Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 70, 1–46.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP801

The ℓ^p-Gaussian-Grothendieck problem with vector spins

Tomas Dominguez*

Abstract

We study the vector spin generalization of the ℓ^p-Gaussian-Grothendieck problem. In other words, given integer $\kappa \geq 1$, we investigate the asymptotic behaviour of the ground state energy associated with the Sherrington-Kirkpatrick Hamiltonian indexed by vector spin configurations in the unit ℓ^p-ball. The ranges $1 \leq p \leq 2$ and $2 < p < \infty$ exhibit significantly different behaviours. When $1 \leq p \leq 2$, the vector spin generalization of the ℓ^p-Gaussian-Grothendieck problem agrees with its scalar counterpart. In particular, its re-scaled limit is proportional to some norm of a standard Gaussian random variable. On the other hand, for $2 < p < \infty$ the re-scaled limit of the ℓ^p-Gaussian-Grothendieck problem with vector spins is given by a Parisi-type variational formula.

Keywords: spin glasses; ground state energy; Parisi formula; vector spins.
MSC2020 subject classifications: 82D30; 82B44; 60K35; 60G15.
Submitted to EJP on September 2, 2021, final version accepted on May 18, 2022.
Supersedes arXiv:2112.05709v1.

1 Introduction and main results

Given an $N \times N$ matrix $A = (a_{ij})$ and some $1 \leq p < \infty$, the ℓ^p-Grothendieck problem consists in maximizing the quadratic form $\sum_{i,j=1}^N a_{ij}\sigma_i\sigma_j$ over all vectors $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \mathbb{R}^N$ with unit ℓ^p-norm, $\|\boldsymbol{\sigma}\|_p^p = \sum_{i=1}^N |\sigma_i|^p = 1$. In other words, it involves computing the quantity

$$\mathrm{GP}_{N,p}(A) = \max\Big\{ \sum_{i,j=1}^N a_{ij}\sigma_i\sigma_j \mid \sum_{i=1}^N |\sigma_i|^p = 1 \Big\}. \tag{1.1}$$

In the case $p = 2$, this is the maximum eigenvalue of the symmetric matrix $(A+A^T)/2$. On the other hand, the limiting case $p = \infty$ has been extensively studied in the mathematics and computer science literature for its applications to combinatorial optimization, graph

*University of Toronto, Canada.
E-mail: tomas.dominguezchiozza@mail.utoronto.ca

theory and correlation clustering [1, 5, 10, 14, 26]. The range $2 < p < \infty$ can be thought of as an interpolation between the spectral and the correlation clustering problems [21], while the range $1 < p < 2$ seems to be unexplored in the literature. Finding an efficient algorithm to solve the ℓ^p-Grothendieck problem whenever $p \neq 2$ is generally difficult [18, 20, 24, 25, 27], so it is natural to study the ℓ^p-Grothendieck problem for random input matrices first. In fact, it should help understand the typical behaviour of (1.1). This leads to the ℓ^p-Gaussian-Grothendieck problem,

$$\mathrm{GP}_{N,p} = \max \Big\{ \sum_{i,j=1}^N g_{ij}\sigma_i\sigma_j \mid \sum_{i=1}^N |\sigma_i|^p = 1 \Big\}, \tag{1.2}$$

where (g_{ij}) are independent standard Gaussian random variables.

The asymptotic behaviour of (1.2) was studied in great detail in [6]. It was discovered that the re-scaled limit of (1.2) exhibits significantly different behaviour in the ranges $1 \leq p \leq 2$ and $2 < p < \infty$; in the former, it is proportional to some norm of a Gaussian random variable, and in the latter, it is given by a Parisi-type variational formula. In this paper, we will show that this behaviour persists in the vector spin generalization of (1.2). Our work is motivated and greatly influenced by [6]; however, new ideas are needed to treat the range $2 < p < \infty$. These will be detailed at a later stage, and they will allow us to avoid the key truncation step in [6] as well as its associated technicalities. Therefore, specializing our arguments to the scalar setting, $\kappa = 1$, yields a simpler proof of the main result in [6].

Before we describe the vector spin analogue of (1.2), let us mention that another motivation for investigating this optimization problem comes from the study of spin glass models. In the language of statistical physics, the quadratic form $\sum_{i,j=1}^N g_{ij}\sigma_i\sigma_j$ is known as the Hamiltonian of the Sherrington-Kirkpatrick (SK) mean-field spin glass model, and the quantity (1.2) corresponds to the ground state energy of the SK model on the unit ℓ^p-sphere. From this perspective, the vector spin generalization of (1.2) which we will study in this paper is very natural; it also appears in the computer-science literature [1, 5, 14, 19, 21] when studying the convex relaxation of (1.2).

Let us now describe the vector spin generalization of (1.2) using the notation and terminology of statistical physics. Fix an integer $\kappa \geq 1$ throughout the remainder of this paper. The Hamiltonian of the vector spin SK model is the random function of the $N \geq 1$ vector spins taking values in $\mathbb{R}^\kappa$,

$$\vec{\boldsymbol{\sigma}} = \big(\vec{\sigma}_1, \ldots, \vec{\sigma}_N\big) \in (\mathbb{R}^\kappa)^N, \tag{1.3}$$

given by the quadratic form

$$H_N^\circ(\vec{\boldsymbol{\sigma}}) = \sum_{i,j=1}^N g_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big), \tag{1.4}$$

where the interaction parameters (g_{ij}) are independent standard Gaussian random variables and $(\cdot,\cdot)$ is the Euclidean inner product on $\mathbb{R}^\kappa$. Denote the coordinates of each spin $\vec{\sigma}_i$ by

$$\vec{\sigma}_i = (\sigma_i(1), \ldots, \sigma_i(\kappa)) \in \mathbb{R}^\kappa, \tag{1.5}$$

write the configuration of the k'th coordinates as

$$\boldsymbol{\sigma}(k) = (\sigma_1(k), \ldots, \sigma_N(k)) \in \mathbb{R}^N, \tag{1.6}$$

and introduce the $\ell^{p,2}$-norm on the Euclidean space $(\mathbb{R}^\kappa)^N$,

$$\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p = \sum_{i=1}^N \|\vec{\sigma}_i\|_2^p. \tag{1.7}$$

The ℓ^p-Gaussian-Grothendieck problem with vector spins consists in maximizing the Hamiltonian (1.4) over the unit $\ell^{p,2}$-sphere. In other words, it involves computing the quantity

$$\mathbf{GP}_{N,p} = \max\big\{H_N^\circ(\vec{\boldsymbol{\sigma}}) \mid \|\vec{\boldsymbol{\sigma}}\|_{p,2} = 1\big\}. \tag{1.8}$$

To handle the range $1 \leq p \leq 2$, we will use the Gaussian Hilbert space approach to the Grothendieck inequality [1, 5, 14, 21] in order to show that for any $N \times N$ matrix $A = (a_{ij})$,

$$\mathrm{GP}_{N,p}(A) = \max\Big\{\sum_{i,j=1}^N a_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big) \mid \|\vec{\boldsymbol{\sigma}}\|_{p,2} = 1\Big\}. \tag{1.9}$$

This identity was mentioned in [19], but no proof was given. Combining (1.9) with theorem 1.1 and theorem 1.2 in [6] will immediately give our main result for $1 \leq p \leq 2$.

Theorem 1.1. *If $1 < p < 2$, then almost surely,*

$$\lim_{N\to\infty} \frac{1}{N^{1/p^*}} \mathbf{GP}_{N,p} = 2^{\frac{1}{2}-\frac{2}{p}}\big(\mathbb{E}|g|^{p^*}\big)^{1/p^*}, \tag{1.10}$$

where p^* is the Hölder conjugate of p and g is a standard Gaussian random variable. On the other hand, if $p = 1$ or $p = 2$, then almost surely,

$$\lim_{N\to\infty} \frac{1}{\sqrt{\log N}} \mathrm{GP}_{N,1} = \sqrt{2} = \lim_{N\to\infty} \frac{1}{\sqrt{N}} \mathrm{GP}_{N,2}. \tag{1.11}$$

The range $2 < p < \infty$ will require substantially more work, and will occupy the majority of this paper. It will be convenient to introduce a re-scaled version of the Hamiltonian (1.4),

$$H_N(\vec{\boldsymbol{\sigma}}) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^N g_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big), \tag{1.12}$$

as well as a normalized version of the $\ell^{p,2}$-norm (1.7),

$$|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p = \frac{1}{N}\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p = \frac{1}{N}\sum_{i=1}^N \|\vec{\sigma}_i\|_2^p. \tag{1.13}$$

If we denote the classical SK Hamiltonian on $\mathbb{R}^N$ by

$$H_N^k(\boldsymbol{\sigma}(k)) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^N g_{ij}\sigma_i(k)\sigma_j(k), \tag{1.14}$$

we may express the vector spin Hamiltonian (1.12) as

$$H_N(\vec{\boldsymbol{\sigma}}) = \sum_{k=1}^\kappa H_N^k(\boldsymbol{\sigma}(k)). \tag{1.15}$$

It is readily verified that for two spin configurations $\vec{\boldsymbol{\sigma}}^l, \vec{\boldsymbol{\sigma}}^{l'} \in (\mathbb{R}^\kappa)^N$ and two integers $1 \leq k, k' \leq \kappa$,

$$\mathbb{E}\, H_N^k(\boldsymbol{\sigma}^l(k)) H_N^{k'}(\boldsymbol{\sigma}^{l'}(k')) = N\big(R_{l,l'}^{k,k'}\big)^2, \tag{1.16}$$

where

$$R_{l,l'}^{k,k'} = \frac{1}{N}\sum_{i=1}^N \sigma_i^l(k)\sigma_i^{l'}(k') \tag{1.17}$$

is the overlap between $\boldsymbol{\sigma}^l(k)$ and $\boldsymbol{\sigma}^{l'}(k')$. We will denote the matrix of all such overlaps by

$$R(\vec{\boldsymbol{\sigma}}^l, \vec{\boldsymbol{\sigma}}^{l'}) = \big(R_{l,l'}^{k,k'}\big)_{k,k' \leq \kappa} = \frac{1}{N}\sum_{i=1}^{N} \vec{\sigma}_i^l \vec{\sigma}_i^{l'T}. \tag{1.18}$$

The covariance structure of the vector spin Hamiltonian (1.12) may therefore be expressed in terms of this matrix-valued overlap as

$$\mathbb{E}\, H_N(\vec{\boldsymbol{\sigma}}^l) H_N(\vec{\boldsymbol{\sigma}}^{l'}) = N \sum_{k,k'=1}^{\kappa} \big(R_{l,l'}^{k,k'}\big)^2 = N \|R_{l,l'}\|_{\mathrm{HS}}^2, \tag{1.19}$$

where $\|\gamma\|_{\mathrm{HS}}^2 = \sum_{k,k'} |\gamma_{k,k'}|^2$ denotes the Hilbert-Schmidt norm on the space of $\kappa \times \kappa$ matrices. Writing

$$S_p^N = \{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N \mid |||\vec{\boldsymbol{\sigma}}|||_{p,2} = 1\} \tag{1.20}$$

for the unit normalized-$\ell^{p,2}$-sphere, the ℓ^p-Gaussian-Grothendieck problem with vector spins may be recast as the task of computing the ground state energy

$$\mathbf{GSE}_{N,p} = N^{\frac{2}{p}-\frac{3}{2}} \mathbf{GP}_{N,p} = \frac{1}{N} \max_{\vec{\boldsymbol{\sigma}} \in S_p^N} H_N(\vec{\boldsymbol{\sigma}}). \tag{1.21}$$

Using Chevet's inequality as in section 3 of [6], it is easy to see that this is the correct scaling of (1.8) when $2 < p < \infty$. To study the constrained optimization problem (1.21), it is natural to remove the normalized-$\ell^{p,2}$-norm constraint by considering the model with an $\ell^{p,2}$-norm potential. For each $t > 0$ define the Hamiltonian

$$H_{N,p,t}(\vec{\boldsymbol{\sigma}}) = H_N(\vec{\boldsymbol{\sigma}}) - t\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p, \tag{1.22}$$

and introduce the unconstrained Lagrangian

$$L_{N,p}(t) = \frac{1}{N} \max_{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N} H_{N,p,t}(\vec{\boldsymbol{\sigma}}). \tag{1.23}$$

Our first noteworthy result in the range $2 < p < \infty$, which we now describe, will relate the asymptotic behaviour of the unconstrained Lagrangian (1.23) to the limit of the ground state energy (1.21).

Consider the space of $\kappa \times \kappa$ Gram matrices,

$$\Gamma_\kappa = \big\{\gamma \in \mathbb{R}^{\kappa \times \kappa} \mid \gamma \text{ is symmetric and non-negative definite}\big\}, \tag{1.24}$$

endowed with the Loewner order $\gamma_1 \leq \gamma_2$ if and only if $\gamma_2 - \gamma_1 \in \Gamma_\kappa$, and denote by Γ_κ^+ the subspace of positive definite matrices in Γ_κ,

$$\Gamma_\kappa^+ = \{\gamma \in \Gamma_\kappa \mid \gamma \text{ is positive definite}\}. \tag{1.25}$$

For each $D \in \Gamma_\kappa$ write

$$\Sigma(D) = \big\{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N \mid R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}}) = D\big\} \tag{1.26}$$

for the set of spin configurations $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ with self-overlap D, and introduce the constrained Lagrangian

$$L_{N,p,D}(t) = \frac{1}{N} \max_{\vec{\boldsymbol{\sigma}} \in \Sigma(D)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}). \tag{1.27}$$

In section 2, we will show that the constrained Lagrangian (1.27) admits a deterministic limit $L_{p,D}(t)$ with probability one, and in section 4, we will establish the following asymptotic formula for the unconstrained Lagrangian (1.23).

Theorem 1.2. *If $2 < p < \infty$, then almost surely the limit $L_p(t) = \lim_{N\to\infty} L_{N,p}(t)$ exists for every $t > 0$. Moreover, with probability one,*

$$L_p(t) = \sup_{D\in\Gamma_\kappa} L_{p,D}(t) = \sup_{D\in\Gamma_\kappa^+} L_{p,D}(t). \tag{1.28}$$

Subsequently, in section 5, we will use the basic properties of convex functions to derive the following key relationship between the limits of (1.23) and (1.21).

Theorem 1.3. *If $2 < p < \infty$, then almost surely the limit $\mathbf{GSE}_p = \lim_{N\to\infty} \mathbf{GSE}_{N,p}$ exists and is given by*

$$\mathbf{GSE}_p = \frac{p}{2}\Big(\frac{p}{2} - 1\Big)^{\frac{2}{p}-1} t^{\frac{2}{p}} L_p(t)^{1-\frac{2}{p}} \tag{1.29}$$

for every $t > 0$.

This result reduces the study of the ground state energy (1.21) to that of the Lagrangian (1.27) with positive definite self-overlaps $D \in \Gamma_\kappa^+$. The main result of this paper will be a Parisi-type variational formula for the limit $L_{p,D}(t)$ of (1.27) when $D \in \Gamma_\kappa^+$. Together with (1.29), (1.28) and (1.21), this will give a Parisi-type variational formula for the ℓ^p-Gaussian-Grothendieck problem with vector spins when $2 < p < \infty$.

Given $D \in \Gamma_\kappa^+$, we now describe the Parisi-type variational formula for the limit $L_{p,D}(t)$ of (1.27). Let us call a path $\pi : [0,1] \to \Gamma_\kappa$ piecewise linear if there exists a partition $0 = q_{-1} \le q_0 \le \ldots \le q_r = 1$ of $[0,1]$ and matrices $(\gamma_j)_{-1\le j\le r} \subset \Gamma_\kappa$ with

$$\pi(s) = \gamma_{j-1} + \frac{s - q_{j-1}}{q_j - q_{j-1}}(\gamma_j - \gamma_{j-1}) \tag{1.30}$$

when $s \in [q_{j-1}, q_j]$ for some $0 \le j \le r$. Denote by Π the space of piecewise linear and non-decreasing functions on $[0,1]$ with values in Γ_κ,

$$\Pi = \big\{\pi : [0,1] \to \Gamma_\kappa \mid \pi \text{ is piecewise linear, } \pi(x) \le \pi(x') \text{ for } x \le x'\big\}, \tag{1.31}$$

and for each $D \in \Gamma_\kappa$ write Π_D for the set of paths in Π that start at 0 and end at D,

$$\Pi_D = \big\{\pi \in \Pi \mid \pi(0) = 0 \text{ and } \pi(1) = D\big\}. \tag{1.32}$$

Notice that any path $\pi \in \Pi_D$ can be identified with two sequences of parameters,

$$0 = q_{-1} \le q_0 \le \ldots \le q_{r-1} \le q_r = 1, \tag{1.33}$$
$$0 = \gamma_{-1} = \gamma_0 \le \gamma_1 \le \ldots \le \gamma_{r-1} \le \gamma_r = D, \tag{1.34}$$

satisfying $\pi(q_j) = \gamma_j$ for $0 \le j \le r$. Explicitly, the path π is given by

$$\pi(s) = \gamma_{j-1} + \frac{s - q_{j-1}}{q_j - q_{j-1}}(\gamma_j - \gamma_{j-1}) \tag{1.35}$$

when $s \in [q_{j-1}, q_j]$ for some $0 \le j \le r$. Denote by $\mathcal{N}^d$ the set of finite measures on $[0,1]$ with finitely many atoms, and given $t > 0$ and $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$ consider the function $f_\lambda^\infty : \mathbb{R}^\kappa \to \mathbb{R}$ defined by

$$f_\lambda^\infty(\vec{x}) = \sup_{\vec{\sigma}\in\mathbb{R}^\kappa} \Big(\big(\vec{\sigma}, \vec{x}\big) + \sum_{k\le k'} \lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big). \tag{1.36}$$

Notice that any discrete measure $\zeta \in \mathcal{N}^d$ may be identified with two sequences of parameters

$$0 = q_{-1} \le q_0 \le \ldots \le q_{r-1} \le q_r = 1, \tag{1.37}$$
$$0 = \zeta_{-1} \le \zeta_0 \le \ldots \le \zeta_{r-1} \le \zeta_r < \infty, \tag{1.38}$$

satisfying $\zeta(\{q_j\}) = \zeta_j - \zeta_{j-1}$ for $0 \le j \le r$. Moreover, the sequences (1.33) and (1.37) can be taken to be the same by duplicating values in (1.34) and (1.38) if necessary. We will often abuse notation and write ζ both for the measure and its cumulative distribution function. Given independent Gaussian vectors $z_j = (z_j(k))_{k\le\kappa}$ for $0 \le j \le r$ with covariance structure

$$\operatorname{Cov} z_j = \gamma_j - \gamma_{j-1}, \tag{1.39}$$

recursively define the sequence $(Y_l^{\lambda,\zeta,\pi})_{0\le l\le r}$ as follows. Let

$$Y_r^{\lambda,\zeta,\pi}((z_j)_{0\le j\le r}) = f_\lambda^\infty\Big(\sqrt{2}\sum_{j=1}^r z_j\Big), \tag{1.40}$$

and for $0 \le l \le r-1$ let

$$Y_l^{\lambda,\zeta,\pi}((z_j)_{0\le j\le l}) = \frac{1}{\zeta_l}\log \mathbb{E}_{z_{l+1}} \exp \zeta_l Y_{l+1}^{\lambda,\zeta,\pi}((z_j)_{0\le j\le l+1}). \tag{1.41}$$

This inductive procedure is well-defined by the growth bounds established in lemma A.2. Introduce the Parisi functional,

$$\mathcal{P}_\infty(\lambda,\zeta,\pi) = Y_0^{\lambda,\zeta,\pi} - \sum_{k\le k'} \lambda_{k,k'} D_{k,k'} - \int_0^1 \zeta(s)\operatorname{Sum}\big(\pi(s)\odot\pi'(s)\big)\,\mathrm{d}\,s, \tag{1.42}$$

where $\operatorname{Sum}(\gamma) = \sum_{k,k'}\gamma_{k,k'}$ is the sum of all elements in a $\kappa\times\kappa$ matrix and $\odot$ denotes the Hadamard product on the space of $\kappa\times\kappa$ matrices. We have made all dependencies on κ, p, t and D implicit for clarity of notation, but we will make them explicit whenever necessary. The following is our main result.

Theorem 1.4. *If $2 < p < \infty$, then for any $D \in \Gamma_\kappa^+$ and every $t > 0$,*

$$L_{p,D}(t) = \inf_{\lambda,\zeta,\pi} \mathcal{P}_\infty(\lambda,\zeta,\pi), \tag{1.43}$$

where the infimum is taken over all $(\lambda,\zeta,\pi) \in \mathbb{R}^{\kappa(\kappa+1)/2}\times\mathcal{N}^d\times\Pi_D$.

We close this section with a brief outline of the paper. Section 2 will be devoted to the range $1 \le p \le 2$ and will include a proof of theorem 1.1. The rest of the paper will focus on the range $2 < p < \infty$. In section 3, we will use the Guerra-Toninelli interpolation [13, 29] and the Gaussian concentration inequality [28, 29] to show that the constrained Lagrangian (1.27) admits a deterministic limit. In section 4, we show that, in a certain sense, the limit of the constrained Lagrangian depends continuously on the constraint. This continuity result is inspired by lemma 7.1 in [6]. Unfortunately, lemma 7.1 in [6] does not extend to the vector spin setting since we can no longer modify overlaps by simply re-scaling spin configurations. To overcome this issue, we will revisit lemma 4 in [33], originally designed to prove a vector spin version of the Ghirlanda-Guerra identities [12], and we will leverage Dudley's entropy inequality [9, 35]. With this continuity result at hand, we will closely follow section 7 and section 8 in [6] to prove theorem 1.2 and theorem 1.3. This will be the content of section 5 and section 6. In section 7, we will introduce a free energy functional that depends on an inverse temperature parameter $\beta > 0$ and is asymptotically equivalent to the constrained Lagrangian (1.27) after letting $\beta\to\infty$. For each finite $\beta > 0$, a simple modification of the arguments in [33], which we will not detail, gives a Parisi-type variational formula for the limit of the free energy functional. This is reviewed in section 8. The rest of the paper is devoted to finding a similar Parisi-type variational formula after letting $\beta\to\infty$. This is where our approach differs substantially from that in [6]. In our attempt to generalize the truncation argument in sections 10-12 of [6] to the vector spin setting, we discovered that by a careful analysis of the terminal condition (1.36) and its positive temperature analogue, the proof for the scalar, $\kappa = 1$, case could be considerably simplified. This simplified proof extended with minor modifications to the vector spin setting and is what we present between section 9 and section 11 of this paper. In particular, our arguments can be used to simplify the proof of the main result in [6]. The careful analysis of the terminal conditions is undertaken in section 9. The resulting bounds are combined with the Auffinger-Chen representation [4, 16] in section 10 to compare the Parisi functional (1.42) and its positive temperature counterpart. The specific form of the Auffinger-Chen representation that we use is a higher dimensional generalization of that in [6, 7]. The proof of theorem 1.4 is finally completed in section 11. For the reader's convenience, we have postponed a number of technical estimates to appendix A, and we have included a review of elementary results in linear algebra in appendix B.

2 The range $1 \leq p \leq 2$

In this section we show that the ℓ^p-Gaussian-Grothendieck problem with vector spins agrees with its scalar counterpart in the range $1 \leq p \leq 2$ by proving (1.9). Recall the definition (1.1) of the ℓ^p-Grothendieck problem $\mathrm{GP}_{N,p}(A)$ for an arbitrary $N \times N$ matrix $A = (a_{ij})$.

Lemma 2.1. *For any $N \times N$ matrix $A = (a_{ij})$ and every $1 \leq p \leq 2$,*

$$\mathrm{GP}_{N,p}(A) = \max\Big\{ \sum_{i,j=1}^N a_{ij}\big(\vec{\boldsymbol{\sigma}}_i, \vec{\boldsymbol{\sigma}}_j\big) \mid \|\vec{\boldsymbol{\sigma}}\|_{p,2} = 1\Big\}. \tag{2.1}$$

Proof. Given $\boldsymbol{\sigma} \in \mathbb{R}^N$ in the unit ℓ^p-sphere, consider the vector spin configuration $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ defined by

$$\vec{\boldsymbol{\sigma}}(k) = \begin{cases} \boldsymbol{\sigma} & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p = \|\boldsymbol{\sigma}\|_p^p = 1$ and $\sum_{i,j=1}^N a_{ij}\sigma_i\sigma_j = \sum_{i,j=1}^N a_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big)$. It follows that

$$\sum_{i,j=1}^N a_{ij}\sigma_i\sigma_j \leq \max\Big\{ \sum_{i,j=1}^N a_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big) \mid \|\vec{\boldsymbol{\sigma}}\|_{p,2} = 1\Big\},$$

and taking the maximum over all such $\boldsymbol{\sigma} \in \mathbb{R}^N$ gives the upper bound in (2.1). To prove the matching lower bound, fix a vector spin configuration $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ in the unit $\ell^{p,2}$-sphere. Let g be a standard Gaussian random vector in $\mathbb{R}^\kappa$ and for each $1 \leq i \leq N$ consider the random variable

$$X_i = \big(g, \vec{\sigma}_i\big).$$

Observe that $\mathbb{E}\, X_iX_j = \sum_{k=1}^\kappa \sigma_i(k)\sigma_j(k) = \big(\vec{\sigma}_i, \vec{\sigma}_j\big)$. Normalizing the random vector $X = (X_i)_{i\leq N}$, it is easy to see that

$$\sum_{i,j=1}^N a_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big) = \mathbb{E} \sum_{i,j=1}^N a_{ij}X_iX_j \leq \mathrm{GP}_{N,p}(A)\, \mathbb{E}\,\Big(\sum_{i=1}^N |X_i|^p\Big)^{2/p}. \tag{2.2}$$

To bound this further, denote by $\|\cdot\|_{L^2}$ the L^2-norm defined by the law of g. Minkowski's integral inequality and the assumption $1 \leq p \leq 2$ imply that

$$\mathbb{E}\,\Big(\sum_{i=1}^N |X_i|^p\Big)^{2/p} = \big\|\|X\|_p\big\|_{L^2}^2 \leq \big\|\|X\|_{L^2}\big\|_p^2 = \Big(\sum_{i=1}^N \big(\mathbb{E}|X_i|^2\big)^{p/2}\Big)^{2/p} = \|\vec{\boldsymbol{\sigma}}\|_{p,2}^2 = 1.$$

Substituting this into (2.2) gives the lower bound in (2.1) and completes the proof. □

Applying this result to the random matrix $G_N = (g_{ij})_{i,j\leq N}$ conditionally on the disorder chaos shows that $\mathbf{GP}_{N,p} = \mathrm{GP}_{N,p}$ for $1 \leq p \leq 2$. Theorem 1.1 is therefore an immediate consequence of theorem 1.1 and theorem 1.2 in [6]. This concludes our discussion of the ℓ^p-Gaussian-Grothendieck problem for $1 \leq p \leq 2$.

3 The limit of the constrained Lagrangian

In this section we begin the proof of theorem 1.2 by combining the Gaussian concentration inequality with the Guerra-Toninelli interpolation to show that the random quantity (1.27) almost surely admits a deterministic limit for every constraint $D \in \Gamma_\kappa$. As usual [13, 29], the proof will come down to proving super-additivity of an appropriate sequence and appealing to the classical Fekete lemma.

Theorem 3.1. *If $2 < p < \infty$ and $t > 0$, then for every $D \in \Gamma_\kappa$ the limit*

$$L_{p,D}(t) = \lim_{N\to\infty} \mathbb{E}\, L_{N,p,D}(t) \tag{3.1}$$

exists. Moreover, with probability one, $L_{p,D}(t) = \lim_{N\to\infty} L_{N,p,D}(t)$.

Proof. We will be working with systems of different sizes, so let us make the dependence of (1.26) on N explicit by writing $\Sigma_N(D)$. Given $\vec{\boldsymbol\sigma} \in \Sigma_N(D)$, the covariance structure of the vector spin Hamiltonian (1.12) established in (1.19) together with lemma B.3 reveal that

$$\mathbb{E}\, H_N(\vec{\boldsymbol\sigma})^2 = N\|R(\vec{\boldsymbol\sigma},\vec{\boldsymbol\sigma})\|_{\mathrm{HS}}^2 \leq N\,\mathrm{tr}(R(\vec{\boldsymbol\sigma},\vec{\boldsymbol\sigma}))^2 = N\,\mathrm{tr}(D)^2.$$

It follows by the Gaussian concentration of the maximum that for any $s > 0$,

$$\mathbb{P}\big\{|L_{N,p,D}(t) - \mathbb{E}\, L_{N,p,D}(t)| \geq s\big\} \leq 2\exp\Big(-\frac{Ns^2}{4\,\mathrm{tr}(D)^2}\Big).$$

Since the right-hand side of this expression is summable in N, the Borel-Cantelli lemma implies that

$$\limsup_{N\to\infty} |L_{N,p,D}(t) - \mathbb{E}\, L_{N,p,D}(t)| = 0$$

with probability one. It is therefore sufficient to prove that the sequence $(\mathbb{E}\, L_{N,p,D}(t))_N$ admits a limit. We will do this through the Fekete lemma by showing that the sequence $(N\,\mathbb{E}\, L_{N,p,D}(t))_N$ is super-additive. This is equivalent to proving that for all integers $N, M \geq 1$,

$$\mathbb{E}\max_{\vec{\boldsymbol\sigma}\in\Sigma_N(D)} H_{N,p,t}(\vec{\boldsymbol\sigma}) + \mathbb{E}\max_{\vec{\boldsymbol\sigma}\in\Sigma_M(D)} H_{M,p,t}(\vec{\boldsymbol\sigma}) \leq \mathbb{E}\max_{\vec{\boldsymbol\sigma}\in\Sigma_{N+M}(D)} H_{N+M,p,t}(\vec{\boldsymbol\sigma}). \tag{3.2}$$

Given a spin configuration $\vec{\boldsymbol\rho} \in (\mathbb{R}^\kappa)^{N+M}$, write $\vec{\boldsymbol\rho} = (\vec{\boldsymbol\sigma},\vec{\boldsymbol\tau})$ for $\vec{\boldsymbol\sigma} \in (\mathbb{R}^\kappa)^N$ and $\vec{\boldsymbol\tau} \in (\mathbb{R}^\kappa)^M$. Consider three independent vector spin SK Hamiltonians $H_{N+M}(\vec{\boldsymbol\rho})$, $H_N(\vec{\boldsymbol\sigma})$ and $H_M(\vec{\boldsymbol\tau})$ defined on $\Sigma_N(D)\times\Sigma_M(D)$, $\Sigma_N(D)$ and $\Sigma_M(D)$ respectively. For each $s \in [0,1]$ introduce the interpolating Hamiltonian on $\Sigma_N(D)\times\Sigma_M(D)$,

$$H_{N+M,s}(\vec{\boldsymbol\rho}) = \sqrt{s}H_{N+M}(\vec{\boldsymbol\rho}) + \sqrt{1-s}\big(H_N(\vec{\boldsymbol\sigma}) + H_M(\vec{\boldsymbol\tau})\big) - t\|\vec{\boldsymbol\sigma}\|_{p,2}^p - t\|\vec{\boldsymbol\tau}\|_{p,2}^p.$$

Given an inverse temperature parameter $\beta > 0$ and two probability measures μ_N and μ_M supported on $\Sigma_N(D)$ and $\Sigma_M(D)$ respectively, denote by

$$\varphi(s) = \frac{1}{\beta(N+M)}\,\mathbb{E}\log\int_{\Sigma_N(D)\times\Sigma_M(D)} \exp\beta H_{N+M,s}(\vec{\boldsymbol\rho})\,\mathrm{d}\,\mu_N(\vec{\boldsymbol\sigma})\,\mathrm{d}\,\mu_M(\vec{\boldsymbol\tau})$$

the interpolating free energy and write $\langle\cdot\rangle_s$ for the Gibbs average with respect to the interpolating Gibbs measure

$$\mathrm{d}\, G_{N+M}(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{\tau}}) = \frac{\exp \beta H_{N+M,s}(\vec{\boldsymbol{\rho}})\,\mathrm{d}\,\mu_N(\vec{\boldsymbol{\sigma}})\,\mathrm{d}\,\mu_M(\vec{\boldsymbol{\tau}})}{\int_{\Sigma_N(D)\times\Sigma_M(D)} \exp \beta H_{N+M,s}(\vec{\boldsymbol{\rho}})\,\mathrm{d}\,\mu_N(\vec{\boldsymbol{\sigma}})\,\mathrm{d}\,\mu_M(\vec{\boldsymbol{\tau}})}.$$

The Gaussian integration by parts formula (see for instance lemma 1.1 in [29]) yields

$$\varphi'(s) = \frac{1}{N+M}\,\mathbb{E}\left\langle \frac{\partial H_{N+M,s}(\vec{\boldsymbol{\sigma}})}{\partial s}\right\rangle_s = \frac{1}{N+M}\,\mathbb{E}\left\langle C(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^1) - C(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^2)\right\rangle_s,$$

where

$$C(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^2) = \frac{\beta(N+M)}{2}\Big(\|R(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^2)\|_{\mathrm{HS}}^2 - \frac{N}{N+M}\|R(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}^2 - \frac{M}{N+M}\|R(\vec{\boldsymbol{\tau}}^1,\vec{\boldsymbol{\tau}}^2)\|_{\mathrm{HS}}^2\Big).$$

Since

$$R(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^2) = \frac{N}{N+M}R(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2) + \frac{M}{N+M}R(\vec{\boldsymbol{\tau}}^1,\vec{\boldsymbol{\tau}}^2),$$

the convexity of the square of a norm implies that $C(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^2)\leq 0$. Combined with the fact that $R(\vec{\boldsymbol{\rho}}^1,\vec{\boldsymbol{\rho}}^1) = R(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^1) = R(\vec{\boldsymbol{\tau}}^1,\vec{\boldsymbol{\tau}}^1) = D$, this shows that $\varphi'(s)\geq 0$ and therefore $\varphi(0)\leq\varphi(1)$. Letting $\beta\to\infty$ in this inequality and remembering that the L^q-norm tends to the L^∞-norm as $q\to\infty$ yields

$$\mathbb{E}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma_N(D)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) + \mathbb{E}\max_{\vec{\boldsymbol{\tau}}\in\Sigma_M(D)} H_{M,p,t}(\vec{\boldsymbol{\tau}}) \leq \mathbb{E}\max_{\vec{\boldsymbol{\rho}}\in\Sigma_N(D)\times\Sigma_M(D)} H_{N+M,p,t}(\vec{\boldsymbol{\rho}}).$$

Since $\Sigma_N(D)\times\Sigma_M(D)\subset\Sigma_{N+M}(D)$, this gives (3.2) and completes the proof. □

The heuristic validity of theorem 1.2 should now be clear. From (1.18), the self-overlap of any vector spin configuration is a Gram matrix in Γ_κ. This means that for every integer $N\geq 1$, the relationship between the unconstrained Lagrangian (1.23) and the constrained Lagrangian (1.27) is

$$L_{N,p}(t) = \sup_{D\in\Gamma_\kappa} L_{N,p,D}(t). \tag{3.3}$$

Formally bringing the limit into the supremum and using the density of positive definite matrices in the space of non-negative definite matrices yields (1.28). To turn this heuristic into a rigorous argument, we will use a compactness argument. This will be done in section 5 and will require the continuity properties of the constrained Lagrangian (1.27) that we explore in the next section.

4 Continuity of the constrained Lagrangian

In this section we prove that, in a certain sense, the limit of the constrained Lagrangian (1.27) is continuous with respect to the constraint $D\in\Gamma_\kappa$ by combining lemma 4 in [33] with the classical Dudley inequality as it is stated in equation (A.23) of [35].

Lemma 4 in [33] was originally designed to modify the vector spin coordinates in the mixed-p-spin model in order to prove the matrix-overlap Ghirlanda-Guerra identities. Using these identities, it is then possible to access the synchronization mechanism [31, 32] and find a tight lower bound for the limit of the free energy through the Aizenman-Sims-Starr scheme [22, 33]. We will apply this lemma for a different purpose, and, as it turns out, we will need a more explicit expression for the constant $L>0$ appearing in the upper bound. For our purposes, it will be important that this constant is uniformly bounded for all $D \in \Gamma_\kappa$ with uniformly bounded trace. We will therefore repeat the proof of this result and carefully track the dependence of constants.

For each $\epsilon > 0$ and $D \in \Gamma_\kappa$ denote by $B_\epsilon(D)$ the open ϵ-neighbourhood of D,

$$B_\epsilon(D) = \{\gamma \in \Gamma_\kappa \mid \|\gamma - D\|_\infty < \epsilon\}, \tag{4.1}$$

with respect to the sup-norm $\|\gamma\|_\infty = \max_{k,k'}|\gamma_{k,k'}|$ on the space of $\kappa \times \kappa$ matrices, and consider the set of spin configurations

$$\Sigma_\epsilon(D) = \big\{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N \mid R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}}) \in B_\epsilon(D)\big\} \tag{4.2}$$

with self-overlap in the ϵ-neighbourhood of D. Denote by $\lambda_1 \geq \cdots \geq \lambda_\kappa$ the real and non-negative eigenvalues of D and let

$$D = Q\Lambda Q^T \tag{4.3}$$

be the eigendecomposition of D with diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_\kappa)$. Given $\epsilon > 0$, let $0 \leq m \leq \kappa$ be such that $\lambda_m \geq \sqrt{\epsilon}$ and $\lambda_{m+1} < \sqrt{\epsilon}$. Introduce the matrix

$$D_\epsilon = Q\Lambda_\epsilon Q^T, \tag{4.4}$$

where $\Lambda_\epsilon = \mathrm{diag}(\lambda_1, \ldots, \lambda_m, 0, \ldots, 0)$. Notice that $D_\epsilon = D$ when $\epsilon > 0$ is smaller than the smallest non-zero eigenvalue of D. Given any $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$, we will construct a $\kappa \times \kappa$ matrix $A_{\vec{\boldsymbol{\sigma}}}$ such that the self-overlap of the configuration $A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} = (A_{\vec{\boldsymbol{\sigma}}}\vec{\sigma}_i)_{i \leq N}$ is equal to D_ϵ and such that, in a certain sense, $A_{\vec{\boldsymbol{\sigma}}}$ has small distortion. Notice that the self-overlap of $A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}$ is given by

$$R(A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}, A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}) = \frac{1}{N}\sum_{i=1}^N (A_{\vec{\boldsymbol{\sigma}}}\vec{\sigma}_i)(A_{\vec{\boldsymbol{\sigma}}}\vec{\sigma}_i)^T = A_{\vec{\boldsymbol{\sigma}}}R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})A_{\vec{\boldsymbol{\sigma}}}^T, \tag{4.5}$$

so we will need a matrix with $A_{\vec{\boldsymbol{\sigma}}}R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})A_{\vec{\boldsymbol{\sigma}}}^T = D_\epsilon$. In this context, small distortion will mean that the overlap of $\vec{\boldsymbol{\sigma}}$ with other configurations should not change much when $\vec{\boldsymbol{\sigma}}$ is replaced by $A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}$. To control this distortion, fix a configuration $\vec{\boldsymbol{\rho}} \in (\mathbb{R}^\kappa)^N$ with $|||\vec{\boldsymbol{\rho}}|||_{2,2}^2 \leq u$ for some $u > 0$, let $\vec{\boldsymbol{\tau}} = A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}$ and observe that by the Cauchy-Schwarz inequality

$$\begin{aligned}
\|R(A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\rho}}) - R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\rho}})\|_{\mathrm{HS}} &= \Big\|\frac{1}{N}\sum_{i=1}^N A_{\vec{\boldsymbol{\sigma}}}\vec{\sigma}_i\vec{\rho}_i^{\,T} - \frac{1}{N}\sum_{i=1}^N \vec{\sigma}_i\vec{\rho}_i^{\,T}\Big\|_{\mathrm{HS}} \\
&\leq \frac{1}{N}\sum_{i=1}^N \|\vec{\tau}_i\vec{\rho}_i^{\,T}\|_{\mathrm{HS}} = \frac{1}{N}\sum_{i=1}^N \|\vec{\tau}_i\|_2\|\vec{\rho}_i\|_2 \\
&\leq \sqrt{u}\Big(\frac{1}{N}\sum_{i=1}^N \|\vec{\tau}_i\|_2^2\Big)^{1/2} = \sqrt{u}\,\mathrm{tr}(R(\vec{\boldsymbol{\tau}}, \vec{\boldsymbol{\tau}}))^{1/2} \\
&= \sqrt{u}\,\mathrm{tr}\left((A_{\vec{\boldsymbol{\sigma}}} - I)R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})(A_{\vec{\boldsymbol{\sigma}}} - I)^T\right)^{1/2},
\end{aligned} \tag{4.6}$$

where the last inequality follows from the fact that $\vec{\boldsymbol{\tau}} = (A_{\vec{\boldsymbol{\sigma}}} - I)\vec{\boldsymbol{\sigma}}$. We therefore need a matrix for which $\mathrm{tr}((A_{\vec{\boldsymbol{\sigma}}} - I)R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})(A_{\vec{\boldsymbol{\sigma}}} - I)^T)$ is small. The construction of the matrix $A_{\vec{\boldsymbol{\sigma}}}$ is precisely the content of lemma 4 in [33].

Lemma 4.1. *Given $0 < \epsilon < \kappa^{-2}$, $D \in \Gamma_\kappa$, and $R \in B_\epsilon(D)$, there exists a matrix $A = A(R)$ such that $ARA^T = D_\epsilon$ and*

$$\mathrm{tr}\left((A - I)R(A - I)^T\right) \leq C(\mathrm{tr}(D) + 1)\sqrt{\epsilon} \tag{4.7}$$

for some constant $C > 0$ that depends only on κ.

Proof. The proof proceeds in two steps: first we reduce the problem to the case when $D = \Lambda$ and then we use Gershgorin's theorem to conclude. For the reader's convenience, Gershgorin's theorem has been transcribed as theorem B.1 in the appendix.

Step 1: reducing to $D = \Lambda$

Let us suppose temporarily that the result holds when D is a diagonal matrix. Since Q is an orthogonal matrix and the Hilbert-Schmidt norm is rotationally invariant,

$$\|Q^T RQ - \Lambda\|_\infty \leq \|R - D\|_{\mathrm{HS}} \leq \kappa\epsilon.$$

We may therefore find a matrix $A(Q^T RQ)$ with $A(Q^T RQ)Q^T RQA(Q^T RQ)^T = \Lambda_\epsilon$ and

$$\mathrm{tr}\left((A(Q^T RQ) - I)Q^T RQ(A(Q^T RQ) - I)^T\right) \leq C\,\mathrm{tr}(\Lambda)\sqrt{\epsilon}.$$

If we set $A = QA(Q^T RQ)Q^T$, it is easy to see that $ARA^T = Q\Lambda_\epsilon Q^T = D_\epsilon$ and

$$\begin{aligned}\mathrm{tr}\left((A - I)R(A - I)^T\right) &= \mathrm{tr}\left(Q(A(Q^T RQ) - I)Q^T RQ(A(Q^T RQ) - I)Q^T\right)\\ &\leq C\,\mathrm{tr}(D)\sqrt{\epsilon}.\end{aligned}$$

The last inequality uses the cyclicity of the trace, the orthogonality of Q and the fact that $\mathrm{tr}(D) = \mathrm{tr}(\Lambda)$. This shows that it suffices to prove the result when $D = \Lambda$ and $R \in B_\epsilon(\Lambda)$.

Step 2: proof for $D = \Lambda$

Introduce the matrices $R_m = (R_{k,k'})_{k,k'\leq m}$ and $\Lambda_m = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ consisting of the first m rows and columns of R and Λ respectively. Consider the matrix $\tilde{R}_m = \Lambda_m^{-1/2} R_m \Lambda_m^{-1/2}$. Since $R_m \in B_\epsilon(\Lambda_m)$ and Λ_m is diagonal with all elements bounded below by $\sqrt{\epsilon}$, it is readily verified that $\|\tilde{R}_m - I\|_\infty \leq \sqrt{\epsilon}$. Gershgorin's theorem implies that all the eigenvalues of $\tilde{R}_m$ are within $m\sqrt{\epsilon}$ from 1. The assumption $\epsilon < \kappa^{-2}$ implies that $\tilde{R}_m$ is invertible and allows us to define the matrix

$$B = B(R_m) = \Lambda_m^{1/2}\tilde{R}_m^{-1/2}\Lambda_m^{-1/2}.$$

Using the fact that $R_m = \Lambda_m^{1/2}\tilde{R}_m\Lambda_m^{1/2}$, it is easy to see that $BR_mB^T = \Lambda_m$ and

$$(B - I)R_m(B - I)^T = \Lambda_m^{1/2}(I - \tilde{R}_m^{1/2})^2\Lambda_m^{1/2}.$$

Since the eigenvalues of $\tilde{R}_m$ are within $m\sqrt{\epsilon}$ from 1, so are the eigenvalues of $\tilde{R}_m^{1/2}$. Observe that $\tilde{R}_m^{1/2}$ is symmetric and non-negative definite, so it admits an eigendecomposition $\tilde{R}_m^{1/2} = \tilde{Q}_m\tilde{\Lambda}_m\tilde{Q}_m^T$. It follows by the orthogonality of $\tilde{Q}_m$ that

$$\|I - \tilde{R}_m^{1/2}\|_{\mathrm{HS}} = \|I - \tilde{\Lambda}_m\|_{\mathrm{HS}} \leq m\|I - \tilde{\Lambda}_m\|_\infty \leq \kappa^2\sqrt{\epsilon}.$$

The cyclicity of the trace, the Cauchy-Schwarz inequality and lemma B.3 now give

$$\begin{aligned}\mathrm{tr}\left((B - I)R_m(B - I)^T\right) &= \mathrm{tr}\left(\Lambda_m(I - \tilde{R}_m^{1/2})^2\right) \leq \|\Lambda_m\|_{\mathrm{HS}}\|I - \tilde{R}_m^{1/2}\|_{\mathrm{HS}}^2\\ &\leq \kappa^4\,\mathrm{tr}(\Lambda_m)\epsilon.\end{aligned}$$

Finally, define the matrix A by filling all rows and columns of B from $m+1$ to κ with zeros. It is clear that $ARA^T = \Lambda_\epsilon$. If we denote by $T = (R_{k,k'})_{k,k'\geq m+1}$ the matrix consisting of the last $\kappa - m$ rows and columns of R, then

$$\begin{aligned}\mathrm{tr}\left((A - I)R(A - I)^T\right) &= \mathrm{tr}\left((B - I)R_m(B - I)^T\right) + \mathrm{tr}(T)\\ &\leq \kappa^4\,\mathrm{tr}(\Lambda_m)\epsilon + (\kappa - m)\epsilon + \sum_{k=m+1}^{\kappa}\lambda_i\\ &\leq \left(\kappa^4\,\mathrm{tr}(\Lambda) + 2\kappa\right)\sqrt{\epsilon}.\end{aligned}$$

We have used the fact that $T \in B_\epsilon(\Lambda)$ in the second inequality. This completes the proof. □

This result allows us to map any spin configuration $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ with self-overlap in the ϵ-neighbourhood of $D \in \Gamma_\kappa$ to a modified spin configuration $A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}$ that is not too far from $\vec{\boldsymbol{\sigma}}$ and has a configuration-independent self-overlap D_ϵ. These two facts will be fundamental to understanding the continuity of the constrained Lagrangian (1.27). We will now quantify the distance between $\vec{\boldsymbol{\sigma}}$ and $A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}$ in two different ways: with respect to the normalized-$\ell^{2,2}$-norm and relative to the canonical metric associated with the Hamiltonian (1.12),

$$\mathrm{d}(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2) = \Big(\mathbb{E}\left(H_N(\vec{\boldsymbol{\sigma}}^1) - H_N(\vec{\boldsymbol{\sigma}}^2)\right)^2\Big)^{1/2}. \tag{4.8}$$

It will be convenient to notice that for any $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$,

$$|||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 = \frac{1}{N}\sum_{i=1}^N \sum_{k=1}^\kappa |\sigma_i(k)|^2 = \sum_{k=1}^\kappa R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})_{k,k} = \mathrm{tr}(R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})), \tag{4.9}$$

and to write

$$B_2^N(u) = \left\{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N \mid |||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 \leq u\right\} \tag{4.10}$$

for the ball of radius $\sqrt{u}$ with respect to the normalized-$\ell^{2,2}$-norm.

Corollary 4.2. *If $0 < \epsilon < \kappa^{-2}$ and $D \in \Gamma_\kappa$, then for any $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$,*

$$|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}|||_{2,2} \leq C(\mathrm{tr}(D) + 1)^{1/2}\epsilon^{1/4}, \tag{4.11}$$

where $C > 0$ is a constant that depends only on κ.

Proof. By (4.9), (4.5) and lemma 4.1,

$$|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}|||_{2,2}^2 = \mathrm{tr}((A_{\vec{\boldsymbol{\sigma}}} - I)R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})(A_{\vec{\boldsymbol{\sigma}}} - I)) \leq C(\mathrm{tr}(D) + 1)\epsilon^{1/2}.$$

This finishes the proof. □

Corollary 4.3. *If $u > 1$ and $\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2 \in B_2^N(u)$, then*

$$\mathrm{d}(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2) \leq 2N^{1/4}u^{3/4}\|\vec{\boldsymbol{\sigma}}^1 - \vec{\boldsymbol{\sigma}}^2\|_{2,2}^{1/2}. \tag{4.12}$$

In particular, if $0 < \epsilon < \kappa^{-2}$ and $D \in \Gamma_\kappa$, then for any $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$,

$$\mathrm{d}(\vec{\boldsymbol{\sigma}}, A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}) \leq CN^{1/2}(\mathrm{tr}(D) + 1)\epsilon^{1/8}, \tag{4.13}$$

where $C > 0$ is a constant that depends only on κ.

Proof. By the reverse triangle inequality,

$$\begin{aligned}
\mathrm{d}(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)^2 &= N\big(\|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^1)\|_{\mathrm{HS}}^2 + \|R(\vec{\boldsymbol{\sigma}}^2, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}^2 - 2\|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}^2\big) \\
&\leq N\big(\|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^1) - R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}(\|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^1)\|_{\mathrm{HS}} + \|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}) \\
&\quad + \|R(\vec{\boldsymbol{\sigma}}^2, \vec{\boldsymbol{\sigma}}^2) - R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}}(\|R(\vec{\boldsymbol{\sigma}}^2, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}} + \|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}})\big).
\end{aligned}$$

To bound this further, notice that by (1.18) and the Cauchy-Schwarz inequality,

$$\begin{aligned}
\|R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^1) - R(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}} &\leq \frac{1}{N}\sum_{i=1}^N \|\vec{\sigma}_i^1(\vec{\sigma}_i^1 - \vec{\sigma}_i^2)^T\|_{\mathrm{HS}} = \frac{1}{N}\sum_{i=1}^N \|\vec{\sigma}_i^1\|_2 \|\vec{\sigma}_i^1 - \vec{\sigma}_i^2\|_2 \\
&\leq |||\vec{\boldsymbol{\sigma}}^1|||_{2,2}|||\vec{\boldsymbol{\sigma}}^1 - \vec{\boldsymbol{\sigma}}^2|||_{2,2}.
\end{aligned}$$

Similarly, $\|R(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2)\|_{\mathrm{HS}} \le |||\vec{\boldsymbol{\sigma}}^1|||_{2,2}|||\vec{\boldsymbol{\sigma}}^2|||_{2,2}$. It follows that for any $\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2 \in B_2^N(u)$,

$$\mathrm{d}(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2)^2 \le 4N^{1/2}u^{3/2}\|\vec{\boldsymbol{\sigma}}^1-\vec{\boldsymbol{\sigma}}^2\|_{2,2}.$$

Taking square roots yields (4.12). To prove (4.13), observe that for any $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$,

$$|||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 = \mathrm{tr}(R(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{\sigma}})) \le \mathrm{tr}(D)+\epsilon\kappa \le \mathrm{tr}(D)+1. \tag{4.14}$$

Invoking corollary 4.2 and (4.12) implies that $\mathrm{d}(\vec{\boldsymbol{\sigma}},A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}) \le CN^{1/2}(\mathrm{tr}(D)+1)\epsilon^{1/8}$. This completes the proof. □

Combining corollary 4.2 and corollary 4.3 with Dudley's entropy inequality, we will now show that, in a certain sense, the constrained Lagrangian (1.27) is continuous with respect to the constraint $D \in \Gamma_\kappa$. To state this continuity result precisely, for each $\epsilon > 0$ and $D \in \Gamma_\kappa$ introduce the relaxed constrained Lagrangian

$$L_{N,p,D,\epsilon}(t) = \frac{1}{N}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma_\epsilon(D)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}). \tag{4.15}$$

Proposition 4.4. *If $2<p<\infty$, then for each $0<\epsilon<\kappa^{-2}$, every $t>0$ and all $D \in \Gamma_\kappa$,*

$$\limsup_{N\to\infty} L_{N,p,D,\epsilon}(t) \le L_{p,D_\epsilon}(t) + C(1+tp)(\mathrm{tr}(D)+1)^{p/2}\epsilon^{1/64} \tag{4.16}$$

for some constant $C>0$ that depends only on κ.

Proof. To simplify notation, let $C>0$ denote a constant that depends only on κ whose value might not be the same at each occurrence. By the Gaussian concentration of the maximum and a simple application of the Borel-Cantelli lemma, it suffices to prove that

$$\limsup_{N\to\infty}\frac{1}{N}\,\mathbb{E}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma_\epsilon(D)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \le L_{p,D_\epsilon}(t) + C(1+tp)(\mathrm{tr}(D)+1)^{p/2}\epsilon^{1/64}. \tag{4.17}$$

To simplify notation, let $u=\mathrm{tr}(D)+1$. Notice that $|||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 \le u$ for every $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$ by (4.14). Invoking corollary 4.3 and corollary 4.2 gives

$$\frac{1}{N}\,\mathbb{E}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma_\epsilon(D)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \le \frac{1}{N}\,\mathbb{E}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma(D_\epsilon)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) + \frac{1}{N}(I) + \frac{t}{N}(II), \tag{4.18}$$

where

$$(I) = \frac{1}{N}\,\mathbb{E}\max_{\vec{\boldsymbol{\sigma}}\in\Sigma_\epsilon(D)}|H_N(\vec{\boldsymbol{\sigma}})-H_N(A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}})| \le \mathbb{E}\max_{\mathrm{d}(\vec{\boldsymbol{\sigma}}^1,\vec{\boldsymbol{\sigma}}^2)\le CuN^{1/2}\epsilon^{1/8}}|H_N(\vec{\boldsymbol{\sigma}}^1)-H_N(\vec{\boldsymbol{\sigma}}^2)|$$
$$(II) = \max_{\vec{\boldsymbol{\sigma}}\in\Sigma_\epsilon(D)}\big(|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2}^p - |||\vec{\boldsymbol{\sigma}}|||_{p,2}^p\big).$$

To bound the first of these terms, for each $\epsilon>0$ denote by $\mathcal{N}(A,d,\epsilon)$ the ϵ-covering number of the set $A \subset (\mathbb{R}^\kappa)^N$ with respect to the metric d on $(\mathbb{R}^\kappa)^N$, and write B_N for the Euclidean unit ball in $(\mathbb{R}^\kappa)^N$. Dudley's entropy inequality and corollary 4.3 imply that

$$\begin{aligned}(I) &\le C\int_0^{CuN^{1/2}\epsilon^{1/8}}\sqrt{\log\mathcal{N}\big(B_2^N(u),\mathrm{d},\delta\big)}\,\mathrm{d}\,\delta\\
&\le C\int_0^{CuN^{1/2}\epsilon^{1/8}}\sqrt{\log\mathcal{N}\big(B_2^N(u),\|\cdot\|_{2,2},2^{-2}u^{-3/2}N^{-1/2}\delta^2\big)}\,\mathrm{d}\,\delta\\
&\le C\int_0^{CuN^{1/2}\epsilon^{1/8}}\sqrt{\log\mathcal{N}\big(B_N,\|\cdot\|_{2,2},2^{-2}u^{-2}N^{-1}\delta^2\big)}\,\mathrm{d}\,\delta.\end{aligned}$$

At this point, recall that the covering number of the Euclidean unit ball B_N in $(\mathbb{R}^\kappa)^N$ satisfies

$$\Big(\frac{1}{\epsilon}\Big)^{N\kappa} \leq \mathcal{N}(B_N, \|\cdot\|_{2,2}, \epsilon) \leq \Big(\frac{2}{\epsilon}+1\Big)^{N\kappa}$$

for every $\epsilon > 0$. A proof of this bound may be found in corollary 4.2.13 of [38]. Combining this with a change of variables reveals that

$$\begin{aligned}(I) &\leq CN^{1/2}\kappa^{1/2}\int_0^{CuN^{1/2}\epsilon^{1/8}} \sqrt{\log(1+8u^2N\delta^{-2})}\,\mathrm{d}\,\delta \\ &= CNu\int_{C\epsilon^{-1/16}}^\infty \frac{\sqrt{\log(1+\delta)}}{\delta^{3/2}}\,\mathrm{d}\,\delta \leq CNu\int_{C\epsilon^{-1/16}}^\infty \frac{\sqrt{\delta^{1/2}}}{\delta^{3/2}}\,\mathrm{d}\,\delta \\ &\leq CNu\epsilon^{1/64}. \end{aligned} \tag{4.19}$$

To bound the term (II), notice that for any $x, y > 0$,

$$(x+y)^p - x^p = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\,t}(x+ty)^p\,\mathrm{d}\,t = p\int_0^1 (x+ty)^{p-1}y\,\mathrm{d}\,t \leq py(x+y)^{p-1}. \tag{4.20}$$

If $\vec{\boldsymbol{\sigma}} \in \Sigma_\epsilon(D)$ is such that $|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2} > |||\vec{\boldsymbol{\sigma}}|||_{p,2}$, then applying this inequality with $x = |||\vec{\boldsymbol{\sigma}}|||_{p,2}$ and $y = |||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2} - |||\vec{\boldsymbol{\sigma}}|||_{p,2}$ gives

$$|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2}^p - |||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \leq p|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}|||_{p,2}|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2}^{p-1} \tag{4.21}$$

$$\leq p|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\sigma}}|||_{2,2}|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{2,2}^{p-1}. \tag{4.22}$$

The second inequality uses the fact that $\ell^{2,2}$ is continuously embedded in $\ell^{p,2}$ for $p > 2$. Since this bound holds trivially when $|||A_{\vec{\boldsymbol{\sigma}}}\vec{\boldsymbol{\sigma}}|||_{p,2} \leq |||\vec{\boldsymbol{\sigma}}|||_{p,2}$, we deduce from corollary 4.2 that

$$(II) \leq Cpu^{p/2}\epsilon^{1/4}. \tag{4.23}$$

Substituting (4.19) and (4.23) into (4.18) and letting $N \to \infty$ yields (4.17). This completes the proof. □

In the heuristic proof of theorem 1.2 given at the end of section 3, we used the density of positive definite matrices in the space of non-negative definite matrices to obtain the second equality in (1.28). When we come to the rigorous proof of this equality, the argument will be more subtle as proposition 4.4 does not quite give continuity. We will instead content ourselves with controlling the limit of the constrained Lagrangian (1.27) for a non-negative definite matrix $D \in \Gamma_\kappa$ by that for some positive definite matrix in Γ_κ^+ through the following bound.

Proposition 4.5. *If $2 < p < \infty$, then for each $0 < \epsilon < \kappa^{-2}$, every $t > 0$ and all $D \in \Gamma_\kappa$,*

$$L_{p,D}(t) \leq L_{p,D+\epsilon I} + C(1+tp)(\mathrm{tr}(D)+1)^{p/2}\epsilon^{1/64} \tag{4.24}$$

for some constant $C > 0$ that depends only on κ.

Proof. Fix $N > 2\kappa$ and $\vec{\boldsymbol{\sigma}} \in \Sigma(D)$. Endow $\mathbb{R}^N$ with the inner product

$$\langle \boldsymbol{\rho}, \boldsymbol{\tau}\rangle = \frac{1}{N}\sum_{i=1}^N \boldsymbol{\rho}_i\boldsymbol{\tau}_i.$$

Since $N > 2\kappa$, there exist mutually orthogonal vectors $\boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(1), \ldots, \boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(\kappa)$ that are also orthogonal to each of the vectors $\boldsymbol{\sigma}(1), \ldots, \boldsymbol{\sigma}(\kappa)$ and satisfy $\langle \boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(k), \boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(k)\rangle = \kappa^{-1}$ for

$1 \leq k \leq \kappa$. Consider the configuration $\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}} \in (\mathbb{R}^\kappa)^N$ defined by $\boldsymbol{\rho}_{\vec{\boldsymbol{\sigma}}}(k) = \boldsymbol{\sigma} + \sqrt{\epsilon}\boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}$. By orthogonality,

$$R(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{\rho}})_{k,k'} = \langle \boldsymbol{\rho}(k), \boldsymbol{\rho}(k') \rangle = \langle \boldsymbol{\sigma}(k), \boldsymbol{\sigma}(k') \rangle + \epsilon \delta_{k,k'} = R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}})_{k,k'} + \epsilon \delta_{k,k'},$$

where $\delta_{k,k'} = 1$ if $k = k'$ and is zero otherwise. This means that $\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}} \in \Sigma(D + \epsilon I)$. Moreover, the normalization of the vectors $\boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(k)$ implies that

$$|||\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}} - \vec{\boldsymbol{\sigma}}|||_{2,2}^2 = \frac{1}{N} \sum_{k=1}^{\kappa} \|\boldsymbol{\rho}_{\vec{\boldsymbol{\sigma}}}(k) - \boldsymbol{\sigma}(k)\|_2^2 = \epsilon \sum_{k=1}^{\kappa} \langle \boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(k), \boldsymbol{\tau}_{\vec{\boldsymbol{\sigma}}}(k) \rangle = \epsilon.$$

If we let $u = \mathrm{tr}(D) + 1$, then (4.12) reveals that

$$\mathrm{d}(\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}}, \vec{\boldsymbol{\sigma}}) \leq 2N^{1/2} u^{3/4} \epsilon^{1/4} \leq 2N^{1/2} u \epsilon^{1/8},$$

while an identical argument to that used to obtain (4.22) yields

$$|||\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}}|||_{p,2}^p - |||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \leq p |||\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}} - \vec{\boldsymbol{\sigma}}|||_{2,2} |||\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}}|||_{2,2}^{p-1} \leq p\sqrt{\epsilon} u^{p/2}.$$

It follows that

$$\begin{aligned}\frac{1}{N} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) &\leq \frac{1}{N} H_{N,p,t}(\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}}) + \frac{1}{N} |H_N(\vec{\boldsymbol{\sigma}}) - H_N(\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}})| + t\big(|||\vec{\boldsymbol{\rho}}_{\vec{\boldsymbol{\sigma}}}|||_{p,2}^p - |||\vec{\boldsymbol{\sigma}}|||_{p,2}^p\big) \\ &\leq L_{N,p,D+\epsilon I}(t) + \frac{1}{N} \max_{\mathrm{d}(\vec{\boldsymbol{\sigma}}^1, \vec{\boldsymbol{\sigma}}^2) \leq 2uN^{1/2}\epsilon^{1/8}} |H_N(\vec{\boldsymbol{\sigma}}^1) - H_N(\vec{\boldsymbol{\sigma}}^2)| + tp\sqrt{\epsilon} u^{p/2}.\end{aligned}$$

Taking the maximum over configurations $\vec{\boldsymbol{\sigma}} \in \Sigma(D)$ and using Dudley's entropy inequality exactly as in the proof of proposition 4.4 gives

$$\mathbb{E}\, L_{N,p,D}(t) \leq \mathbb{E}\, L_{N,p,D+\epsilon I}(t) + C(1 + tp) u^{p/2} \epsilon^{1/64}$$

for some constant $C > 0$ that depends only on κ. Letting $N \to \infty$ completes the proof. □

The results established in this section together with the arguments in section 7 of [6] will allow us to give a rigorous proof of theorem 1.2. The proof will consist of two key steps. First, we will use proposition 4.4 to express a version of the Lagrangian (1.23) localized to a ball of fixed but arbitrary radius $u > 0$ as a supremum of constrained Lagrangians (1.27). Then, we will modify the scaling arguments in section 7 of [6] to show that the unconstrained Lagrangian (1.23) can be obtained by taking the supremum of these localized Lagrangians over all radii $u > 0$. The formula obtained by taking these successive suprema will be equivalent to the first equality in (1.28). As previously mentioned, the second equality will follow immediately from proposition 4.5. The purpose of restricting the supremum to positive definite matrices is technical and will be emphasized when we prove lemma 11.2.

5 The limit of the unconstrained Lagrangian

In this section we combine proposition 4.4 with the arguments in section 7 of [6] to prove theorem 1.2. As explained at the end of section 4, we will first find a formula for the limit of the localized Lagrangian

$$L_{N,p,u}(t) = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 \leq u} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \tag{5.1}$$

defined for each $u > 0$. If $\Gamma_{\kappa,u}$ denotes the set of matrices in Γ_κ with trace at most u,

$$\Gamma_{\kappa,u} = \{D \in \Gamma_\kappa \mid \mathrm{tr}(D) \leq u\}, \tag{5.2}$$

then (4.9) implies that for every $t > 0$,

$$L_{N,p,u}(t) = \sup_{D \in \Gamma_{\kappa,u}} L_{N,p,D}(t). \tag{5.3}$$

A compactness argument similar to that in lemma 3 of [33] can be used to show that this equality is preserved in the limit.

Proposition 5.1. *If $2 < p < \infty$, then for every $t > 0$ and $u > 0$, the limit $L_{p,u}(t) = \lim_{N\to\infty} \mathbb{E}\, L_{N,p,u}(t)$ exists and is given by*

$$L_{p,u}(t) = \sup_{D \in \Gamma_{\kappa,u}} L_{p,D}(t). \tag{5.4}$$

Moreover, with probability one, $L_{p,u}(t) = \lim_{N\to\infty} L_{N,p,u}(t)$.

Proof. Given $\epsilon > 0$, observe that the collection of sets $B_\epsilon(D)$ for $D \in \Gamma_{\kappa,u}$ forms an open cover of the compact set $\Gamma_{\kappa,u}$. It is therefore possible to find $n \in \mathbb{N}$ and $D^1, \ldots, D^n \in \Gamma_{\kappa,u}$ with $\Gamma_{\kappa,u} \subset \bigcup_{i\le n} B_\epsilon(D^i)$, or equivalently $B_2^N(u) \subset \bigcup_{i\le n} \Sigma_\epsilon(D^i)$. With this in mind, given a probability measure μ^N supported on $B_2^N(u)$, an inverse temperature parameter $\beta > 0$ and a subset $S \subset B_2^N(u)$, consider the free energy

$$F_N^\beta(S) = \frac{1}{N\beta} \log \int_S \exp \beta H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \,\mathrm{d}\, \mu^N(\vec{\boldsymbol{\sigma}}).$$

By monotonicity of the logarithm and the inclusion $B_2^N(u) \subset \bigcup_{i\le n} \Sigma_\epsilon(D^i)$,

$$\begin{aligned} F_N^\beta(B_2^N(u)) &\le \frac{\log n}{N\beta} + \frac{1}{N\beta} \log \max_{1\le i\le n} \int_{\Sigma_\epsilon(D^i)} \exp \beta H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \,\mathrm{d}\, \mu^N(\vec{\boldsymbol{\sigma}}) \\ &= \frac{\log n}{N\beta} + \max_{1\le i\le n} F_N^\beta(\Sigma_\epsilon(D^i)). \end{aligned}$$

The Gaussian concentration inequality implies that the free energy $F_N^\beta(S)$ deviates from its expectation by more than $1/\sqrt{N}$ with probability at most $Le^{-N/L}$, where the constant L does not depend on β, N or S. We deduce from this that with probability at least $1 - Le^{-N/L}$,

$$F_N^\beta(B_2^N(u)) \le \frac{1}{\sqrt{N}} + \frac{\log n}{N\beta} + \max_{1\le i\le n} \mathbb{E}\, F_N^\beta(\Sigma_\epsilon(D^i)).$$

Letting $\beta \to \infty$ and remembering that the L^q-norm converges to the L^∞-norm reveals that with probability at least $1 - Le^{-N/L}$,

$$L_{N,p,u}(t) \le \frac{2}{\sqrt{N}} + \max_{1\le i\le n} L_{N,p,D^i,\epsilon}(t).$$

The Borel-Cantelli lemma and proposition 4.4 now give a constant $C > 0$ that depends only on κ with

$$\limsup_{N\to\infty} L_{N,p,u}(t) \le \max_{1\le i\le n} \big(L_{p,D^i_\epsilon}(t) + C(1+tp)(\operatorname{tr}(D^i)+1)^{p/2}\epsilon^{1/64} \big).$$

Since $\operatorname{tr}(D^i_\epsilon) \le \operatorname{tr}(D^i) \le u$, this can be bounded further by

$$\limsup_{N\to\infty} L_{N,p,u}(t) \le \sup_{D\in\Gamma_{\kappa,u}} L_{p,D}(t) + C(1+tp)(u+1)^{p/2}\epsilon^{1/64}.$$

Remembering that $L_{N,p,D}(t) \le L_{N,p,u}(t)$ for every $N \ge 1$ and $D \in \Gamma_{\kappa,u}$, it follows that

$$\begin{aligned} \sup_{D\in\Gamma_{\kappa,u}} L_{p,D}(t) &\le \liminf_{N\to\infty} L_{N,p,u}(t) \\ &\le \limsup_{N\to\infty} L_{N,p,u}(t) \le \sup_{D\in\Gamma_{\kappa,u}} L_{p,D}(t) + C(1+tp)(u+1)^{p/2}\epsilon^{1/64} \end{aligned}$$

Letting $\epsilon \to 0$ and using the Gaussian concentration of the maximum completes the proof. ☐

This result reduces the proof of theorem 1.2 to establishing the asymptotic version of the equality

$$L_{N,p}(t) = \sup_{u>0} L_{N,p,u}(t). \tag{5.5}$$

This will be done using the techniques in section 7 of [6] and relying upon the identity

$$L_{N,p,u}(t) = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 \leq u} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{2,2} \leq 1} \big(uH_N(\vec{\boldsymbol{\sigma}}) - tu^{p/2} \|\vec{\boldsymbol{\sigma}}\|_{p,2}^p\big) \tag{5.6}$$

which holds for every $t, u > 0$ by a change of variables. The absence of such an equality at the level of the constrained Lagrangian (1.27) is the reason we had to develop the results in section 4.

For technical reasons, before we start thinking about proving the asymptotic version of (5.5), we will have to upgrade the statement of proposition 5.1 to show that $L_{p,u}(t)$ is the limit of the localized Lagrangian (5.1) with probability one simultaneously over all $t, u > 0$. Heuristically, this should not be too surprising. As the maximum of a collection of concave functions, the localized Lagrangian (5.6) is concave in the pair (u, t) conditionally on the disorder chaos (g_{ij}). Since a concave function is Lipschitz continuous on compact sets, this suggests that $(u, t) \mapsto L_{N,p,u}(t)$ should be Lipschitz continuous on compact sets. This continuity would immediately promote almost sure convergence for each $t, u > 0$ to a convergence with probability one simultaneously over all $t, u > 0$. To make this argument rigorous, we will use an ℓ^2-boundedness result of the $N \times N$ random matrix

$$G_N = (g_{ij})_{i,j \leq N}. \tag{5.7}$$

Its proof will rely upon Chevet's inequality as it appears in theorem 8.7.1 of [38].

Lemma 5.2. *There exist constants $C, M > 0$ such that with probability at least $1 - Ce^{-N/C}$,*

$$\frac{1}{\sqrt{N}} \|G_N\|_2 \leq M. \tag{5.8}$$

Proof. Since $\|G_N\|_2 = \max_{\|x\|_2=1}(G_N x, x)$ and $\mathbb{E}(G_N x, x)^2 = 1$ whenever $\|x\|_2 = 1$, the Gaussian concentration of the maximum gives a constant $C > 0$ such that with probability at least $1 - Ce^{-N/C}$,

$$\frac{1}{\sqrt{N}} \|G_N\|_2 \leq \frac{1}{\sqrt{N}} \mathbb{E}\|G_N\|_2 + 1.$$

If g is a standard Gaussian vector in $\mathbb{R}^N$, then Chevet's inequality applied with T and S equal to the Euclidean unit ball in $\mathbb{R}^N$ gives an absolute constant $M > 0$ with

$$\mathbb{E}\|G_N\|_2 = \mathbb{E} \max_{\|x\|_2=1} (G_N x, x) \leq M\, \mathbb{E}\|g\|_2. \tag{5.9}$$

We have used the fact that the Gaussian width of the unit ball is $\mathbb{E} \sup_{\|x\|_2=1}(g, x) = \mathbb{E}\|g\|_2$ while its radius is one. Finally, Jensen's inequality reveals that

$$(\mathbb{E}\|g\|_2)^2 \leq \mathbb{E}\|g\|_2^2 = N\, \mathbb{E}|g_1|^2 = N.$$

Substituting this into (5.9) and redefining the constant $M > 0$ completes the proof. □

Lemma 5.3. *If $2 < p < \infty$, then for any $0 < K_1 < K_2$, there exist constants $C, M > 0$ such that with probability at least $1 - Ce^{-N/C}$,*

$$|L_{N,p,u}(t) - L_{N,p,u'}(t')| \leq M\big(|u - u'| + |t - t'|\big) \tag{5.10}$$

for all $t, t', u, u' \in [K_1, K_2]$.

Proof. Let $\vec{\boldsymbol{\rho}} \in B_2^N(1)$ maximize the right-hand side of (5.6), and define the vector spin configuration $\vec{\boldsymbol{\tau}} \in B_2^N(1)$ by $\vec{\tau}_i = (\kappa^{-1/2}, \ldots, \kappa^{-1/2}) \in \mathbb{R}^\kappa$ for $1 \leq i \leq N$. The Cauchy-Schwarz inequality shows that

$$\begin{aligned} uN^{1/2}\|G_N\|_2 - tu^{p/2}\|\vec{\boldsymbol{\rho}}\|_{p,2}^p &\geq \frac{u}{\sqrt{N}}\sum_{k=1}^{\kappa}(G_N\boldsymbol{\rho}(k), \boldsymbol{\rho}(k)) - tu^{p/2}\|\vec{\boldsymbol{\rho}}\|_{p,2}^p \\ &\geq -\frac{u}{\sqrt{N}}\sum_{k=1}^{\kappa}|(G_N\boldsymbol{\tau}(k), \boldsymbol{\tau}(k))| - tu^{p/2}\|\vec{\boldsymbol{\tau}}\|_{p,2}^p \\ &\geq -uN^{1/2}\|G_N\|_2 - tu^{p/2}N. \end{aligned}$$

Rearranging and using the fact that $p > 2$ gives

$$|||\vec{\boldsymbol{\rho}}|||_{p,2}^p \leq \frac{2u^{1-p/2}\|G_N\|_2}{t\sqrt{N}} + 1 \leq \frac{2\|G_N\|_2}{K_1^{p/2}\sqrt{N}} + 1.$$

It follows by (5.6), the Cauchy-Schwarz inequality and the mean value theorem that for any $u', t' \in [K_1, K_2]$,

$$\begin{aligned} L_{N,p,u}(t) - L_{N,p,u'}(t') &\leq N^{-1}|u-u'|H_N(\vec{\boldsymbol{\rho}}) - \big(tu^{p/2} - t'u'^{\,p/2}\big)|||\vec{\boldsymbol{\rho}}|||_{p,2}^p \\ &\leq \frac{\|G_N\|_2}{\sqrt{N}}|u-u'| + |||\vec{\boldsymbol{\rho}}|||_{p,2}^p\big(K_2^{p/2}|t-t'| + pK_2^{1+p/2}K_1^{-1}|u-u'|\big) \\ &\leq M\frac{\|G_N\|_2}{\sqrt{N}}(|u-u'| + |t-t'|) \end{aligned}$$

for some constant $M > 0$ that depends only on K_1, K_2 and p. Interchanging the roles of u, u' and t, t', it is easy to see that

$$|L_{N,p,u}(t) - L_{N,p,u'}(t')| \leq M\frac{\|G_N\|_2}{\sqrt{N}}\big(|u-u'| + |t-t'|\big).$$

Invoking lemma 5.2 and redefining the constant M completes the proof. ☐

Proposition 5.4. *If $2 < p < \infty$, then almost surely*

$$L_{p,u}(t) = \lim_{N\to\infty} L_{N,p,u}(t) \tag{5.11}$$

for every $t, u > 0$.

Proof. By lemma 5.3 and a simple application of the Borel-Cantelli lemma, for any $0 < K_1 < K_2$ there exists some constant $M = M(K_1, K_2)$ such that almost surely

$$\limsup_{N\to\infty}|L_{N,p,u}(t) - L_{N,p,u'}(t')| \leq M\big(|u-u'| + |t-t'|\big) \tag{5.12}$$

for all $u, u', t, t' \in [K_1, K_2]$. Since $L_{p,u}(t)$ is a deterministic quantity, we also have

$$|L_{p,u}(t) - L_{p,u'}(t')| \leq M\big(|u-u'| + |t-t'|\big) \tag{5.13}$$

for all $u, u', t, t' \in [K_1, K_2]$. By countability of rationals and proposition 5.1, we can find a set Ω of probability one where (5.12) holds simultaneously for all rationals $K_1, K_2 \in \mathbb{Q}_+$ and at the same time $L_{p,u}(t) = \lim_{N\to\infty} L_{N,p,u}(t)$ for all $u, t \in \mathbb{Q}_+$. The triangle inequality implies that for any $u, t > 0$ and $u', t' \in \mathbb{Q}_+$,

$$\begin{aligned} |L_{N,p,u}(t) - L_{p,u}(t)| \leq |L_{N,p,u}(t) - L_{N,p,u'}(t')| + |L_{N,p,u'}(t') - L_{p,u'}(t')| \\ + |L_{p,u'}(t') - L_{p,u}(t)|. \end{aligned}$$

It follows by (5.13) that on the set Ω,

$$\limsup_{N\to\infty}|L_{N,p,u}(t) - L_{p,u}(t)| \le 2M\big(|u-u'| + |t-t'|\big).$$

Letting $u' \to u$ and $t' \to t$ along rational points completes the proof. $\square$

In addition to proposition 5.4, the proof of theorem 1.2 will rely on the fact that the $\ell^{p,2}$-norm potential in the definition of the Hamiltonian (1.22) forces the maximizers of this random function to concentrate in a large enough neighbourhood of the origin with overwhelming probability.

Lemma 5.5. *If $2 < p < \infty$, then there exist constants $C, M > 0$ such that with probability at least $1 - Ce^{-N/C}$,*

$$L_{N,p}(t) = L_{N,p,M/t}(t) \tag{5.14}$$

for all $t > 0$.

Proof. Given $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ with $H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \ge 0$, the Cauchy-Schwarz inequality implies that

$$t|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \le \frac{1}{N} H_N(\vec{\boldsymbol{\sigma}}) \le \frac{\|G_N\|_2}{\sqrt{N}} |||\vec{\boldsymbol{\sigma}}|||_{2,2}^2.$$

It follows by Jensen's inequality that

$$|||\vec{\boldsymbol{\sigma}}|||_{2,2}^p \le |||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \le \frac{\|G_N\|_2}{t\sqrt{N}} |||\vec{\boldsymbol{\sigma}}|||_{2,2}^2.$$

Since $L_{N,p}(t) \ge \frac{1}{N} H_{N,p,t}(0) = 0$, rearranging shows that

$$L_{N,p}(t) = \frac{1}{N} \max\Big\{ H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \mid |||\vec{\boldsymbol{\sigma}}|||_{2,2} \le \Big(\frac{\|G_N\|_2}{t\sqrt{N}}\Big)^{1/(p-2)} \Big\}.$$

Invoking lemma 5.2 completes the proof. $\square$

Proof (Theorem 1.2). By lemma 5.5, there exist constants $C, M > 0$ such that with probability at least $1 - Ce^{-N/C}$,

$$L_{N,p}(t) = L_{N,p,M/t}(t)$$

for any $t > 0$. It follows by a simple application of the Borel-Cantelli lemma and proposition 5.4 that with probability one,

$$L_{p,u}(t) = \lim_{N\to\infty} L_{N,p,u}(t) \le \liminf_{N\to\infty} L_{N,p}(t) \le \limsup_{N\to\infty} L_{N,p}(t) = L_{p,M/t}(t) \le \sup_{u>0} L_{p,u}(t)$$

for every $t > 0$ and $u > 0$. Taking the supremum over all $u > 0$ gives the almost sure existence of $L_p(t)$, and invoking proposition 5.1 shows that

$$L_p(t) = \sup_{u>0} L_{p,u}(t) = \sup_{D\in\Gamma_\kappa} L_{p,D}(t). \tag{5.15}$$

To establish the second equality in (1.28), fix a non-negative definite matrix $D \in \Gamma_{\kappa,u}$ as well as $0 < \epsilon < \kappa^{-2}$. By proposition 4.5, there exists a constant $K > 0$ that depends only on κ such that

$$L_{p,D}(t) \le L_{p,D+\epsilon I}(t) + K(1+tp)(u+1)^{p/2}\epsilon^{1/64}.$$

It is readily verified that $D + \epsilon I \in \Gamma_\kappa^+$, so in fact

$$L_{p,D}(t) \le \sup_{D\in\Gamma_\kappa^+} L_{p,D}(t) + K(1+tp)(u+1)^{p/2}\epsilon^{1/64}.$$

Taking the supremum over all $D \in \Gamma_{\kappa,u}$, letting $\epsilon \to 0$ and remembering (5.15) completes the proof. $\square$

6 The ground state energy in terms of the Lagrangian

In section 5 we proved the first noteworthy result of this paper by expressing the unconstrained Lagrangian (1.23) as a supremum of constrained Lagrangians (1.27) in the limit. As we will see in section 7 and section 8, the constrained Lagrangian (1.27) can be understood using the results in [33]. It is for this reason that we constrained the Lagrangian (1.23) in the first place. However, the task that we originally set ourselves is understanding the ℓ^p-Gaussian-Grothendieck problem with vector spins (1.8). In this section we connect the unconstrained Lagrangian (1.23) and the ground state energy (1.21) by proving theorem 1.3. This will reduce the ℓ^p-Gaussian-Grothendieck problem with vector spins to understanding the asymptotic behaviour of the constrained Lagrangian (1.27).

Before we proceed with the proof of theorem 1.3, we give a formal argument that will motivate the results in this section. Given $N \in \mathbb{N}$ and $t > 0$, let $\vec{\boldsymbol{\rho}}(t)$ be a point at which the Hamiltonian $H_{N,p,t}$ defined in (1.23) attains its supremum. Differentiating the expression $L_{N,p}(t) = \frac{1}{N} H_{N,p,t}(\vec{\boldsymbol{\rho}}(t))$ shows that

$$L'_{N,p}(t) = \frac{1}{N}\partial_t H_{N,p,t}(\vec{\boldsymbol{\rho}}(t)) + \frac{1}{N}\big(\vec{\boldsymbol{\rho}}'(t), \nabla_{\vec{\boldsymbol{\sigma}}} H_{N,p,t}(\vec{\boldsymbol{\rho}}(t))\big) = -|||\vec{\boldsymbol{\rho}}(t)|||_{p,2}^p. \tag{6.1}$$

We have used the fact that $\nabla_{\vec{\boldsymbol{\sigma}}} H_{N,p,t}(\vec{\boldsymbol{\rho}}(t)) = 0$. This suggests that

$$L_{N,p}(t) = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p = -L'_{N,p}(t)} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p = -L'_{N,p}(t)} H_N(\vec{\boldsymbol{\sigma}}) + tL'_{N,p}(t), \tag{6.2}$$

and therefore

$$\mathbf{GSE}_{N,p} = \frac{1}{N} \max_{|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p = -L'_{N,p}(t)} H_N\big((-L'_{N,p}(t))^{-1/p}\vec{\boldsymbol{\sigma}}\big) = \frac{L_{N,p}(t) - tL'_{N,p}(t)}{(-L'_{N,p}(t))^{2/p}}. \tag{6.3}$$

To express this ground state energy entirely in terms of the unconstrained Lagrangian (1.23) as in theorem 1.3, we compute the gradient of the Hamiltonian (1.23). Since our calculation will be rigorous, we formulate it as a lemma.

Lemma 6.1. *If $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ and $t, u > 0$, then, conditionally on the disorder chaos (g_{ij}),*

$$\big(\nabla_{\vec{\boldsymbol{\sigma}}} H_{N,p,t}(\vec{\boldsymbol{\sigma}}), \vec{\boldsymbol{\sigma}}\big) = 2H_N(\vec{\boldsymbol{\sigma}}) - tp\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p. \tag{6.4}$$

Proof. Given $1 \le i \le N$ and $1 \le k \le \kappa$, a simple computation shows that

$$\frac{\partial H_{N,p,t}(\vec{\boldsymbol{\sigma}})}{\partial \sigma_i(k)} = \frac{1}{\sqrt{N}} \sum_{j=1}^N (g_{ij} + g_{ji})\sigma_j(k) - tp\sigma_i(k)\|\vec{\sigma}_i\|_2^{p-2}.$$

It follows that

$$\begin{aligned}\big(\nabla_{\vec{\boldsymbol{\sigma}}} H_{N,p,t}(\vec{\boldsymbol{\sigma}}), \vec{\boldsymbol{\sigma}}\big) &= \sum_{k=1}^\kappa \frac{1}{\sqrt{N}} \sum_{i,j=1}^N (g_{ij} + g_{ji})\sigma_j(k)\sigma_i(k) - tp\sum_{i=1}^N \sum_{k=1}^\kappa \|\vec{\sigma}_i\|_2^{p-2}\sigma_i(k)^2 \\ &= 2\sum_{k=1}^\kappa H_N^k(\boldsymbol{\sigma}(k)) - tp\sum_{i=1}^N \|\vec{\sigma}_i\|_2^p = 2H_N(\vec{\boldsymbol{\sigma}}) - tp\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p.\end{aligned}$$

This finishes the proof. □

This simple calculation suggests that

$$0 = \big(\nabla_{\vec{\boldsymbol{\sigma}}} H_{N,p,t}(\vec{\boldsymbol{\rho}}(t)), \vec{\boldsymbol{\rho}}(t)\big) = 2H_N(\vec{\boldsymbol{\rho}}(t)) - tp\|\vec{\boldsymbol{\rho}}(t)\|_{p,2}^p, \tag{6.5}$$

which combined with (6.1) gives

$$L_{N,p}(t) = t\Big(\frac{p}{2}-1\Big)|||\vec{\boldsymbol{\rho}}(t)|||_{p,2}^p = -t\Big(\frac{p}{2}-1\Big)L'_{N,p}(t). \tag{6.6}$$

Substituting this into (6.3) gives (1.29) upon letting $N \to \infty$. The problem with this argument is that the map $t \mapsto \vec{\boldsymbol{\rho}}(t)$ might not be differentiable. To overcome this issue, we will prove (6.6) directly at the points of differentiability of $L_{N,p}(t)$. We will then use a convexity argument to deduce that (6.6) holds for every $t > 0$ in the limit.

Lemma 6.2. *If (g_{ij}) is a realization of the disorder chaos for which the unconstrained Lagrangian $L_{N,p}$ is differentiable at $t > 0$, then*

$$L_{N,p}(t) = -t\Big(\frac{p}{2}-1\Big)L'_{N,p}(t). \tag{6.7}$$

Proof. Fix $\epsilon > 0$ and $\lambda > 0$. For any configuration with $|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \geq -L'_{N,p}(t) + \epsilon$,

$$\begin{aligned}
\frac{1}{N}H_{N,p,t}(\vec{\boldsymbol{\sigma}}) &\leq \frac{1}{N}H_{N,p,t}(\vec{\boldsymbol{\sigma}}) + \lambda\big(|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p + L'_{N,p}(t) - \epsilon\big)\\
&\leq L_{N,p}(t-\lambda) + \lambda L'_{N,p}(t) - \lambda\epsilon\\
&= \lambda\Big(L'_{N,p}(t) - \frac{L_{N,p}(t) - L_{N,p}(t-\lambda)}{\lambda}\Big) - \lambda\epsilon + L_{N,p}(t).
\end{aligned}$$

Similarly, for any configuration with $|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p \leq -L'_{N,p}(t) - \epsilon$,

$$\begin{aligned}
\frac{1}{N}H_{N,p,t}(\vec{\boldsymbol{\sigma}}) &\leq \frac{1}{N}H_{N,p,t}(\vec{\boldsymbol{\sigma}}) + \lambda\big(-|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p - L'_{N,p}(t) - \epsilon\big)\\
&\leq L_{N,p}(t+\lambda) - \lambda L'_{N,p}(t) - \lambda\epsilon\\
&= \lambda\Big(\frac{L_{N,p}(t+\lambda) - L_{N,p}(t)}{\lambda} - L'_{N,p}(t)\Big) - \lambda\epsilon + L_{N,p}(t).
\end{aligned}$$

The differentiability of $L_{N,p}$ at t gives $\lambda = \lambda(\epsilon) > 0$ small enough so that

$$\frac{1}{N}\max_{||||\vec{\boldsymbol{\sigma}}|||_{p,2}^p + L'_{N,p}(t)| \geq \epsilon} H_{N,p,t}(\vec{\boldsymbol{\sigma}}) \leq L_{N,p}(t) - \frac{\lambda\epsilon}{2}.$$

This means that an optimizer $\vec{\boldsymbol{\rho}}(t)$ of $L_{N,p}(t)$ satisfies $||||\vec{\boldsymbol{\rho}}(t)|||_p^p + L'_{N,p}(t)| < \epsilon$ for every $\epsilon > 0$. Letting $\epsilon \to 0$ reveals that $|||\vec{\boldsymbol{\rho}}(t)|||_{p,2}^p = -L'_{N,p}(t)$. It follows by lemma 6.1 that

$$L_{N,p}(t) - tL'_{N,p}(t) = \frac{1}{N}H_N(\vec{\boldsymbol{\rho}}(t)) = \frac{tp}{2}|||\vec{\boldsymbol{\rho}}(t)|||_{p,2}^p = -\frac{tp}{2}L'_{N,p}(t).$$

Rearranging completes the proof. ☐

Lemma 6.3. *If $2 < p < \infty$, then the function $L_p(t)$ is differentiable on $(0, \infty)$ with*

$$L_p(t) = -t\Big(\frac{p}{2}-1\Big)L'_p(t). \tag{6.8}$$

Proof. Using theorem 1.2, fix a realization (g_{ij}) of the disorder chaos for which $L_{N,p}(t)$ converges to $L_p(t)$ for all $t > 0$. Notice that $L_{N,p}$ and L_p are convex functions. In particular, they are continuous everywhere on $(0, \infty)$ and differentiable almost everywhere on $(0, \infty)$. If $0 < t_1 < s < t_2$ are such that $L_{N,p}$ is differentiable at s, then the convexity of $L_{N,p}$ gives

$$\frac{L_{N,p}(s) - L_{N,p}(t_1)}{s - t_1} \leq L'_{N,p}(s) \leq \frac{L_{N,p}(t_2) - L_{N,p}(s)}{t_2 - s},$$

and lemma 6.2 yields

$$\frac{L_{N,p}(s) - L_{N,p}(t_1)}{s - t_1} \le -\frac{L_{N,p}(s)}{(\frac{p}{2}-1)s} \le \frac{L_{N,p}(t_2) - L_{N,p}(s)}{t_2 - s}.$$

By continuity of $L_{N,p}$ and density of the points of differentiability of $L_{N,p}$ in $(0,\infty)$, this inequality implies that for all $0 < t_1 < t < t_2 < \infty$,

$$\frac{L_{N,p}(t) - L_{N,p}(t_1)}{t - t_1} \le -\frac{L_{N,p}(t)}{(\frac{p}{2}-1)t} \le \frac{L_{N,p}(t_2) - L_{N,p}(t)}{t_2 - t}.$$

Letting $N \to \infty$ and then letting $t_1 \nearrow t$ and $t_2 \searrow t$ shows that at any point $t \in (0,\infty)$ of differentiability of L_p,

$$L_p'(t) = -\frac{L_p(t)}{(\frac{p}{2}-1)t}. \tag{6.9}$$

We will now use this equality to show that L_p is differentiable everywhere on $(0,\infty)$. By convexity of L_p and theorem 25.1 in [34], it suffices to prove that the sub-differential $\partial L_p(t)$ consists of a single point for every $t > 0$. Fix $t \in (0,\infty)$ as well as $a \in \partial L_p(t)$, and let (s_k) and (t_k) be points of differentiability of L_p with $t_k \nearrow t$ and $s_k \searrow t$. By definition of the sub-differential,

$$L_p'(t_k) \le \frac{L_p(t) - L_p(t_k)}{t - t_k} \le a \le \frac{L_p(s_k) - L_p(t)}{s_k - t} \le L_p'(s_k)$$

for every integer $k \ge 1$. Letting $k \to \infty$ and combining (6.9) with the continuity of L_p yields

$$-\frac{L_p(t)}{(\frac{p}{2}-1)t} = \limsup_{k\to\infty} L_p'(t_k) \le a \le \liminf_{k\to\infty} L_p'(s_k) = -\frac{L_p(t)}{(\frac{p}{2}-1)t}.$$

This completes the proof. $\square$

To leverage this result into a proof of theorem 1.3, we must verify the legitimacy of the change of variables used in (6.3). In other words, we must show that $L_p'(t)$ does not vanish on $(0,\infty)$. Our proof will rely upon the properties of the eigenvalues and eigenvectors of the Gaussian orthogonal ensemble discussed in chapter 2 of [2]. Recall the definition of the random matrix G_N in (5.7), and notice that the $N \times N$ random matrix

$$\bar{G}_N = \frac{G_N + G_N^T}{\sqrt{2}} \tag{6.10}$$

is distributed according to the Gaussian orthogonal ensemble.

Lemma 6.4. *If $2 < p < \infty$, then the function L_p is strictly positive on $(0,\infty)$. In particular, $L_p'(t) < 0$ for every $t > 0$.*

Proof. Given $\boldsymbol{\sigma} \in \mathbb{R}^N$, consider the vector spin configuration $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^N$ defined by

$$\vec{\boldsymbol{\sigma}}(k) = \begin{cases} \boldsymbol{\sigma} & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p = \|\boldsymbol{\sigma}\|_p^p = 1$ and $\sum_{i,j=1}^N g_{ij}\sigma_i\sigma_j = \sum_{i,j=1}^N g_{ij}\big(\vec{\sigma}_i, \vec{\sigma}_j\big)$. It follows that

$$L_{N,p}(t) \ge \frac{1}{N}\big(H_N(\vec{\boldsymbol{\sigma}}) - t\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p\big) = \frac{\big(\bar{G}_N\boldsymbol{\sigma}, \boldsymbol{\sigma}\big)}{\sqrt{2}N^{3/2}} - |||\boldsymbol{\sigma}|||_p^p. \tag{6.11}$$

With this in mind, let v denote the ℓ^2-normalized eigenvector associated with the largest eigenvalue λ_N^N of the Gaussian orthogonal ensemble $\bar{G}_N$. Given $\delta > 0$, applying (6.11) to the spin configuration $\boldsymbol{\sigma}_\delta = \sqrt{N\delta}v$ reveals that

$$L_{N,p}(t) \geq \frac{\big(\bar{G}_N\boldsymbol{\sigma}_\delta, \boldsymbol{\sigma}_\delta\big)}{\sqrt{2}N^{3/2}} - t|||\boldsymbol{\sigma}_\delta|||_p^p = \frac{\delta}{\sqrt{2}}\frac{\lambda_N^N}{\sqrt{N}} - t\delta^{p/2}N^{p/2-1}\|v\|_p^p.$$

By corollary 2.5.4 in [2], the eigenvector v is equal in distribution to $g/\|g\|_2$ for a standard Gaussian random vector g in $\mathbb{R}^N$. Moreover, by the strong law of large numbers,

$$N^{\frac{p}{2}-1}\frac{\|g\|_p^p}{\|g\|_2^p} = \frac{\frac{1}{N}\sum_{i\leq N}|g_i|^p}{(\frac{1}{N}\sum_{i\leq N}|g_i|^2)^{p/2}} \longrightarrow \frac{\mathbb{E}|g_1|^p}{(\mathbb{E}|g_1|^2)^{p/2}} = \mathbb{E}|g_1|^p$$

almost surely. Together with the asymptotics of λ_N^N established in theorem 2.1.22 of [2], this implies that

$$L_p(t) \geq \sqrt{2}\delta - t\delta^{p/2}\,\mathbb{E}|g_1|^p.$$

Taking $\delta > 0$ small enough and using the fact that $p > 2$ shows that L_p is strictly positive on $(0,\infty)$. Invoking lemma 6.3 completes the proof. ☐

Proof (Theorem 1.3). Using theorem 1.2, fix a realization (g_{ij}) of the disorder chaos for which $L_{N,p}(t)$ converges to $L_p(t)$ for all $t > 0$. Let $\Omega \subset (0,\infty)$ be the collection of points at which $L_{N,p}$ is differentiable for all $N \geq 1$. Fix $t \in \Omega$, and notice that by convexity of $L_{N,p}$,

$$L_{N,p}(t+h) \geq L_{N,p}(t) + L'_{N,p}(t)h \tag{6.12}$$

for every $h \in \mathbb{R}$. By lemma 6.2, the sequence $(L'_{N,p}(t))_N$ is uniformly bounded. It therefore admits a subsequential limit a. Letting $N \to \infty$ in (6.12) shows that a belongs to the sub-differential $\partial L_p(t)$. Invoking lemma 6.3 shows that $a = L'_p(t)$, and therefore $L'_{N,p}(t) \to L'_p(t)$. It follows by lemma 6.4 that $L_{N,p}(t) < 0$ for large enough N, so

$$\mathbf{GSE}_{N,p} = \frac{1}{N}\max_{|||\vec{\boldsymbol{\sigma}}|||_{p,2}^p = -L'_{N,p}(t)} H_N\big((-L'_{N,p}(t))^{-1/p}\vec{\boldsymbol{\sigma}}\big) = \frac{L_{N,p}(t) - tL'_{N,p}(t)}{(-L'_{N,p}(t))^{2/p}}.$$

Since Ω is dense in $(0,\infty)$ and $L'_{N,p}$ is continuous, this equality extends to all $t > 0$. Letting $N \to \infty$ and using lemma 6.3 completes the proof. ☐

7 Replacing the constrained Lagrangian by a free energy

So far, we have reduced the ℓ^p-Gaussian-Grothendieck problem with vector spins to understanding the asymptotic behaviour of the constrained Lagrangian (1.27) with positive definite constraints. This task will occupy the remainder of the paper. The starting point of our analysis will be the Parisi-type variational formula for free energy functionals established in [33]. To access this result, we must first replace the constrained Lagrangian by a free energy functional. In this section, given a constraint $D \in \Gamma_\kappa$ which is fixed throughout, we introduce a free energy functional that depends on an inverse temperature parameter $\beta > 0$ and is asymptotically equivalent to the constrained Lagrangian (1.27) upon letting $\beta \to \infty$.

For each inverse temperature parameter $\beta > 0$ and every $\epsilon > 0$, consider the free energy

$$\tilde{F}_{N,\epsilon}(\beta) = \frac{1}{\beta N}\log\int_{\Sigma_\epsilon(D)}\exp\beta H_{N,p,t}(\vec{\boldsymbol{\sigma}})\,\mathrm{d}\,\vec{\boldsymbol{\sigma}} \tag{7.1}$$

and the quenched free energy

$$F_{N,\epsilon}(\beta) = \frac{1}{\beta N}\,\mathbb{E}\log\int_{\Sigma_\epsilon(D)}\exp\beta H_{N,p,t}(\vec{\boldsymbol{\sigma}})\,\mathrm{d}\,\vec{\boldsymbol{\sigma}}. \tag{7.2}$$

Recall the definition of the relaxed constrained Lagrangian in (4.15). Since the L^q-norm converges to the L^∞-norm, it is clear that

$$\lim_{\epsilon\to 0}\lim_{N\to\infty}\lim_{\beta\to\infty} F_{N,\epsilon}(\beta) = \lim_{\epsilon\to 0}\lim_{N\to\infty} \mathbb{E}\, L_{N,p,D,\epsilon}(t). \tag{7.3}$$

We will now use the continuity result in proposition 4.4 to show that the right-hand side of this equation coincides with the limit of the constrained Lagrangian (1.27). Subsequently, we will prove that (7.3) still holds if the limit in β is taken after the limits in ϵ and N. The benefit of exchanging these limits is that the main result in [33] gives a Parisi-type variational formula for the limit in ϵ and N of the quenched free energy (7.2) for each fixed $\beta > 0$. In section 9 and section 10 we will study this formula in the limit $\beta \to \infty$ to finally prove theorem 1.4 in section 11.

Proposition 7.1. *If $2 < p < \infty$ and $t > 0$, then*

$$\lim_{\epsilon\to 0}\lim_{N\to\infty} L_{N,p,D,\epsilon}(t) = \lim_{\epsilon\to 0}\lim_{N\to\infty} \mathbb{E}\, L_{N,p,D,\epsilon}(t) = L_{p,D}(t) \tag{7.4}$$

almost surely.

Proof. By the Gaussian concentration of the maximum, it suffices to prove that

$$\lim_{\epsilon\to 0}\lim_{N\to\infty} \mathbb{E}\, L_{N,p,D,\epsilon}(t) = L_{p,D}(t). \tag{7.5}$$

Given $0 < \epsilon < \kappa^{-2}$ smaller than the smallest non-zero eigenvalue of D, the equality $D_\epsilon = D$ and proposition 4.4 imply that

$$\lim_{\epsilon\to 0}\limsup_{N\to\infty} \mathbb{E}\, L_{N,p,D,\epsilon}(t) \leq L_{p,D}(t). \tag{7.6}$$

On the other hand, the Gaussian concentration of the maximum reveals that for every $\epsilon > 0$,

$$L_{p,D}(t) = \lim_{N\to\infty} L_{N,p,D}(t) \leq \liminf_{N\to\infty} L_{N,p,D,\epsilon}(t) = \liminf_{N\to\infty} \mathbb{E}\, L_{N,p,D,\epsilon}(t).$$

Letting $\epsilon \to 0$ and remembering (7.6) establishes (7.5) and completes the proof. $\square$

Lemma 7.2. *If $2 < p < \infty$ and $t > 0$, then*

$$\limsup_{\beta\to\infty}\lim_{\epsilon\to 0}\lim_{N\to\infty} F_{N,\epsilon}(\beta) \leq L_{p,D}(t). \tag{7.7}$$

Proof. Fix $\delta \in (0,t)$, and for each $t > 0$ let $\vec{\boldsymbol{\rho}}_t$ be a maximizer of the relaxed constrained Lagrangian (4.15). By Fubini-Tonelli and a change of variables,

$$\begin{aligned}\tilde{F}_{N,\epsilon}(\beta) &\leq L_{N,p,D,\epsilon}(t-\delta) + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa} e^{-\beta\delta\|\vec{\sigma}\|_2^p}\,\mathrm{d}\,\vec{\sigma} \\ &= \frac{1}{N}H_{N,p,t}(\vec{\boldsymbol{\rho}}_{t-\delta}) + \delta|||\vec{\boldsymbol{\rho}}_{t-\delta}|||_{p,2}^p - \frac{\kappa\log\beta\delta}{p\beta} + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p}\,\mathrm{d}\,\vec{\sigma} \\ &\leq L_{N,p,D,\epsilon}(t) + \delta|||\vec{\boldsymbol{\rho}}_{t-\delta}|||_{p,2}^p - \frac{\kappa\log\beta\delta}{p\beta} + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p}\,\mathrm{d}\,\vec{\sigma}. \end{aligned} \tag{7.8}$$

To bound this further, let $A \in \mathbb{R}^{\kappa\times\kappa}$ be a symmetric and non-negative definite matrix with $AA^T = \kappa D$, and denote by $\vec{\sigma}_i \in \mathbb{R}^\kappa$ the i'th column of A. Consider the subsequence $M = N\kappa$, and define the κ-periodic vector spin configuration $\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^M$ by $\vec{\sigma}_j = \vec{\sigma}_i$ whenever $j \equiv i \bmod \kappa$. From (1.18), it is clear that $\vec{\boldsymbol{\sigma}} \in \Sigma(D)$. Indeed,

$$R(\vec{\boldsymbol{\sigma}},\vec{\boldsymbol{\sigma}}) = \frac{1}{M}\sum_{i=1}^M \vec{\sigma}_i\vec{\sigma}_i^T = \frac{1}{\kappa}\sum_{i=1}^\kappa \vec{\sigma}_i\vec{\sigma}_i^T = \frac{1}{\kappa}AA^T = D.$$

If G_M denotes the $M \times M$ random matrix in (5.7), then the Cauchy-Schwarz inequality implies that for each $t > 0$,

$$H_{M,p,t}(\vec{\boldsymbol{\rho}}_t) \leq \sqrt{M}\|G_M\|_2 |||\vec{\boldsymbol{\rho}}_t|||_{2,2}^2 - t\|\vec{\boldsymbol{\rho}}_t\|_{p,2}^p,$$

and similarly,

$$H_{M,p,t}(\vec{\boldsymbol{\rho}}_t) \geq H_{M,p,t}(\vec{\boldsymbol{\sigma}}) \geq -\sqrt{M}\|G_M\|_2 |||\vec{\boldsymbol{\sigma}}|||_{2,2}^2 - t\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p.$$

Rearranging and remembering (4.9) gives

$$\begin{aligned} |||\vec{\boldsymbol{\rho}}_t|||_{p,2}^p &\leq \frac{2\|G_M\|_2(\mathrm{tr}(D) + \epsilon\kappa)}{t\sqrt{M}} + \max_{1\leq i\leq\kappa} \|\vec{\sigma}_i\|_2^p \\ &= \frac{2(\mathrm{tr}(D) + \epsilon\kappa)\lambda_M^M}{\sqrt{2M}t} + \max_{1\leq i\leq\kappa} \|\vec{\sigma}_i\|_2^p, \end{aligned} \tag{7.9}$$

where λ_M^M denotes the largest eigenvalue of the Gaussian orthogonal ensemble $\bar{G}_M$ in (6.10). Substituting this into (7.8), appealing to the Gaussian concentration of the free energy and leveraging the asymptotics of λ_M^M established in theorem 2.1.22 of [2] shows that

$$\begin{aligned} \lim_{N\to\infty} F_{N,\epsilon}(\beta) \leq \lim_{N\to\infty} L_{N,p,D,\epsilon}(t) + \delta\Big(\frac{2\sqrt{2}(\mathrm{tr}(D) + \epsilon\kappa)}{t - \delta} + \max_{1\leq i\leq\kappa} \|\vec{\sigma}_i\|_2^p\Big) \\ - \frac{\kappa \log \beta\delta}{p\beta} + \frac{1}{\beta} \log \int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma}. \end{aligned}$$

We have implicitly used the fact that the limit of $F_{N,\epsilon}(\beta)$ exists and therefore coincides with that of $F_{M,\epsilon}(\beta)$. This can be shown using a Guerra-Toninelli argument as in theorem 3.1, or by appealing to the results in [33] as we will do in section 8. Letting $\epsilon \to 0$, then $\beta \to \infty$ and finally $\delta \to 0$ completes the proof upon invoking proposition 7.1. □

Theorem 7.3. *If $2 < p < \infty$ and $t > 0$, then*

$$\limsup_{\beta\to\infty} \lim_{\epsilon\to 0} \lim_{N\to\infty} F_{N,\epsilon}(\beta) \leq L_{p,D}(t) \leq \liminf_{\beta\to\infty} \lim_{N\to\infty} F_{N,\beta^{-1}}(\beta). \tag{7.10}$$

Proof. By lemma 7.2, it suffices to prove the upper bound in (7.10). Fix $\epsilon \in (0,1)$, and let $\delta = \epsilon/K$ for a large enough $K > 0$ to be determined. Consider the subsequence $M = N\kappa$ as in the proof of lemma 7.2, and let $\vec{\boldsymbol{\rho}} \in \Sigma_\delta(D)$ be a maximizer of the relaxed constrained Lagrangian $L_{M,p,D,\delta}(t)$ in (4.15). Introduce the $\delta/\sqrt{\kappa}$-neighbourhood,

$$\mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}}) = \vec{\boldsymbol{\rho}} + [-\delta/\sqrt{\kappa}, \delta/\sqrt{\kappa}]^{M\kappa} \subset \big\{\vec{\boldsymbol{\sigma}} \in (\mathbb{R}^\kappa)^M \mid |||\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\rho}}|||_{2,2} \leq \delta\big\},$$

and observe that $\mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}}) \subset \Sigma_\epsilon(D)$. Indeed, the same argument used to obtain (4.6) implies that for any $\vec{\boldsymbol{\sigma}} \in \mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}})$,

$$\begin{aligned} \|R(\vec{\boldsymbol{\sigma}}, \vec{\boldsymbol{\sigma}}) - R(\vec{\boldsymbol{\rho}}, \vec{\boldsymbol{\rho}})\|_\infty &\leq |||\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\rho}}|||_{2,2}\big(|||\vec{\boldsymbol{\sigma}}|||_{2,2} + |||\vec{\boldsymbol{\rho}}|||_{2,2}\big) \leq \delta\big(1 + 2\sqrt{\mathrm{tr}(D) + \kappa}\big) \\ &< \frac{\epsilon}{2} \end{aligned}$$

provided that $K = K(D,\kappa)$ is large enough. The second inequality uses (4.9). This means that

$$\begin{aligned} \tilde{F}_{M,\epsilon}(\beta) &\geq \frac{1}{\beta M} \int_{\mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}})} \exp \beta H_{M,p,t}(\vec{\boldsymbol{\sigma}}) \,\mathrm{d}\,\vec{\boldsymbol{\sigma}} \\ &\geq L_{M,p,D,\delta}(t) + \frac{1}{M} \inf_{\vec{\boldsymbol{\sigma}} \in \mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}})} \big(H_{M,p,t}(\vec{\boldsymbol{\sigma}}) - H_{M,p,t}(\vec{\boldsymbol{\rho}})\big) + \frac{\kappa}{\beta} \log \frac{2\delta}{\sqrt{\kappa}}. \end{aligned} \tag{7.11}$$

To bound this further, fix $\vec{\boldsymbol{\sigma}} \in \mathcal{C}_{\delta/\sqrt{\kappa}}(\vec{\boldsymbol{\rho}})$ and recall the definition of the $M \times M$ random matrix G_M in (5.7). The Cauchy-Schwarz inequality implies that

$$\begin{aligned} H_N(\vec{\boldsymbol{\sigma}}) - H_N(\vec{\boldsymbol{\rho}}) &= \frac{1}{2\sqrt{M}} \sum_{k=1}^{\kappa} \big((G_M + G_M^T)(\boldsymbol{\sigma}(k) - \boldsymbol{\rho}(k)), \boldsymbol{\sigma}(k) + \boldsymbol{\rho}(k)\big) \\ &\geq -\frac{\|G_M\|_2}{\sqrt{M}} \sum_{k=1}^{\kappa} \|\boldsymbol{\sigma}(k) - \boldsymbol{\rho}(k)\|_2 \big(\|\boldsymbol{\sigma}(k)\|_2 + \|\boldsymbol{\rho}(k)\|_2\big) \\ &\geq -\sqrt{M}\|G_M\|_2 |||\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\rho}}|||_{2,2}(|||\vec{\boldsymbol{\sigma}}|||_{2,2} + |||\vec{\boldsymbol{\rho}}|||_{2,2}) \\ &\geq -M\delta \frac{\|G_M\|_2}{\sqrt{M}}(1 + 2\sqrt{\operatorname{tr}(D) + \kappa}). \end{aligned} \tag{7.12}$$

On the other hand, an identical argument as that used to obtain (4.21) shows that

$$\|\vec{\boldsymbol{\sigma}}\|_{p,2}^p - \|\vec{\boldsymbol{\rho}}\|_{p,2}^p \leq Mp|||\vec{\boldsymbol{\sigma}} - \vec{\boldsymbol{\rho}}|||_{p,2}|||\vec{\boldsymbol{\sigma}}|||_{p,2}^{p-1} \leq Mp\delta\big(1 + \|\vec{\boldsymbol{\rho}}\|_{p,2}\big)^{p-1}.$$

Together with (7.12), (7.9) and lemma 5.2, this gives constants $C, K' > 0$ that depend only on κ, D, p and t such that with probability at least $1 - Ce^{-M/C}$,

$$H_{M,p,t}(\vec{\boldsymbol{\sigma}}) - H_{M,p,t}(\vec{\boldsymbol{\rho}}) \geq -M\delta K'.$$

Substituting this lower bound into (7.11) and combining the Gaussian concentration of the free energy with the Borel-Cantelli lemma to let $N \to \infty$ yields

$$\lim_{N\to\infty} F_{N,\epsilon}(\beta) \geq \lim_{N\to\infty} L_{N,p,D,\epsilon/K}(t) - \epsilon K^{-1}K' + \frac{\kappa}{\beta}\log\frac{2\epsilon}{K\sqrt{\kappa}}.$$

Taking $\epsilon = \beta^{-1}$ and letting $\beta \to \infty$ completes the proof upon invoking proposition 7.1. □

8 The limit of the free energy

In this section we describe the implications of the main result in [33] on the asymptotic representation of the constrained Lagrangian (1.27) established in theorem 7.3. Given a constraint $D \in \Gamma_\kappa$, some $t > 0$ and an inverse temperature parameter $\beta > 0$, all of which will remain fixed throughout this section, consider the measure on $\mathbb{R}^\kappa$ defined by

$$\mathrm{d}\,\mu(\vec{\sigma}) = \exp\big(-t\beta\|\vec{\sigma}\|_2^p\big)\,\mathrm{d}\,\vec{\sigma}. \tag{8.1}$$

Notice that the quenched free energy (7.2) may be written as

$$F_{N,\epsilon}(\beta) = \frac{1}{\beta N}\,\mathbb{E}\log\int_{\Sigma_\epsilon(D)} \exp\beta H_N(\vec{\boldsymbol{\sigma}})\,\mathrm{d}\,\mu^{\otimes N}(\vec{\boldsymbol{\sigma}}). \tag{8.2}$$

If it were not for the fact that the measure μ in (8.1) is not compactly supported, this free energy functional would fall into the class of free energy functionals studied in [33]. Fortunately, the compact support assumption in [33] is not necessary. Instead, it is a convenient assumption that ensures all objects introduced are well-defined and spin configurations in the set $\Sigma_\epsilon(D)$ remain bounded. Replicating the arguments in [33], it is not hard to use that the measure (8.1) exhibits super-Gaussian decay in the range $2 < p < \infty$ to show that the analogue of the Parisi formula with vector spins proved in [33] for compactly supported measures also holds for the free energy functional (8.2). We will not give any more details than this, and simply content ourselves with stating the asymptotic formula for (8.2) which we will use between section 9 and section 11 to prove theorem 1.4.

Denote by $\mathcal{M}^d$ the set of probability distributions on $[0,1]$ with finitely many atoms. Notice that any discrete measure $\alpha \in \mathcal{M}^d$ may be identified with two sequences of parameters

$$0 = q_{-1} \leq q_0 \leq \ldots \leq q_{r-1} \leq q_r = 1, \tag{8.3}$$

$$0 = \alpha_{-1} \leq \alpha_0 \leq \ldots \leq \alpha_{r-1} \leq \alpha_r = 1, \tag{8.4}$$

satisfying $\alpha(\{q_j\}) = \alpha_j - \alpha_{j-1}$ for $0 \leq j \leq r$. For each Lagrange multiplier $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$, define the function $f_\lambda^\beta : \mathbb{R}^\kappa \to \mathbb{R}$ by

$$f_\lambda^\beta(\vec{x}) = \frac{1}{\beta} \log \int_{\mathbb{R}^\kappa} \exp \beta \Big((\vec{\sigma}, \vec{x}) + \sum_{k \leq k'} \lambda_{k,k'} \sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p \Big) \, \mathrm{d}\, \vec{\sigma}. \tag{8.5}$$

Given a path $\pi \in \Pi_D$ defined by the sequences (1.33) and (1.34), for each $0 \leq j \leq r$ consider an independent Gaussian vector $z_j = (z_j(k))_{k \leq \kappa}$ with covariance structure (1.39). Define the sequence $(X_l^{\lambda,\alpha,\pi,\beta})_{0 \leq l \leq r}$ recursively as follows. Let

$$X_r^{\lambda,\alpha,\pi,\beta}((z_j)_{0 \leq j \leq r}) = f_\lambda^\beta\Big(\sqrt{2} \sum_{j=1}^r z_j\Big), \tag{8.6}$$

and for $0 \leq l \leq r-1$ let

$$X_l^{\lambda,\alpha,\pi,\beta}((z_j)_{0 \leq j \leq l}) = \frac{1}{\beta \alpha_l} \log \mathbb{E}_{z_{l+1}} \exp \beta \alpha_l X_{l+1}^{\lambda,\alpha,\pi,\beta}((z_j)_{0 \leq j \leq l+1}). \tag{8.7}$$

This inductive procedure is well-defined by the growth bounds in lemma A.1. Introduce the Parisi functional,

$$\mathcal{P}_\beta(\lambda,\alpha,\pi) = X_0^{\lambda,\alpha,\pi,\beta} - \sum_{k \leq k'} \lambda_{k,k'} D_{k,k'} - \frac{\beta}{2} \sum_{0 \leq j \leq r-1} \alpha_j \big(\|\gamma_{j+1}\|_{\mathrm{HS}}^2 - \|\gamma_j\|_{\mathrm{HS}}^2 \big). \tag{8.8}$$

Observe that

$$\begin{aligned} \sum_{0 \leq j \leq r-1} \alpha_j \big(\|\gamma_{j+1}\|_{\mathrm{HS}}^2 - \|\gamma_j\|_{\mathrm{HS}}^2 \big) &= \sum_{0 \leq j \leq r-1} \int_{q_j}^{q_{j+1}} \alpha(s) \frac{\mathrm{d}}{\mathrm{d}\, s} \|\pi(s)\|_{\mathrm{HS}}^2 \, \mathrm{d}\, s \\ &= 2 \int_0^1 \alpha(s) \, \mathrm{Sum}\, \big(\pi(s) \odot \pi'(s)\big) \, \mathrm{d}\, s, \end{aligned} \tag{8.9}$$

where we have abused notation by writing α both for the measure and its cumulative distribution function. The Parisi functional may therefore be expressed as

$$\mathcal{P}_\beta(\lambda,\alpha,\pi) = X_0^{\lambda,\alpha,\pi,\beta} - \sum_{k \leq k'} \lambda_{k,k'} D_{k,k'} - \int_0^1 \beta \alpha(s) \, \mathrm{Sum}\, \big(\pi(s) \odot \pi'(s)\big) \, \mathrm{d}\, s. \tag{8.10}$$

We have made all dependencies on D, p and t implicit for clarity of notation, but we will make them explicit whenever necessary. The proof of theorem 1.4 will leverage the following consequence of the main result in [33].

Theorem 8.1. *If $2 < p < \infty$, then*

$$\lim_{\epsilon \to 0} \lim_{N \to \infty} F_{N,\epsilon}(\beta) = \inf_{\lambda,\alpha,\pi} \mathcal{P}_\beta(\lambda,\alpha,\pi), \tag{8.11}$$

where the infimum is taken over all $(\lambda,\alpha,\pi) \in \mathbb{R}^{\kappa(\kappa+1)/2} \times \mathcal{M}^d \times \Pi_D$.

This result can be viewed as a positive temperature analogue of theorem 1.4. Together with theorem 7.3, it essentially reduces the proof of theorem 1.4 to showing that

$$\lim_{\beta\to\infty} \inf_{\lambda,\alpha,\pi} \mathcal{P}_\beta(\lambda,\alpha,\pi) = \inf_{\lambda,\zeta,\pi} \mathcal{P}_\infty(\lambda,\zeta,\pi). \tag{8.12}$$

Notice the similarity between the Parisi functionals (8.10) and (1.42). If it were not for the terms X_0 and Y_0 in (8.7) and (1.41), there would be a natural correspondence between these two functionals obtained by setting $\zeta = \beta\alpha$. The proof of (8.12) will therefore consist in showing that, when evaluated at almost minimizers, the terms X_0 and Y_0 in (8.7) and (1.41) differ by a quantity that vanishes as $\beta \to \infty$. To control this difference, we will compare the terminal conditions (8.5) and (1.36) in section 9. We will then use the Auffinger-Chen representation [4, 16] in section 10 to translate the bounds on the terminal conditions into control on X_0 and Y_0. This analysis will be exploited in section 11 to establish (8.12) and therefore prove theorem 1.4. This strategy is considerably different to that in [6], where the free energy functional (8.2) is truncated at some level $M > 0$. This truncation simplifies much of the analysis for fixed $M > 0$, but requires a lot of care when sending $M \to \infty$. By not truncating the free energy, we simplify and shorten the proof of theorem 1.4 even in the scalar case, $\kappa = 1$, studied in [6].

9 Comparison of the terminal conditions

In this section we prove quantitative bounds on the difference between the terminal conditions (8.5) and (1.36). Although the analysis in this section uses only elementary concepts, it is the key to proving theorem 1.4; the rest of the paper will use tools from the literature to transform the bounds established in this section into a proof of theorem 1.4. To alleviate notation, the inverse temperature parameter $\beta > 0$, the Lagrange multiplier $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$ and the parameters $t > 0$ and $2 < p < \infty$ will be fixed throughout this section. We will write $C > 0$ for a constant that depends only on κ, p and t whose value might not be the same at each occurrence.

We begin by bounding f_λ^β from above by f_λ^∞ up to a small error. It will be necessary to make the dependence of these terminal conditions on $t > 0$ explicit by writing $f_{\lambda,t}^\beta$ and $f_{\lambda,t}^\infty$.

Proposition 9.1. *If $2 < p < \infty$, $\vec{x} \in \mathbb{R}^\kappa$ and $\delta \in (0,t)$, then*

$$f_{\lambda,t}^\beta(\vec{x}) \le f_{\lambda,t-\delta}^\infty(\vec{x}) - \frac{\kappa \log \beta\delta}{p\beta} + \frac{1}{\beta}\log \int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma}. \tag{9.1}$$

Proof. By a change of variables,

$$\begin{aligned} f_{\lambda,t}^\beta(\vec{x}) &\le f_{\lambda,t-\delta}^\infty(\vec{x}) + \frac{1}{\beta}\log \int_{\mathbb{R}^\kappa} e^{-\beta\delta\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma} \\ &= f_{\lambda,t-\delta}^\infty(\vec{x}) - \frac{\kappa \log \beta\delta}{p\beta} + \frac{1}{\beta}\log \int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma}. \end{aligned}$$

This finishes the proof. □

This result will play its part when we prove the upper bound in (8.12), at which point we will have to replace $f_{\lambda,t-\delta}^\infty$ in (9.1) by $f_{\lambda,t}^\infty$. This will be achieved through a continuity result that is an immediate consequence of the following bound on any maximizer $\vec{\sigma}^*_{\vec{x},\lambda,t}$ of (1.36).

Lemma 9.2. *If $2 < p < \infty$ and $\vec{x} \in \mathbb{R}^\kappa$, then there exists $\vec{\sigma}^*_{\vec{x},\lambda,t} \in \mathbb{R}^\kappa$ which attains the maximum in* (1.36). *Moreover,*

$$\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2 \le \max\Big(\Big(\frac{2\|\lambda\|_\infty}{t}\Big)^{\frac{1}{p-2}}, \Big(\frac{2\|\vec{x}\|_2}{t}\Big)^{\frac{1}{p-1}}\Big). \tag{9.2}$$

Proof. Consider the function $g : \mathbb{R}^\kappa \to \mathbb{R}$ defined by

$$g(\vec{\sigma}) = \big(\vec{\sigma}, \vec{x}\big) + \sum_{k\leq k'} \lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p. \tag{9.3}$$

By the Cauchy-Schwarz inequality,

$$g(\vec{\sigma}) \leq \|\vec{\sigma}\|_2\|\vec{x}\|_2 + \|\lambda\|_\infty\|\vec{\sigma}\|_2^2 - t\|\vec{\sigma}\|_2^p.$$

Since $p > 2$, it follows that $\lim_{\|\vec{\sigma}\|_2\to\infty} g(\vec{\sigma}) = -\infty$. Remembering that a continuous function attains its maximum on each compact set, there must exist $\vec{\sigma}^*_{\vec{x},\lambda,t} \in \mathbb{R}^\kappa$ which attains the maximum in (1.36). If we had

$$t\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2^p > \max\big(2\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2\|\vec{x}\|_2, 2\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2^2\|\lambda\|_\infty\big), \tag{9.4}$$

then we would have

$$0 = g(0) \leq f^\infty_\lambda(\vec{x}) = g(\vec{\sigma}^*_{\vec{x},\lambda,t}) < \frac{t\|\vec{\sigma}^*_{x,\lambda,t}\|_2^p}{2} + \frac{t\|\vec{\sigma}^*_{x,\lambda,t}\|_2^p}{2} - t\|\vec{\sigma}^*_{x,\lambda,t}\|_2^p = 0$$

which is not possible. Rearranging the reverse of (9.4) gives (9.2) and completes the proof. □

Proposition 9.3. *If $2 < p < \infty$ and $\vec{x} \in \mathbb{R}^\kappa$, then*

$$\lim_{\delta\to 0} f^\infty_{\lambda,t-\delta}(\vec{x}) = f^\infty_{\lambda,t}(\vec{x}). \tag{9.5}$$

Proof. Fix $\delta \in (0, t/2)$. It is clear that $f^\infty_{\lambda,t}(\vec{x}) \leq f^\infty_{\lambda,t-\delta}(\vec{x})$. On the other hand, lemma 9.2 implies that

$$\begin{aligned} f^\infty_{\lambda,t-\delta}(\vec{x}) &= \big(\vec{\sigma}^*_{\vec{x},\lambda,t-\delta}, \vec{x}\big) + \sum_{k\leq k'} \lambda_{k,k'}\vec{\sigma}^*_{\vec{x},\lambda,t-\delta}(k)\vec{\sigma}^*_{\vec{x},\lambda,t-\delta}(k') - (t-\delta)\|\vec{\sigma}^*_{\vec{x},\lambda,t-\delta}\|_2^p \\ &\leq f^\infty_{\lambda,t}(x) + \delta\max\Big(\Big(\frac{4\|\lambda\|_\infty}{t}\Big)^{\frac{1}{p-2}}, \Big(\frac{4\|\vec{x}\|_2}{t}\Big)^{\frac{1}{p-1}}\Big)^p. \end{aligned}$$

Letting $\delta \to 0$ completes the proof. □

We now turn our attention to bounding f^∞_λ from above by f^β_λ up to a small error. Once again, we drop the dependence of these terminal conditions on $t > 0$. Through a simple calculation detailed in the proof of proposition 9.5, this essentially comes down to bounding the average of the function (9.3) on a cube by its value at the centre of the cube. In other words, we need a type of mean-value property for the function (9.3). There are two main issues to address: the function (9.3) is not necessarily convex and, for technical reasons, we would like this mean-value property on cubes instead of balls. We will deal with the lack of convexity by adding a convex perturbation to (9.3). Replacing balls by cubes will be done by applying Jensen's inequality to a function defined on a cube of side-length $\delta > 0$ centred at some $\vec{\rho} \in \mathbb{R}^\kappa$,

$$\mathcal{C}_\delta(\vec{\rho}) = \vec{\rho} + [-\delta, \delta]^\kappa. \tag{9.6}$$

When we prove (8.12), the error incurred by these two fixes will vanish upon letting $\beta \to \infty$. The mean-value property for cubes takes the following form, and uses corollary B.2 to establish convexity of the function to which Jensen's inequality is applied.

Lemma 9.4. *If $2 < p < \infty$, $\vec{x} \in \mathbb{R}^\kappa$ and $\delta > 0$, then*

$$\begin{aligned} f^\infty_\lambda(\vec{x}) \leq \frac{1}{|\mathcal{C}_\delta(\vec{\sigma}^*_{\vec{x},\lambda,t})|}\int_{\mathcal{C}_\delta(\vec{\sigma}^*_{\vec{x},\lambda,t})} \Big(\big(\vec{\sigma}, \vec{x}\big) + \sum_{k\leq k'} \lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma} \\ + C\delta^2\big(\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2^{p-2} + \delta^{p-2} + \|\lambda\|_\infty\big) \end{aligned} \tag{9.7}$$

Proof. To simplify notation, for $k > k'$ let $\lambda_{k,k'} = \lambda_{k',k}$. Recall the function $g : \mathbb{R}^\kappa \to \mathbb{R}$ in (9.3). Given $\delta > 0$ and $\vec{\rho} \in \mathbb{R}^\kappa$, consider the function $f : \mathcal{C}_{2\delta}(\vec{\rho}) \to \mathbb{R}$ defined by

$$f(\vec{\sigma}) = g(\vec{\sigma}) + \sum_{k=1}^{\kappa} \frac{1}{2}\sigma(k)^2 h_k(\vec{\rho}),$$

where the constant $h_k(\vec{\rho})$ depends on $\vec{\rho}$ and is given by

$$h_k(\vec{\rho}) = \big(tp(p-1) + tp(p-2)\kappa\big)\big((2\|\vec{\rho}\|_2)^{p-2} + (2\sqrt{\kappa}\delta)^{p-2}\big) + 2|\lambda_{k,k}| + \sum_{k'\neq k} |\lambda_{k,k'}|.$$

Fix $\vec{\sigma} \in \mathcal{C}_{2\delta}(\vec{\rho})$ and $1 \leq k \leq \kappa$. A direct computation shows that

$$\begin{aligned}
\partial_{\sigma(k)\sigma(k)} f(\vec{\sigma}) &= 2\lambda_{k,k} - tp\|\vec{\sigma}\|_2^{p-2} - tp(p-2)\|\vec{\sigma}\|_2^{p-4}\sigma(k)^2 + h_k(\vec{\rho}) \\
&\geq 2(\lambda_{k,k} + |\lambda_{k,k}|) + tp(p-1)\big((2\|\vec{\rho}\|_2)^{p-2} + (2\sqrt{\kappa}\delta)^{p-2} - \|\vec{\sigma}\|_2^{p-2}\big) \\
&\quad + tp(p-2)\kappa\big((2\|\vec{\rho}\|_2)^{p-2} + (2\sqrt{\kappa}\delta)^{p-2}\big) + \sum_{k'\neq k}|\lambda_{k,k'}| \\
&\geq tp(p-2)\kappa\big((2\|\vec{\rho}\|_2)^{p-2} + (2\sqrt{\kappa}\delta)^{p-2}\big) + \sum_{k'\neq k}|\lambda_{k,k'}|.
\end{aligned}$$

Similarly, for $1 \leq k \neq k' \leq \kappa$,

$$\partial_{\sigma(k)\sigma(k')} f(\vec{\sigma}) = \lambda_{k,k'} - tp(p-2)\|\vec{\sigma}\|_2^{p-4}\sigma(k)\sigma(k').$$

It follows that

$$\sum_{k'\neq k} |\partial_{\sigma(k)\sigma(k')} f(\vec{\sigma})| \leq \sum_{k'\neq k}|\lambda_{k,k'}| + tp(p-2)\kappa\|\vec{\sigma}\|_2^{p-2} \leq \partial_{\sigma(k)\sigma(k)} f(\vec{\sigma}).$$

Invoking corollary B.2 shows that the Hessian of f is non-negative definite, and therefore f is convex. With this in mind, let $X = (X_i)_{i\leq\kappa}$ be a vector of independent random variables with X_i uniformly distributed on the interval $[\rho(i) - \delta, \rho(i) + \delta]$. Jensen's inequality implies that

$$f(\vec{\rho}) = f(\mathbb{E}\, X) \leq \mathbb{E}\, f(X) = \frac{1}{|\mathcal{C}_\delta(\vec{\rho})|}\int_{\mathcal{C}_\delta(\vec{\rho})} f(\vec{\sigma}) \,\mathrm{d}\,\vec{\sigma}$$

Substituting the definition of f into the right-hand side of this inequality and integrating yields

$$f(\vec{\rho}) \leq \frac{1}{|\mathcal{C}_\delta(\vec{\rho})|}\int_{\mathcal{C}_\delta(\vec{\rho})} g(\vec{\sigma}) \,\mathrm{d}\,\vec{\sigma} + \sum_{k=1}^{\kappa}\frac{1}{2}\rho(k)^2 h_k(\vec{\rho}) + \frac{\delta^2}{6}\sum_{k=1}^{\kappa} h_k(\vec{\rho}).$$

Rearranging and taking $\vec{\rho} = \vec{\sigma}^*_{\vec{x},\lambda,t}$ completes the proof. ☐

Proposition 9.5. *If $2 < p < \infty$ and $L > 0$, then for any $\vec{x} \in \mathbb{R}^\kappa$ with $\|\vec{x}\|_2 \leq L$ and every $0 < \delta < L^{\frac{1}{p-1}}$,*

$$f_\lambda^\infty(\vec{x}) \leq f_\lambda^\beta(\vec{x}) + C\delta^2\Big(\|\lambda\|_\infty + L^{\frac{p-2}{p-1}}\Big) - \frac{\kappa \log 2\delta}{\beta}. \tag{9.8}$$

Proof. Given $\vec{\rho} \in \mathbb{R}^\kappa$, Jensen's inequality implies that

$$\begin{aligned}
f_\lambda^\beta(\vec{x}) &\geq \frac{1}{\beta}\log\int_{\mathcal{C}_\delta(\vec{\rho})} \exp\beta\Big(\big(\vec{\sigma}, \vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma} \\
&\geq \frac{1}{|\mathcal{C}_\delta(\vec{\rho})|}\int_{\mathcal{C}_\delta(\vec{\rho})}\Big(\big(\vec{\sigma}, \vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma} + \frac{\kappa\log 2\delta}{\beta}.
\end{aligned} \tag{9.9}$$

Applying this with $\vec{\rho} = \vec{\sigma}^*_{\vec{x},\lambda,t}$ and invoking lemma 9.4 yields

$$f^\beta_\lambda(\vec{x}) \geq f^\infty_\lambda(\vec{x}) - C\delta^2\big(\|\vec{\sigma}^*_{\vec{x},\lambda,t}\|_2^{p-2} + \delta^{p-2} + \|\lambda\|_\infty\big) + \frac{\kappa \log 2\delta}{\beta}.$$

The result now follows by lemma 9.2. □

This result will play its part when we prove the lower bound in (8.12), at which point we will have to carefully deal with the fact that it only gives a bound of f^∞_λ by f^β_λ for values of $\vec{x}$ in a (possibly large) neighbourhood of the origin. Fortunately, this will not be a problem. It turns out that the bound (9.8) will be applied to one of the Auffinger-Chen control processes introduced in the next section. The generalization of the moment bound in lemma 12.3 of [6] to the vector spin setting, which corresponds to lemma 10.4 in this paper, will be used to show that dominating f^∞_λ by f^β_λ around the origin is sufficient for our purposes.

10 The Auffinger-Chen representation

In this section we extend the Auffinger-Chen stochastic control representation established for $\kappa = 2$ and Lipschitz terminal conditions in [7] to the setting of arbitrary integer $\kappa \geq 1$ and terminal conditions with sub-quadratic growth such as (8.5) and (1.36). The results in this section will be combined with the bounds obtained in section 9 to compare the quantities X_0 and Y_0 in (8.7) and (1.41). This will lead to a proof of theorem 1.4 in section 11.

Throughout this section, a constraint $D \in \Gamma_\kappa$, an inverse temperature parameter $\beta > 0$, a Lagrange multiplier $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$, a κ-dimensional Brownian motion $\boldsymbol{W} = (W_1, \ldots, W_\kappa)$ and parameters $t > 0$ and $2 < p < \infty$ will be fixed. We will also give ourselves a piecewise linear path $\pi \in \Pi_D$ defined by the sequences (1.33) and (1.34), as well as a discrete probability distribution $\alpha \in \mathcal{M}^d$ defined by the sequences (1.33) and (8.4). To prove the Auffinger-Chen representation, it will be convenient to replace the Gaussian random vectors z_j with covariance structure (1.39) appearing in the definition of the Parisi functional (8.10) by a continuous time stochastic process $\boldsymbol{B} = (\boldsymbol{B}(s))_{s\geq 0}$ that plays the same role,

$$\boldsymbol{B}(s) = \sqrt{2}\int_0^s \pi'(r)^{\frac{1}{2}}\,\mathrm{d}\,\boldsymbol{W}(r). \tag{10.1}$$

Since $\pi'(r) = (q_j - q_{j-1})^{-1}(\gamma_j - \gamma_{j-1}) \in \Gamma_\kappa$ for $r \in (q_{j-1}, q_j)$, this process is well-defined. Moreover, the Ito isometry shows that

$$\mathrm{Cov}\,\big(\boldsymbol{B}(q_j) - \boldsymbol{B}(q_{j-1})\big) = 2\int_{q_{j-1}}^{q_j} \pi'(r)\,\mathrm{d}\,r = 2(\gamma_j - \gamma_{j-1}). \tag{10.2}$$

If we introduce the function $\Phi : [0,1] \times \mathbb{R}^\kappa \to \mathbb{R}$ defined recursively by

$$\begin{cases} \Phi(1, \vec{x}) = f^\beta_\lambda(\vec{x}), \\ \Phi(s, \vec{x}) = \dfrac{1}{\beta\alpha_j}\log \mathbb{E}\exp \beta\alpha_j \Phi(q_{j+1}, \vec{x} + \boldsymbol{B}(q_{j+1}) - \boldsymbol{B}(s)),\ s \in [q_j, q_{j+1}), \end{cases} \tag{10.3}$$

then the independence of the increments of $\boldsymbol{B}$ and (10.2) imply that the Parisi functional (8.10) may be written as

$$\mathcal{P}_\beta(\lambda, \alpha, \pi) = \Phi(0,0) - \sum_{k\leq k'} \lambda_{k,k'} D_{k,k'} - \int_0^1 \beta\alpha(s)\,\mathrm{Sum}\,\big(\pi(s) \odot \pi'(s)\big)\,\mathrm{d}\,s. \tag{10.4}$$

We have made all dependencies of Φ on D, β, λ, α, π, p and t implicit for clarity of notation, but we will make them explicit whenever necessary. To obtain the Auffinger-Chen representation, we first use Gaussian integration by parts (see for instance lemma 1.1 in [29]) to show that Φ satisfies a non-linear parabolic PDE.

Lemma 10.1. *If $2 < p < \infty$ and $(s, \vec{x}) \in [0,1] \times \mathbb{R}^\kappa$, then*

$$\partial_s \Phi(s, \vec{x}) = -\Big(\big(\pi'(s), \nabla^2 \Phi(s, \vec{x})\big) + \beta\alpha(s)\big(\pi'(s)\nabla\Phi(s, \vec{x}), \nabla\Phi(s, \vec{x})\big)\Big), \tag{10.5}$$

where $\partial_s \Phi$ is understood as the right-derivative at the points of discontinuity of α.

Proof. Introduce the process $\boldsymbol{Y}(s) = \vec{x} + \boldsymbol{B}(q_{j+1}) - \boldsymbol{B}(s)$, and fix $s \in [q_j, q_{j+1})$. A direct computation shows that

$$\Phi_{x_l}(s, \vec{x}) = \mathbb{E}\, \Phi_{x_l}(q_{j+1}, \boldsymbol{Y}(s)) Z(s)$$

for the process $Z(s) = \exp \beta\alpha_j(\Phi(q_{j+1}, \boldsymbol{Y}(s)) - \Phi(s, \vec{x}))$. Differentiating again yields

$$\begin{aligned}\Phi_{x_l x_{l'}}(s, \vec{x}) = \mathbb{E}\,&\big(\Phi_{x_l x_{l'}}(q_{j+1}, \boldsymbol{Y}(s)) + \beta\alpha_j \Phi_{x_l}(q_{j+1}, \boldsymbol{Y}(s))\Phi_{x_{l'}}(q_{j+1}, \boldsymbol{Y}(s))\big) Z(s)\\ &- \beta\alpha_j \Phi_{x_l}(s, \vec{x})\Phi_{x_{l'}}(s, \vec{x}).\end{aligned}$$

To compute the time derivative of Φ, let g be a standard Gaussian vector in $\mathbb{R}^\kappa$ and consider the function

$$v(s) = \sqrt{\frac{2(q_{j+1} - s)}{q_{j+1} - q_j}} (\gamma_{j+1} - \gamma_j)^{1/2}.$$

Since $\Phi(s, \vec{x}) = \frac{1}{\beta\alpha_j} \log \mathbb{E} \exp \beta\alpha_j \Phi(q_{j+1}, \vec{x} + v(s)g)$, the Gaussian integration by parts formula gives

$$\begin{aligned}\Phi_s(s, x) &= \sum_{l,l'=1}^{\kappa} v'(s)_{l,l'}\, \mathbb{E}\, g_{l'} \Phi_{x_l}(q_{j+1}, \vec{x} + v(s)g) \exp \beta\alpha_j(\Phi(q_{j+1}, \vec{x} + v(s)g) - \Phi(s, \vec{x}))\\ &= \sum_{l,l',i=1}^{\kappa} v'(s)_{l,l'} v(s)_{il'}\, \mathbb{E}\, \big(\Phi_{x_l x_i}(q_{j+1}, \boldsymbol{Y}(s))\\ &\qquad\qquad + \beta\alpha_j \Phi_{x_l}(q_{j+1}, \boldsymbol{Y}(s))\Phi_{x_i}(q_{j+1}, \boldsymbol{Y}(s))\big) Z(s)\\ &= -\sum_{l,l'=1}^{\kappa} \frac{(\gamma_{j+1} - \gamma_j)_{l,l'}}{q_{j+1} - q_j}\, \mathbb{E}\, \big(\Phi_{x_l x_{l'}}(q_{j+1}, \boldsymbol{Y}(s))\\ &\qquad\qquad + \beta\alpha_j \Phi_{x_l}(q_{j+1}, \boldsymbol{Y}(s))\Phi_{x_{l'}}(q_{j+1}, \boldsymbol{Y}(s))\big) Z(s)\\ &= -\Big(\big(\pi'(s), \nabla^2\Phi(s, \vec{x})\big) + \beta\alpha_j\big(\pi'(s)\nabla\Phi(s, \vec{x}), \nabla\Phi(s, \vec{x})\big)\Big).\end{aligned}$$

Remembering that $\alpha(s) = \alpha_j$ completes the proof. □

The Hamilton-Jacobi equation (10.5) is the vector spin analogue of the Parisi PDE [29]. We now use similar ideas to those in [4, 16, 29] to obtain the vector spin analogue of the Auffinger-Chen representation from (10.5). To overcome the lack of Lipschitz continuity in the terminal condition (8.5), we will rely upon three classical results in stochastic analysis: the Ito formula, the Girsanov theorem and the Novikov condition [11, 17]. Given a filtration $\mathcal{F} = (\mathcal{F}_s)_{0 \le s \le 1}$, it will be convenient to denote by $\mathcal{A}$ the class of admissible control processes,

$$\begin{aligned}\mathcal{A} = \Big\{\boldsymbol{v} = (v_1, \dots, v_\kappa) \mid \boldsymbol{v} = (\boldsymbol{v}(s))_{0 \le s \le 1} \text{ is progressively measurable}\\ \text{and } \mathbb{E}\int_0^1 \|\boldsymbol{v}(s)\|_2^2 \,\mathrm{d}\, s < \infty\Big\}\end{aligned} \tag{10.6}$$

Proposition 10.2. *If $2 < p < \infty$, then there exists a probability space $(\Omega, \mathcal{F}_1, \mathbb{P})$, a filtration $\mathcal{F} = (\mathcal{F}_s)_{0\leq s\leq 1}$, a Brownian motion $\boldsymbol{W} = (\boldsymbol{W}_s)_{0\leq s\leq 1}$ and a continuous adapted process $\boldsymbol{X} = (\boldsymbol{X}_s)_{0\leq s\leq 1}$ which together form a weak solution to the stochastic differential equation*

$$\mathrm{d}\,\boldsymbol{X}(s) = 2\beta\alpha(s)\pi'(s)\nabla\Phi(s, \boldsymbol{X}(s))\,\mathrm{d}\,s + \sqrt{2}\pi'(s)^{1/2}\,\mathrm{d}\,\boldsymbol{W}(s), \quad \boldsymbol{X}(0) = 0. \tag{10.7}$$

Moreover,

$$\begin{aligned}\Phi(0,0) = \sup_{\boldsymbol{v}\in\mathcal{A}} \mathbb{E}\Big[f_\lambda^\beta\Big(2\int_0^1 \beta\alpha(s)\pi'(s)\boldsymbol{v}(s)\,\mathrm{d}\,s + \boldsymbol{B}(1)\Big)& \\ - \int_0^1 \beta\alpha(s)\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s\Big]&\end{aligned} \tag{10.8}$$

with the supremum attained by the admissible process $\boldsymbol{v}(s) = \nabla\Phi(s, \boldsymbol{X}(s))$.

Proof. To alleviate notation, let $C > 0$ denote a constant that depends only on κ, λ, α, π, β, D, p and t whose value might not be the same at each occurrence. An induction based on lemma A.1 and lemma A.3 can be used to show that for any $(s, \vec{x}) \in (0,1] \times \mathbb{R}^\kappa$,

$$\|\nabla\Phi(s, \vec{x})\|_2 \leq C\Big(1 + \|\vec{x}\|_2^{\frac{1}{p-1}}\Big). \tag{10.9}$$

With this in mind, consider the process $\boldsymbol{L} = (\boldsymbol{L}(s))_{0\leq s\leq 1}$,

$$\boldsymbol{L}(s) = \sqrt{2}\int_0^s \beta\alpha(r)\pi'(r)^{1/2}\nabla\Phi(r, \boldsymbol{B}(r))\,\mathrm{d}\,r.$$

The growth bound (10.9) and the assumption $\frac{1}{p-1} < 1$ imply that

$$\mathbb{E}\exp\int_0^1 \|\boldsymbol{L}(s)\|_2^2\,\mathrm{d}\,s \leq C\,\mathbb{E}\exp\big(C\sup_{0\leq s\leq 1}\|\boldsymbol{B}(s)\|_2^2\big) \leq C\,\mathbb{E}\exp C\big(\sup_{0\leq s\leq 1}\|\boldsymbol{W}(s)\|_2^2\big),$$

where the last inequality uses the fact that π' is piecewise constant. Combining this with Doob's maximal inequality reveals that

$$\mathbb{E}\exp\int_0^1 \|\boldsymbol{L}(s)\|_2^2\,\mathrm{d}\,s \leq C\,\mathbb{E}\sup_{0\leq s\leq 1}\exp C\|\boldsymbol{W}(s)\|_2^2 \leq C\,\mathbb{E}\exp C\|\boldsymbol{W}(1)\|_2^2 < \infty.$$

It follows by the Novikov condition that the stochastic exponential

$$\begin{aligned}\mathcal{E}(\boldsymbol{L})_s = \exp\Big(\int_0^s \sqrt{2}\beta\alpha(r)\pi'(r)^{1/2}\nabla\Phi(r, \boldsymbol{B}(r))\cdot\mathrm{d}\,\boldsymbol{W}(r)& \\ - \beta^2\int_0^s \alpha(r)^2\big(\pi'(r)\nabla\Phi(r, \boldsymbol{B}(r)), \nabla\Phi(r, \boldsymbol{B}(r))\big)\,\mathrm{d}\,r\Big)&\end{aligned}$$

is a martingale. If we denote by $\mathbb{Q}$ the measure under which $\boldsymbol{W}$ is a Brownian motion and introduce the measure $\mathrm{d}\,\mathbb{P} = \mathcal{E}(\boldsymbol{L})_1\,\mathrm{d}\,\mathbb{Q}$, then Girsanov's theorem implies that

$$\tilde{\boldsymbol{W}}(s) = \boldsymbol{W}(s) - \sqrt{2}\int_0^s \beta\alpha(r)\pi'(r)^{1/2}\nabla\Phi(s, \boldsymbol{B}(r))\,\mathrm{d}\,r$$

is a $\mathbb{P}$-Brownian motion. Rearranging shows that $(\boldsymbol{B}, \tilde{\boldsymbol{W}})$ is a weak solution to (10.7). Henceforth, we will write $\mathcal{F} = (\mathcal{F}_s)_{0\leq s\leq 1}$, $\boldsymbol{W} = (\boldsymbol{W}_s)_{0\leq s\leq 1}$ and $\boldsymbol{X} = (\boldsymbol{X}_s)_{0\leq s\leq 1}$ for a filtration, a Brownian motion and a continuous adapted process which together form a weak solution to (10.7). Given $\boldsymbol{v} \in \mathcal{A}$, let $\boldsymbol{Y}(s) = 2\int_0^s \beta\alpha(r)\pi'(r)\boldsymbol{v}(r)\,\mathrm{d}\,r + \boldsymbol{B}(s)$. By Ito's formula and the Parisi PDE (10.5),

$$\begin{aligned}\mathrm{d}\,\Phi &= \Phi_s\,\mathrm{d}\,s + 2\beta\alpha(s)\big(\nabla\Phi, \pi'(s)\boldsymbol{v}(s)\big)\,\mathrm{d}\,s + \big(\nabla^2\Phi, \pi'(s)\big)\,\mathrm{d}\,s + \sqrt{2}\big(\nabla\Phi, \pi(s)\,\mathrm{d}\,\boldsymbol{W}(s)\big)\\ &= -\beta\alpha(s)\Big(\big(\pi'(s)(\nabla\Phi - \boldsymbol{v}(s)), \nabla\Phi - \boldsymbol{v}(s)\big) - \big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\Big)\,\mathrm{d}\,s\\ &\quad + \sqrt{2}\big(\nabla\Phi, \pi(s)\,\mathrm{d}\,\boldsymbol{W}(s)\big),\end{aligned}$$

where Φ and its derivatives are evaluated at $(s, \boldsymbol{Y}(s))$. The growth bound (10.9), the Cauchy-Schwarz inequality, the boundedness of π' and the Ito isometry reveal that

$$\mathbb{E}\int_0^1 \|\nabla\Phi(s, \boldsymbol{Y}(s))\|_\infty^2\,\mathrm{d}\,s \le C\big(1 + \sup_{0\le s\le 1}\mathbb{E}\|\boldsymbol{Y}(s)\|_2^2\big) \le C\Big(1 + \mathbb{E}\int_0^1 \|\boldsymbol{v}(s)\|_2^2\,\mathrm{d}\,s\Big) < \infty.$$

This means that $(\sqrt{2}\int_0^s(\nabla\Phi, \pi(s)\,\mathrm{d}\,\boldsymbol{W}(s)))_{s\le 1}$ is a martingale. Together with the non-negative definiteness of π', this implies that

$$\mathbb{E}\,\Phi(1, \boldsymbol{Y}(1)) - \Phi(0,0) \le \int_0^1 \beta\alpha(s)\,\mathbb{E}\,\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s$$

with equality for the process $\boldsymbol{v}(s) = \nabla\Phi(s, \boldsymbol{X}(s))$. Rearranging gives the lower bound in (10.8). To prove the matching upper bound, it suffices to show that $\boldsymbol{v}(s) = \nabla\Phi(s, \boldsymbol{X}(s))$ belongs to the admissible class $\mathcal{A}$. Fix $0 < s \le r \le 1$. By the triangle inequality, the Cauchy-Schwarz inequality and the growth bound (10.9),

$$\sup_{0\le s\le r}\|\boldsymbol{X}(s)\|_2^2 \le C\Big(1 + \int_0^r \sup_{0\le s\le w}\|\boldsymbol{X}(s)\|_2^2\,\mathrm{d}\,w + \sup_{0\le s\le r}\|\boldsymbol{B}(s)\|_2^2\Big). \tag{10.10}$$

On the other hand, Doob's maximal inequality and the Ito isometry yield

$$\mathbb{E}\sup_{0\le s\le r}\|\boldsymbol{B}(s)\|_2^2 \le \mathbb{E}\|\boldsymbol{B}(r)\|_2^2 \le C\,\mathrm{tr}\int_0^r \pi'(w)\,\mathrm{d}\,w \le C\,\mathrm{tr}(D). \tag{10.11}$$

Substituting this into (10.10) and applying Gronwall's inequality to the resulting bound shows that $\mathbb{E}\sup_{0\le s\le 1}\|\boldsymbol{X}(s)\|_2^2 \le C$. Invoking (10.9) one last time completes the proof. □

Of course, an analogous result holds for the random variable Y_0 in (1.41). Given a discrete measure $\zeta \in \mathcal{N}^d$ defined by the sequences (1.33) and (1.38), introduce the function $\Psi : [0,1] \times \mathbb{R}^\kappa \to \mathbb{R}$ defined recursively by

$$\begin{cases}\Psi(1, \vec{x}) = f_\lambda^\infty(\vec{x}),\\ \Psi(s, \vec{x}) = \dfrac{1}{\zeta_j}\log\mathbb{E}\exp\zeta_j\Psi(q_{j+1}, \vec{x} + \boldsymbol{B}(q_{j+1}) - \boldsymbol{B}(s)),\ s \in [q_j, q_{j+1}).\end{cases} \tag{10.12}$$

The Parisi functional (1.42) may be written as

$$\mathcal{P}_\infty(\lambda, \zeta, \pi) = \Psi(0,0) - \sum_{k\le k'}\lambda_{k,k'}D_{k,k'} - \int_0^1 \zeta(s)\,\mathrm{Sum}\,\big(\pi(s)\odot\pi'(s)\big)\,\mathrm{d}\,s, \tag{10.13}$$

and the Gaussian integration by parts formula can be used as in lemma 10.1 to show that (10.12) satisfies the Parisi PDE,

$$\partial_s\Psi(s, \vec{x}) = -\Big(\big(\pi'(s), \nabla^2\Psi(s, \vec{x})\big) + \zeta(s)\big(\pi'(s)\nabla\Psi(s, \vec{x}), \nabla\Psi(s, \vec{x})\big)\Big), \tag{10.14}$$

where $\partial_s\Psi$ is understood as the right derivative at the points of discontinuity of ζ. An identical argument to that in proposition 10.2 gives the following weak form of the Auffinger-Chen representation.

Proposition 10.3. *If $2 < p < \infty$, then there exists a probability space $(\Omega, \mathcal{F}_1, \mathbb{P})$, a filtration $\mathcal{F} = (\mathcal{F}_s)_{0\leq s\leq 1}$, a Brownian motion $\boldsymbol{W} = (\boldsymbol{W}_s)_{0\leq s\leq 1}$ and a continuous adapted process $\boldsymbol{X} = (\boldsymbol{X}_s)_{0\leq s\leq 1}$ which together form a weak solution to the stochastic differential equation*

$$\mathrm{d}\,\boldsymbol{X}(s) = 2\zeta(s)\pi'(s)\nabla\Psi(s,\boldsymbol{X}(s))\,\mathrm{d}\,s + \sqrt{2}\pi'(s)\,\mathrm{d}\,\boldsymbol{W}(s), \quad \boldsymbol{X}(0) = 0. \tag{10.15}$$

Moreover,

$$\begin{aligned}\Psi(0,0) = \sup_{\boldsymbol{v}\in\mathcal{A}} \mathbb{E}\Big[f_\lambda^\infty\Big(2\int_0^1 \zeta(s)\pi'(s)\boldsymbol{v}(s)\,\mathrm{d}\,s + \boldsymbol{B}(1)\Big)\\ - \int_0^1 \zeta(s)\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s\Big]\end{aligned} \tag{10.16}$$

with the supremum attained by the admissible process $\boldsymbol{v}(s) = \nabla\Psi(s,\boldsymbol{X}(s))$.

We close this section with a moment bound on the weak solution to the stochastic differential equation (10.15) which will allow us to deal with the fact that proposition 9.5 only holds for bounded values of $\vec{x}$.

Lemma 10.4. *If $(\boldsymbol{X}(s))_{0\leq s\leq 1}$ is a weak solution to* (10.15) *and $\eta = \max(1 + \frac{1}{p-1}, \frac{2}{p-1}) \in (1,2)$, then*

$$\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \leq C\Big(1 + \|\zeta\|_\infty(1+\|\lambda\|_\infty)^{1+\frac{2}{p-2}}\Big)^{\frac{2}{2-\eta}} \tag{10.17}$$

for some constant $C > 0$ that depends only on κ, p, t and D.

Proof. To alleviate notation, let $C > 0$ denote a constant that depends only on κ, p, t and D whose value might not be the same at each occurrence. If $\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 < 1$ the result is trivial, so assume without loss of generality that $\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \geq 1$. Introduce the process $\boldsymbol{v}(s) = \nabla\Psi(s,\boldsymbol{X}(s))$ in such a way that

$$\boldsymbol{X}(1) = \int_0^1 2\zeta(s)\pi'(s)\boldsymbol{v}(s)\,\mathrm{d}\,s + \boldsymbol{B}(1).$$

The triangle inequality and (10.11) reveal that

$$\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \leq C\Big(1 + \mathbb{E}\Big\|\int_0^1 \zeta(s)\pi'(s)\boldsymbol{v}(s)\,\mathrm{d}\,s\Big\|_2^2\Big). \tag{10.18}$$

With this in mind, fix $1 \leq l \leq \kappa$. The Cauchy-Schwarz inequality implies that

$$\begin{aligned}\Big(\int_0^1 \zeta(s)(\pi'(s)\boldsymbol{v}(s))_l\,\mathrm{d}\,s\Big)^2 &\leq C\sum_{k=1}^\kappa\Big(\int_0^1 \zeta(s)\pi'(s)_{lk}^{1/2}\sum_{i=1}^\kappa \pi'(s)_{ki}^{1/2}v_i(s)\,\mathrm{d}\,s\Big)^2\\ &\leq C\sum_{k=1}^\kappa \int_0^1 \zeta(s)\pi'(s)_{lk}^{1/2}\pi'(s)_{lk}^{1/2}\,\mathrm{d}\,s\\ &\qquad\int_0^1 \zeta(s)\Big(\sum_{i=1}^\kappa \pi'(s)_{ki}^{1/2}v_i(s)\Big)^2\,\mathrm{d}\,s\\ &\leq C\|\zeta\|_\infty \int_0^1 \zeta(s)\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s.\end{aligned}$$

Substituting this back into (10.18) yields

$$\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \leq C\Big(1 + \|\zeta\|_\infty\int_0^1 \zeta(s)\,\mathbb{E}\,\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s\Big).$$

On the other hand, taking the zero process in proposition 10.3 shows that

$$\mathbb{E}\left[f_\lambda^\infty(\boldsymbol{X}(1)) - \int_0^1 \zeta(s)\big(\pi'(s)\boldsymbol{v}(s), \boldsymbol{v}(s)\big)\,\mathrm{d}\,s\right] = \Psi(0,0) \geq \mathbb{E}\, f_\lambda^\infty(\boldsymbol{B}(1)) \geq 0,$$

and therefore

$$\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \leq C\big(1 + \|\zeta\|_\infty\, \mathbb{E}\, f_\lambda^\infty(\boldsymbol{X}(1))\big). \tag{10.19}$$

To bound this further, observe that by (A.6) in appendix A and Jensen's inequality,

$$\begin{aligned}\mathbb{E}\, f_\lambda^\infty(\boldsymbol{X}(1)) &\leq C(1+\|\lambda\|_\infty)^{1+\frac{2}{p-2}}\Big((\mathbb{E}\|\boldsymbol{X}(1)\|_2^2)^{\frac{1}{2}+\frac{1}{2(p-1)}} + (\mathbb{E}\|\boldsymbol{X}(1)\|_2^2)^{\frac{1}{2}} \\ &\qquad + (\mathbb{E}\|\boldsymbol{X}(1)\|_2^2)^{\frac{1}{p-1}} + 1\Big) \\ &\leq C(1+\|\lambda\|_\infty)^{1+\frac{2}{p-2}}(\mathbb{E}\|\boldsymbol{X}(1)\|_2^2)^{\eta/2},\end{aligned}$$

where we have used the assumption that $\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \geq 1$. Substituting this back into (10.19) and again using the fact that $\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \geq 1$ gives

$$\mathbb{E}\|\boldsymbol{X}(1)\|_2^2 \leq C\Big(1 + \|\zeta\|_\infty(1+\|\lambda\|_\infty)^{1+\frac{2}{p-2}}\Big)(\mathbb{E}\|\boldsymbol{X}(1)\|_2^2)^{\eta/2}.$$

Rearranging completes the proof. □

11 Proof of the main result

In this section we finally prove theorem 1.4. The proof of the upper bound will follow section 12.2 of [6]. On the other hand, the proof of the lower bound will be considerably shorter and less involved than its one-dimensional analogue in [6]. In particular, it will leverage the results of section 9 to avoid the technicalities associated with truncating. Specializing our arguments to the scalar, $\kappa = 1$, case gives a shorter and more direct proof of the main result in [6] when $2 < p < \infty$.

Lemma 11.1. *If $2 < p < \infty$, $D \in \Gamma_\kappa$ and $t > 0$, then*

$$L_{p,D}(t) \leq \inf_{\lambda,\zeta,\pi} \mathcal{P}_\infty(\lambda,\zeta,\pi), \tag{11.1}$$

where the infimum is taken over all $(\lambda,\zeta,\pi) \in \mathbb{R}^{\kappa(\kappa+1)/2} \times \mathcal{N}^d \times \Pi_D$.

Proof. By theorem 7.3, it suffices to show that

$$\liminf_{\beta\to\infty} \lim_{N\to\infty} F_{N,\beta^{-1}}(\beta) \leq \inf_{\lambda,\zeta,\pi} \mathcal{P}_\infty(\lambda,\zeta,\pi). \tag{11.2}$$

Fix a Lagrange multiplier $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$, a piecewise linear path $\pi \in \Pi_D$ and a discrete measure $\zeta \in \mathcal{N}^d$ defined by the sequences (1.37) and (1.38). Given an inverse temperature parameter $\beta > 0$, introduce the measure

$$\alpha^\beta(s) = \beta^{-1}\zeta(s)\mathbf{1}_{[0,1)}(s) + \mathbf{1}_{\{1\}}(s)$$

on $[0,1]$. It is clear that $\alpha^\beta \in \mathcal{M}^d$ for β large enough. Moreover, $\alpha^\beta(\{q_j\}) = \alpha_j^\beta - \alpha_{j-1}^\beta$ for the sequence of parameters (8.4) defined by $\alpha_j^\beta = \beta^{-1}\zeta_j$. The Guerra replica symmetry breaking bound in lemma 2 of [33] implies that

$$\lim_{N\to\infty} F_{N,\beta^{-1}}(\beta) \leq \mathcal{P}_\beta(\lambda,\alpha^\beta,\pi) + \beta^{-1}\|\lambda\|_1 + L\beta^{-1}. \tag{11.3}$$

for some constant $L > 0$ independent of β. To bound this further, recall that by proposition 9.1,

$$f_{\lambda,t}^\beta(\vec{x}) \leq f_{\lambda,t-\delta}^\infty(\vec{x}) - \frac{\kappa\log\beta\delta}{p\beta} + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p}\,\mathrm{d}\,\vec{\sigma}$$

for any $\delta \in (0, t/2)$. If we make the dependence of the functions Φ and Ψ in (10.3) and (10.12) on the underlying measure and parameter $t > 0$ explicit by writing $\Phi^{\alpha^\beta, t}$ and $\Psi^{\zeta, t}$, then a simple induction yields

$$\Phi^{\alpha^\beta,t}(0,0) \leq \Psi^{\zeta,t-\delta}(0,0) - \frac{\kappa \log \beta\delta}{p\beta} + \frac{1}{\beta} \log \int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma}.$$

It follows that

$$\begin{aligned}\mathcal{P}_\beta(\lambda, \alpha^\beta, \pi) \leq \Psi^{\zeta,t-\delta}(0,0) &- \sum_{k \leq k'} \lambda_{k,k'} D_{k,k'} - \int_0^1 \zeta(s) \,\mathrm{Sum}\,\big(\pi(s) \odot \pi'(s)\big) \,\mathrm{d}\, s \\ &- \frac{\kappa \log \beta\delta}{p\beta} + \frac{1}{\beta} \log \int_{\mathbb{R}^\kappa} e^{-\|\vec{\sigma}\|_2^p} \,\mathrm{d}\,\vec{\sigma}.\end{aligned} \tag{11.4}$$

Substituting this into (11.3) and letting $\beta \to \infty$ gives

$$\lim_{\beta\to\infty} \lim_{N\to\infty} F_{N,\beta^{-1}}(\beta) \leq \Psi^{\zeta,t-\delta}(0,0) - \sum_{k \leq k'} \lambda_{k,k'} D_{k,k'} - \int_0^1 \zeta(s) \,\mathrm{Sum}\,\big(\pi(s) \odot \pi'(s)\big) \,\mathrm{d}\, s.$$

By proposition 9.3 and an induction that combines the dominated convergence theorem with (A.6) in appendix A, it is readily verified that $\lim_{\delta\to 0} \Psi^{\zeta,t-\delta}(0,0) = \Psi^{\zeta,t}(0,0)$. Letting $\delta \to 0$ in the above inequality and taking the infimum over λ, ζ and π establishes (11.2) and completes the proof. □

The proof of the matching lower bound in (1.43) requires more work. Given an inverse temperature parameter $\beta > 0$, denote by $(\lambda^\beta, \alpha^\beta, \pi^\beta)$ a triple of almost minimizers defined by the condition

$$\mathcal{P}_\beta(\lambda^\beta, \alpha^\beta, \pi^\beta) \leq \inf_{\lambda,\alpha,\pi} \mathcal{P}_\beta(\lambda, \alpha, \pi) + \beta^{-1}. \tag{11.5}$$

It will be convenient to control the magnitude of λ^β. It is at this point that we have to specialize the claim of theorem 1.4 to positive definite matrices $D \in \Gamma_\kappa^+$. The author's inability to extend the following result to the space of non-negative definite matrices is the reason for proving the second equality in (1.28) and extending section 4 beyond proposition 4.4.

Lemma 11.2. *If $D \in \Gamma_\kappa^+$ and $(\lambda^\beta, \alpha^\beta, \pi^\beta) \in \mathbb{R}^{\kappa(\kappa+1)/2} \times \mathcal{M}^d \times \Pi_D$ satisfies (11.5) for some $\beta > 1$, then there exists a constant $C > 0$ that depends only on p, κ, D and t with*

$$\|\lambda^\beta\|_\infty \leq C\beta. \tag{11.6}$$

Proof. To alleviate notation, let $C > 0$ denote a constant that depends only on p, κ, D and t whose value might not be the same at each occurrence. For each pair $k > k'$, let $\lambda_{k,k'}^\beta = \lambda_{k',k}^\beta$. Consider the $\kappa \times \kappa$ symmetric matrix $\Lambda^\beta = (\lambda_{kk'}^\beta)$ as well as the $\kappa \times \kappa$ symmetric matrix $\mathrm{sgn}\,\Lambda^\beta = (\mathrm{sgn}\,\lambda_{kk'}^\beta)$ containing its signs. We adopt the convention that $\mathrm{sgn}(0) = 0$. Since D is positive definite and $\|\mathrm{sgn}\,\Lambda^\beta\|_\infty \leq 1$, using lemma B.5 it is possible to find $\epsilon > 0$ small enough that depends only on D and κ such that

$$D' = D + \epsilon \,\mathrm{sgn}\,\Lambda^\beta$$

is positive definite. Introduce the Gaussian measure

$$\mathrm{d}\,\mu(\vec{\sigma}) = \Big(\frac{1}{\sqrt{2\pi}}\Big)^\kappa \frac{1}{\sqrt{\det D'}} \exp\Big(-\frac{1}{2} \vec{\sigma}^T D'^{-1} \vec{\sigma}\Big) \,\mathrm{d}\,\vec{\sigma}$$

associated with a centred Gaussian random vector X having covariance matrix D'. Denote by Φ the function (10.3) corresponding to the parameters $\lambda^\beta, \alpha^\beta, \pi^\beta$ and $\beta > 0$.

Since the terminal condition (8.5) is convex, taking the zero process in proposition 10.2 and invoking Jensen's inequality shows that

$$\Phi(0,0) \geq \mathbb{E}\, f^\beta_{\lambda^\beta}(\boldsymbol{B}(1)) \geq f^\beta_{\lambda^\beta}(\mathbb{E}\,\boldsymbol{B}(1)) = f^\beta_{\lambda^\beta}(0).$$

Another application of Jensen's inequality gives

$$\begin{aligned}\Phi(0,0) &\geq \frac{\log\sqrt{\det D'}}{\beta} + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa}\exp\beta\Big(\sum_{k\leq k'}\lambda^\beta_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\mu(\vec{\sigma})\\ &\geq \frac{\log\sqrt{\det D'}}{\beta} + \sum_{k\leq k'}\lambda^\beta_{k,k'}\,\mathbb{E}\,X_kX_{k'} - t\,\mathbb{E}\|X\|_2^p.\end{aligned}$$

It follows by definition of the Parisi functional in (8.8) and the equality $D'_{k,k'} - D_{k,k'} = \epsilon\,\mathrm{sgn}\,\lambda^\beta_{k,k'}$ that

$$\begin{aligned}\mathcal{P}_\beta(\lambda^\beta,\alpha^\beta,\pi^\beta) \geq \epsilon\sum_{k\leq k'}|\lambda^\beta_{k,k'}| - t\,\mathbb{E}\|X\|_2^p + \frac{\log\sqrt{\det D'}}{\beta}\\ - \frac{\beta}{2}\sum_{0\leq j\leq r-1}\alpha_j^\beta\big(\|\gamma^\beta_{j+1}\|^2_{\mathrm{HS}} - \|\gamma^\beta_j\|^2_{\mathrm{HS}}\big).\end{aligned}$$

To bound this further, observe that

$$\sum_{0\leq j\leq r-1}\alpha_j^\beta\big(\|\gamma^\beta_{j+1}\|^2_{\mathrm{HS}} - \|\gamma^\beta_j\|^2_{\mathrm{HS}}\big) = \alpha^\beta_{r-1}\|\gamma^\beta_r\|^2_{\mathrm{HS}} - \sum_{j=1}^r(\alpha^\beta_j - \alpha^\beta_{j-1})\|\gamma^\beta_j\|^2_{\mathrm{HS}} \leq \|D\|^2_{\mathrm{HS}},$$

where the last inequality uses lemma B.4, and denote by $\Psi^{\lambda,\zeta,\pi,t}$ the function (10.12) associated with the Lagrange multiplier $\lambda\in\mathbb{R}^{\kappa(\kappa+1)/2}$, the discrete measure $\zeta\in\mathcal{N}^d$, the piecewise linear path $\pi\in\Pi_D$ and the parameter $t>0$. By (11.5) and (11.4),

$$\begin{aligned}\mathcal{P}_\beta(\lambda^\beta,\alpha^\beta,\pi^\beta) \leq 1 - \frac{\kappa\log\beta t/2}{p\beta} + \frac{1}{\beta}\log\int_{\mathbb{R}^\kappa}e^{-\|\vec{\sigma}\|_2^p}\,\mathrm{d}\,\vec{\sigma} + \inf_{\lambda,\zeta,\pi}\Big(\Psi^{\lambda,\zeta,\pi,t/2}(0,0)\\ - \sum_{k\leq k'}\lambda_{k,k'}D_{k,k'} - \int_0^1\zeta(s)\,\mathrm{Sum}\,\big(\pi(s)\odot\pi'(s)\big)\,\mathrm{d}\,s\Big).\end{aligned}$$

Combining these three bounds and rearranging yields

$$\epsilon\sum_{k\leq k'}|\lambda^\beta_{k,k'}| \leq C\beta + t\,\mathbb{E}\|X\|_2^p - \frac{\log\sqrt{\det D'}}{\beta}. \tag{11.7}$$

Notice that the matrix D' depends only on D,κ and $\mathrm{sgn}\,\Lambda^\beta$. Since there are only finitely many choices for the matrix $\mathrm{sgn}\,\Lambda^\beta$ and $\beta^{-1}\leq 1$, we can absorb the term $\beta^{-1}\log\sqrt{\det D'}$ into the constant $C>0$. To deal with the term involving the random vector X, let $M\in\mathbb{R}^{\kappa\times\kappa}$ be a positive definite matrix with $M^TM = D'$. If g is a standard Gaussian vector in $\mathbb{R}^\kappa$, then the Cauchy-Schwarz inequality implies that

$$\mathbb{E}\|X\|_2^p = \mathbb{E}\|Mg\|_2^p \leq \|M\|^p_{\mathrm{HS}}\,\mathbb{E}\|g\|_2^p = \mathrm{tr}(D')^{p/2}\,\mathbb{E}\|g\|_2^p \leq (\mathrm{tr}(D)+\epsilon)^{p/2}\,\mathbb{E}\|g\|_2^p.$$

Substituting this into (11.7) completes the proof. □

Proof (Theorem 1.4). To alleviate notation, let $C>0$ denote a constant that depends only on p,κ,D and t whose value might not be the same at each occurrence. By theorem 7.3, lemma 11.1, theorem 8.1 and the choice of the minimizing sequence $(\lambda^\beta, \alpha^\beta, \pi^\beta)$ satisfying (11.5), it suffices to show that

$$\inf_{\lambda,\zeta,\pi} \mathcal{P}_\infty(\lambda,\zeta,\pi) \leq \limsup_{\beta\to\infty} \mathcal{P}_\beta(\lambda^\beta,\alpha^\beta,\pi^\beta). \tag{11.8}$$

Fix $\beta > 1$, $L > 0$ and $0 < \delta < \min(L^{1/(p-1)}, 1/2)$. Consider the measure $\zeta^\beta = \beta\alpha^\beta \in \mathcal{N}^d$, and denote by Ψ the function (10.12) associated with the Lagrange multiplier λ^β, the measure ζ^β and the path π^β. Let $\boldsymbol{X}^\beta$ be a weak solution to the stochastic differential equation (10.15), and write $\boldsymbol{v}^\beta(s) = \nabla\Psi(s, \boldsymbol{X}^\beta(s))$ for its corresponding optimal control process. Consider the set on which $\boldsymbol{X}^\beta(1)$ lies within the ball of radius $L > 0$ around the origin,

$$B = \big\{\|\boldsymbol{X}^\beta(1)\|_2 \leq L\big\},$$

and notice that by proposition 9.5,

$$\mathbb{E}\, f^\infty_{\lambda^\beta}(\boldsymbol{X}^\beta(1))\mathbf{1}_B \leq \mathbb{E}\, f^\beta_{\lambda^\beta}(\boldsymbol{X}^\beta(1))\mathbf{1}_B + C\delta^2\Big(\|\lambda^\beta\|_\infty + L^{\frac{p-2}{p-1}}\Big) - \frac{\kappa\log 2\delta}{\beta}. \tag{11.9}$$

To bound this further, observe that by (9.9) and symmetry,

$$\begin{aligned} f^\beta_{\lambda^\beta}(\vec{x}) &\geq \frac{1}{|\mathcal{C}_\delta(0)|}\int_{\mathcal{C}_\delta(0)} \Big((\vec{\sigma},\vec{x}) + \sum_{k\leq k'} \lambda^\beta_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma} + \frac{\kappa\log 2\delta}{\beta} \\ &\geq \sum_{k=1}^\kappa \frac{\delta^2\lambda^\beta_{k,k}}{3} - \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa} \|\vec{\sigma}\|_2^p\,\mathrm{d}\,\vec{\sigma} + \frac{\kappa\log 2\delta}{\beta}. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbb{E}\, f^\beta_{\lambda^\beta}(\boldsymbol{X}^\beta(1))\mathbf{1}_B &\leq \mathbb{E}\Big(f^\beta_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) - \sum_{k=1}^\kappa \frac{\delta^2\lambda^\beta_{k,k}}{3} + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec{\sigma}\|_2^p\,\mathrm{d}\,\vec{\sigma} \\ &\qquad - \frac{\kappa\log 2\delta}{\beta}\Big)\mathbf{1}_B + \kappa\delta^2\|\lambda^\beta\|_\infty + \frac{\kappa|\log 2\delta|}{\beta} \\ &\leq \mathbb{E}\, f^\beta_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) + C\delta^2\|\lambda^\beta\|_\infty + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec{\sigma}\|_2^p\,\mathrm{d}\,\vec{\sigma} - \frac{2\kappa\log 2\delta}{\beta}. \end{aligned}$$

Together with (11.9) and lemma 11.2, this implies that

$$\begin{aligned} \mathbb{E}\, f^\infty_{\lambda^\beta}(\boldsymbol{X}^\beta(1))\mathbf{1}_B &\leq \mathbb{E}\, f^\beta_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) + C\delta^2\Big(\beta + L^{\frac{p-2}{p-1}}\Big) \\ &\qquad + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec{\sigma}\|_2^p\,\mathrm{d}\,\vec{\sigma} - \frac{3\kappa\log 2\delta}{\beta}. \end{aligned} \tag{11.10}$$

On the other hand, if $\eta = \max(1 + \frac{1}{p-1}, \frac{2}{p-1}) \in (1,2)$ as in lemma 10.4, then Hölder's inequality and Chebyshev's inequality give

$$\begin{aligned} \mathbb{E}\, f^\infty_{\lambda^\beta}(\boldsymbol{X}^\beta(1))\mathbf{1}_{B^c} &\leq (\mathbb{E}|f^\infty_{\lambda^\beta}(\boldsymbol{X}^\beta(1))|^{2/\eta})^{\eta/2}\mathbb{P}(B^c)^{1-\eta/2} \\ &\leq \frac{1}{L^{2-\eta}}(\mathbb{E}|f^\infty_{\lambda^\beta}(\boldsymbol{X}^\beta(1))|^{2/\eta})^{\eta/2}(\mathbb{E}\|\boldsymbol{X}^\beta(1)\|_2^2)^{1-\eta/2}. \end{aligned} \tag{11.11}$$

Remembering that $\zeta^\beta = \beta\alpha^\beta$ and invoking lemma 10.4 as well as lemma 11.2 shows that

$$\mathbb{E}\|\boldsymbol{X}^\beta(1)\|_2^2 \leq C\Big(1 + \|\zeta^\beta\|_\infty(1+\|\lambda^\beta\|_\infty)^{1+\frac{2}{p-2}}\Big)^{\frac{2}{2-\eta}} \leq C\beta^K, \tag{11.12}$$

where $K>0$ is a constant that depends only on p whose value might not be the same at each occurrence. Leveraging (A.6) in appendix A and lemma 11.2 yields

$$\mathbb{E}|f^{\infty}_{\lambda^\beta}(\boldsymbol{X}^\beta(1))|^{2/\eta} \leq C\Big(\mathbb{E}\|\boldsymbol{X}^\beta(1)\|_2^2 + \|\lambda^\beta\|_\infty^K\,\mathbb{E}\|\boldsymbol{X}^\beta(1)\|_2^2 + \|\lambda^\beta\|_\infty^K\,\mathbb{E}\|\boldsymbol{X}^\beta(1)\|_2^2 + \|\lambda^\beta\|_\infty^K\Big) \leq C\beta^K. \tag{11.13}$$

Combining the bound resulting from substituting (11.12) and (11.13) into (11.11) with (11.10) reveals that

$$\mathbb{E}\, f^{\infty}_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) \leq \mathbb{E}\, f^{\beta}_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) + C\delta^2\Big(\beta + L^{\frac{p-2}{p-1}}\Big) + \frac{C\beta^K}{L^{2-\eta}} + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec\sigma\|_2^p \,\mathrm{d}\,\vec\sigma - \frac{3\kappa\log 2\delta}{\beta}.$$

If we write Φ for the function (10.3) associated with the inverse temperature $\beta>0$, the Lagrange multiplier λ^β, the measure α^β and the path π^β, then proposition 10.2 and proposition 10.3 imply that

$$\begin{aligned}\Psi(0,0) &= \mathbb{E}\Big[f^{\infty}_{\lambda^\beta}(\boldsymbol{X}^\beta(1)) - \int_0^1 \zeta^\beta(s)\big((\pi^\beta)'(s)\boldsymbol{v}^\beta(s),\boldsymbol{v}^\beta(s)\big)\,\mathrm{d}\,s\Big]\\ &\leq \Phi(0,0) + C\delta^2\Big(\beta + L^{\frac{p-2}{p-1}}\Big) + \frac{C\beta^K}{L^{2-\eta}} + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec\sigma\|_2^p\,\mathrm{d}\,\vec\sigma - \frac{3\kappa\log 2\delta}{\beta}.\end{aligned}$$

It follows that

$$\inf_{\lambda,\zeta,\pi}\mathcal{P}_\infty(\lambda,\zeta,\pi) \leq \mathcal{P}_\infty(\lambda^\beta,\zeta^\beta,\pi^\beta) \leq \mathcal{P}_\beta(\lambda^\beta,\alpha^\beta,\pi^\beta) + C\delta^2\Big(\beta + L^{\frac{p-2}{p-1}}\Big) + \frac{C\beta^K}{L^{2-\eta}} + \frac{t}{(2\delta)^\kappa}\int_{[-\delta,\delta]^\kappa}\|\vec\sigma\|_2^p\,\mathrm{d}\,\vec\sigma - \frac{3\kappa\log 2\delta}{\beta}.$$

Taking $L=\beta^m$ and $\delta = \frac{1}{\beta^{(m+1)/2}}$ for $m=\frac{K+1}{2-\eta}$ and letting $\beta\to\infty$ establishes (11.8) and completes the proof. ☐

A Terminal conditions growth bounds

In this appendix we include the technical bounds on the terminal conditions f^β_λ and f^∞_λ in (8.5) and (1.36) that make it possible to define (1.41) and (8.7). We also show that these bounds are preserved by the iterative procedure used to define X_0 and Y_0 in (8.7) and (1.41), which plays an instrumental role in the proofs of proposition 10.2 and proposition 10.3.

Lemma A.1. *If $2<p<\infty$, $\beta>0$, $\vec x\in\mathbb{R}^\kappa$, $\lambda\in\mathbb{R}^{\kappa(\kappa+1)/2}$ and $t>0$, then*

$$|f^\beta_\lambda(\vec x)| \leq C\Big(1 + \|\vec x\|_2^{1+\frac{1}{p-1}}\Big), \tag{A.1}$$

$$\|\nabla f^\beta_\lambda(\vec x)\|_\infty \leq C\Big(1 + \|\vec x\|_2^{\frac{1}{p-1}}\Big) \tag{A.2}$$

for some constant $C>0$ that depends only on β,κ,λ,p and t.

Proof. Fix $1\leq i\leq\kappa$. Since $p>2$, a direct computation shows that

$$\partial_{x_i} f^\beta_\lambda(\vec x) = \frac{\int_{\mathbb{R}^\kappa}\sigma(i)\exp\beta\Big((\vec\sigma,\vec x) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec\sigma\|_2^p\Big)\,\mathrm{d}\,\vec\sigma}{\int_{\mathbb{R}^\kappa}\exp\beta\Big((\vec\sigma,\vec x) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec\sigma\|_2^p\Big)\,\mathrm{d}\,\vec\sigma}.$$

To simplify notation, write (I) for the numerator and (II) for the denominator in this expression. Introduce the set

$$A = \Big\{\vec{\sigma} \in \mathbb{R}^\kappa \mid \|\vec{x}\|_2 < \frac{t}{2}\|\vec{\sigma}\|_2^{p-1}\Big\} = \Big\{\vec{\sigma} \in \mathbb{R}^\kappa \mid \|\vec{\sigma}\|_2 > \Big(\frac{2\|\vec{x}\|_2}{t}\Big)^{\frac{1}{p-1}}\Big\}.$$

On the one hand, the Cauchy-Schwarz inequality implies that

$$\begin{aligned}
(I) &\leq \int_A \|\vec{\sigma}\|_2 \exp\beta\Big(\|\vec{\sigma}\|_2\|\vec{x}\|_2 + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}\\
&\quad + \int_{A^c}\|\vec{\sigma}\|_2 \exp\beta\Big(\big(\vec{\sigma},\vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}\\
&\leq \int_{\mathbb{R}^\kappa}\|\vec{\sigma}\|_2 \exp\beta\Big(\sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - \frac{t}{2}\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}\\
&\quad + \Big(\frac{2\|\vec{x}\|_2}{t}\Big)^{\frac{1}{p-1}}\int_{\mathbb{R}^\kappa}\exp\beta\Big(\big(\vec{\sigma},\vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}.
\end{aligned}$$

On the other hand, Jensen's inequality and symmetry yield

$$\begin{aligned}
(II) &\geq \int_{[-\frac{1}{2},\frac{1}{2}]^\kappa}\exp\beta\Big(\big(\vec{\sigma},\vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}\\
&\geq \exp\beta\int_{[-\frac{1}{2},\frac{1}{2}]^\kappa}\Big(\big(\vec{\sigma},\vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}\\
&= \exp\beta\int_{[-\frac{1}{2},\frac{1}{2}]^\kappa}\Big(\sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p\Big)\,\mathrm{d}\,\vec{\sigma}.
\end{aligned}$$

Combining these two bounds gives (A.2). The fundamental theorem of calculus reveals that

$$\begin{aligned}
|f_\lambda^\beta(\vec{x})| - |f_\lambda^\beta(0)| &\leq \int_0^1 |\partial_s f_\lambda^\beta(s\vec{x})|\,\mathrm{d}\,s \leq \int_0^1 |(\nabla f_\lambda^\beta(s\vec{x}),\vec{x})|\,\mathrm{d}\,s\\
&\leq \int_0^1 \|\nabla f_\lambda^\beta(s\vec{x})\|_2\|\vec{x}\|_2\,\mathrm{d}\,s \leq C\Big(1 + \|\vec{x}\|_2^{1+\frac{1}{p-1}}\Big)
\end{aligned}$$

which establishes (A.1) and completes the proof. □

Lemma A.2. *If $2 < p < \infty$, $\vec{x} \in \mathbb{R}^\kappa$, $\lambda \in \mathbb{R}^{\kappa(\kappa+1)/2}$ and $t > 0$, then*

$$|f_\lambda^\infty(\vec{x})| \leq C\Big(1 + \|\vec{x}\|_2^{1+\frac{1}{p-1}}\Big) \tag{A.3}$$

for some constant $C > 0$ that depends only on λ, p and t. Moreover, the function f_λ^∞ is differentiable for almost every $\vec{x} \in \mathbb{R}^\kappa$ with

$$\|\nabla f_\lambda^\infty(\vec{x})\|_\infty \leq C\Big(1 + \|\vec{x}\|_2^{\frac{1}{p-1}}\Big) \tag{A.4}$$

for a possibly different constant $C > 0$ that depends only on λ, p and t.

Proof. Consider the function $g : \mathbb{R}^\kappa \times \mathbb{R}^\kappa \to \mathbb{R}$ defined by

$$g(\vec{x},\vec{\sigma}) = \big(\vec{\sigma},\vec{x}\big) + \sum_{k\leq k'}\lambda_{k,k'}\sigma(k)\sigma(k') - t\|\vec{\sigma}\|_2^p.$$

Notice that $f_\lambda^\infty(\vec{x}) = \sup_{\vec{\sigma}\in\mathbb{R}^\kappa} g(\vec{x},\vec{\sigma})$. By lemma 9.2, there exists $\vec{\sigma}(\vec{x}) \in \mathbb{R}^\kappa$ with $f_\lambda^\infty(\vec{x}) = g(\vec{x},\vec{\sigma}(\vec{x}))$ and

$$\|\vec{\sigma}(\vec{x})\|_2 \leq \max\Big(\Big(\frac{2\|\lambda\|_\infty}{t}\Big)^{\frac{1}{p-2}}, \Big(\frac{2\|\vec{x}\|_2}{t}\Big)^{\frac{1}{p-1}}\Big). \tag{A.5}$$

It follows by the Cauchy-Schwarz inequality that

$$\begin{aligned} 0 = g(\vec{x},0) \leq f_\lambda^\infty(\vec{x}) &\leq \|\vec{\sigma}(\vec{x})\|_2\|\vec{x}\|_2 + \|\lambda\|_\infty\|\vec{\sigma}(\vec{x})\|_2^2 \\ &\leq \Big(\frac{2}{t}\Big)^{\frac{1}{p-1}}\|\vec{x}\|_2^{1+\frac{1}{p-1}} + \Big(\frac{2\|\lambda\|_\infty}{t}\Big)^{\frac{1}{p-2}}\|\vec{x}\|_2 + \\ &\quad + \Big(\frac{2}{t}\Big)^{\frac{2}{p-1}}\|\lambda\|_\infty\|\vec{x}\|_2^{\frac{2}{p-1}} + \Big(\frac{2\|\lambda\|_\infty}{t}\Big)^{\frac{2}{p-2}}\|\lambda\|_\infty \\ &\leq C\Big(1+\|\vec{x}\|_2^{1+\frac{1}{p-1}}\Big), \end{aligned} \tag{A.6}$$

where the last inequality uses the fact that $\frac{1}{p-1} < 1$. To establish (A.4), notice that $\vec{x} \mapsto g(\vec{x},\vec{\sigma})$ is convex for each $\vec{\sigma} \in \mathbb{R}^\kappa$. As the pointwise supremum of a family of convex functions, the function f_λ^∞ is also convex and therefore differentiable almost everywhere. If $\vec{x} \in \mathbb{R}^\kappa$ is a point of differentiability of f_λ^∞, then for any other $\vec{x}' \in \mathbb{R}^\kappa$,

$$f_\lambda^\infty(\vec{x}') - f_\lambda^\infty(\vec{x}) = \big(\vec{x}'-\vec{x}, \nabla f_\lambda^\infty(\vec{x})\big) + o\big(|\vec{x}'-\vec{x}|\big).$$

Combining this with the fact that

$$f_\lambda^\infty(\vec{x}') - f_\lambda^\infty(\vec{x}) \geq g(\vec{x}',\vec{\sigma}(\vec{x})) - g(\vec{x},\vec{\sigma}(\vec{x})) = \big(\vec{\sigma}(\vec{x}), \vec{x}'-\vec{x}\big)$$

yields

$$\big(\vec{x}'-\vec{x}, \nabla f_\lambda^\infty(\vec{x}) - \vec{\sigma}(\vec{x})\big) \geq o\big(|\vec{x}'-\vec{x}|\big).$$

This is only possible if $\nabla f_\lambda^\infty(\vec{x}) = \vec{\sigma}(\vec{x})$ so the result follows by (A.5). ☐

Lemma A.3. *Let $f : \mathbb{R}^\kappa \to \mathbb{R}$ be a convex and differentiable function with*

$$-M \leq f(\vec{x}) \leq C\big(1+\|\vec{x}\|_2^{a+1}\big) \quad \textit{and} \quad \|\nabla f(\vec{x})\|_2 \leq C\big(1+\|\vec{x}\|_2^a\big) \tag{A.7}$$

for some $a \in (0,1)$ and some constants $C, M > 0$. If $F : [0,1)\times\mathbb{R}^\kappa \to \mathbb{R}$ is defined by

$$F(s,\vec{x}) = \frac{1}{m}\log\mathbb{E}\exp mf(\vec{x}+A(s)g) \tag{A.8}$$

for some $m > 0$, a standard Gaussian vector g in $\mathbb{R}^\kappa$ and a non-negative definite matrix $A(s)$ with uniformly bounded norm, $\|A(s)\|_{\mathrm{HS}}^2 \leq C$, then there exists $C' > 0$ that depends on κ, a, C, m and M such that

$$\|\nabla F(s,\vec{x})\|_2 \leq C'\big(1+\|\vec{x}\|_2^a\big) \tag{A.9}$$

for all $(s,\vec{x}) \in [0,1)\times\mathbb{R}^\kappa$.

Proof. To simplify notation, write $C' > 0$ for a constant that depends on κ, a, C, m and M whose value might not be the same at each occurrence. Fix $s \in [0,1)\times\mathbb{R}^\kappa$ and $1 \leq l \leq \kappa$. Since $a \in (0,1)$,

$$\partial_{x_l}F(s,\vec{x}) = \frac{\mathbb{E}\,\partial_{x_l}f(\vec{x}+A(s)g)\exp mf(\vec{x}+A(s)g)}{\mathbb{E}\exp mf(\vec{x}+A(s)g)}.$$

With this in mind, consider the set $B = \big\{\|\vec{x}\|_2 \le \|A(s)g\|_2\big\}$. On the one hand,

$$\begin{aligned}\mathbb{E}|\partial_{x_l} f(\vec{x} + A(s)g)|e^{mf(\vec{x}+A(s)g)}\mathbf{1}_B &\le C'\Big(1 + \mathbb{E}\|A(s)g\|_2^a e^{mC'\|A(s)g\|_2^{a+1}}\Big)\\ &\le C'\Big(1 + \|A(s)\|_{\mathrm{HS}}^a\,\mathbb{E}\|g\|_2^a e^{mC'\|A(s)\|_{\mathrm{HS}}^{a+1}\|g\|_2^{a+1}}\Big)\\ &\le C'\Big(1 + \mathbb{E}\|g\|_2^a e^{C'\|g\|_2^{a+1}}\Big) \le C',\end{aligned}$$

where the last inequality uses the fact that $a + 1 < 2$. On the other hand, Jensen's inequality and the convexity of f imply that

$$e^{-mM} \le e^{mf(\vec{x})} = e^{mf(\vec{x}+\mathbb{E}\,A(s)g)} \le \mathbb{E}\,e^{mf(\vec{x}+A(s)g)}.$$

It follows that

$$\begin{aligned}|\partial_{x_l} F(x, \vec{x})| &\le C' + \frac{\mathbb{E}|\partial_{x_l} f(\vec{x} + A(s)g)|e^{mf(\vec{x}+A(s)g)}\mathbf{1}_B}{\mathbb{E}\,e^{mf(\vec{x}+A(s)g)}}\\ &\le C'\Big(1 + \frac{\mathbb{E}\|\vec{x}\|_2^a e^{mf(\vec{x}+A(s)g)}}{\mathbb{E}\,e^{mf(\vec{x}+A(s)g)}}\Big) \le C'\big(1 + \|\vec{x}\|_2^a\big).\end{aligned}$$

This completes the proof. □

B Background material

In this appendix we collect a number of elementary results from linear algebra.

Theorem B.1 (Gershgorin)**.** *If $A \in \mathbb{R}^{n\times n}$ and $R_i = \sum_{j\neq i}|a_{ij}|$ is the sum of the absolute values of the non-diagonal entries in the i'th row of A, then the eigenvalues of A are all contained in the union of the Gershgorin discs,*

$$G(A) = \bigcup_{i=1}^{n}\{z \in \mathbb{C} \mid |z - a_{ii}| \le R_i\}. \tag{B.1}$$

Proof. See theorem 6.1.1 in [15]. □

Corollary B.2. *If $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix with non-negative diagonal entries satisfying*

$$a_{ii} = |a_{ii}| \ge \sum_{j\neq i}|a_{ij}| \tag{B.2}$$

for $1 \le i \le n$, then A is non-negative definite.

Proof. In the notation of Gershgorin's theorem, condition (B.2) may be written as $a_{ii} \ge R_i$. It follows by the symmetry of A and Gershgorin's theorem that all the eigenvalues of A are non-negative. This completes the proof. □

Lemma B.3. *If $A \in \mathbb{R}^{n\times n}$ is a non-negative definite and symmetric matrix, then*

$$\|A\|_{\mathrm{HS}} \le \mathrm{tr}(A). \tag{B.3}$$

Proof. Let $\lambda_n \ge \lambda_{n-1} \ge \cdots \ge \lambda_1 \ge 0$ be the real and non-negative eigenvalues of the matrix A. Since the trace of a matrix is the sum of its eigenvalues,

$$\|A\|_{\mathrm{HS}}^2 = \mathrm{tr}(AA^T) = \mathrm{tr}(A^2) = \sum_{i=1}^{n}\lambda_i^2 \le \Big(\sum_{i=1}^{n}\lambda_i\Big)^2 = \mathrm{tr}(A)^2.$$

We have used the fact that the eigenvalues of A^2 are $\lambda_n^2 \ge \lambda_{n-1}^2 \ge \cdots \ge \lambda_1^2 \ge 0$ in the third equality and the non-negativity of the eigenvalues of A in the inequality. □

Lemma B.4. *If $A, B, C \in \mathbb{R}^{n\times n}$ are non-negative definite and symmetric matrices with $B \le C$, then $\mathrm{tr}(AB) \le \mathrm{tr}(AC)$. In particular, $\|B\|_{\mathrm{HS}} \le \|C\|_{\mathrm{HS}}$ whenever $B \le C$.*

Proof. Since A is symmetric and non-negative definite, there exists a symmetric and non-negative definite matrix M with $M^TM = A$. If $M = (m_1, \ldots, m_n)$, where $m_i \in \mathbb{R}^n$ denotes the i'th column of M, then

$$A = M^TM = \sum_{i=1}^{n} m_i m_i$$

The linearity and cyclicity of the trace imply that

$$\mathrm{tr}(AB) = \sum_{i=1}^{n} \mathrm{tr}(m_i m_i^T B) = \sum_{i=1}^{n} \mathrm{tr}(m_i^T B m_i) = \sum_{i=1}^{n} m_i^T B m_i.$$

Invoking the assumption that $B \le C$ yields $\mathrm{tr}(AB) \le \sum_{i=1}^{n} m_i C m_i^T = \mathrm{tr}(AC)$. To complete the proof observe that

$$\|B\|_{\mathrm{HS}}^2 = \mathrm{tr}(B^TB) \le \mathrm{tr}(B^TC) = \mathrm{tr}(C^TB) \le \mathrm{tr}(C^TC) = \|C\|_{\mathrm{HS}}^2,$$

where we have used the fact that the trace of a matrix coincides with the trace of its transpose. □

Lemma B.5. *If $A \in \mathbb{R}^{n\times n}$ is a symmetric and positive definite matrix and $P \in \mathbb{R}^{n\times n}$ is a symmetric matrix, then there exists $\epsilon^* = \epsilon^*(A, \|P\|_\infty, n) > 0$ such that*

$$A + \epsilon P \tag{B.4}$$

is symmetric and positive definite for every $\epsilon < \epsilon^*$.

Proof. Denote by λ_1 the smallest eigenvalue of A. For any $x \in \mathbb{R}^n$ and every $\epsilon > 0$,

$$x^T(A + \epsilon P)x \ge x^TAx - \epsilon\|P\|_\infty \|x\|_1^2 \ge (\lambda_1 - \epsilon n\|P\|_\infty)\|x\|_2^2.$$

Since $\lambda_1 > 0$ by positive definiteness of A, setting $\epsilon^* = \frac{\lambda_1}{n\|P\|_\infty}$ completes the proof. □

References

[1] Alon, N. and Naor, A. (2006). Approximating the cut-norm via Grothendieck's inequality. *SIAM journal on computing* **35** 787–803. MR2203567

[2] Anderson, G.W. and Guionnet, A. and Zeitouni, O. (2010). *An Introduction to Random Matrices*. Cambridge University Press. MR2760897

[3] Auffinger, A. and Chen, W.-K (2017). Parisi formula for the ground state energy in the mixed p-spin model. *Annals of Probability* **45** 4617–4631. MR3737919

[4] Auffinger, A. and Chen, W.-K (2015). The Parisi formula has a unique minimizer. *Communications in Mathematical Physics* **335** 1429–1444. MR3320318

[5] Charikar, M. and Wirth, A. (2004). Maximizing quadratic programs: extending Grothendieck's inequality. In 45'th *Annual IEEE Symposium on Foundations of Computer Science* 54–60.

[6] Chen, W.-K. and Sen, A. (2020). On ℓ_p-Gaussian-Grothendieck problem. Prepint. Available at arXiv:2012.09343.

[7] Chen, W.-K. (2017). Variational representations for the Parisi functional and the two-dimensional Guerra–Talagrand bound. *Annals of Probability* **45** 3929–3966. MR3729619

[8] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Springer-Verlag, New York. MR1619036

[9] Dudley, R.M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis* **1** 290–330. MR0220340

[10] Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal* **23** 298–305. MR0318007

[11] Gall, J.F.L. (2016). *Brownian Motion, Martingales, and Stochastic Calculus*. Springer International Publishing. MR3497465

[12] Ghirlanda, S. and Guerra, F. (1998). General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *Journal of Physics A* **31** 9149–9155. MR1662161

[13] Guerra, F. and Toninelli, F. (2002). The thermodynamic limit in mean field spin glass models. *Communications in Mathematical Physics* **230** 71–79. MR1930572

[14] Guruswami, V. and Raghavendra, P. and Saket, R. and Yi, W. (2016). Bypassing UGC from some optimal geometric inapproximability results. *ACM Transactions on Algorithms* **12** 1–25. MR3462842

[15] Horn, R.A. and Johnson, C.R. (2013). *Matrix Analysis*. Cambridge University Press. MR2978290

[16] Jagannath, A. and Tobasco, I. (2016). A dynamic programming approach to the Parisi functional. *Proceedings of the American Mathematical Society* **144** 3135–3150. MR3487243

[17] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York. MR1121940

[18] Kashin, B. and Szarek, S. (2003). On the Gram matrices of systems of uniformly bounded functions. *Tr. Mat. Inst. Steklova* **243** 237–243. MR2054436

[19] Khot, S. and Naor, A. (2012). Grothendieck-type inequalities in combinatorial optimization. *Communications on Pure and Applied Mathematics* **65** 992–1035. MR2922372

[20] Khot, S. and Safra, M. (2013). A two prover one round game with strong soundness. *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science* **9** 863–887. MR3142437

[21] Kindler, G. and Naor A. and Schechtman G. (2010). The UGC hardness threshold of the L_p Grothendieck problem. *Math. Oper. Res.* **35** 267–283. MR2674720

[22] Ko, J. (2020). Free energy of multiple systems of spherical spin glasses with constrained overlaps. *Electronic Journal of Probability* **25** 1–34. MR4073689

[23] Ko, J. (2020) *The free energy of spherical vector spin glasses*. PhD thesis, University of Toronto. MR4197608

[24] Megretski, A. (2001). Relaxation of quadratic programs in operator theory and system analysis. In *Systems, Approximation, Singular Integral Operators, and Related Topics* 365–392. MR1882703

[25] Montanari, A. (2019). Optimization of the Sherrington-Kirkpatrick Hamiltonian. In *IEEE 60th Annual Symposium on Foundations of Computer Science* 1417–1433. MR4228234

[26] Motzkin, T. and Straus, E. (1965). Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics* **17** 533–540. MR0175813

[27] Nemirovski, A. and Roos, C. and Terlaky, T. (1999). On maximization of quadratic form over intersection of ellipsoids with common centre *Math. Program.* **86** 463–473. MR1733748

[28] Panchenko, D. (2019). *Lecture Notes on Probability Theory*.

[29] Panchenko, D. (2013). *The Sherrington-Kirkpatrick model*. Springer-Verlag, New York. MR2988963

[30] Panchenko, D. (2005). Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.* **17** 793–857. MR2159369

[31] Panchenko, D. (2015). The free energy in a multi-species Sherrington-Kirkpatrick model. *Annals of Probability* **43** 3494–3513. MR3433586

[32] Panchenko, D. (2018). Free energy in the Potts spin glass. *Annals of Probability* **46** 829–864. MR3773375

[33] Panchenko, D. (2018). Free energy in the mixed p-spin models with vector spins. *Annals of Probability* **46** 865–896. MR3773376

[34] Rockafellar, R.T. (1970). *Convex analysis*. Princeton University Press. MR0274683

[35] Talagrand, M. (2014). *Upper and lower bounds for stochastic processes: modern methods and classical problems*. Springer Berlin Heidelberg. MR3184689

[36] Talagrand, M. (2010). *Mean Field Models for Spin Glasses: Bolume I: Basic Examples*. Springer Berlin Heidelberg. MR3024566

[37] Talagrand, M. (2011). *Mean Field Models for Spin Glasses: Volume II: Advanced Mean Field Models for Spin Glasses: Volume II: Advanced Replica-Symmetry and Low Temperature*. Springer Berlin Heidelberg. MR3024566

[38] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press. MR3837109