\mathscr{V} \subset \mathrm{Sym}(d)$, we have*

$$p(\mathscr{V}) \geq \sup_{U \in \mathscr{V} \cap \mathscr{U}} \inf_{(x,\mathscr{Q},\Lambda)} \left[f(x,\mathscr{Q},\Lambda,U) + \mathscr{R}(x,\mathscr{Q},U) \right], \tag{5.18}$$

where the infimum runs over all x satisfying (1.8)*; all $\mathscr{Q}$ satisfying* (1.7) *and all $\Lambda \in \mathrm{Sym}(d)$.*

Proof. The proof is identical to the one of Theorem 1.2. The only new ingredient is Lemma 5.1 needed to recover Guerra's remainder term (5.7). □

5.4 The filtered d-dimensional GREM

Given $U \in \mathrm{Sym}^+(d)$ non-negative definite, denote by $\mathscr{Q}(U,d)$ the set of all cÃădlÃăg (right continuous with left limits) $\mathrm{Sym}^+(d)$-valued non-decreasing paths which end in matrix U, i.e.,

$$\mathscr{Q}(U,d) \equiv \{\rho : [0;1] \to \mathrm{Sym}^+(d) \mid \rho(0) = 0; \rho(1) = U; \rho(t) \preceq \rho(s), \text{ for } t \leq s; \rho \text{ is cádlág}\}. \tag{5.19}$$

Define the natural inverse $\rho^{-1} : \mathrm{Im}\,\rho \to [0;1]$ as

$$\rho^{-1}(Q) \equiv \inf\{t \in [0;1] \mid \rho(t) \succeq Q\},$$

where $\mathrm{Im}\,\rho \equiv \rho([0;1])$. Let $x \equiv \rho^{-1} \circ \rho \in \mathscr{Q}(1,1)$.

Let also $\mathscr{Q}'(U,d) \subset \mathscr{Q}(U,d)$ be the space of all piece-wise constant paths in $\mathscr{Q}(U,d)$ with finite (but arbitrary) number of jumps with an additional requirement that they have a jump at $x = 1$. Given some $\rho \in \mathscr{Q}'(U,d)$, we enumerate its jumps and define the finite collection of matrices $\{Q^{(k)}\}_{k=0}^{n+1} \equiv \mathrm{Im}\,\rho \subset \mathbb{R}^d$. This implies that there exist $\{x_k\}_{k=0}^{n+1} \subset \mathbb{R}$ such that

$$0 \equiv x_0 < x_1 < \ldots < x_n < x_{n+1} \equiv 1,$$
$$0 \equiv Q^{(0)} \preceq Q^{(1)} \preceq Q^{(2)} \preceq \cdots \preceq Q^{(n+1)} \equiv U,$$

where $\rho(x_k) = Q^{(k)}$. Let us associate to $\rho \in \mathscr{Q}'(U,d)$ a new path $\tilde{\rho} \in \mathscr{Q}(U,d)$ which is obtained by the linear interpolation of the path ρ. Namely, let

$$\tilde{\rho}(t) \equiv Q^{(k)} + (Q^{(k+1)} - Q^{(k)}) \frac{t - x_k}{x_{k+1} - x_k}, t \in [x_k; x_{k+1}).$$

Let $g : \mathbb{R}^d \to \mathbb{R}$ be a function satisfying Assumption 5.1. Let us introduce the *filtered d-dimensional GREM process W*. Let

$$W \equiv \left\{ \{W_k(t, [\alpha]_k)\}_{t \in \mathbb{R}_+} : \alpha \in \mathscr{A}, k \in [0;n] \cap \mathbb{N} \right\}$$

be the collection of independent (for different α and k) $\mathbb{R}^d$-valued correlated Brownian motions satisfying

$$W_k(t, [\alpha]_k) \sim (Q^{(k+1)} - Q^{(k)})^{1/2} W \left(\frac{t - x_k}{x_{k+1} - x_k} \right),$$

where $\{W(t)\}_{t \in \mathbb{R}_+}$ is the standard (uncorrelated) $\mathbb{R}^d$-valued Brownian motion. Now, for $k \in [0;n] \cap \mathbb{N}$, we define the $\mathbb{R}^d$-valued process $\{Y(t,\alpha) \mid \alpha \in \mathscr{A}, t \in [0;1]\}$ by

$$Y(t,\alpha) \equiv \sum_{k=0}^{n} \mathbb{1}_{[x_k;1]}(t) W_k(t \wedge x_{k+1}, [\alpha]_k).$$

Lemma 5.3. *For $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}$, we have*

$$\operatorname{Cov}\left[Y(t_1, \alpha^{(1)}), Y(t_2, \alpha^{(2)})\right] = \tilde{\rho}\left(t_1 \wedge t_2 \wedge x_{q_L(\alpha^{(1)}, \alpha^{(2)})}\right).$$

Proof. The proof is straightforward. □

Assumption 5.1. *Suppose that the function $g : \mathbb{R}^d \to \mathbb{R}$ satisfies $g \in C^{(2)}(\mathbb{R}^d)$ and, for any $c > 0$, we have $\int_{\mathbb{R}^d} \exp\left(g(y) - c\|y\|_2^2\right) \mathrm{d}y < \infty$ and also*

$$\sup_{y \in \mathbb{R}^d} \left(\|\nabla g(y)\|_2 + \|\nabla^2 g(y)\|_2\right) < +\infty, \tag{5.20}$$

where $\nabla^2 g(y)$ denotes the matrix of second derivatives of the function g at $y \in \mathbb{R}^d$.

Assume g satisfies the above assumption. Let $f \equiv f_\rho : [0;1] \times \mathbb{R}^d \to \mathbb{R}$ be the function satisfying the following (backward) recursive definition

$$f(t, y) \equiv \begin{cases} g(y), & t = 1, \\ \frac{1}{x_k} \log \mathbb{E}\left[\exp\left\{x_k f(x_{k+1}, y + Y(x_{k+1}, \alpha) - Y(t, \alpha))\right\}\right], & t \in [x_k; x_{k+1}), \end{cases} \tag{5.21}$$

where $k \in [0; n] \cap \mathbb{N}$, $\alpha \in \mathscr{A}$ is arbitrary and fixed.

Remark 5.3. *It is easy to recognise that the definition of f is a continuous "algorithmisation" of* (1.11). *Namely, $X_k(x, \mathscr{Q}, \Lambda, U) = f(x_k, 0)$, where*

$$f(1, y) = g(y) \equiv \log \int_\Sigma \exp\left(\sqrt{2}\beta \langle y, \sigma \rangle + \langle \Lambda\sigma, \sigma \rangle\right) \mathrm{d}\mu(\sigma). \tag{5.22}$$

5.5 A computation of the remainder term

Recall the equivalence relation (5.2). In words, the equivalence $i \underset{k}{\sim} j$ means that the atoms of the RPC ξ with ranks i and j have the same ancestors up to the k-th generation. Varying the k in (5.2), we get a family of equivalences on $\mathbb{N}$ which possesses important Markovian properties, see [6].

Lemma 5.4. *For all $k \in [0; n-1] \cap \mathbb{N}$, we have*

$$\mathbb{E}\left[\sum_{\substack{i \underset{k}{\sim} j \\ i \underset{k+1}{\nsim} j}} \mathscr{N}(\xi)(i) \mathscr{N}(\xi)(j)\right] = x_{k+1} - x_k, \tag{5.23}$$

and also

$$\mathbb{E}\left[\sum_i \mathscr{N}(\xi)(i)^2\right] = 1 - x_n. \tag{5.24}$$

Proof. 1. To prove (5.23) we notice that

$$\mathbb{E}\left[\sum_{\substack{i \underset{k}{\sim} j \\ i \underset{k+1}{\nsim} j}} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right] = \mathbb{E}\left[\sum_{i \underset{k+1}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j) - \sum_{i \underset{k}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right]$$
$$= x_{k+1} - x_k,$$

where the last equality is due to Proposition 5.1.

2. Similarly, (5.24) follows from the following observation

$$\mathbb{E}\left[\sum_{i} \mathcal{N}^2(\xi)(i)\right] = \mathbb{E}\left[\sum_{i,j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j) - \sum_{i \underset{n}{\nsim} j} \mathcal{N}(\xi)(i)\mathcal{N}(\xi)(j)\right]$$
$$= 1 - x_n,$$

where the last equality is due to Proposition 5.1.

□

Note that, using the above notations, we readily have

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

where $\{Y^{(i)} \equiv \{Y^{(i)}(1, \alpha)\}_{\alpha \in \mathscr{A}}\}_{i=1}^{N}$ are i.i.d. copies of $\{Y(1, \alpha)\}_{\alpha \in \mathscr{A}}$. Consider the following weights

$$\tilde{\xi}^{(t)}(\alpha) \equiv \xi(\alpha) \exp\left(f(t, Y(t, \alpha))\right).$$

As in [6], the above weights induce the permutation $\tilde{\pi}^{(t)} : \mathbb{N} \to \mathscr{A}$ such that, for all $i \in \mathbb{N}$, the following holds

$$\tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i)) > \tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i+1)). \tag{5.25}$$

In what follows, we shall use the short-hand notations $\tilde{\xi}^{(t)}(i) \equiv \tilde{\xi}^{(t)}(\tilde{\pi}^{(t)}(i))$, $\tilde{Y}^{(t)}(s, i) \equiv Y(s, \tilde{\pi}^{(t)}(i))$ and $\tilde{Q}^{(t)} \equiv \{\tilde{Q}^{(t)}(i, j) \equiv Q(\tilde{\pi}^{(t)}(i), \tilde{\pi}^{(t)}(j))\}_{i,j \in \mathbb{N}}$.

Theorem 5.3. *Given a discrete order parameter $x \in \mathscr{Q}'(1, 1)$, we have*

1. Independence #1. *The normalised RPC point process $\mathcal{N}(\xi)$ is independent from the corresponding randomised limiting GREM overlaps q.*

2. Independence #2. *The reordered filtered limiting GREM $\tilde{Y}$ is independent from the corresponding reordered weights $\tilde{\xi}$.*

3. The reordering change of measure. *Given $I \Subset \mathbb{N}$, let $\nu_I(\cdot|Q)$ be the joint distribution of $\{Y(1,i)\}_{i\in I}$, and $\tilde{\nu}_I(\cdot|Q)$ be the joint distribution of $\{\tilde{Y}^{(1)}(1,i)\}_{i\in I}$ both conditional on Q. Then*

$$\frac{\mathrm{d}\tilde{\nu}_I(\cdot|Q)}{\mathrm{d}\nu_I(\cdot|Q)} = \prod_{k=0}^{n} \prod_{i\in\left(I/\underset{k}{\sim}\right)} \exp\left(x_k\left\{f(x_{k+1},Y(x_{k+1},i)) - f_k(x_k Y(x_k,i))\right\}\right), \tag{5.26}$$

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

4. The averaging property. *For all $s,t\in[0;1]$, we have*

$$\left(\left\{\xi^{(t)}(\alpha)\right\}_{\alpha\in\mathscr{A}}, \tilde{Q}^{(t)}\right) \sim \left(\left\{\xi^{(s)}(\alpha)\right\}_{\alpha\in\mathscr{A}}, \tilde{Q}^{(s)}\right). \tag{5.27}$$

Proof. The proof is the same as in the case of the one-dimensional SK model, see [6; 3]. □

Keeping in mind (5.26), we define, for $k\in[0;n-1]\cap\mathbb{N}$, the following random variables

$$T_k(\alpha) \equiv \exp\left(x_k\left[f(x_{k+1},Y(x_{k+1},\alpha)) - f(x_k,Y(x_k,\alpha))\right]\right).$$

Given $k\in[1;n]\cap\mathbb{N}$, assume that $\alpha^{(1)},\alpha^{(2)}\in\mathscr{A}$ satisfy $q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}) = k$. We introduce, for notational convenience, the (random) measure $\mu_k(t,\mathscr{U})$ – an element of $\mathscr{M}_1(\Sigma_N)$ – by demanding the following

$$\mu_k(t,\mathscr{U})\left[g\right] \equiv \mathbb{E}\Big[T_1(\alpha^1)\cdots T_k(\alpha^1)T_{k+1}(\alpha^1)T_{k+1}(\alpha^2)\cdots T_n(\alpha^1)T_n(\alpha^2) \\ \mathscr{G}(t,\alpha^{(1)},\mathscr{U})\otimes\mathscr{G}(t,\alpha^{(2)},\mathscr{U})\left[g\right]\Big], \tag{5.28}$$

where $g:\mathscr{U}^2\to\mathbb{R}$ is an arbitrary measurable function such that (5.28) is finite. Using this notation, we can state the following lemma.

Lemma 5.5. *For any $i,j\in\mathbb{N}$, satisfying $i\underset{k}{\sim}j$, $i\underset{k+1}{\nsim}j$, we have*

$$\mathbb{E}\left[\mathscr{G}(t,i,\mathscr{U})\otimes\mathscr{G}(t,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_F^2\right]\right] = \mu_k(t,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q^{(k)}\|_F^2\right]. \tag{5.29}$$

Proof. This is a direct consequence of (5.26) and the fact that under the assumptions of the theorem $Q(i,j)=Q^{(k)}$.

□

Remark 5.4. *It is obvious from the previous theorem that μ_k is a probability measure.*

The main result of this subsection is an "analytic projection" of the probabilistic RPC representation which integrates out the dependence on the RPC. Comparing to (1.20), it has a more analytic flavor which will be exploited in the remainder estimates (Section 7). This is also a drawback in some sense, since the initial beauty of the RPCs is lost.

Theorem 5.4. *In the case of Guerra's interpolation* (1.23)*, we have*

$$\begin{aligned}\mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon))) = &\frac{1}{2}\sum_{k=0}^{n-1}(x_{k+1}-x_k)\mu_k(t,\Sigma_N(B(U,\varepsilon)))\left[\|R(\sigma^1,\sigma^2)-Q^{(k)}\|_F^2\right]\\ &+\mathscr{O}(\varepsilon)+\mathscr{O}(1-x_n),\end{aligned} \tag{5.30}$$

as $\varepsilon\to 0$ *and* $x_n\to 1$.

Proof. Recalling (5.7) and (5.6), we write

$$\begin{aligned}\mathscr{R}(t,x,Q,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\mathbb{E}\Big[\sum_{i,j}&\mathscr{N}(\tilde{\xi})(i)\mathscr{N}(\tilde{\xi})(j)\\ &\times\mathscr{G}(t,x,Q,i,\mathscr{U})\otimes\mathscr{G}(t,x,Q,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\right]\Big].\end{aligned}$$

Using Theorem 5.3, we arrive to

$$\begin{aligned}\mathscr{R}(t,x,Q,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\sum_{i,j}&\mathbb{E}\left[\mathscr{N}(\tilde{\xi})(i)\mathscr{N}(\tilde{\xi})(j)\right]\\ &\times\mathbb{E}\left[\mathscr{G}(t,x,Q,i,\mathscr{U})\otimes\mathscr{G}(t,x,Q,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\right]\right].\end{aligned}$$

(We can interchange the summation and expectation since all summands are non-negative.) The averaging property (see Theorem 5.3) then gives

$$\mathscr{R}(t,\Sigma(U,\varepsilon)) = \frac{\beta^2}{2}\sum_{i,j}\mathbb{E}\left[\mathscr{N}(\xi)(i)\mathscr{N}(\xi)(j)\right]\mathbb{E}\left[\mathscr{G}(t,i,\mathscr{U})\otimes\mathscr{G}(t,j,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q(i,j)\|_{\mathrm{F}}^2\right]\right]. \tag{5.31}$$

For each $k\in[1;n-1]\cap\mathbb{N}$, we fix any indexes $i_0, i_0^{(k)}, j_0^{(k)}\in\mathbb{N}$ such that $i\underset{k}{\sim}j$ and $i\underset{k+1}{\nsim}j$. Rearranging the terms in (5.31), we get

$$\begin{aligned}\mathscr{R}(t,\Sigma(U,\varepsilon)) = &\frac{\beta^2}{2}\sum_{k=1}^{n}\mathbb{E}\left[\mathscr{G}(t,i_0^{(k)},\mathscr{U})\otimes\mathscr{G}(t,j_0^{(k)},\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-Q^{(k)}\|_{\mathrm{F}}^2\right]\right]\\ &\times\sum_{\substack{i\underset{k}{\sim}j\\ i\underset{k+1}{\nsim}j}}\mathbb{E}\left[\mathscr{N}(\xi)(i)\mathscr{N}(\xi)(j)\right]\\ &+\frac{\beta^2}{2}\mathbb{E}\left[\mathscr{G}(t,i_0,\mathscr{U})\otimes\mathscr{G}(t,i_0,\mathscr{U})\left[\|R(\sigma^1,\sigma^2)-U\|_{\mathrm{F}}^2\right]\right]\sum_{i}\mathbb{E}\left[\mathscr{N}(\xi)(i)^2\right].\end{aligned} \tag{5.32}$$

Finally, applying Lemmata 5.4 and 5.5 to (5.32), we arrive at (5.30).

□

6 The Parisi functional in terms of differential equations

In this section, we study the properties of the multidimensional Parisi functional. We derive the multidimensional version of the Parisi PDE. This allows to represent the Parisi functional as a solution of a PDE evaluated at the origin. We also obtain a variational representation of the Parisi functional in terms of a HJB equation for a linear problem of diffusion control. As a by-product, we arrive at the strict convexity of the Parisi functional in 1-D which settles a problem of uniqueness of the optimal Parisi order parameter posed by [31; 20].

Lemma 6.1. *Consider the function $B : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ defined as*

$$B(y,t) \equiv \frac{1}{x} \log \mathbb{E}\left[\exp\left\{x f(y + z(t))\right\}\right],$$

where $f : \mathbb{R}^d \to \mathbb{R}$ satisfies Assumption 5.1 and $\{z(t)\}_{t\in[0;1]}$ is a Gaussian $\mathbb{R}^d$-valued process with $\mathrm{Cov}\left[z(t)\right] \equiv Q(t) \in \mathrm{Sym}(d)$ such that $Q(t)_{u,v}$ is differentiable, for all u,v. Then

$$\partial_t B(y,t) = \frac{1}{2}\sum_{u,v=1}^{d} \dot{Q}_{u,v}(t)\left(\partial^2_{y_u y_v} B(y,t) + x\,\partial_{y_u} B(y,t)\partial_{y_v} B(y,t)\right), \quad (t,y) \in (0;1)\times\mathbb{R}^d. \tag{6.1}$$

In particular, the function B is differentiable with respect to the t-variable on $(0;1)$ and $C^2(\mathbb{R}^d)$ with respect to the y-variable.

Proof. Denote $Z \equiv \mathbb{E}\left[\mathrm{e}^{xf(y+z(t))}\right]$. By [2, Lemma A.1], we have

$$\partial_t B(y,t) = \frac{1}{2x}\left(\frac{1}{Z}\mathbb{E}\Big[\sum_{u,v=1}^{d}\dot{Q}_{u,v}(t)\partial^2_{z_u z_v}\mathrm{e}^{xf(z)}|_{z=y+z(t)}\Big]\right).$$

A straightforward calculation then gives

$$\partial_t B(y,t) = \frac{1}{2x}\left(\frac{1}{Z}\mathbb{E}\Big[\sum_{u,v=1}^{d}\dot{Q}_{u,v}(t)\left(x^2\partial_{z_u} f(z)\partial_{z_v} f(z) + x\partial^2_{z_u z_v} f(z)\right)\mathrm{e}^{xf(z)}|_{z=y+z(t)}\Big]\right). \tag{6.2}$$

We also have

$$\partial_{y_u} B(y,t) = \frac{1}{xZ}\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_u} f(z)|_{z=y+z(t)}\right], \tag{6.3}$$

and

$$\begin{aligned}\partial^2_{y_u y_v} B(y,t) = \frac{1}{x}\Big(&\frac{1}{Z}\mathbb{E}\left[\mathrm{e}^{xf(z)}\left(x^2\partial_{z_u} f(z)\partial_{z_v} f(z) + \partial^2_{z_u z_v} f(z)\right)|_{z=y+z(t)}\right]\\ &- \frac{1}{Z^2}\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_u} f(z)|_{z=y+z(t)}\right]\mathbb{E}\left[x\mathrm{e}^{xf(z)}\partial_{z_v} f(z)|_{z=y+z(t)}\right]\Big).\end{aligned} \tag{6.4}$$

Combining (6.2), (6.3) and (6.4), we get (6.1). $\square$

Proposition 6.1. *Denote $D \equiv \bigcup_{k=0}^{n}(x_k;x_{k+1})$. The function $f = f_\rho$ defined in* (5.21) *satisfies the final-value problem for the controlled semi-linear parabolic Parisi-type PDE*

$$\begin{cases} \partial_t f(y,t) + \frac{1}{2}\sum_{u,v=1}^{d} \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}_{u,v}(t)\left(\partial^2_{y_u y_v} f(y,t) + x(t)\partial_{y_u} f(y,t)\partial_{y_v} f(y,t)\right) = 0, & (t,y) \in D \times \mathbb{R}^d, \\ f(1,y) = g(y), & y \in \mathbb{R}^d, \\ f(y,x_k-0) = f(y,x_k+0), \quad k \in [1;n] \cap \mathbb{N}, & y \in \mathbb{R}^d. \end{cases} \tag{6.5}$$

Note that $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\rho}(t) = \frac{Q^{(k+1)}-Q^{(k)}}{x_{k+1}-x_k}$, for $t \in (x_k;x_{k+1})$.

Proof. A successive application of Lemma 6.1 to (5.21) on the intervals D starting from $(x_n;1)$ gives (6.5). □

Remark 6.1. *Note that a straightforward inspection of* (5.21)*, using* (6.2)*,* (6.3) *and* (6.4)*, shows that the function f defined in* (5.21) *is $C^1(D) \cap C([0;1])$ with respect to the t-variable and $C^2(\mathbb{R}^d)$ with respect to the y-variable.*

Lemma 6.2. *Given $\rho \in \mathscr{Q}'(U,d)$, the function* (5.21) *satisfies the following:*

$$f_\rho(0,0) = \mathbb{E}\left[\log \sum_{\alpha \in \mathscr{A}} \xi(\alpha)\exp\left\{g(Y(1,\alpha))\right\}\right]. \tag{6.6}$$

Proof. This is an immediate consequence of the RPC averaging property (5.27). □

Lemma 6.3.

1. *Given $k \in [1;n] \cap \mathbb{N}$ and a non-negative definite matrix $Q \in \mathrm{Sym}(d)$, we have*

$$\partial_{Q^{(k)} \rightsquigarrow Q} f_\rho(0,0) = -\frac{1}{2}(x_k - x_{k-1})\mathbb{E}\left[\langle Q, M\rangle\right], \tag{6.7}$$

   where $M \in \mathbb{R}^{d\times d}$ is defined as

$$\begin{aligned} M_{u,v} \equiv & T_1(\alpha^{(1)})\cdots T_k(\alpha^{(1)})T_{k+1}(\alpha^{(1)})T_{k+1}(\alpha^{(2)})\cdots T_n(\alpha^{(1)})T_n(\alpha^{(2)}) \\ & \partial_{z_u} g(z)|_{z=Y(1,\alpha^{(1)})}\partial_{z_v} g(z)|_{z=Y(1,\alpha^{(2)})} \end{aligned}$$

   with $q_L(\alpha^{(1)},\alpha^{(2)}) = k$. Moreover, (6.7) *does not depend on the choice of $\alpha^{(1)},\alpha^{(2)} \in \mathscr{A}$ but only on k.*

2. *Given a non-negative definite matrix $Q \in \mathrm{Sym}(d)$, we have*

$$\partial_{U \rightsquigarrow Q} f_\rho(0,0) = \frac{1}{2}\mathbb{E}\left[\langle Q, M'\rangle\right], \tag{6.8}$$

   where $M' \in \mathrm{Sym}(d)$ is satisfies

$$M'_{u,v} = T_1(\alpha)\cdots T_n(\alpha)\left(\partial^2_{z_u z_v} g(z) + \partial_{z_u} g(z)\partial_{z_v} g(z)\right)\Big|_{z=Y(1,\alpha)} + \mathscr{O}(1-x_n),$$

   as $x_n \to 1$. Note that (6.8) *obviously does not depend on the choice of $\alpha \in \mathscr{A}$.*

Proof. Applying [2, Lemma A.1] to (6.6), we obtain

$$\partial_s \mathbb{E}\Big[\log \sum_{\alpha\in\mathscr{A}} \exp\{g(Y(1,\alpha))\}\Big|_{Q^{(k)}=Q^{(k)}+sQ}\Big]$$
$$= \frac{1}{2}\mathbb{E}\Big[\sum_{u,v=1}^{d} \mathscr{N}(\tilde{\xi})\otimes\mathscr{N}(\tilde{\xi})\Big[\partial_s\left(Q(\alpha^{(1)},\alpha^{(2)})_{u,v}|_{Q^{(k)}=Q^{(k)}+sQ}\right)$$
$$\Big\{\mathbb{1}_{\alpha^{(1)}=\alpha^{(2)}}(\alpha^{(1)},\alpha^{(2)})\Big(\partial^2_{z_u z_v} g(z)+\partial_{z_u} g(z)\partial_{z_v} g(z)\Big)\Big|_{z=Y(1,\alpha^{(1)})}$$
$$-\partial_{z_u} g(z)|_{z=Y(1,\alpha^{(1)})}\partial_{z_v} g(z)|_{z=Y(1,\alpha^{(2)})}\Big\}\Big|_{Q^{(k)}=Q^{(k)}+sQ}\Big]\Big].$$

Note that

$$\partial_s\left(Q(\alpha^{(1)},\alpha^{(2)})_{u,v}|_{Q^{(k)}=Q^{(k)}+sQ}\right)=\begin{cases}Q_{u,v}, & q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})=k,\\ 0, & q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})\neq k.\end{cases}$$

1. Define $M(\alpha^{(1)},\alpha^{(2)})\in\mathbb{R}^{d\times d}$ as

$$M(\alpha^{(1)},\alpha^{(2)})_{u,v}\equiv \partial_{z_u} g(z)|_{z=Y(1,\alpha^{(1)})}\partial_{z_v} g(z)|_{z=Y(1,\alpha^{(2)})}.$$

   Hence, we arrive at

$$\partial_{Q^{(k)}\rightsquigarrow Q} f_\rho(0,0)=-\frac{1}{2}\mathbb{E}\Big[\sum_{\alpha^{(1)}\alpha^{(2)}\in\mathscr{A}} \mathbb{1}_{q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)})=k}\,\xi(\alpha^{(1)})\xi(\alpha^{(2)})(\alpha^{(1)},\alpha^{(2)})\langle Q, M(\alpha^{(1)},\alpha^{(2)})\rangle\Big].$$

   The proof is concluded similarly to the proof of Theorem 5.4 by using the properties of the RPC (Theorem 5.3 and Lemma 5.4).

2. The proof is the same as in (1).

□

The following is a multidimensional version of [30, Lemma 4.3].

Lemma 6.4. *For any $\alpha\in\mathscr{A}$, we have*

1.

$$\partial_{x_k} f_\rho(0,0)|_{x_k=x_{k-1}}=\frac{1}{x_{k-1}}\mathbb{E}\Big[T_1(\alpha)\cdots T_{k-2}(\alpha)T_{k-1}(\alpha)|_{x_k=x_{k-1}}$$
$$\Big(\mathbb{E}\Big[f(x_{k+1},Y(x_{k+1},\alpha))T_k(\alpha)|_{x_k=x_{k-1}}\Big]-f(x_k,Y(x_k,\alpha))\Big)\Big].$$

2. Let $M\in\mathrm{Sym}(d)$ with $M_{u,v}\equiv\partial_{z_u} f(x_k,Y(x_k,\alpha))\partial_{z_v} f(x_k,Y(x_k,\alpha))$, then

$$\partial^2_{Q^{(k)}\rightsquigarrow Q,x_k} f_\rho(0,0)=\frac{1}{2}\mathbb{E}\Big[T_1(\alpha)\cdots T_{k-2}(\alpha)\langle Q,M\rangle\Big].$$

Proof. This proof is the same as in [30]. □

We now generalise the PDE (6.5). Given a piece-wise continuous $x \in \mathscr{Q}(1,1)$ and $Q \in \mathscr{Q}(U,d)$, consider the following terminal value problem

$$\begin{cases} \partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (0,1), \\ f(y,1) = g(y). \end{cases} \tag{6.9}$$

We say that $f \in C([0;1] \times \mathbb{R}^d \to \mathbb{R})$ is a piece-wise viscosity solution of (6.9), if there exists the partition of the unit segment $0 =: x_0 < x_1 < \ldots < x_{n+1} \equiv 1$ such that, for each $k \in [0,n] \cap \mathbb{N}$, $f : (x_k; x_{k+1}) \times \mathbb{R}^d \to \mathbb{R}$ is a viscosity solution (see, e.g., [9]) of

$$\begin{cases} \partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (x_k, x_{k+1}), \\ f(y, x_{k+1}+0) = f(y, x_{k+1}-0), \\ f(y,1) = g(y). \end{cases}$$

Proposition 6.2. *For any $\rho^{(1)}, \rho^{(2)} \in \mathscr{Q}'(U,d)$, we have*

$$|f_{\rho^{(1)}}(0,0) - f_{\rho^{(2)}}(0,0)| \leq \frac{C}{2}\int_0^1 \|\rho^{(1)}(t) - \rho^{(2)}(t)\|_F \mathrm{d}t,$$

where $C = C(\Sigma) \equiv \mathbb{E}\left[\|M\|_F\right]$.

Proof. This is an adaptation of the proof of [31, Theorem 3.1] to the multidimensional case. Assume without loss of generality that the paths $\rho^{(1)}$ and $\rho^{(2)}$ have same jump times $\{x_k\}_{k=0}^{n+1}$. Denote the corresponding overlap matrices as $\{Q^{(1,k)}\}_{k=0}^{n+1}$ and $\{Q^{(2,k)}\}_{k=0}^{n+1}$. Given $s \in [0;1]$, define the new path $\rho(s) \in \mathscr{Q}'(U,d)$ by assuming that it has the same jump times $\{x_k\}_{k=0}^{n+1}$ as the paths $\rho^{(1)}, \rho^{(2)}$ and defining its overlap matrices as $Q^{(k)}(s) \equiv sQ^{(1,k)} + (1-s)Q^{(2,k)}$. On the one hand, we readily have

$$\int_0^1 \|\rho^{(1)}(t) - \rho^{(2)}(t)\|_{\mathrm{F}} \mathrm{d}t = \sum_{k=1}^{n} (x_k - x_{k-1})\|Q^{(1,k)} - Q^{(2,k)}\|_{\mathrm{F}}.$$

On the other hand, using Lemma 6.3, we have

$$|\partial_s f_{\rho(s)}(0,0)| \leq \frac{C}{2}\sum_{k=1}^{n}(x_k - x_{k-1})\|Q^{(1,k)} - Q^{(2,k)}\|_{\mathrm{F}}.$$

Finally, we have

$$|f_{\rho^{(1)}}(0,0) - f_{\rho^{(2)}}(0,0)| \leq \int_0^1 |\partial_s f_{\rho(s)}(0,0)| \mathrm{d}s.$$

Combining the last three formulae, we get the theorem. ☐

Remark 6.2. *Note that using the same argument and notations as in the previous theorem we get that, for any $(y,t) \in \mathbb{R}^d \times [0;1]$,*

$$|f_{\rho^{(1)}}(y,t) - f_{\rho^{(2)}}(y,t)| \leq \frac{C(\Sigma)}{2}\int_t^1 \|\rho^{(1)}(s) - \rho^{(2)}(s)\|_F \mathrm{d}s.$$

Remark 6.3. *Note that we can associate to each $\rho \in \mathscr{Q}(U,d)$ a $\mathrm{Sym}^+(d)$-valued countably additive vector measure $\nu_\rho \in \mathscr{M}([0;1], \mathrm{Sym}^+(d))$ by the following standard procedure. Given $[a;b) \subset [0;1]$, define*

$$\nu_\rho([a;b)) \equiv \rho(b) - \rho(a)$$

and then extend the measure, e.g., to all Borell subsets of $[0;1]$.

Theorem 6.1. *Given $U \in \mathrm{Sym}^+(d)$, we have*

1. *The set $\mathscr{Q}(U,d)$ is compact under the topology induced by the following norm*

$$\|\rho\| \equiv \int_0^1 \|\rho(t)\|_F \mathrm{d}t, \quad \rho \in \mathscr{Q}(U,d). \tag{6.10}$$

2. *The functional $\mathscr{Q}'(U,d) \ni \rho \mapsto f_\rho(0,0)$ is Lipschitzian and can be uniquely extended by continuity to the whole $\mathscr{Q}(U,d)$.*

Proof. 1. The topology induced by the norm (6.10) coincides with the topology of weak convergence of the above-defined vector measures. Since $\mathscr{Q}(U,d)$ is a bounded set, it is compact in the weak topology.

2. This is an immediate consequence of Proposition 6.2.

$\square$

In the next result, we summarise some results on the PDE (6.9) for the non-discrete parameters, cf. Proposition 6.1.

Theorem 6.2.

1. Existence. *Assume that Q is in $\mathscr{Q}(U,d)$ and is piece-wise $C^{(1)}$. Assume also that x is in $\mathscr{Q}(1,1)$ and is piece-wise continuous. Then the terminal value problem* (6.9) *has a unique continuous, piece-wise viscosity solution $f_{Q,x} \in C([0;1] \times \mathbb{R}^d)$.*

2. Monotonicity with respect to x. *Assume $Q \in \mathscr{Q}(U,d)$. Assume also that $x^{(1)}, x^{(2)} \in \mathscr{Q}(1,1)$ are such that $x^{(1)}(t) \leq x^{(2)}(t)$, almost everywhere for $t \in [0;1]$. Let $f_{Q,x^{(1)}}$ and $f_{Q,x^{(2)}}$ be the corresponding solutions of* (6.9). *Then $f_{Q,x^{(1)}} \leq f_{Q,x^{(2)}}$.*

3. Monotonicity with respect to g. *Assume $g_1, g_2 : \mathbb{R}^d \to \mathbb{R}$ satisfy Assumption 5.1 and also $g_1 \leq g_2$ almost everywhere. Let $f_{g_1}, f_{g_2} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ be the corresponding solutions of* (6.9) *with $g = g_1$, $g = g_2$, respectively. Then $f_{g_1} \leq f_{g_2}$.*

Proof. 1. Due to the assumptions, the diffusion matrix $\dot{Q}(t) = \dot{\rho}(t)$ in (6.9) is non-negative definite. Applying [9, Proposition 8] to the PDE (6.9) successively on the intervals $[x_k; x_{k+1})$, where the $\dot{\rho}$ is continuous, gives the existence of the solutions in viscosity sense and, moreover, gives their continuity. Uniqueness is ensured by [12, Theorem 1.1].

2. By the approximation argument (cf. Theorem 6.1), it is enough to assume that $x^{(1)}, x^{(2)} \in \mathscr{Q}'(1,1)$ and $Q \in \mathscr{Q}'(U,d)$. Then Proposition 6.1 gives the existence of the corresponding piece-wise classical solutions of (6.9): $f_{Q,x^{(1)}}, f_{Q,x^{(2)}}$. These solutions are obviously also the (unique) piece-wise viscosity solutions of (6.9). The comparison result [9, Theorem 5] and the non-linear Feynman-Kac formula [9, Proposition 8] give then the claim.

3. This can be seen either from the representation (6.6) and an approximation argument, or exactly as in (2) by invoking the results of [9].

□

6.1 The Parisi functional

We consider now a specific terminal condition in the system (6.5) given in (5.22).

Given $\rho \in \mathscr{Q}(U,d)$, let $f_\rho : [0;1] \times \mathbb{R}^d \to \mathbb{R}$ be the value of (the continuous extension onto $\mathscr{Q}(U,d)$ of) the solution of (6.5) with the specific terminal condition given by (5.22). Following the ideas in the physical literature, we now define the *Parisi functional* $\mathscr{P}(\beta,\rho,\Lambda) : \mathbb{R}_+ \times \mathscr{Q}'(U,d) \times \mathrm{Sym}^+(d) \times \mathrm{Sym}(d) \to \mathbb{R}$ in as

$$\mathscr{P}(\beta,\rho,\Lambda) \equiv f_\rho(0,0) - \frac{\beta^2}{2}\int_0^1 x(t)\mathrm{d}\left(\|\rho(t)\|_{\mathrm{F}}^2\right) - \langle U, \Lambda\rangle. \tag{6.11}$$

The integral in (6.11) is understood in the usual Lebesgue-Stiltjes sense.

Remark 6.4. *Note that the path integral term in* (6.11) *equals* $f(0,0)$*, where* $f(t,y)$ *is the solution of* (6.9) *with the following boundary condition*

$$g(y) \equiv \beta\langle y, \mathbb{1}\rangle = \beta\sum_{u=1}^{d} y_u, \quad y \in \mathbb{R}^d.$$

Obviously $\mathscr{Q}'(d)$ is dense in $\mathscr{Q}(d)$.

Theorem 6.3. *We have*

$$p(\beta) \le \sup_{U\in\mathrm{Sym}^+(d)} \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda). \tag{6.12}$$

Proof. The bound (6.12) is a straightforward consequence of Theorem 5.1. □

6.2 On strict convexity of the Parisi functional and its variational representation

In this subsection, we derive a variational representation for Parisi's functional. As a consequence, for $d = 1$, we prove that the functional is strictly convex with respect to the $x \in \mathscr{Q}(1,1)$, if the terminal condition g (cf. (6.9)) is strictly convex and increasing. This result is related to the problem of strict convexity of the Parisi functional in the case of the SK model.

Let $W \equiv \{W(s)\}_{s\in\mathbb{R}_+}$ be the standard $\mathbb{R}^d$-valued Brownian motion and let $\{\mathscr{F}_t\}_{t\in\mathbb{R}_+}$ be the correspondent filtration. Define

$$\mathscr{U}[t;T] \equiv \{u : [t;T] \to \mathbb{R}^d \mid u \text{ is } \{\mathscr{F}_t\}_{t\in\mathbb{R}_+} \text{ progressively measurable }\}.$$

Given $u \in \mathscr{U}[t;1]$, $Q \in \mathscr{Q}(U,d)$ and $x \in \mathscr{Q}(1,1)$, consider the following $\mathbb{R}^d$-valued and adapted to $\{\mathscr{F}_t\}_{t\in\mathbb{R}^+}$ diffusion

$$Y^{(Q,x,u,t,y)}(s) = y - \int_t^s \left(x(s)\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^s \left(\dot{Q}(s)\right)^{1/2} \mathrm{d}W(s), \quad s \in [t;1].$$

Given some function $g : \mathbb{R}^d \to \mathbb{R}$ satisfying Assumption 5.1, define $f_{Q,x} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ as

$$f_{Q,x}(y,t) \equiv \sup_{u\in\mathscr{U}[t;1]} \mathbb{E}\left[g(Y^{(Q,x,u,t,y)}(1)) - \frac{1}{2}\int_t^1 \|u(s)\|_2^2 \mathrm{d}s\right]. \tag{6.13}$$

Proposition 6.3. *Let $d = 1$. If g is strictly convex and increasing, then the functional $\mathscr{Q}(1,1) \ni x \mapsto f_{Q,x}$ is strictly convex.*

Proof. We have

$$Y^{(Q,x,u,t,y)}(1) = y - \int_t^1 \left(x(s)\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s).$$

By an approximation argument, it is enough to prove the strict convexity for the continuous $x_1, x_2 \in \mathscr{Q}(1,1)$ $(x_1 \neq x_2)$. For any $\gamma \in (0;1)$, we have

$$\begin{aligned}
Y^{(Q,\gamma x_1+(1-\gamma)x_2,u,t,y)}(1) &= -\int_t^1 \left(\gamma x_1 + (1-\gamma)x_2\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s)\\
&< -\gamma\int_t^1 \left(x_1\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s - (1-\gamma)\int_t^1 \left(x_2\dot{Q}(s)\right)^{1/2} u(s)\mathrm{d}s\\
&\quad + \int_t^1 \left(\dot{Q}(s)\right)^{1/2} W(s)\\
&= \gamma Y^{(Q,x_1,u,t,y)}(1) + (1-\gamma)Y^{(Q,x_2,u,t,y)}(1),
\end{aligned} \tag{6.14}$$

where the strict inequality above is due to the strict concavity of the square root function. The strict convexity and monotonicity of g combined with the representation (6.14) implies that (6.13) is strictly convex as a function of x, since a supremum of a family of convex functions is convex. ☐

Proposition 6.4. *Given a piece-wise continuous $x \in \mathscr{Q}(1,1)$ and a $Q \in \mathscr{Q}(U,d)$ which is piece-wise in $C^1(0;1)$, the function $f_{Q,x} : \mathbb{R}^d \times [0;1] \to \mathbb{R}$ defined by* (6.13) *is a unique, continuous, piece-wise viscosity solution of the following terminal value problem*

$$\begin{cases}\partial_t f + \frac{1}{2}\left(\langle \dot{Q}, \nabla^2 f\rangle + x\langle \dot{Q}\nabla f, \nabla f\rangle\right) = 0, & (y,t) \in \mathbb{R}^d \times (0,1),\\ f(y,1) = g(y). & \end{cases}$$

Proof. In a way similar to the proof of Theorem 6.2, we successively use [12, Theorem 2.1] on the intervals $(x_k; x_{k+1})$, where the data of the PDE are continuous. □

Theorem 6.4. *Assume $d = 1$. Suppose also that g satisfies the assumptions of Proposition 6.3. For any $u \in \mathbb{R}$, the generalised Parisi functional given by* (6.11) *with $f_\rho(0,0)$ corresponding to the terminal condition g is strictly convex on $Q(u,1)$. Consequently, there exists a unique optimising order parameter.*

Proof. In 1-D, we can choose the coordinates such that $Q \equiv Ut$, on $[0;1]$. Consequently, $\dot{Q} \equiv U \equiv$ const on $[0;1]$. Hence, it is enough check the strict convexity with respect to $x \in \mathscr{Q}(1,1)$. The result follows by approximation in the norm (6.10) of an arbitrary pair of different elements of $\mathscr{Q}(U,d)$ by a pair of elements of $\mathscr{Q}'(U,d)$ and Propositions 6.1, 6.3 and 6.4. □

Remark 6.5. *Due to the monotonicity assumption on g, Theorem 6.4 does not cover the case of the SK model, where the terminal value g is given by* (5.22).

6.3 Simultaneous diagonalisation scenario

In the setups with highly symmetric state spaces Σ_N (such as the spherical spin models of [23] or the Gaussian spin models, see Section 8 below), less complex order parameter spaces as $Q(U,d)$ suffice.

Given some orthogonal matrix $O \in \mathscr{O}(d)$, we briefly discuss the case $\rho \in \mathscr{Q}_{\text{diag}}(U,O,d)$, where

$$\mathscr{Q}_{\text{diag}}(U,O,d) \equiv \{\rho \in \mathscr{Q}(U,d) \mid \text{for all } t \in [0;1], \text{ the matrix } O\rho(t)O^* \text{ is diagonal}\}.$$

The space $\mathscr{Q}_{\text{diag}}(U,O,d)$ is obviously isomorphic to the space of “paths” with the non-decreasing coordinate functions in $\mathbb{R}^d$, starting from the origin and ending at u, i.e.,

$$\bar{\mathscr{Q}}(u,d) \equiv \{\rho : [0;1] \to \mathbb{R}^d \mid \bar{\rho}(0) = 0; \bar{\rho}(1) = u; \bar{\rho}(t) \preceq \bar{\rho}(s), \text{ for } t \leq s; \bar{\rho} \text{ is cádlág}\},$$

where $u = OUO^* \in \mathbb{R}^d$. The isomorphism is then given by

$$\bar{\mathscr{Q}}(u,d) \ni \bar{\rho} \mapsto O\rho O^* \in \mathscr{Q}_{\text{diag}}(U,O,d). \tag{6.15}$$

7 Remainder estimates

In this section, we partially extend Talagrand's remainder estimates to the multidimensional setting. Due to Proposition 5.2, to prove the validity of Parisi's formula it is enough to show that all the μ_k terms in (5.30) almost vanish for the almost optimal parameters of the optimisation problem in (5.16). This can be done if the free energy of two coupled replicas of the system (7.3) is strictly smaller than twice the free energy of the uncoupled single system (5.4), see inequality (7.2). However, the systems involved in (7.2) are effectively at least as complex as the SK model itself. In Section 7.2, we again apply Guerra's scheme to obtain the upper bounds on (7.3) in terms of the free energy of the corresponding comparison GREM-inspired model. One might then hope that by a careful choice of the comparison model one can prove inequality (7.2). In Sections 7.3 and 7.4, we formulate some conditions on the comparison system which would suffice to get inequality (7.2), giving, hence, the conditional proof of the Parisi formula, see Theorem 7.1.

7.1 A sufficient condition for μ_k-terms to vanish

In this subsection, we are going to establish a sufficient condition for the measures μ_k to vanish. This condition states roughly the following. Whenever the free energy of a certain replicated system uniformly in N strictly less then twice the free energy of the single system, the measure μ_k vanishes in $N \to +\infty$ limit (see Lemma 7.2).

Keeping in mind the definition of μ_k (cf. (5.28)) and of the Hamiltonian $H_t(\sigma, \alpha)$ (cf. (5.3)), we define, for $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}^{(2),k}$, the corresponding replicated Hamiltonian as

$$H_t^{(2)}(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) \equiv H_t(\sigma^{(1)}, \alpha^{(1)}) + H_t(\sigma^{(2)}, \alpha^{(2)}). \tag{7.1}$$

Remark 7.1. *We note here that the distribution of the Hamiltonian $H_t(k, \sigma^{(1)}, \sigma^{(2)})$ depends only on k and not on the choice of the indices $\alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}^{(2),k}$.*

Remark 7.2. *The superscript* (2) *in* (7.1) *(and in what follows) indicates that the quantity is related to the twice replicated objects.*

Define

$$\mathscr{A}^{(2),k} \equiv \{(\alpha^{(1)}, \alpha^{(2)}) \in \mathscr{A}^2 : q_{\mathrm{L}}(\alpha^{(1)}, \alpha^{(2)}) = k\}.$$

Additionally, for any $\mathscr{V} \subset \Sigma(B(U, \varepsilon))^2$ and any suitable Gaussian process,

$$\{F(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) : \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N, \alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}\},$$

we define the *local remainder comparison functional* as

$$\Phi_{\mathscr{V}}^{(2),k,x}[F] \equiv \frac{1}{N}\mathbb{E}\Big[\log \iint_{\mathscr{V}} \iint_{\mathscr{A}^{(2),k}} \exp\left\{\beta\sqrt{N} F(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\right\} \\ \mathrm{d}\mu^{\otimes N}(\sigma^{(1)}) \mathrm{d}\mu^{\otimes N}(\sigma^{(2)}) \mathrm{d}\xi(\alpha^{(1)}) \mathrm{d}\xi(\alpha^{(2)})\Big]. \tag{7.2}$$

Define

$$\varphi_N^{(2)}(k, t, x, Q, \mathscr{V}) \equiv \Phi_{\mathscr{V}}^{(2),k}\left[H_t^{(2)}\right]. \tag{7.3}$$

Lemma 7.1. *Recalling the definition* (5.4)*, for any $\mathscr{V} \subset \Sigma(B(U, \varepsilon))^2$, we have*

$$\varphi_N^{(2)}(k, t, x, Q, \mathscr{V}) \leq \varphi_N^{(2)}(k, t, x, Q, \Sigma(B(U, \varepsilon))^2) = 2\varphi_N(t, x, Q, \Sigma(B(U, \varepsilon))). \tag{7.4}$$

Proof. The first inequality in (7.4) is obvious, since the expression under the integral in (7.2) is positive. The equality in (7.4) is an immediate consequence of the RPC averaging property (5.27). □

In what follows, we shall be looking for the sharper (in particular, *strict*) versions of the inequality (7.4) because of the following observation due to Talagrand [30].

Lemma 7.2. *Fix an arbitrary $\mathscr{V} \subset \Sigma_N(B(U,\varepsilon))^2$. Suppose that, for some $\varepsilon > 0$, the following inequality holds*

$$\varphi_N^{(2)}(k,t,x,Q,\mathscr{V}) \leq 2\varphi_N(t,x,Q,\Sigma_N(B(U,\varepsilon))) - \varepsilon. \tag{7.5}$$

Then, for some $K > 0$, we have

$$\mu_k(\mathscr{V}) \leq K \exp\left(-\frac{N}{K}\right).$$

Proof. The proof is based on Theorem 2.2 and follows the lines of [21, Lemma 7]. ☐

7.2 Upper bounds on $\varphi^{(2)}$: Guerra's scheme revisited

In this subsection, we shall develop a mechanism to obtain upper bounds on $\varphi^{(2)}$ defined in (7.3). This will be achieved in the full analogy to Guerra's scheme by using a suitable Gaussian comparison system.

Given $U \in \mathrm{Sym}^+(d)$, we say that $V \in \mathbb{R}^{d\times d}$ is an *admissible mutual overlap matrix for U*, if

$$\mathfrak{U} \equiv \begin{bmatrix} U & V \\ V^* & U \end{bmatrix} \in \mathrm{Sym}^+(2d). \tag{7.6}$$

Furthermore, define

$$\mathscr{V}(U) \equiv \{V \in \mathbb{R}^{d\times d} : V \text{ is an admissible mutual overlap matrix for } U\}.$$

Hereinafter without further notice we assume that $\mathfrak{U} \in \mathrm{Sym}^+(2d)$ has the form (7.6), where V is some admissible mutual overlap matrix for U.

Let $\mathfrak{Q} \in \mathscr{Q}(\mathfrak{U}, 2d)$. Let $\mathfrak{x} \equiv \{\mathfrak{x}_l \in [0;1]\}_{l=1}^{\mathrm{n}}$ be the "jump times" of the path ϱ. We assume that the "times" are increasingly ordered, i.e.,

$$0 = \mathfrak{x}_0 < \mathfrak{x}_1 < \ldots < \mathfrak{x}_{\mathrm{n}} < \mathfrak{x}_{\mathrm{n}+1} = 1.$$

Consider the following collection of matrices

$$\mathfrak{Q} \equiv \{\mathfrak{Q}_l \equiv \mathfrak{Q}(\mathfrak{x}_l) \subset \mathrm{Sym}^+(2d)\}_{l=0}^{\mathrm{n}+1}.$$

We obviously then have

$$0 = \mathfrak{Q}^{(0)} \prec \mathfrak{Q}^{(1)} \prec \ldots \prec \mathfrak{Q}^{(\mathrm{n})} \prec \mathfrak{Q}^{(\mathrm{n}+1)} = \mathfrak{U}. \tag{7.7}$$

Such a path $\mathfrak{Q}$ induces in the usual way the "doubled" GREM overlap kernel $\mathfrak{Q} \equiv \{\mathfrak{Q}(\alpha^{(1)},\alpha^{(2)}) \in \mathrm{Sym}^+(2d) \mid \alpha^{(1)},\alpha^{(2)} \in \mathscr{A}_{\mathrm{n}}\}$, defined as

$$\mathfrak{Q}(\alpha^{(1)},\alpha^{(2)}) \equiv \mathfrak{Q}^{(q_{\mathrm{L}}(\alpha^{(1)},\alpha^{(2)}))}.$$

We also need the $d \times d$ submatrices of the above overlap such that

$$\mathfrak{Q}(\alpha^{(1)},\alpha^{(2)}) = \begin{bmatrix} \mathfrak{Q}|_{11}(\alpha^{(1)},\alpha^{(2)}) & \mathfrak{Q}|_{12}(\alpha^{(1)},\alpha^{(2)}) \\ \mathfrak{Q}|_{12}(\alpha^{(1)},\alpha^{(2)})^* & \mathfrak{Q}|_{22}(\alpha^{(1)},\alpha^{(2)}) \end{bmatrix}. \tag{7.8}$$

Remark 7.3. *For $\sigma^{(1)}\sigma^{(2)} \in \Sigma_N$, we shall use the notation $\sigma^{(1)} \shortparallel \sigma^{(2)} \in \left(\mathbb{R}^{2d}\right)^N$ to denote the vector obtained by the following concatenation of the vectors $\sigma^{(1)}$ and $\sigma^{(2)}$*

$$\sigma^{(1)} \shortparallel \sigma^{(2)} \equiv \left(\sigma_i^{(1)}\sigma_i^{(2)} \in \Sigma \times \Sigma \subset \mathbb{R}^{2d}\right)_{i=1}^N.$$

Let us observe that the process

$$X^{(2)} \equiv \left\{X^{(2)}(\tau) = X(\sigma^{(1)}) + X(\sigma^{(1)}) \mid \tau = \sigma^{(1)} \shortparallel \sigma^{(2)}; \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N\right\}$$

is actually an instance of the $2d$-dimensional Gaussian process defined in (1.1). Hence, it has the following correlation structure, for $\tau^1, \tau^2 \in \Sigma_N^{(2)}$,

$$\operatorname{Cov}\left[X^{(2)}(\tau^1), X^{(2)}(\tau^2)\right] = \|R^{(2)}(\tau^1, \tau^2)\|_{\mathrm{F}}^2.$$

The path ϱ induces also the following two new (independent of everything before) comparison process $Y^{(2)} \equiv \left\{Y^{(2)}(\alpha) \in \mathbb{R}^{2d} \mid \alpha \in \mathscr{A}_{\mathfrak{n}}\right\}$, with the following correlation structures

$$\operatorname{Cov}\left[Y^{(2)}(\alpha^{(1)}), Y^{(2)}(\alpha^{(2)})\right] = \mathfrak{Q}(\alpha^{(1)}, \alpha^{(2)}) \in \operatorname{Sym}^+(d).$$

As usual, let $\{Y_i^{(2)}\}_{i=1}^N$ be the independent copies of $Y^{(2)}$. For the purposes of new Guerra's scheme we define a GREM-like process (cf. (1.17))

$$A^{(2)} = \{A^{(2)}(\tau, \alpha) : \tau = \sigma^{(1)} \shortparallel \sigma^{(2)}; \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N; \alpha \in \mathscr{A}_{\mathfrak{n}}\}$$

as

$$A^{(2)}(\tau, \alpha) \equiv \left(\frac{2}{N}\right)^{1/2} \sum_{i=1}^N \langle Y_i^{(2)}(\alpha), \tau_i\rangle.$$

We fix some $t \in [0;1]$. We would now like to apply Guerra's scheme to the comparison functional (7.2) and the following two processes

$$\left\{H_t^{(2)}(\sigma^{(1)}, \sigma^{(2)}, \alpha)\right\}_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N, \alpha \in \mathscr{A}}, \left\{\sqrt{t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha)\right\}_{\sigma^{(1)}, \sigma^{(2)} \in \Sigma_N, \alpha \in \mathscr{A}}.$$

These two processes are, respectively, the counterparts of the processes $X(\sigma)$ and $A(\sigma, \alpha)$ in Guerra's scheme.

Consider a path $\widetilde{Q} \in \mathscr{Q}'(U, d)$ with the following jumps

$$0 =: \widetilde{Q}^{(0)} \prec \widetilde{Q}^{(1)} \prec \ldots \prec \widetilde{Q}^{(\mathfrak{n})} \prec \widetilde{Q}^{(\mathfrak{n}+1)}.$$

Let $\widetilde{A} \equiv \left\{\widetilde{A}(\sigma, \alpha) : \sigma \in \Sigma_N; \alpha \in \mathscr{A}_{\mathfrak{n}}\right\}$ be a Gaussian process (independent of all random objects around) with the following covariance structure

$$\mathbb{E}\left[\widetilde{A}(\sigma^{(1)}, \alpha^{(1)})\widetilde{A}(\sigma^{(2)}, \alpha^{(2)})\right] = 2\langle R(\sigma^{(1)}, \sigma^{(2)}), \widetilde{Q}(\alpha^{(1)}, \alpha^{(2)})\rangle.$$

For notational convenience, we introduce also the following process

$$\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) \equiv \widetilde{A}(\sigma^{(1)}, \alpha^{(1)}) + \widetilde{A}(\sigma^{(2)}, \alpha^{(2)}). \tag{7.9}$$

Recalling the replicated Hamiltonian (7.1) and following Guerra's scheme, we introduce, for $s \in [0;1]$, the following interpolating Hamiltonian

$$H_{t,s}^{(2)}(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) \equiv \sqrt{st} X^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}) + \sqrt{(1-s)t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) + \sqrt{1-t} \widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}). \tag{7.10}$$

Given $\varepsilon, \delta > 0$ and $\mathfrak{L} \in \mathrm{Sym}(2d)$, define (cf. (5.8))

$$\mathscr{V}^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta) \equiv \{\mathfrak{U}' \in \mathrm{Sym}^+(2d) : \|\mathfrak{U}' - \mathfrak{U}\|_{\mathrm{F}} < \varepsilon, \langle \mathfrak{U}' - \mathfrak{U}, \mathfrak{L} \rangle < \delta\}.$$

We consider the following set of the local configurations

$$\Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta) \equiv \left\{(\sigma^{(1)}, \sigma^{(2)}) \in \Sigma_N \times \Sigma_N : R_N^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \sigma^{(1)} \shortparallel \sigma^{(2)}) \in \mathscr{V}^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)\right\}. \tag{7.11}$$

Note that $\Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta) \subset \Sigma_N(B(U, \varepsilon))^2$. We consider also the RPC $\zeta = \zeta(\mathfrak{x})$ generated by the vector $\mathfrak{x}$ and, for any suitable Gaussian process

$$F \equiv \{F(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) \mid \sigma^{(1)}, \sigma^{(2)} \in \Sigma_N; \alpha^{(1)}, \alpha^{(2)} \in \mathscr{A}_n\},$$

define the corresponding local comparison functional (cf. (7.2)) as follows

$$\Phi_{\mathscr{V}}^{(2),k,\mathfrak{x}}[F] \equiv \frac{1}{N} \mathbb{E}\Big[\log \iint_{\mathscr{V}} \iint_{\mathscr{A}^{(2),k}} \exp\left\{\beta\sqrt{N} F(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\right\} \mathrm{d}\mu^{\otimes N}(\sigma^{(1)}) \mathrm{d}\mu^{\otimes N}(\sigma^{(2)}) \mathrm{d}\zeta(\alpha^{(1)}) \mathrm{d}\zeta(\alpha^{(2)})\Big].$$

Define the corresponding local free energy-like quantity as (cf. (5.4))

$$\chi(s, t, k, \mathfrak{x}, \mathfrak{Q}, \widetilde{\mathfrak{Q}}, \Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)) \equiv \Phi_{\Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)}^{(2),k,\mathfrak{x}}\left[H_{t,s}^{(2)}\right]. \tag{7.12}$$

To lighten the notation, we indicate hereinafter only the dependence of χ on s. Denote

$$B^{\mathfrak{x},\mathfrak{Q}} \equiv \frac{t\beta^2}{2} \sum_{l=1}^{\mathfrak{n}} \mathfrak{x}_l \left(\|\mathfrak{Q}^{(l+1)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(l)}\|_{\mathrm{F}}^2\right).$$

Lemma 7.3. *There exists $C = C(\Sigma) > 0$ such that, for any $\mathfrak{U}$ as above, we have*

$$\frac{\partial}{\partial s} \chi(s, t, k, \mathfrak{x}, \mathfrak{Q}, \widetilde{\mathfrak{Q}}, \Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)) \leq -B^{\mathfrak{x},\mathfrak{Q}} + C\varepsilon, \tag{7.13}$$

Consequently,

$$\varphi_N^{(2)}(k, t, x, Q, \Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)) \leq \Phi_{\Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)}^{(2),k,\mathfrak{x}}\Big[\sqrt{t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) + \sqrt{1-t} \widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\Big] - B^{\mathfrak{x},\mathfrak{Q}} + C\varepsilon. \tag{7.14}$$

Proof. The idea is the same as in the proof of Theorem 5.1 and is based on Proposition 2.5. Since we are considering the localised free energy-like quantities (7.12), the variance terms induced by the interpolation (7.10) in (2.14) cancel out (up to the correction $\mathscr{O}(\varepsilon)$) and we are left with the non-positive contribution of the covariance terms. □

Given $\mathfrak{L} \in \mathrm{Sym}(2d)$, we consider the following stencil of the Legendre transform

$$\widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}}[F] \equiv -\langle \mathfrak{L}, \mathfrak{U} \rangle - B^{\mathfrak{x},\mathfrak{Q}} + \frac{1}{N}\mathbb{E}[\log \iint_{\Sigma_N^2} \iint_{\mathscr{A}^{(2),k}} \exp\{\beta\sqrt{N} F(\sigma^{(1)}, \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)}) + \langle \mathfrak{L}(\sigma^{(1)} \shortparallel \sigma^{(2)}), \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle\} \mathrm{d}\mu^{\otimes N}(\sigma^{(1)})\mathrm{d}\mu^{\otimes N}(\sigma^{(2)})\mathrm{d}\zeta(\alpha^{(1)})\mathrm{d}\zeta(\alpha^{(2)})\Big]. \tag{7.15}$$

Definition 7.1. *Let $F : \mathrm{Sym}(2d) \to \mathbb{R}$. Given $\delta > 0$, we call $\mathfrak{L}^{(0)} \in \mathrm{Sym}(2d)$* δ-minimal for F, *if*

$$F(\Lambda^{(0)}) \leq \inf_{\Lambda \in \mathrm{Sym}(2d)} F(\Lambda) + \delta.$$

Lemma 7.4. *There exists $C = C(\Sigma) > 0$ such that, for all $\mathfrak{U}$ and $\mathfrak{Q} \in \mathscr{Q}'(\mathfrak{U}, 2d)$ as above, all $\varepsilon, \delta > 0$, there exists a δ-minimal Lagrange multiplier $\mathfrak{L} = \mathfrak{L}(\mathfrak{U}, \varepsilon, \delta) \in \mathrm{Sym}(2d)$ for (7.15) such that, for all $k \in [1; n] \cap \mathbb{N}$, all $t \in [0; 1]$, and all $(x, \mathscr{Q})$, we have*

$$\varphi_N^{(2)}(k, t, x, Q, \Sigma_N^{(2)}(\mathfrak{L}, \mathfrak{U}, \varepsilon, \delta)) \leq \inf_{\mathfrak{L} \in \mathrm{Sym}(2d)} \widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}} \Big[\sqrt{t} A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) + \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\Big] + C(\varepsilon + \delta). \tag{7.16}$$

Proof. The argument is the same as in the proof of Theorem 1.1. $\square$

Consider the family of matrices $\widetilde{\mathfrak{Q}} \equiv \left\{\widetilde{\mathfrak{Q}}^{(l)} \in \mathrm{Sym}^+(2d) \mid l \in [0; \mathfrak{n}+1] \cap \mathbb{N}\right\}$, defined as

$$\widetilde{\mathfrak{Q}}^{(l)} \equiv \begin{bmatrix} \widetilde{Q}^{(l)} & \widetilde{Q}^{(l)} \\ \widetilde{Q}^{(l)} & \widetilde{Q}^{(l)} \end{bmatrix}, \tag{7.17}$$

for $l \in [0; k] \cap \mathbb{N}$, and as

$$\widetilde{\mathfrak{Q}}^{(l)} \equiv \begin{bmatrix} \widetilde{Q}^{(l)} & \widetilde{Q}^{(k)} \\ \widetilde{Q}^{(k)} & \widetilde{Q}^{(l)} \end{bmatrix}, \tag{7.18}$$

for $l \in [k+1; \mathfrak{n}+1] \cap \mathbb{N}$. Additionally we define, for $l \in [0; \mathfrak{n}+1]$, the matrices

$$\widehat{\mathfrak{Q}}^{(l)}(t) \equiv t\mathfrak{Q} + (1-t)\widetilde{\mathfrak{Q}}.$$

Let $\widehat{Z}^{(l)} \in \mathbb{R}^{2d \times 2d}$, for $l \in [0; \mathfrak{r}]$, be independent Gaussian vectors with

$$\mathrm{Cov}\left[\widehat{Z}^{(l)}\right] = 2\beta^2 \Big(\widehat{\mathfrak{Q}}^{(l+1)}(t) - \widehat{\mathfrak{Q}}^{(l)}(t)\Big).$$

Given $\widehat{y} \in \mathbb{R}^{2d}$, $\mathfrak{L} \in \mathrm{Sym}(2d)$, consider the random variable

$$X_{\mathfrak{n}+1}^{(2)}(\widehat{y}, \mathfrak{r}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) \equiv \log \int_\Sigma \int_\Sigma \exp\Big(\langle \widehat{y}, \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle + \langle \mathfrak{L}(\sigma^{(1)} \shortparallel \sigma^{(2)}), \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle\Big) \mathrm{d}\mu(\sigma^{(1)})\mathrm{d}\mu(\sigma^{(2)}). \tag{7.19}$$

Define recursively, for $l \in [\mathfrak{n}; 0] \cap \mathbb{N}$, the following quantities

$$X_l^{(2)}(\widehat{y}, k, \mathfrak{r}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) \equiv \frac{1}{\mathfrak{r}_l} \log \mathbb{E}^{\widehat{Z}^{(l)}} \left[\exp\left(\mathfrak{r}_l X_{l+1}^{(2)}(\widehat{y} + \widehat{Z}^{(l)}, k, \mathfrak{r}, \widehat{\mathfrak{Q}}^{(l)}(t), \mathfrak{L})\right)\right]. \tag{7.20}$$

Lemma 7.5. *We have*

$$\widetilde{\Phi}^{(2),k,\mathfrak{x},\mathfrak{L}}\left[\sqrt{t}A^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}) + \sqrt{1-t}\widetilde{A}^{(2)}(\sigma^{(1)} \shortparallel \sigma^{(2)}, \alpha^{(1)}, \alpha^{(2)})\right]$$
$$= -\langle \mathfrak{L}, \mathfrak{U}\rangle + X_0^{(2)}(0, \mathfrak{x}, \widehat{\mathfrak{Q}}^{(l)}(t), \mathfrak{L}). \tag{7.21}$$

Proof. This is an immediate consequence of the RPC averaging property (5.27). □

Proposition 7.1. *Under the conditions of Lemma 7.4, we have*

$$\varphi_N^{(2)}(k,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \le \inf_{\mathfrak{L}\in \mathrm{Sym}(2d)}\Big(-\langle \mathfrak{L},\mathfrak{U}\rangle + X_0^{(2)}(0,\mathfrak{x},\widehat{\mathfrak{Q}}(t),\mathfrak{L})\Big) - B^{\mathfrak{x},\mathfrak{Q}} + C(\varepsilon+\delta).$$

Remark 7.4. *Similarly to* (5.15)*, there exists $C = C(\Sigma,\mu) > 0$, such that, for any $\mathfrak{L} \in \mathrm{Sym}(2d)$,*

$$\varphi_N^{(2)}(k,t,x,Q,\Sigma_N^{(2)}(B(\mathfrak{U},\varepsilon)) \le -\langle \mathfrak{L},\mathfrak{U}\rangle - B^{\mathfrak{x},\mathfrak{Q}} + X_0^{(2)}(0,\mathfrak{x},\widehat{\mathfrak{Q}}(t),\mathfrak{L})\Big) + C\|\mathfrak{L}\|_F\varepsilon.$$

Proof. Immediately follows from Lemmata 7.4 and 7.5. □

7.3 Adjustment of the upper bounds on $\varphi^{(2)}$

Proposition 2.1 implies that there exists $r \in [1;n] \cap \mathbb{N}$ such that

$$\|Q^{(r-1)}\|_{\mathrm{F}}^2 < \|V\|_{\mathrm{F}}^2 < \|Q^{(r)}\|_{\mathrm{F}}^2. \tag{7.22}$$

Assume $r = k$. (Other cases are similar or easier as shown for 1-D in [30].) We make the following tuning of the upper bounds of the previous subsection. Set $\mathfrak{n} \equiv n+1$. Let $w \in [x_{r-1}/2; x_r]$. Define

$$\mathfrak{x}_l \equiv \mathfrak{x}_l(w) \equiv \begin{cases} \frac{x_l}{2}, & l \in [0;k-1]\cap\mathbb{N},\\ w, & l = k,\\ x_l, & l \in [k+1;n+1]\cap \mathbb{N}.\end{cases} \tag{7.23}$$

Let

$$\widetilde{Q}^{(l)} \equiv \begin{cases} Q^{(l)}, & l \in [0;k-1]\cap\mathbb{N},\\ Q^{(l-1)}, & l\in[k;n+2]\cap\mathbb{N}.\end{cases}$$

Moreover, suppose $\mathfrak{Q} \equiv \{\mathfrak{Q}^{(l)}\}_{l=0}^{n+2}$ satisfy

$$\|\mathfrak{Q}^{(l)}\|_{\mathrm{F}}^2 = \begin{cases} 4\|Q^{(l)}\|_{\mathrm{F}}^2, & l\in[0;k-1]\cap\mathbb{N},\\ 4\|V\|_{\mathrm{F}}^2, & l = k,\\ 2\Big(\|Q^{(l-1)}\|_{\mathrm{F}}^2 + \|V\|_{\mathrm{F}}^2\Big), & l\in[k+1;n+2]\cap\mathbb{N}.\end{cases} \tag{7.24}$$

Such $\mathfrak{Q}$ exists due to (7.22). Moreover, if $d \ge 2$, then it is obviously non-unique.

Lemma 7.6. *In the above setup, we have*

$$B^{\mathfrak{x},\mathfrak{Q}} \equiv t\beta^2\Big\{(w - x_{l-1})\left(\|Q^{(k)}\|_F^2 - \|V\|_F^2\right) + \sum_{l=1}^{n} x_l\left(\|Q^{(l+1)}\|_F^2 - \|Q^{(l)}\|_F^2\right)\Big\}.$$

Proof. The claim is a straightforward consequence of (7.23) and (7.24). □

Define the matrix $\mathfrak{D}^{(n+1)} \in \mathrm{Sym}^+(2d)$ block-wise as

$$\begin{aligned}
\mathfrak{D}^{(n+1)}|_{11} &\equiv \beta^2 t(U - \mathfrak{Q}^{(n+1)}|_{11}) + \beta^2(1-t)(U - Q^{(n)}) + \mathfrak{L}|_{11},\\
\mathfrak{D}^{(n+1)}|_{12} &\equiv \beta^2 t(V - \mathfrak{Q}^{(n+1)}|_{12}) + \mathfrak{L}|_{12},\\
\mathfrak{D}^{(n+1)}|_{21} &\equiv \beta^2 t(V - \mathfrak{Q}^{(n+1)}|_{12})^* + \mathfrak{L}|_{12}^*,\\
\mathfrak{D}^{(n+1)}|_{22} &\equiv \beta^2 t(U - \mathfrak{Q}^{(n+1)}|_{22}) + \beta^2(1-t)(U - Q^{(n)}) + \mathfrak{L}|_{22}.
\end{aligned}$$

Furthermore, we define

$$\mathrm{Sym}^+(2d) \ni \widetilde{\mathfrak{D}}^{(n+1)} \equiv \begin{bmatrix} \beta^2(U - Q^{(n)}) + \Lambda & 0 \\ 0 & \beta^2(U - Q^{(n)}) + \Lambda \end{bmatrix}.$$

Lemma 7.7. *We have*

$$X^{(2)}_{n+1}(\widehat{y}, k, \mathfrak{x}, \widehat{\mathfrak{Q}}(t), \mathfrak{L}) \equiv \log \int_\Sigma \int_\Sigma \exp\Big(\langle \widehat{y}, \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle + \langle \mathfrak{D}^{(n+1)}(\sigma^{(1)} \shortparallel \sigma^{(2)}), \sigma^{(1)} \shortparallel \sigma^{(2)} \rangle \Big) \times \mathrm{d}\mu(\sigma^{(1)}) \mathrm{d}\mu(\sigma^{(2)}).$$

Proof. Since $\mathfrak{x}_{n+2} = 1$, the result follows from a straightforward calculation of the Gaussian integrals in (7.20) for $l = n+1$. □

Define

$$\widetilde{\mathfrak{L}} \equiv \begin{bmatrix} \Lambda & 0 \\ 0 & \Lambda \end{bmatrix}, \widetilde{\mathfrak{U}} \equiv \begin{bmatrix} U & Q^{(k)} \\ Q^{(k)} & U \end{bmatrix}.$$

Lemma 7.8. *For any $y \in \mathbb{R}^d$, $l \in [0; n+2] \cap \mathbb{N}$, we have*

$$X^{(2)}_l(y \shortparallel y, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w = x_{k-1}} = \begin{cases} 2X_{l-1}(y, x, \mathscr{Q}, U, \Lambda), & l \in [k; n+2] \cap \mathbb{N}, \\ 2X_l(y, x, \mathscr{Q}, U, \Lambda), & l \in [0; k-1] \cap \mathbb{N}. \end{cases}$$

Proof. A straightforward (decreasing) induction argument on l gives the result. Indeed: for $l = n+2$, an inspection of (7.19) and (1.9) immediately yields

$$X^{(2)}_{n+2}(y^{(1)} \shortparallel y^{(2)}, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) = X_{n+1}(y^{(1)}, x, \mathscr{Q}, U, \Lambda) + X_{n+1}(y^{(2)}, x, \mathscr{Q}, U, \Lambda),$$

where $y^{(1)}, y^{(2)} \in \mathbb{R}^d$. Let $\widetilde{Z}^{(l)}$ be a Gaussian $2d$-dimensional vector with

$$\mathrm{Cov}\Big[\widetilde{Z}^{(l)}\Big] = 2\beta^2(\widetilde{\mathfrak{Q}}^{(l+1)} - \widetilde{\mathfrak{Q}}^{(l)}).$$

Define two Gaussian d-dimensional vectors $\widetilde{Z}^{(l),1}$ and $\widetilde{Z}^{(l),2}$ by demanding that

$$\widetilde{Z}^{(l)} = \widetilde{Z}^{(l),1} \shortparallel \widetilde{Z}^{(l),2}.$$

Due to (7.17) and (7.18), the vectors $\widetilde{Z}^{(l),1}$ and $\widetilde{Z}^{(l),2}$ are independent, for $l \in [k; n+1]$. We have $\widetilde{Z}^{(l),1} \sim \widetilde{Z}^{(l),2}$, for $l \in [0; k-1]$. Assume that $l \in [k; n+1] \cap \mathbb{N}$ and

$$X^{(2)}_{l+1}(y^{(1)} \shortparallel y^{(2)}, \mathfrak{x}(w), \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) = X_l(y^{(1)}, x, \mathcal{Q}, U, \Lambda) + X_l(y^{(2)}, x, \mathcal{Q}, U, \Lambda).$$

By definition (7.19), we have

$$\begin{aligned} X^{(2)}_{l}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) &= \frac{1}{\mathfrak{x}_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \left[\exp\left(\mathfrak{x}_l X^{(2)}_{l+1}(y^{(1)} \shortparallel y^{(2)} + \widehat{Z}^{(l)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \mathfrak{L})\right)\right] \\ &= \frac{1}{x_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \Big[\exp\Big\{x_l \Big(X_l(y^{(1)} + \widetilde{Z}^{(l),1}, x, \mathcal{Q}, U, \Lambda) \\ &\quad + X_l(y^{(2)} + \widetilde{Z}^{(l),2}, x, \mathcal{Q}, U, \Lambda)\Big)\Big\}\Big] \\ &= X_{l-1}(y^{(1)}, x, \mathcal{Q}, U, \Lambda) + X_{l-1}(y^{(2)}, x, \mathcal{Q}, U, \Lambda). \end{aligned}$$

By the construction and previous formula, for $l = k-1$, we have

$$\begin{aligned} X^{(2)}_{k-1}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w=x_{k-1}} &= X^{(2)}_{k}(y^{(1)} \shortparallel y^{(2)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}}) \\ &= X_{k-1}(y^{(1)}, x, \mathcal{Q}, U, \Lambda) + X_{k-1}(y^{(2)}, x, \mathcal{Q}, U, \Lambda). \end{aligned}$$

Finally, for $l \in [0; k-2]$, we recursively obtain

$$\begin{aligned} X^{(2)}_{l}(y^{(1)} \shortparallel y^{(1)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \widetilde{\mathfrak{L}})|_{w=x_{k-1}} &= \frac{1}{\mathfrak{x}_l} \log \mathbb{E}^{\widetilde{Z}^{(l)}} \left[\exp\left(\mathfrak{x}_l X^{(2)}_{l+1}(y^{(1)} \shortparallel y^{(1)} + \widehat{Z}^{(l)}, k, \mathfrak{x}, \widetilde{\mathfrak{Q}}, \mathfrak{L})|_{w=x_{k-1}}\right)\right] \\ &= \frac{2}{x_l} \log \mathbb{E}^{\widetilde{Z}^{(l),1}} \Big[\exp\Big\{\frac{x_l}{2} \Big(X_{l+1}(y^{(1)} + \widetilde{Z}^{(l),1}, x, \mathcal{Q}, U, \Lambda) \\ &\quad + X_{l+1}(y^{(1)} + \widetilde{Z}^{(l),1}, x, \mathcal{Q}, U, \Lambda)\Big)\Big\}\Big] \\ &= 2X_l(y^{(1)}, x, \mathcal{Q}, U, \Lambda). \end{aligned}$$

□

Remark 7.5. *Motivated by Lemmata 7.2 and 7.8 (see also Section 7.4), we pose the following problem. Is it true that, as in 1-D (see [30; 21]), there exists $\mathfrak{Q} \in \mathcal{Q}'(\mathfrak{U}, 2d)$ satisfying the assumption* (7.24) *such that the following inequality holds*

$$\begin{aligned} &\inf_{\mathfrak{L} \in \mathrm{Sym}(2d)} \left(-\langle \mathfrak{L}, \mathfrak{U} \rangle + X^{(2)}_0(0, \mathfrak{x}(w), \widehat{\mathfrak{Q}}(t), \mathfrak{L})|_{w=x_{k-1}}\right) \\ &\overset{?}{\leq} 2 \inf_{\Lambda \in \mathrm{Sym}(d)} \Big(-\langle \Lambda, U \rangle + X_0(0, x, \mathcal{Q}, U, \Lambda)\Big)? \end{aligned} \tag{7.25}$$

Similar problems have at first been posed in [33]. The resolution of the above problem seems to require more detailed information on the behaviour of the Parisi functional (6.11) *or, equivalently, of the solution of* (6.9) *as a function of $Q \in \mathcal{Q}(U, d)$.*

7.4 Talagrand's a priori estimates

We start from defining a class of the almost optimal paths for the optimisation problem in (6.12). Recall the following convenient definition from [21].

Definition 7.2. *Given* $U \in \mathrm{Sym}^+(d)$*, we shall call the triple* $(n, \rho^*, \Lambda^*) \in \mathbb{N} \times \mathcal{Q}'_n(U,d) \times \mathbb{R}^d$ *a* θ*-* optimiser *of the Parisi functional* (6.11)*, if it satisfies the following two conditions*

$$\mathscr{P}(\beta, \rho^*, \Lambda^*) \leq \inf_{\substack{\rho \in \mathcal{Q}'(U,d) \\ \Lambda \in \mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda) + \theta. \tag{7.26}$$

$$\mathscr{P}(\beta, \rho^*, \Lambda^*) = \inf_{\substack{\rho \in \mathcal{Q}'_n(U,d) \\ \Lambda \in \mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda). \tag{7.27}$$

Remark 7.6. *It is obvious that for any* $\theta > 0$ *such a* θ*-optimiser exists. The main convenient feature of this definition (as pointed out in [30]) is that* n *(the number of jumps of* ρ^**) is finite and fixed.*

Recalling (5.13), we set

$$\phi^{(x,\mathcal{Q},\Lambda)}(t) \equiv -\langle U, \Lambda \rangle - \frac{t\beta^2}{2} \sum_{k=1}^{n} x_k \left(\|Q^{(k+1)}\|_{\mathrm{F}}^2 - \|Q^{(k)}\|_{\mathrm{F}}^2 \right) + X_0(x, \mathcal{Q}, U, \Lambda). \tag{7.28}$$

Under the following assumption (at first proposed in 1-D in [30]), we shall effectively prove that remainder term almost vanishes on the θ minimisers of (6.11), see Theorem 7.1.

Assumption 7.1. *Let* $\mathfrak{U} \in \mathrm{Sym}^+(2d)$ *be defined by* (7.6)*. We fix arbitrary* $t_0 \in [0;1)$*,* $\varepsilon > 0$ *and* $\delta > 0$*. There exists* $K = K(t_0, \varepsilon, \delta, \mathfrak{U}) > 0$*,* $\theta(t_0, \varepsilon, \delta, \mathfrak{U}) > 0$*, and* $N_0 = N_0(t_0, \varepsilon, \delta, \mathfrak{U}) \in \mathbb{N}$ *and* $\mathfrak{L}^* \in \mathrm{Sym}(2d)$ *with the following property:*

If (n, ρ^*, Λ^*) *is a* θ*-optimiser, for some* $\theta \in (0; \theta(t_0, \varepsilon, \delta, \mathfrak{U})]$*, then uniformly, for all* $t \in [0; t_0)$*,* $N > N_0$ *and all* $k \in [1; n] \cap \mathbb{N}$*, we have*

$$\varphi_N^{(2)}(k, t, x^*, Q^*, \Sigma_N^{(2)}(\mathfrak{L}^*, \mathfrak{U}, \varepsilon, \delta)) \leq 2\phi^{(x^*, \mathcal{Q}^*, \Lambda)}(t) - \frac{1}{K} \|Q^{*(k)} - V\|_F^2 + C(\varepsilon + \delta). \tag{7.29}$$

Remark 7.7. *The validity of the above assumption for general a priori measures is an open problem. However, in the particular case of the Gaussian a priori distribution the assumption is indeed effectively satisfied. See Section 8 and Theorem 8.1, in particular. This gives a complete proof of the Parisi formula for the case of Gaussian spins.*

Remark 7.8. *If the bound* (7.25) *holds then Lemma 7.6 with* $w = x_{r-1}$ *would imply that*

$$\varphi_N^{(2)}(k, t, \Sigma_N^{(2)}(\mathfrak{L}^*, \mathfrak{U}, \varepsilon, \delta)) \overset{?}{\leq} 2\phi^{(x^*, \mathcal{Q}^*, \Lambda^*)}(t) + C(\varepsilon + \delta). \tag{7.30}$$

The above inequality would then be a starting point for the a priori estimates in the spirit of Talagrand [30] which might lead to the proof of Assumption 7.1.

7.5 Gronwall's inequality and the Parisi formula

Theorem 7.1. *Suppose Assumption 7.1 holds.*

Then we have

$$\lim_{N \uparrow +\infty} p_N(\beta) = \sup_{U \in \mathrm{Sym}^+(d)} \inf_{\substack{\rho \in \mathcal{Q}'(U,d) \\ \Lambda \in \mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, \Lambda).$$

Proof. The proof follows the argument of [30] (see also [21]) with the adaptations to the case of multidimensional spins. The main ingredients are the Gronwall inequality and Lemma 7.2. Theorem 5.1 implies that

$$\lim_{N\uparrow+\infty} p_N(\beta) \le \sup_{U\in\mathrm{Sym}^+(d)} \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda).$$

We now turn to the proof of the matching lower bound. As in the proof of Theorem 1.2, it is enough to show that

$$\lim_{\varepsilon\downarrow+0}\lim_{N\uparrow+\infty} \varphi_N(1,x,Q,B(U,\varepsilon)) \ge \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,\Lambda). \tag{7.31}$$

1. We fix an arbitrary $U\in\mathrm{Sym}^+(d)$. Fix also some $t_0\in[0;1)$. By Assumption 7.1, we can find the corresponding $\theta(t_0,V,U)>0$ with the properties listed in the assumption. We pick any $\theta\in(0;\theta(t_0,V,U)]$ and let (n,ρ^*,Λ^*) be a correspondent θ-optimiser. Note that, by definition (7.28), we have

   $$\phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(1) = \mathscr{P}(\beta,\rho^*,U,\Lambda^*)$$

   and, by Definition 7.2,

   $$|\phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(1) - \inf_{\substack{\rho\in\mathscr{Q}'(U,d)\\ \Lambda\in\mathrm{Sym}(d)}} \mathscr{P}(\beta,\rho,U,\Lambda)| \le \theta. \tag{7.32}$$

2. We denote

   $$\Delta_N(t) \equiv \phi^{(x^*,\mathscr{Q}^*,\Lambda^*)}(t) - \varphi_N(t,x^*,Q^*,B(U,\varepsilon)).$$

   Note that, due to (5.12), we obviously have

   $$\Delta_N(t) \ge -C\varepsilon. \tag{7.33}$$

   Define

   $$\Delta(t) \equiv \lim_{N\uparrow+\infty} \Delta_N(t).$$

   The definition (7.28) and Theorem 5.4 yield

   $$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t) \le \frac{1}{2}\sum_{k=0}^{n-1}(x_{k+1}-x_k)\mu_k\left[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2\right] + C\varepsilon. \tag{7.34}$$

3. Let us set $D\equiv\sup_{\sigma\in\Sigma}\|\sigma\|_2$. We note that, for any $\sigma^{(1)},\sigma^{(2)}\in\Sigma_N$, we have

   $$R(\sigma^{(1)},\sigma^{(2)}) \in [-D^2;D^2]^{d\times d}.$$

   Given the constant K from (7.29), for any $c>0$, we define the set

   $$\Sigma_N^{(2),k}(U,\varepsilon) \equiv \left\{(\sigma^{(1)},\sigma^{(2)})\in\Sigma_N(B(U,\varepsilon))^2 : \|R(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2 \ge 2K\left(\Delta_N(t)+c\right)\right\}. \tag{7.35}$$

It is easy to see that by compactness we can find a finite covering of $\Sigma_N^{(2),k}(U,\varepsilon)$ by the neighbourhoods (7.11) with centres, e.g., in the corresponding set of admissible overlap matrices

$$\mathcal{V}_N^{(k)}(U,\varepsilon) \equiv \left\{ R(\sigma^{(1)},\sigma^{(1)}) \in [-D^2;D^2]^{d\times d} : (\sigma^{(1)},\sigma^{(2)}) \in \Sigma_N^{(2),k}(U,\varepsilon) \right\}.$$

That is, there exists $M = M(\varepsilon,\delta) \in \mathbb{N}$ and the finite collections of matrices $\{V(i)\}_{i=1}^M \subset \mathcal{V}_N^{(k)}(U,\varepsilon)$ and $\{U(i)\}_{i=1}^M \subset B(U,\varepsilon) \cap \mathrm{Sym}^+(d)$ such that

$$\Sigma_N^{(2),k}(U,\varepsilon) \subset \bigcup_{i=1}^{M} \Sigma_N^{(2)}(\mathfrak{L}^*(i),\mathfrak{U}(i),\varepsilon,\delta), \tag{7.36}$$

where

$$\mathfrak{U}(i) \equiv \begin{bmatrix} U(i) & V(i) \\ V^*(i) & U(i) \end{bmatrix} \in \mathrm{Sym}^+(2d),$$

and $\mathfrak{L}^*(i)$ is the corresponding δ-minimal Lagrange multiplier.

4. Given $i \in [1;M]\cap\mathbb{N}$, let $(n(i),x^*(i),Q^*(i),\Lambda^*(i))$ be the corresponding to $U(i)$ $\theta(i)$-optimisers. Due to Lipschitzianity of the Parisi functional (Proposition 6.2) and the fact that $U(i) \in B(U,\varepsilon)$ we can assume that $n(i) = n$. Using the bound (7.29) and the definition (7.35), we obtain

$$\begin{aligned}\varphi_N^{(2)}(k,t,x_i^*,Q_i^*,\Sigma_N^{(2)}(\mathfrak{L}^*(i),\mathfrak{U}(i),\varepsilon,\delta)) &\le 2\phi^{(x^*(i),\mathcal{Q}^*(i),\Lambda^*(i))}(t) - \frac{1}{K}\|Q^{(k)} - V(i)\|_{\mathrm{F}}^2 + C(\varepsilon+\delta)\\ &\le 2\varphi_N(t,x^*,Q^*,B(U,\varepsilon)) - c + C(\varepsilon+\delta),\end{aligned}$$

where the last inequality is again due to Lipschitzianity of the Parisi functional (Proposition 6.2) which allows to approximate functional's value at $(x^*(i),Q^*(i),\Lambda^*(i))$ by the value at (x^*,Q^*,Λ^*) paying the cost of at most $C\varepsilon$. Choose $c > C(\varepsilon+\delta)$. Then Lemma 7.2 implies that there exists $L = L(\varepsilon,\delta,c) > 0$ such that

$$\mu_k\left(\Sigma_N^{(2)}(\mathfrak{L}^*,\mathfrak{U},\varepsilon,\delta)\right) \le L\exp\left(-\frac{N}{L}\right).$$

Therefore, the inclusion (7.36) gives

$$\mu_k\left(\Sigma_N^{(2),k}(U,\varepsilon)\right) \le LM\exp\left(-\frac{N}{L}\right). \tag{7.37}$$

Hence, for each $k \in [1;n]\cap\mathbb{N}$, we have

$$\begin{aligned}\mu_k\left[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2\right] &= \mu_k\left[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2 \mathbb{1}_{\Sigma_N^{(2),k}(U,\varepsilon)}(\sigma^{(1)},\sigma^{(2)})\right]\\ &\quad + \mu_k\left[\|R_N(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2\left(1 - \mathbb{1}_{\Sigma_N^{(2),k}(U,\varepsilon)}(\sigma^{(1)},\sigma^{(2)})\right)\right]\\ &=: \mathrm{I} + \mathrm{II}.\end{aligned} \tag{7.38}$$

For all $(\sigma^{(1)},\sigma^{(2)}) \in \left(\Sigma_N(B(U,\varepsilon))^2 \setminus \Sigma_N^{(2),k}(U,\varepsilon,\delta)\right)$, we have by definition

$$\|R(\sigma^{(1)},\sigma^{(2)}) - Q^{(k)}\|_{\mathrm{F}}^2 < 2K\left(\Delta_N(t) + c\right).$$

Therefore, using Remark 5.4, we arrive to

$$\mathrm{II} \le 2K\left(\Delta_N(t)+c\right). \tag{7.39}$$

The bound (7.37) assures that

$$\mathrm{I} \le LM\exp\left(-\frac{N}{L}\right). \tag{7.40}$$

5. Combining (7.39) and (7.40) with (7.38) and (7.34), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t) \le 2K\left(\Delta_N(t)+c\right) + LM\exp\left(-\frac{N}{L}\right) + C(\varepsilon+\delta).$$

Hence,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big((\Delta_N(t)+c)\exp(-2Kt)\Big) &= \exp(-2Kt)\Big(\frac{\mathrm{d}}{\mathrm{d}t}(\Delta_N(t)+c) - 2K(\Delta_N(t)+c)\Big)\\ &\le \exp(-2Kt)\Big(\frac{\mathrm{d}}{\mathrm{d}t}(LM\exp\left(-\frac{N}{L}\right) + C(\varepsilon+\delta)\Big).\end{aligned}$$

Integrating the above inequality and noting that due to (5.12) $|\Delta_N(0)| \le C\varepsilon$, we arrive to

$$\begin{aligned}\Delta_N(t)+c \le &(C\varepsilon+c)\exp(-2Kt) + LM\exp\left(-\frac{N}{L}\right)\\ &+ C(\varepsilon+\delta)(\exp(-2Kt)-1) + C(\varepsilon+\delta).\end{aligned}$$

Passing consequently to the limits $N \uparrow +\infty$, $\varepsilon \downarrow +0$, $\delta \downarrow +0$ and finally $c \downarrow +0$ in the above inequality, we get

$$\lim_{\varepsilon\downarrow+0} \Delta(t) \le 0, \quad \text{for all } t \in [0;t_0].$$

The existence of the $N \uparrow +\infty$ limits is guaranteed by the general result of Guerra and Toninelli [19]. The limits $\varepsilon \downarrow +0$, $\delta \downarrow +0$ exist due to monotonicity. Finally, combining the above inequality with (7.33), we get

$$\lim_{\varepsilon\downarrow+0} \Delta(t) = 0, \quad \text{for all } t \in [0;t_0]. \tag{7.41}$$

6. Now, it is easy to extend the validity of (7.41) onto the whole interval $[0;1]$. Indeed, due to the boundedness of the derivatives of φ_N and ϕ, we have, for any $t \in [0;1]$,

$$\begin{aligned}\Delta_N(t) &\le \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t)\mathrm{d}t\\ &= \left(\int_0^{t_0} + \int_{t_0}^1\right)\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t)\mathrm{d}t\\ &\le \left(\Delta_N(t_0) - \Delta_N(0)\right) + \int_{t_0}^1 \left|\frac{\mathrm{d}}{\mathrm{d}t}\Delta_N(t)\right|\mathrm{d}t\\ &\le \Delta_N(t_0) + L(1-t_0).\end{aligned} \tag{7.42}$$

Passing to the $N \uparrow +\infty$ limit, applying (7.41), and then to $t_0 \to 1$ limit in (7.42), we get

$$\lim_{\varepsilon\downarrow+0} \Delta(t) = 0, \quad \text{for all } t \in [0;1].$$

7. In particular, the previous formula yields

$$0 = \lim_{\varepsilon \downarrow +0} \Delta(1) = \phi^{(x^*, \mathscr{Q}^*, \Lambda^*)}(1) - \lim_{\varepsilon \downarrow +0} \varphi_N(1, x^*, Q^*, B(U, \varepsilon)).$$

Note that $\varphi_N(1, x, Q, B(U, \varepsilon))$ does not depend on the choice of x and Q. Hence, by (7.32), we obtain

$$|\lim_{\varepsilon \downarrow +0} \varphi_N(1, x^*, Q^*, B(U, \varepsilon)) - \inf_{\substack{\rho \in \mathscr{Q}'(U,d) \\ \Lambda \in \mathrm{Sym}(d)}} \mathscr{P}(\beta, \rho, U, \Lambda)| \leq \theta.$$

The proof of (7.31) is finished by noticing that the θ can be made arbitrary small.

□

8 Proof of the local Parisi formula for the SK model with multidimensional Gaussian spins

In this section, we prove Theorem 1.3. The rich symmetries of the Gaussian a priori distribution allow rather explicit computations of the X_0 terms (see (1.11)). This allows us to prove that the analogon of Assumption 7.1 is satisfied, implying the Parisi formula for the local free energy (Theorem 1.3).

Remark 8.1. *The case of Gaussian spins is very tractable due to the (unusually) good symmetry (i.e., the rotational invariance) of the Gaussian measure. Therefore, it is not surprising that in this case the calculus resembles the one for the spherical SK model, cf. [23; 29].*

We start from the estimates under a generic (i.e., no simultaneous diagonalisation, cf. Section 6.3) scenario.

8.1 The case of positive increments

Let, for $k \in [0; n] \cap \mathbb{N}$,

$$\Delta Q^{(k)} \equiv Q^{(k+1)} - Q^{(k)}.$$

We define, for $\Lambda \in \mathrm{Sym}(d)$, a family of matrices $\left\{D^{(l)} \in \mathbb{R}^{d \times d}\right\}_{l=0}^{n+1}$ as follows

$$D^{(n+1)} \equiv C,$$

and, further, for $k \in [0; n] \cap \mathbb{N}$,

$$D^{(k)} \equiv C - \Lambda - 2\beta^2 \sum_{l=k}^{n} x_l \Delta Q^{(l)}. \tag{8.1}$$

We assume that the matrices Λ and C are such that, for all $l \in [1; n+1] \cap \mathbb{N}$, we have

$$D^{(l)} \succ 0.$$

We need the following two small (and surely known) technical Lemmata which exploit the symmetries of our Gaussian setting. We include their statements for reader's convenience.

Lemma 8.1. *Fix some vector $h \in \mathbb{R}^d$ and a Gaussian random vector $z \in \mathbb{R}^d$ with* $\operatorname{Var} z = C^{-1} \in \mathbb{R}^{d\times d}$. *Then we have*

$$\mathbb{E}^z \left[\exp\left(\langle z,h\rangle + \langle \Lambda\sigma,\sigma\rangle\right)\right] = \left(\det\left[C\,(C-\Lambda)^{-1}\right]\right)^{1/2} \times \exp\left(\frac{1}{2}\left\langle (C-\Lambda)^{-1}h,h\right\rangle\right).$$

Proof. This is a standard Gaussian averaging argument. □

Lemma 8.2. *For a positive definite matrix $\Delta Q \in$ Sym(d), let $z \sim \mathcal{N}(0,\Delta Q)$. We fix also another positive definite matrix $D \in$ Sym(d) such that $\Delta Q^{-1} \succ D^{-1}$.*
Then we have

$$\mathbb{E}^z\left[\exp\left(\frac{1}{2}\langle D^{-1}(z+h),z+h\rangle\right)\right] = \left(\det\left[D(D-\Delta Q)^{-1}\right]\right)^{-1/2} \times \int_{\mathbb{R}^d} \exp\left(\frac{1}{2}\langle (D-\Delta Q)^{-1}h,h\rangle\right).$$

Proof. This is a standard Gaussian averaging argument. See, e.g., [29] for an argument in 1-D. □

Now we are ready to compute the term $X_0(x,\mathcal{Q},U,\Lambda)$ (see (1.11)) corresponding to the a priori distribution (1.26) in a rather explicit way.

Lemma 8.3. *We have*

$$X_0(x,\mathcal{Q},U,\Lambda) = \frac{1}{2}\left(\langle [D^{(1)}]^{-1},\Delta Q^{(0)}\rangle + \langle [D^{(1)}]^{-1}h,h\rangle + \sum_{l=1}^{n}\frac{1}{x_l}\log\left(\frac{\det D^{(l+1)}}{\det D^{(l)}}\right)\right).$$

Proof. 1. We start from computing the following quantity

$$X_{n+1} \equiv \log\int_{\mathbb{R}^d}\exp\left(\sum_{l=0}^{n}\langle Y^{(l)},\sigma\rangle + \langle\Lambda\sigma,\sigma\rangle\right)\mathrm{d}\mu(\sigma), \tag{8.2}$$

where $Y^{(l)} \in \mathbb{R}^d$ are independent Gaussian vectors with variance

$$\operatorname{Var}\left[Y^{(l)}\right] = 2\beta^2\Delta Q^{(l)}.$$

We denote

$$\widetilde{h} \equiv h + \sum_{l=0}^{n} Y^{(l)}.$$

Lemma 8.1 gives

$$\int_{\mathbb{R}^d}\exp\left(\sum_{l=0}^{n}\langle Y^{(l)},\sigma\rangle + \langle\Lambda\sigma,\sigma\rangle\right)\mathrm{d}\mu(\sigma) = \left(\det\left[C\,(C-\Lambda)^{-1}\right]\right)^{1/2} \times \exp\left(\frac{1}{2}\left\langle (C-\Lambda)^{-1}\widetilde{h},\widetilde{h}\right\rangle\right).$$

2. Next, we define, for $l \in [0; n] \cap \mathbb{N}$, recursively the following quantities

$$X_l \equiv \frac{1}{x_l} \log \mathbb{E}^{Y_l} \left[\exp \left(x_l X_{l+1} \right) \right].$$

Applying the Lemma 8.2 to (8.2) recursively, we obtain

$$X_1 \equiv \frac{1}{2} \langle [(D^{(1)}]^{-1} \left(Y^{(0)} + h \right), Y^{(0)} + h \rangle + \frac{1}{2} \sum_{l=1}^{n} \frac{1}{x_l} \log \left(\frac{\det D^{(l+1)}}{\det D^{(l)}} \right). \tag{8.3}$$

Recall that we have

$$\begin{aligned} X_0 &= \lim_{x \to +0} \frac{1}{x} \log \mathbb{E}^{Y_0} \left[\exp \left(x X_1 \right) \right] \\ &= \mathbb{E}^{Y_0} \left[X_1 \right] \end{aligned} \tag{8.4}$$

and note that

$$\mathbb{E}^{Y_0} \left[\langle [D^{(1)}]^{-1} (Y^{(0)} + h), Y^{(0)} + h \rangle \right] = 2\beta^2 \langle [D^{(1)}]^{-1}, \Delta Q^{(0)} \rangle + \langle [D^{(1)}]^{-1} h, h \rangle. \tag{8.5}$$

Hence, combining (8.4) and (8.5) with (8.3), we obtain the theorem.

□

8.2 Simultaneous diagonalisation scenario

In what follows, we employ the simultaneous diagonalisation scenario introduced in Section 6.3. Suppose that, for $l \in [0; n+1] \cap \mathbb{N}$, and some matrix $O \in \mathcal{O}(d)$, we have

$$D^{(l)} \equiv O^* d^{(l)} O,$$

where the vectors $d^{(l)} \in \mathbb{R}^d$, for $l \in [0; n] \cap \mathbb{N}$, satisfy

$$0 \prec d^{(l)} \prec d^{(l+1)}.$$

That is, the vectors $d^{(l)}$ are (component-wise) increasingly ordered and non-negative.

Lemma 8.4. *We have*

$$X_0(x, \mathcal{Q}, U, \Lambda) = \frac{1}{2} \sum_{v=1}^{d} \left(\frac{2\beta^2 q_v^{(1)} + h_v^2}{d_v^{(1)}} + \sum_{l=1}^{n} \frac{1}{x_l} \log \left(\frac{d_v^{(l+1)}}{d_v^{(l)}} \right) \right), \tag{8.6}$$

$$\frac{\beta^2}{2} \sum_{k=1}^{n} x_k \left(\|Q^{(k+1)}\|_F^2 - \|Q^{(k)}\|_F^2 \right) = \frac{\beta^2}{2} \sum_{k=1}^{n} x_l \left(\|q^{(k+1)}\|_2^2 - \|q^{(k)}\|_2^2 \right). \tag{8.7}$$

Proof. This is a standard argument which relies on the standard invariance properties of the determinant and the matrix trace.

□

Define the 1-D Parisi functional for the case (1.26) as

$$\mathscr{P}(\rho,\lambda)\equiv -\lambda u+\frac{2\beta^2 q^{(1)}+h^2}{d^{(1)}}+\sum_{l=1}^{n}\frac{1}{x_l}\log\left(\frac{d^{(l+1)}}{d^{(l)}}\right)$$
$$-\beta^2\sum_{l=1}^{n}x_l\left([q^{(l+1)}]^2-[q^{(l)}]^2\right). \tag{8.8}$$

Proposition 8.1. *There exists $C=C(\Sigma)>0$ such that, for all $u\in\mathbb{R}_+^d$ and all $\varepsilon,\delta>0$, there exists an δ-minimal Lagrange multiplier $\lambda=\lambda(U,\varepsilon,\delta)\in\mathbb{R}^d$ in* (1.12) *such that, for all $t\in[0;1]$ and all (x,ρ), we have*

$$p_N(\Sigma_N(\mathscr{V}(\Lambda,U,\varepsilon,\delta)))\leq\frac{1}{2}\inf_{\rho,\lambda}\left(\sum_{v=1}^{d}\mathscr{P}(\rho_v,\lambda_v)\right)+C(\varepsilon+\delta) \tag{8.9}$$

and

$$\lim_{N\uparrow+\infty}p_N(\Sigma_N(B(U,\varepsilon)))\geq\frac{1}{2}\inf_{\rho,\lambda}\left(\sum_{v=1}^{d}\mathscr{P}(\rho_v,\lambda_v)+\lim_{N\uparrow+\infty}\int_0^1\mathscr{R}(t,x,Q,\Sigma_N(B(U,\varepsilon)))\mathrm{d}t\right)$$
$$+C(\varepsilon+\delta). \tag{8.10}$$

Proof. We combine (8.6) and (8.7) and the Proposition 5.2 to get (8.9) and (8.10). □

8.3 The Crisanti-Sommers functional in 1-D

In this subsection, we adapt the proof of [29] to obtain the equivalence between the (very tractable) Crisanti-Sommers functional [11] and the Parisi one (8.8) in the case of the Gaussian a priori measure (1.26). Similar ideas based on the symmetry of the a priori measure were exploited in the case of the spherical models by [4; 23].

We restrict the consideration to 1-D situation for a moment. Given $u\geq 0$, consider $\rho\in\mathscr{Q}_n'(u,1)$, $\lambda\in\mathbb{R}$, $h\in\mathbb{R}$ and let $\{d^{(l)}\in\mathbb{R}\}_{l=1}^{n+1}$ be the scalars playing the role of matrices $D^{(l)}$ (cf. (8.1)). That is,

$$d^{(l)}\equiv c-\lambda-2\beta^2\sum_{k=l}^{n}x_k\left(q^{(k+1)}-q^{(k)}\right),$$
$$d^{(n+1)}\equiv c.$$

We define, for $k\in[1;n]\cap\mathbb{N}$, the family of vectors $\{s^{(k)}\in\mathbb{R}^d\}_{k=0}^{n}$ by

$$s^{(k)}\equiv\sum_{l=k}^{n}x_l\left(q^{(l+1)}-q^{(l)}\right). \tag{8.11}$$

We also define the Crisanti-Sommers functional as follows

$$\mathscr{CS}(\rho)\equiv 1-cu+h^2 s^{(1)}+\frac{q^{(1)}}{s^{(1)}}+\sum_{l=1}^{n-1}\frac{1}{x_l}\log\left(\frac{s^{(l)}}{s^{(l+1)}}\right)+\log\left[c(u-q^{(n)})\right]$$
$$+\beta^2\sum_{l=1}^{n}x_l\left([q^{(l+1)}]^2-[q^{(l)}]^2\right). \tag{8.12}$$

Lemma 8.5. *If (ρ, λ) is an optimiser for* (8.8)*, that is,*

$$\mathscr{P}(\rho, \lambda) = \inf_{(\rho', \lambda')} \mathscr{P}(\rho', \lambda'), \tag{8.13}$$

then, for all $k \in [1; n] \cap \mathbb{N}$, the pair (ρ, λ) satisfies

$$q^{(k)} = \frac{h^2 + 2\beta^2 q^{(1)}}{[d^{(1)}]^2} + \sum_{l=1}^{k-1} \frac{1}{x_l} \left(\frac{1}{d^{(l)}} - \frac{1}{d^{(l+1)}} \right). \tag{8.14}$$

Moreover,

$$\lambda = c - 2\beta^2 (u - q^{(n)}) - (u - q^{(n)})^{-1}, \tag{8.15}$$

and, for all $k \in [1; n] \cap \mathbb{N}$, we have

$$\frac{1}{s^{(k+1)}} - \frac{1}{s^{(k)}} = 2\beta^2 x_k \left(q^{(k+1)} - q^{(k)} \right), \tag{8.16}$$

and also

$$s^{(k)} = \frac{1}{d^{(k)}}. \tag{8.17}$$

Remark 8.2. *In the formulation of the theorem (as well as elsewhere), it is implicit that $d^{(k)} = d^{(k)}(\rho, \lambda)$ and $s^{(k)} = s^{(k)}(\rho, \lambda)$.*

Proof. 1. Rearranging the terms in (8.8), we observe that

$$\begin{aligned} \mathscr{P}(\rho', \lambda') = & -\lambda u + \frac{2\beta^2 q^{(1)} + h^2}{d^{(1)}} + \sum_{l=2}^{n} \log d^{(l)} \left(\frac{1}{x_{l-1}} - \frac{1}{x_l} \right) + \frac{1}{x_n} \log d^{(n+1)} - \frac{1}{x_1} \log d^{(1)} \\ & - \beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right). \end{aligned} \tag{8.18}$$

We compute, for $k, l \in [1; n] \cap \mathbb{N}$,

$$\frac{\partial d^{(l)}}{\partial q^{(k)}} = \begin{cases} 0, & k < l, \\ 2\beta^2 x_k, & l = k, \\ 2\beta^2 \left(x_k - x_{k-1} \right), & k > l. \end{cases} \tag{8.19}$$

Using (8.19) and the representation (8.18), we compute the necessary condition for (q, λ) satisfy (7.27), for $k \in [2; n] \cap \mathbb{N}$,

$$\begin{aligned} 0 = \frac{\partial}{\partial q^{(k)}} \mathscr{P}(q, \lambda) = & 2\beta^2 \left(x_k - x_{k-1} \right) \Bigg[-\frac{2\beta^2 q^{(1)} + h^2}{[d^{(1)}]^2} + \sum_{l=2}^{k-1} \frac{1}{d^{(l)}} \left(\frac{1}{x_{l-1}} - \frac{1}{x_l} \right) \\ & + \frac{1}{d^{(k)} x_{k-1}} - \frac{1}{x_1 d^{(1)}} + q_k \Bigg]. \end{aligned} \tag{8.20}$$

We also have (for $k=1$)

$$0=\frac{\partial}{\partial q^{(1)}}\mathscr{P}(q,\lambda)=2\beta^2\left[\frac{d^{(1)}-x_1\left(q^{(1)}+h^2\right)}{[d^{(1)}]^2}-\frac{x_1}{x_1 d^{(1)}}+x_1 q^{(1)}\right]$$
$$=2\beta^2 x_1\left[q^{(1)}-\frac{\left(q^{(1)}+h^2\right)}{[d^{(1)}]^2}\right]. \tag{8.21}$$

Relations (8.20) and (8.21) then imply (8.14).

2. Using the fact that

$$\frac{\partial d^{(l)}}{\partial\lambda}=-1,$$

we obtain

$$\frac{\partial}{\partial\lambda}\mathscr{P}(q,\lambda)=-u+\frac{h^2+2\beta^2 q^{(1)}}{[d^{(1)}]^2}+\sum_{l=1}^{n-1}\frac{1}{x_l}\left(\frac{1}{d^{(l)}}-\frac{1}{d^{(l+1)}}\right)+\frac{1}{d^{(n)}}. \tag{8.22}$$

Applying (8.14) with $k=n$ in (8.22), we obtain that the necessary condition for λ to satisfy (8.13) is as follows

$$0=\frac{\partial}{\partial\lambda}\mathscr{P}(q,\lambda)=-u+q^{(n)}+\frac{1}{x_n}\left(\frac{1}{d^{(n)}}-\frac{1}{d^{(n+1)}}\right)$$
$$=-u+q^{(n)}+\frac{1}{d^{(n)}}=-u+q^{(n)}+\left(c-\lambda-2\beta^2(u-q^{(n)})\right)^{-1} \tag{8.23}$$

which implies (8.15).

3. Relation (8.16) is proved as follows. Subtracting the relations (8.14), we obtain, for $k\in[1;n-1]\cap\mathbb{N}$,

$$x_k\left(q^{(k+1)}-q^{(k)}\right)=\frac{1}{d^{(k)}}-\frac{1}{d^{(k+1)}}. \tag{8.24}$$

By (8.23), we have

$$x_n\left(q^{(n+1)}-q^{(n)}\right)=u-q^{(n)}=\frac{1}{d^{(n)}}.$$

(That is, (8.24) is valid also for $k=n$.) Combining the previous two relations, we get, for $k\in[1;n]\cap\mathbb{N}$,

$$s^{(k)}=\frac{1}{d^{(k)}}. \tag{8.25}$$

Using (8.25) and (8.24), we get

$$\begin{aligned}2\beta^2 x_k\left(q^{(k+1)}-q^{(k)}\right)&=d^{(k+1)}-d^{(k)}\\ \text{(by (8.24))}\quad&=d^{(k+1)}d^{(k)}x_k\left(q^{(k+1)}-q^{(k)}\right)=d^{(k+1)}d^{(k)}\left(s^{(k)}-s^{(k+1)}\right)\\ \text{(by (8.25))}\quad&=\frac{1}{s^{(l+1)}}-\frac{1}{s^{(l)}}\end{aligned}$$

which is (8.16).

□

Lemma 8.6. *If ρ is an optimiser of* (8.12)*, that is,*

$$\mathscr{C}\mathscr{S}(\rho)=\inf_{\rho'}\mathscr{C}\mathscr{S}(\rho'),$$

then, for all $l\in[1;n]\cap\mathbb{N}$, (8.16) *holds.*

Proof. The strategy is the same as in the previous lemma. We rearrange the summands in (8.12) to get

$$\begin{aligned}\mathscr{C}\mathscr{S}(\rho)=&h^2s^{(1)}+\frac{q^{(1)}}{s^{(1)}}+\frac{\log s^{(1)}}{x_1}-\frac{\log s^{(n)}}{x_{n-1}}+\sum_{l=2}^{n-1}\left(\frac{1}{x_l}-\frac{1}{x_{l+1}}\right)\log s^{(l)}\\&+\log\left(c(u-q^{(n)})\right)+\beta^2\sum_{l=1}^{n}x_l\left([q^{(l+1)}]^2-[q^{(l)}]^2\right).\end{aligned}\tag{8.26}$$

We have, for $k,l\in[1;n]\cap\mathbb{N}$,

$$\frac{\partial s^{(l)}}{\partial q^{(k)}}=\begin{cases}0, & k<l,\\ -x_k, & k=l,\\ x_{k-1}-x_k, & k>l.\end{cases}\tag{8.27}$$

1. Relation (8.27) implies, for $k\in[2;n-1]\cap\mathbb{N}$,

$$\begin{aligned}\frac{\partial}{\partial q^{(k)}}\mathscr{C}\mathscr{S}(\rho)=&h^2(x_{k-1}-x_k)-\frac{q^{(1)}}{[s^{(1)}]^2}(x_{k-1}-x_k)+\frac{x_{k-1}-x_k}{x_1s^{(1)}}\\&+\sum_{l=2}^{k-1}\frac{x_{k-1}-x_k}{s^{(l)}}\left(\frac{1}{x_l}-\frac{1}{x_{l-1}}\right)-\frac{x_k}{s^{(k)}}\left(\frac{1}{x_k}-\frac{1}{x_{k-1}}\right)\\&+2\beta^2q^{(k)}\left(x_{k-1}-x_k\right)=0.\end{aligned}$$

   Hence,

$$\begin{aligned}2\beta^2q^{(k)}&=-h^2+\frac{q^{(1)}}{[s^{(1)}]^2}-\frac{1}{x_1s^{(1)}}+\frac{1}{x_{k-1}s^{(k)}}-\sum_{l=2}^{k-1}\frac{1}{s^{(l)}}\left(\frac{1}{x_l}-\frac{1}{x_{l-1}}\right)\\&=-h^2+\frac{q^{(1)}}{[s^{(1)}]^2}-\sum_{l=1}^{k-1}\frac{1}{x_l}\left(\frac{1}{s^{(l)}}-\frac{1}{s^{(l+1)}}\right).\end{aligned}\tag{8.28}$$

2. To handle the case $k=n$, we note that

$$\log\left(1+c(u-q^{(n)})\right)=\frac{1}{x_n}\log\left(\frac{s^{(n)}}{s^{(n+1)}}\right),$$

   and, hence, the argument in the previous item shows that (8.28) is also valid for $k=n$.

3. Differentiating the representation (8.26) with respect to $q^{(1)}$ and using (8.27), we obtain

$$\frac{\partial}{\partial q^{(1)}}\mathscr{C}\mathscr{S}(\rho) = -x_1 h^2 + \frac{1}{s^{(1)}} + \frac{x_1 q^{(1)}}{[s^{(1)}]^2} - \frac{x_1}{x_1 s^{(1)}} - 2\beta^2 x_1 q^{(1)} = 0.$$

   Therefore,

$$2\beta^2 q^{(1)} = -h^2 + \frac{q^{(1)}}{[s^{(1)}]^2}$$

   which is (8.28), for $k = 1$.

4. Subtracting equations (8.28), we arrive to (8.16), for all $k \in [1;n] \cap \mathbb{N}$.

□

Proposition 8.2. *The functionals* (8.12) *and* (8.8) *are equivalent in the following sense*

$$\inf_{\rho',\lambda'} \mathscr{P}(\rho',\lambda') = \inf_{\rho'} \mathscr{C}\mathscr{S}(\rho').$$

Proof. 1. Let (ρ,λ) be the solutions of equations (8.16) and (8.15). Lemma 8.6 guarantees that ρ is the optimiser of the Crisnati-Sommers functional and Lemma 8.5 assures that (ρ,λ) is the optimiser of the Parisi functional.

2. We have

$$\begin{aligned}\mathscr{P}(\rho,\lambda) - \mathscr{C}\mathscr{S}(\rho) = & -\lambda u + 2\beta^2 q^{(1)} s^{(1)} - \frac{q^{(1)}}{s^{(1)}} + cu - 1 \\ & - 2\beta^2 \sum_{l=1}^{n} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right). \end{aligned} \tag{8.29}$$

   We can simplify the $\Phi[B]$-like term (that is the summation) in (8.29), using (8.16) and (8.15). Indeed,

$$\begin{aligned} 2\beta^2 \sum_{l=1}^{n-1} x_l \left([q^{(l+1)}]^2 - [q^{(l)}]^2 \right) &= 2\beta^2 \sum_{l=1}^{n-1} x_l \left(q^{(l+1)}[q^{(l+1)} - q^{(l)}] + q^{(l)}[q^{(l+1)} - q^{(l)}] \right) \\ \text{(by (8.16) and (8.11))} \; &= \sum_{l=1}^{n-1} \left(2\beta^2 q^{(l+1)} \left[s^{(l)} - s^{(l+1)} \right] + q^{(l)} \left[\frac{1}{s^{(l+1)}} - \frac{1}{s^{(l)}} \right] \right). \end{aligned} \tag{8.30}$$

   Regrouping the summands in (8.30), we get

$$\begin{aligned} (8.30) \; = & 2\beta^2 \sum_{l=1}^{n-1} s^{(l)} \left(q^{(l+1)} - q^{(l)} \right) + 2\beta^2 \left(q^{(1)} s^{(1)} - q^{(n)} s^{(n)} \right) \\ & + \sum_{l=1}^{n-1} \frac{q^{(l)} - q^{(l+1)}}{s^{(l+1)}} + \left(\frac{q^{(n)}}{s^{(n)}} - \frac{q^{(1)}}{s^{(1)}} \right). \end{aligned} \tag{8.31}$$

Due to (8.16), we have

$$2\beta^2\left(q^{(l+1)}-q^{(l)}\right)=\frac{s^{(l)}-s^{(l+1)}}{x_l s^{(l)}s^{(l+1)}}=\frac{q^{(l+1)}-q^{(l)}}{s^{(l)}s^{(l+1)}}.$$

Applying the previous relation, we get that the both summations in (8.31) cancel out and we end up with

$$(8.31)\ =2\beta^2\left(q^{(1)}s^{(1)}-q^{(n)}s^{(n)}\right)+\frac{q^{(n)}}{s^{(n)}}-\frac{q^{(1)}}{s^{(1)}}.$$

Now, turning back to (8.29), we get

$$\begin{aligned}\mathscr{P}(\rho,\lambda)-\mathscr{C}\mathscr{S}(\rho)&=-\lambda u-2\beta^2\left(u^2-[q^{(n)}]^2\right)+2\beta^2 q^{(n)}s^{(n)}-\frac{q^{(n)}}{s^{(n)}}+cu-1\\ \text{(by (8.15)) and (8.11)}&=-u\left(c-2\beta^2(u-q^{(n)})-(u-q^{(n)})^{-1}\right)-2\beta^2\left(u^2-[q^{(n)}]^2\right)\\ &\quad-\frac{q^{(n)}}{u-q^{(n)}}+2\beta^2 q^{(n)}\left(u-q^{(n)}\right)+cu-1\\ &=0.\end{aligned}$$

□

8.4 Replica symmetric calculations

In this subsection, we shall consider the one dimensional case of the a priori measure (1.26) with $h=0$. We shall also restrict the computations to the case $n=1$ which is often referred to in physical literature as the replica symmetric scenario. It is indeed the right scenario under the above assumptions, as shows Theorem 1.3.

Lemma 8.7. *Let μ satisfy* (1.26) *with $h=0$. Assume $d=1$, $n=1$ and $c>0$. Given $u\geq 0$, we have*

$$\inf_{\rho\in\mathscr{Q}(u,1)}\mathscr{C}\mathscr{S}(\rho)=\inf_{q\in[0;u]}\left(1-cu+\log\left(c(u-q)\right)+\frac{q}{u-q}+\beta^2\left(u^2-q^2\right)\right)=f(c,u), \tag{8.32}$$

where $f(c,u)$ is defined in (1.27).

Proof. Using the definitions, we obtain

$$\frac{\partial}{\partial q}\mathscr{C}\mathscr{S}(\rho)=\frac{\partial}{\partial q}\left[\log\left(u-q\right)+\frac{q}{u-q}+\beta^2\left(u^2-q^2\right)\right]=\frac{q}{(u-q)^2}-2\beta^2 q.$$

Hence, the critical points of $q\mapsto\mathscr{C}\mathscr{S}(q,u)$ are

$$q_0=0,q_{1,2}=u\pm\frac{\sqrt{2}}{2\beta}.$$

Furthermore, we also have

$$\frac{\partial^2}{\partial q^2}\mathscr{C}\mathscr{S}(q,u)=\frac{1}{(u-q)^2}+\frac{2q}{(u-q)^3}-2\beta^2.$$

Hence, as a simple calculation shows, the infima in (8.32) are attained on

$$q^* = \begin{cases} 0, & u \leq \frac{\sqrt{2}}{2\beta}, \\ u - \frac{\sqrt{2}}{2\beta}, & u > \frac{\sqrt{2}}{2\beta} \end{cases} \tag{8.33}$$

which implies (8.32). ☐

Lemma 8.8. *Under the assumptions of Lemma 8.7, we have*

1. *For $c \geq 2\sqrt{2}\beta$, we have*

$$\sup_{u\geq 0} \inf_{q\in[0;u]} \mathscr{C}\mathscr{S}(q,u) = \mathscr{C}\mathscr{S}(0,u^*) = \beta^2(u^*)^2 + \log cu^* - cu^* + 1,$$

where

$$u^* \equiv \frac{1}{4\beta^2}\left(c - \sqrt{c^2 - 8\beta^2}\right).$$

2. *For $c < 2\sqrt{2}\beta$, we have*

$$\sup_{u\geq 0} \inf_{q\in[0;u]} \mathscr{C}\mathscr{S}(q,u) = +\infty.$$

Remark 8.3. *Under the assumptions, the above theorem says that from the point of view of the global free energy, the system can only exist in the "high temperature" scenario, cf. (1.27). The threshold at $c_0 = 2\sqrt{2}\beta$ could be easily understood from the perspective of the norms of random matrices.*

Proof. 1. Suppose $c \geq 2\sqrt{2}\beta$. Recalling (1.27), for $u \in (0; \frac{\sqrt{2}}{2\beta}]$, we introduce the following function

$$f(u) \equiv \log(cu) + \beta^2 u^2 - cu + 1.$$

We have

$$\frac{\partial}{\partial u} f(u) = \frac{1}{u} + 2\beta^2 u - c.$$

Hence, the critical points of the function f are

$$u_{1,2} = \frac{c \pm \sqrt{c^2 - 8\beta^2}}{4\beta^2}.$$

Furthermore, we have

$$\frac{\partial^2}{\partial u^2} f(u) = 2\beta^2 - \frac{1}{u^2}.$$

We notice that $u^* \leq \frac{\sqrt{2}}{2\beta}$ and, hence, due to (1.27)

$$\mathscr{C}\mathscr{S}(0,u^*) = \beta^2(u^*)^2 + \log cu^* - cu^* + 1.$$

2. If $c < 2\sqrt{2}\beta$, then the function

$$u \mapsto (2\sqrt{2}\beta - c)u + \log\frac{c}{\beta} - \frac{1}{2}\left(1 + \log 2\right)$$

is unbounded on $(\frac{\sqrt{2}}{2\beta}; +\infty)$.

□

8.5 The multidimensional Crisanti-Sommers functional

Recall the definition (7.28).

Proposition 8.3. *Assume $d = 1$. Given $u > 0$, we have*

$$2\phi^{(x*,\mathscr{Q}*,\Lambda*)}(t) = \begin{cases} \left(3\sqrt{2}\beta - c\right) u + \log\frac{c}{\beta} - 1 - \frac{\log 2}{2} - t\left(\sqrt{2}u\beta - \frac{1}{2}\right), & u > \frac{\sqrt{2}}{2\beta}, \\ 2\beta^2(u)^2 + \log(cu) - cu + 1 - t\beta^2(u)^2, & u \le \frac{\sqrt{2}}{2\beta}. \end{cases} \tag{8.34}$$

Proof. Combining (8.8), (8.12) with Lemma 8.7 and Proposition 8.2, we get the claim.

□

8.6 Talagrand's a priori estimates

In this subsection, we prove that Assumption 7.1 is satisfied in the case of the Gaussian a priori distribution (1.26) with $h = 0$.

Theorem 8.1. *Let μ satisfy* (1.26) *with $h = 0$, assume $U \in \mathrm{Sym}^+(d)$ is such that $\min_v u_v > \frac{\sqrt{2}}{2\beta}$ and suppose $C \succ 0$. Let $Q = Q^*$ and $\Lambda = \Lambda^*$.*
Then, for any $t_0 \in (0;1)$ and any $t \in (0;t_0]$, we have (cf. (7.29) *with $k = 1$)*

$$\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \le 2\phi^{(x,\mathscr{Q},\Lambda)}(t) - \frac{1}{K}\|Q^{(1)} - V\|_F^2 + \mathscr{O}(\varepsilon + \delta). \tag{8.35}$$

Proof. 1. We employ the notations of Section 7.2. Let $\mathfrak{n} = 1$. Given $\mathfrak{U} \in \mathrm{Sym}(2d)$ (cf. (7.6)), choose arbitrary matrices $\left\{\mathfrak{Q}^{(l)} \in \mathrm{Sym}(2d) \,|\, l \in [0;2] \cap \mathbb{N}\right\}$ satisfying (7.7). Define $\mathfrak{x} \equiv x$ which, in particular, implies that $\zeta = \xi$. Finally, we set, for $l \in [0; n+1] \cap \mathbb{N}$, $\widetilde{Q}^{(l)} \equiv Q^{(l)}$.

2. Proposition 7.1 implies that, for any δ-minimal $\mathfrak{L} \in \mathbb{R}^{2d\times 2d}$, we have

$$\begin{aligned}\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \le -\langle \mathfrak{L},\mathfrak{U}\rangle - \frac{t\beta^2}{2}\left(\|\mathfrak{Q}^{(2)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(1)}\|_{\mathrm{F}}^2\right) \\ + X_0^{(2)}(1,\mathfrak{x},\widehat{\mathfrak{Q}}^{(l)}(t),\mathfrak{L}) + \mathscr{O}(\varepsilon + \delta).\end{aligned} \tag{8.36}$$

3. We define a matrix $\mathfrak{C} \in \mathbb{R}^{2d\times 2d}$ as follows

$$\mathfrak{C} \equiv \begin{bmatrix} C & 0 \\ 0 & C \end{bmatrix}.$$

Recalling (8.1), we define also the following matrices $\mathfrak{D}^{(2)} \equiv \mathfrak{C}$ and

$$\mathfrak{D}^{(1)} \equiv \mathfrak{C} - \mathfrak{L} - \left(\widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t)\right). \tag{8.37}$$

Applying Proposition 8.1 to (8.36), we get

$$\begin{aligned}\varphi_N^{(2)}(1,t,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \le \frac{1}{2}\Big[&-\langle \mathfrak{L},\mathfrak{U}\rangle - t\beta^2\left(\|\mathfrak{Q}^{(2)}\|_{\mathrm{F}}^2 - \|\mathfrak{Q}^{(1)}\|_{\mathrm{F}}^2\right) \\ &+2\beta^2\langle [\mathfrak{D}^{(1)}]^{-1}, \widehat{\mathfrak{Q}}^{(1)}(t)\rangle + \log\left(\frac{\det \mathfrak{D}^{(2)}}{\det \mathfrak{D}^{(1)}}\right)\Big] + \mathcal{O}(\varepsilon) \\ =: \widetilde{\Phi}^{(2),k,\mathfrak{r},\mathfrak{L}} + \mathcal{O}(\varepsilon).\end{aligned} \tag{8.38}$$

4. Assume that the matrices

$$\mathfrak{Q}^{(1)}, \mathfrak{Q}^{(2)}, \mathfrak{D}^{(1)} \in \mathbb{R}^{2d\times 2d} \tag{8.39}$$

are simultaneously diagonalisable in the same basis which is given by the orthogonal matrix $\mathfrak{O} \in \mathbb{R}^{2d\times 2d}$. Let the vectors

$$\mathfrak{q}^{(1)}, \mathfrak{q}^{(2)}, \mathfrak{d}^{(1)} \in \mathbb{R}^{2d} \tag{8.40}$$

be the corresponding spectra of the matrices (8.39). That is, we assume that

$$\begin{aligned}\mathfrak{Q}^{(1)} = \mathfrak{O}^* \operatorname{diag} \mathfrak{q}^{(1)}\mathfrak{O}, \mathfrak{Q}^{(2)} = \mathfrak{O}^* \operatorname{diag} \mathfrak{q}^{(2)}\mathfrak{O}, \\ \mathfrak{D}^{(1)} = \mathfrak{O}^* \mathfrak{d}^{(1)}\mathfrak{O}, \widetilde{\mathfrak{Q}}^{(1)} = \mathfrak{O}^* \operatorname{diag} \widetilde{\mathfrak{Q}}'^{(1)}\mathfrak{O},\end{aligned}$$

where we have introduced the matrix $\widetilde{\mathfrak{Q}}'^{(1)}(t) \in \mathrm{Sym}^+(2d)$. By (8.33), we have, $Q^{(2)} - Q^{(1)} = \frac{\sqrt{2}}{2\beta} I$, where I denotes the unit matrix of the suitable dimension. The definitions (7.17) and (7.18) then imply

$$\widetilde{\mathfrak{Q}}^{(2)} - \widetilde{\mathfrak{Q}}^{(1)} = \frac{\sqrt{2}}{2\beta} I. \tag{8.41}$$

Using the definitions and the above relation, we obtain

$$\begin{aligned}\widehat{\mathfrak{Q}}_\nu^{(1)}(t) &= \mathfrak{O}^*\left(t \operatorname{diag} \mathfrak{q}^{(1)} + (1-t)\widetilde{\mathfrak{Q}}'^{(1)}\right)\mathfrak{O}, \\ \widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t) &= \mathfrak{O}^*\left(t \operatorname{diag}(\mathfrak{q}^{(2)} - \mathfrak{q}^{(1)}) + (1-t)\frac{\sqrt{2}}{2\beta} I\right)\mathfrak{O}.\end{aligned} \tag{8.42}$$

Motivated by (8.17), we set

$$\mathfrak{d}_\nu^{(1)} \equiv \left(\mathfrak{u}_\nu - \mathfrak{q}_\nu^{(1)}\right)^{-1}. \tag{8.43}$$

In view of (8.37), the above choice necessarily yields (cf. (8.15))

$$\begin{aligned}\mathfrak{L} &= \mathfrak{C} - \mathfrak{D}^* \operatorname{diag}(\mathfrak{u}_\nu - \mathfrak{q}_\nu^{(1)})^{-1}\mathfrak{D} - \left(\widehat{\mathfrak{Q}}^{(2)}(t) - \widehat{\mathfrak{Q}}^{(1)}(t)\right) \\ &= \mathfrak{C} - \mathfrak{D}^*\left(\operatorname{diag}(\mathfrak{u}_\nu - \mathfrak{q}_\nu^{(1)})^{-1} + t\operatorname{diag}(\mathfrak{q}^{(2)} - \mathfrak{q}^{(1)}) + (1-t)\frac{\sqrt{2}}{2\beta}I\right)\mathfrak{D}. \qquad (8.44)\end{aligned}$$

Applying Lemma 8.4 to (8.38) and using (8.44), (8.43), (8.42), we get the following diagonalised representation of (8.36)

$$\begin{aligned}\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) \leq& \frac{1}{2}\log\det\mathfrak{C} - \frac{1}{2}\langle\mathfrak{C},\mathfrak{U}\rangle \\ &+ \frac{1}{2}\sum_{\nu=1}^{2d}\Big\{\mathfrak{u}_\nu\left[(\mathfrak{u}_\nu - \mathfrak{q}_\nu^{(1)})^{-1} + 2\beta^2\left(t(\mathfrak{q}_\nu^{(2)} - \mathfrak{q}_\nu^{(1)}) + (1-t)\frac{\sqrt{2}}{2\beta}\right)\right] \\ &\quad + 2\beta^2(\mathfrak{u}_\nu - \mathfrak{q}_\nu^{(1)})\left(t\mathfrak{q}_\nu^{(1)} + (1-t)\widetilde{\mathfrak{q}}_\nu^{(1)}\right) + \log(\mathfrak{u}_\nu - \widetilde{\mathfrak{Q}}_{\nu,\nu}^{\prime(1)}) \\ &\quad - t\beta^2\left((\mathfrak{q}_\nu^{(2)})^2 - (\mathfrak{q}_\nu^{(1)})^2\right)\Big\} + \mathscr{O}(\varepsilon). \qquad (8.45)\end{aligned}$$

Using the definitions, we get

$$\begin{aligned}\langle\mathfrak{C},\mathfrak{U}\rangle &= 2\langle C,U\rangle = 2\sum_{\nu=1}^{d} c_\nu u_\nu, \\ \log\det\mathfrak{C} &= 2\log\det C = 2\sum_{\nu=1}^{d}\log c_\nu. \qquad (8.46)\end{aligned}$$

Motivated by (8.41) (or by (8.33)), we define

$$\mathfrak{q}_\nu^{(1)} := \mathfrak{u}_\nu - \frac{\sqrt{2}}{2\beta}. \qquad (8.47)$$

In this case, as a straightforward calculation shows, the expression in the curly brackets in (8.45) equals

$$2\sqrt{2}\beta\mathfrak{u}_\nu + \beta\sqrt{2}\widetilde{\mathfrak{Q}}_{\nu,\nu}^{\prime(1)}(1-t) - \log\beta - \frac{1}{2}(\log 2 - t). \qquad (8.48)$$

By the definitions and the general properties of matrix trace, we have

$$\begin{aligned}\sum_{\nu=1}^{2d}\widetilde{\mathfrak{Q}}_{\nu,\nu}^{\prime(1)} &= \sum_{\nu=1}^{2d}\widetilde{\mathfrak{Q}}_{\nu,\nu}^{(1)} = 2\sum_{\nu=1}^{d} Q_{\nu,\nu}^{(1)}, \\ \sum_{\nu=1}^{2d}\mathfrak{u}_\nu &= 2\sum_{\nu=1}^{d} U_{\nu,\nu}. \qquad (8.49)\end{aligned}$$

Combining (8.45) with (8.48), (8.49) and (8.46), we obtain

$$
\begin{aligned}
\varphi_N^{(2)}(1,t,x,Q,\Sigma_N^{(2)}(\mathfrak{L},\mathfrak{U},\varepsilon,\delta)) &\leq \sum_{v=1}^{d}\Big(-c_v u_v + \log c_v + 3\sqrt{2}u\beta \\
&\qquad -\frac{1}{2}(\log 2 - t) - \sqrt{2}\beta t u - \log\beta - 1\Big) + \mathcal{O}(\varepsilon) \\
&= 2\sum_{v=1}^{d}\phi(t)|_{\substack{c=c_v,\\ u=u_v}} + \mathcal{O}(\varepsilon), \qquad (8.50)
\end{aligned}
$$

where in the last line we have used the relation (8.34).

5. To get the version of the a priori bound (8.50) with the quadratic correction term as stated in (8.35), we perturb the r.h.s of (8.36) around our choice of $\mathfrak{D}^{(1)}$ in (8.43), i.e.,

$$
\mathfrak{D}^{(1)} = \left(\mathfrak{U}_v - \mathfrak{Q}_v^{(1)}\right)^{-1} = \sqrt{2}\beta I,
$$

where in the last equality we used (8.47).

□

8.7 The local low temperature Parisi formula

Proof of Theorem 1.3. The result follows from Theorem 8.1 and Theorem 7.1. Note that the proof of Theorem 7.1 requires a minor modification to cope with the fact that the a priori distribution (1.26) is unbounded. This minor problem can be fixed by considering the pruned Gaussian distribution and using the elementary estimates to bound the tiny Gaussian tails. □

A

The general result of Guerra and Toninelli [19] implies that the thermodynamic limit of the local free energy (1.6) exists almost surely and in L^1. The following existence of the limiting average overlap is an immediate consequence of this.

Proposition A.1. *We have*

$$
\mathbb{E}\left[\mathcal{G}_N(\beta)\otimes\mathcal{G}_N(\beta)\left[\operatorname{Var} H_N(\sigma) - \mathbb{E}\left[H_N(\sigma)H_N(\sigma')\right]\right]\right] \xrightarrow[N\uparrow+\infty]{} C(\beta) \geq 0,
$$

where $C:\mathbb{R}_+\to\mathbb{R}_+$.

Proof. The free energy is a convex function of β (a consequence of the HÃűlder inequality). Hence, by a result in [16] the following holds

$$
\lim_{N\uparrow\infty}\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p_N(\beta)\right] = \frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p(\beta)\right].
$$

Proposition 2.4 implies

$$
\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}\left[p_N(\beta)\right] = \beta\mathbb{E}\left[\mathcal{G}_N(\beta)\otimes\mathcal{G}_N(\beta)\left[\operatorname{Var} H_N(\sigma) - \mathbb{E}\left[H_N(\sigma)H_N(\sigma')\right]\right]\right].
$$

□

The following super-additivity result is an application of the Gaussian comparison inequalities obtained in Subection 2.3. Note that the result does not provide enough information for the cavity-like argument of [1].

Proposition A.2. *For any* $\mathcal{V} \equiv B(U, \varepsilon) \subset \mathcal{U}$, *we have*

$$N\mathbb{E}\left[p_N(\mathcal{V})\right] + M\mathbb{E}\left[p_M(\mathcal{V})\right] \leq (N+M)\mathbb{E}\left[p_{N+M}(\mathcal{V})\right] + (N+M)\mathcal{O}(\varepsilon),$$

as $\varepsilon \downarrow +0$.

Proof. Define the process $Y_{N,M} \equiv \{Y(\sigma) : \sigma = \alpha \shortparallel \tau; \alpha \in \Sigma_N, \tau \in \Sigma_M\}$ as follows

$$Y(\alpha \shortparallel \tau) \equiv \left(\frac{N}{N+M}\right)^{1/2} X_N^{(1)}(\alpha) + \left(\frac{M}{N+M}\right)^{1/2} X_M^{(2)}(\tau),$$

where $X^{(1)}$ and $X^{(2)}$ are two independent copies of the process X. Given some Gaussian process $\{C(\sigma)\}_{\sigma\in\Sigma_N}$, let us introduce the functional $\Phi_N(\beta)[C]$ as follows

$$\Phi_{N,M}(\beta)[C] \equiv \mathbb{E}\left[\log \mu^{\otimes(N+M)}\left[\mathbb{1}_{\Sigma_N(\mathcal{V})}\mathbb{1}_{\Sigma_M(\mathcal{V})}\exp(\beta\sqrt{N+M}C)\right]\right].$$

Now, set $\varphi(t) \equiv \Phi_{N+M}(\beta)\left[\sqrt{t}X_{N+M} + \sqrt{1-t}Y_{N,M}\right]$. Applying Proposition 2.5, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \frac{\beta^2(N+M)}{2}&\mathbb{E}\left[\mathcal{G}(t)\otimes\mathcal{G}(t)\right[\\ &\left(\operatorname{Var}X_{N+M}(\sigma^{(1)}) - \operatorname{Var}Y_{N,M}(\sigma^{(1)})\right)\\ &-\left(\operatorname{Cov}\left[X_{N+M}(\sigma^{(1)}), X_{N+M}(\sigma^{(2)})\right] - \operatorname{Cov}\left[Y_{N,M}(\sigma^{(1)}), Y_{N,M}(\sigma^{(2)})\right]\right)\left]\right].\end{aligned} \tag{A.1}$$

Note that we have

$$\begin{aligned}\varphi(0) &= N\mathbb{E}\left[p_N(\mathcal{V})\right] + M\mathbb{E}\left[p_M(\mathcal{V})\right],\\ \varphi(1) &\leq (N+M)\mathbb{E}\left[p_{N+M}(\mathcal{V})\right],\end{aligned} \tag{A.2}$$

where the last inequality is due to the fact that, for all $\alpha \in \Sigma_N(\mathcal{V})$ and all $\tau \in \Sigma_N(\mathcal{V})$, we have

$$\alpha \shortparallel \tau \in \Sigma_{N+M}(\mathcal{V}).$$

Moreover, for $\sigma = \alpha \shortparallel \tau$ with $\alpha \in \Sigma_N(\mathcal{V})$ and $\sigma \in \Sigma_M(\mathcal{V})$ we have

$$\begin{aligned}\operatorname{Var}X_{N+M}(\sigma) - \operatorname{Var}Y_{N,M}(\sigma) = &\left\|\frac{N}{N+M}R_N(\alpha,\alpha) + \frac{M}{N+M}R_M(\tau,\tau)\right\|_2^2 - \frac{N}{N+M}\|R_N(\alpha,\alpha)\|_2^2\\ &-\frac{M}{N+M}\|R_M(\tau,\tau)\|_2^2 = \mathcal{O}(\varepsilon).\end{aligned}$$

Also, due to convexity of the norm, we have

$$\begin{aligned}&\operatorname{Cov}\left[X_{N+M}(\sigma^{(1)}), X_{N+M}(\sigma^{(2)})\right] - \operatorname{Cov}\left[Y_{N,M}(\sigma^{(1)}), Y_{N,M}(\sigma^{(2)})\right]\\ &= \left\|\frac{N}{N+M}R_N(\alpha^{(1)},\alpha^{(2)}) + \frac{M}{N+M}R_M(\tau^{(1)},\tau^{(2)})\right\|_2^2 - \frac{N}{N+M}\|R_N(\alpha^{(1)},\alpha^{(2)})\|_2^2\\ &\quad - \frac{M}{N+M}\|R_M(\tau^{(1)},\tau^{(2)})\|_2^2 \leq 0.\end{aligned}$$

Applying $\int_0^1 \mathrm{d}t$ to (A.1) and using the previous two formulae, we get the claim. □

References

[1] Michael Aizenman, Robert Sims, and Shannon L. Starr. An Extended Variational Principle for the SK Spin-Glass Model. *Phys. Rev. B*, 68:214403, 2003.

[2] Michael Aizenman, Robert Sims, and Shannon L. Starr. Mean-field spin glass models from the cavity-ROSt perspective. In *Prospects in mathematical physics*, volume 437 of *Contemp. Math.*, pages 1–30. Amer. Math. Soc., Providence, RI, 2007. arXiv:math-ph/0607060. MR2354653

[3] Louis-Pierre Arguin. Spin glass computations and Ruelle's probability cascades. *J. Stat. Phys.*, 126(4-5):951–976, 2007. arXiv:math-ph/0608045v1. MR2311892

[4] Gérard Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001. MR1856194

[5] Vladimir I. Bogachev. *Gaussian measures*, volume 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998. MR1642391

[6] Erwin Bolthausen and Alain-Sol Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197(2):247–276, 1998.

[7] Anton Bovier. *Statistical mechanics of disordered systems*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2006. A mathematical perspective. MR2252929

[8] Anton Bovier and Irina Kurkova. Much ado about Derrida's GREM. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 81–115. Springer, Berlin, 2007. MR2309599

[9] Philippe Briand and Ying Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields*, pages –, 2007. arXiv:math/0703423. MR2391164

[10] Francis Comets. Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probab. Theory Related Fields*, 80(3):407–432, 1989. MR0976534

[11] Andrea Crisanti and Hans-Jürgen Sommers. The spherical p-spin interaction spin glass model: the statics. *Zeitschrift für Physik B Condensed Matter*, 87(3):341–354, 1992.

[12] Francesca Da Lio and Olivier Ley. Uniqueness Results for Second-Order Bellman–Isaacs Equations under Quadratic Growth Assumptions and Applications. *SIAM J. Control Optim.*, 45(1):74–106, 2006. MR2225298

[13] Frank den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000. MR1739680

[14] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998. MR1625845

[15] Jürg Fröhlich and Bogusław Zegarliński. Some comments on the Sherrington-Kirkpatrick model of spin glasses. *Comm. Math. Phys.*, 112(4):553–566, 1987. MR0910578

[16] Robert B. Griffiths. A proof that the free energy of a spin system is extensive. *J. Mathematical Phys.*, 5:1215–1222, 1964. MR0169626

[17] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003. MR1957729

[18] Francesco Guerra. Spin glasses. Preprint, 2005. arXiv:cond-mat/0507581v1. MR2185777

[19] Francesco Guerra and Fabio Lucio Toninelli. The infinite volume limit in generalized mean field disordered models. *Markov Process. Related Fields*, 9(2):195–207, 2003.

[20] Dmitry Panchenko. A question about the Parisi functional. *Electron. Commun. Probab.*, 10:155–166, 2005. MR2162815

[21] Dmitry Panchenko. Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.*, 17(7):793–857, 2005. MR2159369

[22] Dmitry Panchenko and Michel Talagrand. Guerra's interpolation using Derrida-Ruelle cascades. *Preprint*, 2007. arXiv:0708.3641v2 [math.PR].

[23] Dmitry Panchenko and Michel Talagrand. On the overlap in the multiple spherical SK models. *Ann. Probab.*, 35(6):2321–2355, 2007. MR2353390

[24] David Ruelle. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.*, 108(2):225–239, 1987. MR0875300

[25] David Sherrington. Spin glasses: a perspective. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 45–62. Springer, Berlin, 2007. MR2309597

[26] David Sherrington and Scott Kirkpatrick. Solvable Model of a Spin-Glass. *Physical Review Letters*, 35(26):1792–1796, 1975.

[27] Michel Talagrand. Large deviation principles and generalized Sherrington-Kirkpatrick models. *Ann. Fac. Sci. Toulouse Math. (6)*, 9(2):203–244, 2000. MR1813802

[28] Michel Talagrand. *Spin glasses: a challenge for mathematicians*, volume 46 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2003. Cavity and mean field models. MR1993891

[29] Michel Talagrand. Free energy of the spherical mean field model. *Probab. Theory Related Fields*, 134(3):339–382, 2006. MR2226885

[30] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006. MR2195134

[31] Michel Talagrand. Parisi measures. *J. Funct. Anal.*, 231(2):269–286, 2006. MR2195333

[32] Michel Talagrand. Large Deviations, Guerra's and A.S.S. Schemes, and the Parisi Hypothesis. *Journal of Statistical Physics*, 126(4):837–894, 2007. MR2311888

[33] Michel Talagrand. Mean field models for spin glasses: some obnoxious problems. In *Spin glasses*, volume 1900 of *Lecture Notes in Math.*, pages 63–80. Springer, Berlin, 2007. MR2309598

[34] Alain Toubol. High temperature regime for a multidimensional Sherrington-Kirkpatrick model of spin glass. *Probab. Theory Related Fields*, 110(4):497–534, 1998. MR1626959