EXTENDING THE PARISI FORMULA ALONG A HAMILTON-JACOBI EQUATION

J.-C. MOURRAT AND D. PANCHENKO

ABSTRACT. We study the free energy of mixed p-spin spin glass models enriched with an additional magnetic field given by the canonical Gaussian field associated with a Ruelle probability cascade. We prove the conjecture in [15] that this free energy converges to the Hopf-Lax solution of a certain Hamilton-Jacobi equation. Using this result, we give a new representation of the free energy of mixed p-spin models with soft spins.

1. INTRODUCTION

Let $(\beta_p)_{p \geqslant 2}$ be a sequence of real numbers and let $\xi(r) \coloneqq \sum_{p \geqslant 2} \beta_p^2 r^p$ for $r \in \mathbb{R}$. We will assume that the sequence $(\beta_p)_{p \geqslant 2}$ is such that ξ is well defined on the entire real line. This assumption can be relaxed if needed by restricting the parameters of the models we will be working with. Denote by $(H_N(\sigma))_{\sigma \in \mathbb{R}^N}$ the centered Gaussian field with covariance

$$\mathbb{E}\left[H_N(\sigma) H_N(\tau)\right] = N\xi\Big(\frac{\sigma \cdot \tau}{N}\Big), \qquad \sigma, \tau \in \mathbb{R}^N.$$

Let $P_N \coloneqq P_1^{\otimes N}$ denote the N-fold product of P_1, a probability measure on $\mathbb{R}$ with bounded support. We aim to study the Gibbs measure built with respect to the energy function $H_N(\sigma)$ and the reference measure P_N. A quantity of fundamental interest is the limit free energy

$$(1.1) \qquad \qquad \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log \int \exp\left(H_N(\sigma)\right) \mathrm{d}P_N(\sigma).$$

When the support of P_1 is $\{\pm 1\}$ and $\xi(r) = \beta^2 r^2$ (the Sherrington-Kirkpatrick model [27]), this limit was discovered by Parisi in a celebrated work [23, 24]; see also [12]. The formula was then proved rigorously for general ξ in [8, 30, 18], and was later extended to the current setting where we only assume that the support of P_1 is bounded in [16, 22].

In order to further our understanding of this object, it was proposed in [15] (following [13, 14]) to recast the limit free energy (1.1) as a particular value of the solution of a Hamilton-Jacobi equation. This solution depends on two parameters $t \geqslant 0$ and $\mu \in \mathcal{M}(\mathbb{R}_+)$, where $\mathcal{M}(\mathbb{R}_+)$ denotes the set of Borel probability measures over $\mathbb{R}_+$. It was conjectured that an enriched version of the free energy, which would depend additionally on the parameters $t \geqslant 0$ and $\mu \in \mathcal{M}(\mathbb{R}_+)$, may converge to the same solution evaluated at these parameters.

The main purpose of this paper is to prove this conjecture. In order to state the result, we start by defining the enriched model precisely. We denote by $\mathcal{M}_b(\mathbb{R}_+)$ the subset of $\mathcal{M}(\mathbb{R}_+)$ of measures with bounded support. By [19, Theorem 2.17], one

can associate a Ruelle probability cascade [25] to each probability measure on $[0,1]$; this Ruelle probability cascade is a random probability measure on the unit ball of a Hilbert space. We denote by $\mathfrak{R}$ the Ruelle probability cascade corresponding to the uniform distribution over $[0,1]$, and by $(\alpha^\ell)_{\ell \geqslant 1}$ an i.i.d. sample from $\mathfrak{R}$ (that is, the law of $(\alpha^\ell)_{\ell \geqslant 1}$ is $\mathfrak{R}^{\otimes \infty}$). In particular, the law of the overlap $\alpha^1 \cdot \alpha^2$ under $\mathbb{E}\mathfrak{R}^{\otimes 2}$ is the uniform distribution over $[0,1]$. Given a measure $\mu \in \mathcal{M}_b(\mathbb{R}_+)$, and conditionally on $\mathfrak{R}$, let $z^\mu(\alpha)$ be a Gaussian process indexed by $\alpha \in \mathrm{supp}(\mathfrak{R})$ with covariance

$$\mathbb{E}\left[z^\mu(\alpha^1)z^\mu(\alpha^2)\right] = \mu^{-1}(\alpha^1 \cdot \alpha^2), \qquad \alpha^1, \alpha^2 \in \mathrm{supp}(\mathfrak{R}).$$

In the expression above and throughout the paper, we use the shorthand notation, for every $r \in [0,1]$,

$$(1.2) \qquad\qquad \mu^{-1}(r) := \inf\{s \geqslant 0 \ : \ \mu([0,s]) \geqslant r\}.$$

To check that the Gaussian process $z^\mu(\alpha)$ exists, it suffices to verify that $\mu^{-1}(\alpha^1 \cdot \alpha^2)$ is a positive semidefinite kernel on $(\mathrm{supp}\,\mathfrak{R})^2$, and this follows from the fact that the support of $\mathfrak{R}$ is ultrametric. Moreover, $\mu^{-1}(r)$ is left-continuous and, thus, continuous at $r = 1$, which implies that the process $z^\mu(\alpha)$ is stochastically continuous on $\mathrm{supp}(\mathfrak{R})$. As a result, it is jointly measurable (see e.g. [6, Theorem 3.3.1]) and we can define, for every $t \geqslant 0$ and $\mu \in \mathcal{M}_b(\mathbb{R}_+)$,

$$(1.3) \quad F_N(t,\mu) := \frac{1}{N}\mathbb{E}\log \iint \exp\Big(\sqrt{t}H_N(\sigma) + \sum_{i=1}^{N}\sigma_i z_i^\mu(\alpha)$$
$$- \frac{1}{2}Nt\xi(N^{-1}|\sigma|^2) - \frac{1}{2}\mu^{-1}(1)|\sigma|^2\Big)\,\mathrm{d}P_N(\sigma)\mathrm{d}\mathfrak{R}(\alpha),$$

where $z_i^\mu(\alpha)$ are independent copies of $z^\mu(\alpha)$ for $i \geqslant 1$ (conditionally on $\mathfrak{R}$ and independent of H_N). For a measure of the form

$$(1.4) \qquad\qquad \mu = \sum_{\ell=0}^{k}(\zeta_{\ell+1} - \zeta_\ell)\delta_{q_\ell}$$

with

$$(1.5) \qquad 0 = \zeta_0 < \zeta_1 < \ldots < \zeta_k < \zeta_{k+1} = 1, \qquad 0 \leqslant q_0 < q_1 < \ldots < q_k < \infty,$$

one can rewrite $F_N(t,\mu)$ in the more familiar form

$$(1.6) \quad F_N(t,\mu) = \frac{1}{N}\mathbb{E}\log \int_{\mathbb{R}^N} \sum_{\alpha \in \mathbb{N}^k} \exp\Big(\sqrt{t}H_N(\sigma) + \sum_{i=1}^{N}\sigma_i z_{\alpha,i}$$
$$- \frac{1}{2}Nt\xi(N^{-1}|\sigma|^2) - \frac{1}{2}\mu^{-1}(1)|\sigma|^2\Big)v_\alpha\,\mathrm{d}P_N(\sigma),$$

where $(v_\alpha)_{\alpha \in \mathbb{N}^k}$ are the weights of the Ruelle probability cascade with parameters $(\zeta_\ell)_{1 \leqslant \ell \leqslant k}$, and $(z_{\alpha,i})_{i \geqslant 1}$ are independent copies of the Gaussian process with the covariance $\mathbb{E}z_{\alpha^1}z_{\alpha^2} = q_{\alpha^1 \wedge \alpha^2}$, where $\alpha^1 \wedge \alpha^2 = \max\{\ell \geqslant 0 \ : \ \alpha_j^1 = \alpha_j^2 \text{ for } j \leqslant \ell\}$. The quantities (1.3) and (1.6) are equal in this case, because, by (the proof of) [19, Theorem 1.3] and standard properties of the Ruelle probability cascades, see [19, Theorem 4.4], both quantities are equal to the same continuous functional of the distribution of the array $(\mu^{-1}(\alpha^\ell \cdot \alpha^{\ell'}))_{\ell,\ell' \geqslant 1}$ under $\mathbb{E}\mathfrak{R}^{\otimes \infty}$ and correspondingly of the array $(q_{\alpha^\ell \wedge \alpha^{\ell'}})_{\ell,\ell' \geqslant 1}$ under $\mathbb{E}(\sum_{\alpha \in \mathbb{N}^k} v_\alpha \delta_\alpha)^{\otimes \infty}$; and these distributions are equal, due to the property of the Ruelle probability cascades that the distribution of an overlap array is determined by the distribution of one overlap. Moreover, denoting by $D \geqslant 0$ the smallest real number such that the support of P_1 is contained in $[-\sqrt{D}, \sqrt{D}]$,

one can check (see for instance [29], [15, Proposition 2.1], or Subsection 3.5 below) that for every $\mu, \nu \in \mathcal{M}_b(\mathbb{R}_+)$,

$$(1.7) \qquad |F_N(t, \mu) - F_N(t, \nu)| \leqslant \frac{D}{2} \int_{\mathbb{R}} |\mu(s) - \nu(s)| \, \mathrm{d}s.$$

In view of this, we can whenever convenient replace the measure μ by an atomic measure. Finally, using again standard properties of the Ruelle probability cascades (see e.g. [31, Theorem 14.2.1] or [19, Theorem 2.9]), one can verify that $F_N(0, \mu)$ does not depend on N; we denote this quantity by

$$(1.8) \qquad \psi(\mu) \coloneqq F_N(0, \mu) = F_1(0, \mu).$$

We will recall a somewhat more explicit expression for $\psi(\mu)$ in (2.3) below. Denote by U a uniform random variable over $[0, 1]$, and for every probability measure μ on $\mathbb{R}_+$, define $X_\mu \coloneqq \mu^{-1}(U)$, where we recall that μ^{-1} is defined in (1.2). We also define, for every $s \in \mathbb{R}$,

$$\xi^*(s) \coloneqq \sup_{r \geqslant 0} \left(rs - \xi(r) \right).$$

Our first goal is to prove the following conjecture from [15] (specialized to the case where P_N is a product measure).

Theorem 1.1. *For every $t \geqslant 0$ and $\mu \in \mathcal{M}_b(\mathbb{R}_+)$,*

$$(1.9) \qquad \lim_{N \to \infty} F_N(t, \mu) = \inf_{\nu \in \mathcal{M}_b(\mathbb{R}_+)} \left(\psi(\nu) + \frac{t}{2} \mathbb{E} \xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right).$$

The motivation in [15] for this statement is that the right side of (1.9), seen as a function of (t, μ), solves the formal Hamilton-Jacobi equation

$$(1.10) \qquad \begin{cases} 2\partial_t f + \int \xi(-2\partial_\mu f) \, \mathrm{d}\mu = 0 & \text{on } \mathbb{R}_+ \times \mathcal{M}(\mathbb{R}_+), \\ f(0, \cdot) = \psi & \text{on } \mathcal{M}(\mathbb{R}_+). \end{cases}$$

For discrete μ as in (1.4), one can check that

$$(1.11) \quad 2\partial_t F_N + \int \xi(-2\partial_\mu F_N) \, \mathrm{d}\mu = \mathbb{E} \langle \xi(R_{1,2}) \rangle_t - \sum_{\ell=0}^{k} p_\ell \, \xi \left(p_\ell^{-1} \mathbb{E} \left\langle R_{1,2} \mathbb{1}_{\{\alpha^1 \wedge \alpha^2 = \ell\}} \right\rangle_t \right),$$

where $R_{1,2} \coloneqq N^{-1} \sigma^1 \cdot \sigma^2$ is the overlap of σ^1 and σ^2,

$$p_\ell \coloneqq \mu(\{q_\ell\}) = \zeta_{\ell+1} - \zeta_\ell = \mathbb{E} \left\langle \mathbb{1}_{\{\alpha^1 \wedge \alpha^2 = \ell\}} \right\rangle$$

(see e.g. [15, Lemma 2.3]), and $\langle \cdot \rangle$ denotes the average with respect to the Gibbs measure

$$\mathrm{d}G_N(\sigma, \alpha) \sim \exp \left(\sqrt{t} H_N(\sigma) + \sum_{i=1}^{N} \sigma_i z_i^\mu(\alpha) \right.$$
$$\left. - \frac{1}{2} N t \xi(N^{-1}|\sigma|^2) - \frac{1}{2} \mu^{-1}(1)|\sigma|^2 \right) v_\alpha \, \mathrm{d}P_N(\sigma).$$

When ξ is the square function, the right side of (1.11) can be interpreted as the conditional variance of the σ-overlap $R_{1,2}$ given the α-overlap $\alpha^1 \cdot \alpha^2$. More generally, the right side of (1.11) is small if and only if the conditional distribution of the overlap $R_{1,2}$ given $\alpha^1 \cdot \alpha^2$ is concentrated. This evokes the synchronization phenomenon used in the proof of the Parisi formula by Talagrand in [30] along Guerra's interpolation [8] with nearly optimal parameters; see also [31]. The idea of using the Hamilton-Jacobi techniques to study replica symmetric solution of the SK model was already utilized in [7], and one-step replica symmetry breaking analogues of the equation (1.11) were derived and studied in various models in [3, 1].

The main step in the proof of Theorem 1.1, which is to pass to the limit $N \to \infty$ for the left side of (1.9) and get some expression for the limit, is almost identical to the argument in [22] (specialized to the one-dimensional case), so we only outline the necessary modifications. The main tool is the synchronization mechanism developed in [20, 21, 22] based on the overlap ultrametricity proved in [17] for measures that satisfy the Ghirlanda-Guerra identities (and the fact that one has a lot of flexibility in enforcing these identities by way of small perturbations). The synchronization has been applied in a variety of situations, e.g. [9, 4, 11], and here we demonstrate another application. A particular synchronization that will be needed here is the one that forces the overlaps $\mu^{-1}(\alpha^1 \cdot \alpha^2)$ and $R_{1,2} = N^{-1}\sigma^1 \cdot \sigma^2$ to be deterministic functions of their sum in the thermodynamic limit. Notice that we need to use a synchronization argument here even in the case of Ising spins.

The reader may rightfully wonder what to make of the term $\xi(N^{-1}|\sigma|^2)$ appearing in the exponential in (1.3), which was introduced for convenience but is otherwise a nuisance (except in the case of Ising spins, where it is deterministic and therefore causes no harm). The second goal of this paper is to explain how to remove this term and deduce from Theorem 1.1 the limit of the "untampered" free energy in (1.1). At present this is perhaps not as interseting as it sounds, since the proof of Theorem 1.1 could be modified to obtain the limit of the quantity without the term $\xi(N^{-1}|\sigma|^2)$ directly. However, it is likely that a more direct proof of Theorem 1.1 exists, in which case it is important to notice that Theorem 1.1 is indeed all the information needed to conclude. Moreover, we obtain in this way a somewhat different expression for the limit in (1.1) than that obtained in [16, 22].

In order to state this second result, we introduce two more parameters to the energy and write, for every $s, t \geqslant 0$, $\mu \in \mathcal{M}_b(\mathbb{R}_+)$ and $h \in \mathbb{R}$,

$$(1.12) \quad \mathsf{F}_N(s, t, \mu, h) := \frac{1}{N}\mathbb{E}\log \iint \exp\Big(\sqrt{t}H_N(\sigma) + \sum_{i=1}^{N}\sigma_i z_i^\mu(\alpha)$$
$$- \frac{1}{2}N(t-s)\xi(N^{-1}|\sigma|^2) - \frac{1}{2}\mu^{-1}(1)|\sigma|^2 + h|\sigma|^2\Big)\,\mathrm{d}P_N(\sigma)\mathrm{d}\mathfrak{R}(\alpha).$$

Notice that when $s = 0$, this quantity is of the form covered by Theorem 1.1, up to a redefinition of P_N to absorb the term $\exp(h|\sigma|^2)$. We denote

$$(1.13) \qquad\qquad \Psi(\mu, h) := \mathsf{F}_1(0, 0, \mu, h) = \mathsf{F}_N(0, 0, \mu, h).$$

Theorem 1.2. *For every $s, t \geqslant 0$, $\mu \in \mathcal{M}_b(\mathbb{R}_+)$ and $h \in \mathbb{R}$, we have*

$$\lim_{N\to\infty}\mathsf{F}_N(s, t, \mu, h) = \sup_{h'\in\mathbb{R}}\inf_{\nu\in\mathcal{M}_b(\mathbb{R}_+)}\left(\Psi(\nu, h') + \frac{t}{2}\mathbb{E}\xi^*\left(\frac{X_\nu - X_\mu}{t}\right) - \frac{s}{2}\xi^*\left(\frac{2(h'-h)}{s}\right)\right).$$

The intuition for this result is simple, and consists in writing the Hopf-Lax formula for the equation

$$(1.14) \qquad\qquad\qquad 2\partial_s\mathsf{F}_N - \xi(\partial_h\mathsf{F}_N) \simeq 0.$$

By setting $s = t = 1$, $\mu = \delta_0$, and $h = 0$ in Theorem 1.2, we thus get the following new representation for the free energy of models with soft spins.

Corollary 1.3. *The limit free energy can be written as*

$$(1.15) \quad \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\log\int\exp\left(H_N(\sigma)\right)\,\mathrm{d}P_N(\sigma)$$
$$= \sup_{h\in\mathbb{R}}\inf_{\nu\in\mathcal{M}_b(\mathbb{R}_+)}\left(\Psi(\nu, h) + \frac{1}{2}\int_{\mathbb{R}_+}\xi^*(r)\,\mathrm{d}\nu(r) - \frac{1}{2}\xi^*(2h)\right).$$

Organization of the paper. In order to prove Theorem 1.1, we first state a different expression for the left side of (1.9) in Proposition 2.1 below. We then rewrite it in the form of the right side of (1.9) in Section 3, by reasoning similarly to what was done in [15] in the case $\mu = \delta_0$. We next turn to the proof of Proposition 2.1 in Section 4. Finally, we provide the proof of Theorem 1.2 in Section 5.

2. Parisi formula

In this section, we present the structure of the argument for identifying the limit on the left side of (1.9) in the more "classical" form in which Parisi formulas are usually stated. As a preparation for stating the formula we will obtain, we provide with an alternative description of the quantities ψ and Ψ appearing in the main statements of Section 1. Given a probability measure ν on $\mathbb{R}_+$, we write $\nu(s) := \nu([0, s])$. For every $\nu \in \mathcal{M}_b(\mathbb{R}_+)$ and $\lambda \in \mathbb{R}$, we denote by $\Phi_{\nu,\lambda} = \Phi_{\nu,\lambda}(t, x)$ the solution of the equation

$$(2.1) \qquad \begin{cases} \partial_t \Phi_{\nu,\lambda} = -\dfrac{1}{2}\big(\partial_x^2 \Phi_{\nu,\lambda} + \nu(t)(\partial_x \Phi_{\nu,\lambda})^2\big) & \text{on } [0, \nu^{-1}(1)] \times \mathbb{R}, \\[2mm] \Phi_{\nu,\lambda}(\nu^{-1}(1), x) = \log \displaystyle\int_{\mathbb{R}} \exp\big(\sigma x + \lambda \sigma^2\big)\, \mathrm{d}P_1(\sigma) & \text{for } x \in \mathbb{R}, \end{cases}$$

and we set

$$(2.2) \qquad \mathcal{P}(\nu, \lambda) := \Phi_{\nu,\lambda}(0, 0).$$

Using classical properties of Ruelle probability cascades, one can verify that the functions ψ and Ψ defined in (1.8) and (1.13) respectively satisfy, for every $\mu \in \mathcal{M}_b(\mathbb{R}_+)$ and $h \in \mathbb{R}$,

$$(2.3) \qquad \Psi(\mu, h) = \mathcal{P}\left(\mu, h - \frac{1}{2}\mu^{-1}(1)\right) \quad \text{and} \quad \psi(\mu) = \Psi(\mu, 0).$$

Given a probability measure ζ on $\mathbb{R}_+$, let ζ_μ denote the probability measure on $\mathbb{R}_+$ whose cumulative distribution function satisfies

$$(2.4) \qquad \zeta_\mu^{-1}(x) := \xi'(\zeta^{-1}(x)) + \mu^{-1}(x).$$

In other words, the c.d.f. of ζ_μ is

$$(2.5) \qquad \zeta_\mu := (\xi' \circ \zeta^{-1} + \mu^{-1})^{-1}.$$

Finally, let $\mathcal{M}_{0,u} = \mathcal{M}([0, u])$ denote the space of probability measures on $[0, u]$.

Notice that it suffices to prove Theorem 1.1 for $t = 1$. Indeed, once the result is known in this case, we recover the general statement by replacing ξ with $t\xi$. The main step towards the proof of Theorem 1.1 is the following result. For σ_1 is distributed according to P_1, we denote by d and D the smallest and largest points of the support of the distribution of σ_1^2. We also write, for every $r \in \mathbb{R}$,

$$\theta(r) := r\xi'(r) - \xi(r).$$

Proposition 2.1. *For every $\mu \in \mathcal{M}_b(\mathbb{R}_+)$, we have*

$$(2.6) \quad \lim_{N \to \infty} F_N(1, \mu) = \sup_{u \in [d, D]} \inf_{\substack{\zeta \in \mathcal{M}_{0,u} \\ \lambda \in \mathbb{R}}} \left[-\lambda u + \mathcal{P}\left(\zeta_\mu, \lambda + \frac{\xi'(u) - \xi'(\zeta^{-1}(1))}{2}\right) \right.$$

$$\left. - \frac{1}{2}\int_0^u \zeta(s)\, \mathrm{d}\theta(s) - \frac{1}{2}\xi(u) - \frac{1}{2}\mu^{-1}(1)u \right].$$

We now outline the structure of the argument for obtaining Proposition 2.1. By the definitions of d and D, when $\sigma \sim P_N = P_1^{\otimes N}$, we have that $N^{-1}|\sigma|^2 \in [d, D]$, and any point $u \in [d, D]$ can be approximated by some $N^{-1}|\sigma|^2$ for large N and $\sigma \in \text{supp}\, P_N$. For every $u \in [d, D]$ and $\varepsilon > 0$, let

$$(2.7) \qquad \Omega_N^\varepsilon(u) = \left\{ \sigma \in \mathbb{R}^N \,:\, N^{-1}|\sigma|^2 \in (u - \varepsilon, u + \varepsilon) \right\},$$

and consider

$$F_N^\varepsilon(u) := \frac{1}{N} \mathbb{E} \log \iint_{\Omega_N^\varepsilon(u)} \exp\Big(H_N(\sigma) + \sum_{i=1}^N \sigma_i z_i^\mu(\alpha)$$
$$- \frac{1}{2} N \xi(N^{-1}|\sigma|^2) - \frac{1}{2}\mu^{-1}(1)|\sigma|^2 \Big) \, \mathrm{d}P_N(\sigma) \, \mathrm{d}\mathfrak{R}(\alpha).$$

The measure μ will be fixed throughout, so we keep the dependency of $F_N^\varepsilon(u)$ on μ implicit in the notation. It is clear that, denoting

$$(2.8) \qquad p_N^\varepsilon(u) := \frac{1}{N} \mathbb{E} \log \iint_{\Omega_N^\varepsilon(u)} \exp\Big(H_N(\sigma) + \sum_{i=1}^N \sigma_i z_i^\mu(\alpha) \Big) \, \mathrm{d}P_N(\sigma) \, \mathrm{d}\mathfrak{R}(\alpha),$$

we have that as $\varepsilon > 0$ tends to zero,

$$(2.9) \qquad F_N^\varepsilon(u) := p_N^\varepsilon(u) - \frac{1}{2}\xi(u) - \frac{1}{2}\mu^{-1}(1)u + \mathcal{O}(\varepsilon).$$

Proposition 2.1 is a direct consequence of the following result.

Theorem 2.2. *For every $u \in [d, D]$,*

$$\lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} p_N^\varepsilon(u) = \inf_{\substack{\zeta \in \mathcal{M}_{0,u} \\ \lambda \in \mathbb{R}}} \Big[-\lambda u + \mathcal{P}\Big(\zeta_\mu, \lambda + \frac{\xi'(u) - \xi'(\zeta^{-1}(1))}{2} \Big) - \frac{1}{2} \int_0^u \zeta(s) \, \mathrm{d}\theta(s) \Big].$$

The proof of Theorem 2.2 will be given in Section 4. Before doing this, we show in the next section how to deduce Theorem 1.1 from Proposition 2.1.

3. Hopf-Lax representation

In this section, we take the validity of Proposition 2.1 for granted, and show that it implies Theorem 1.1. We decompose the argument into five subsections.

3.1. Change of variables. In (2.6), let us make the change of variables

$$\lambda \to \lambda - \frac{\xi'(u) - \xi'(\zeta^{-1}(1))}{2} - \frac{1}{2}\zeta_\mu^{-1}(1) = \lambda - \frac{1}{2}\xi'(u) - \frac{1}{2}\mu^{-1}(1),$$

where the equality follows from (2.4) with $x = 1$. By the definition of Ψ in (2.3), under this change of variables the second term in (2.6) becomes

$$\mathcal{P}\Big(\zeta_\mu, \lambda - \frac{1}{2}\zeta_\mu^{-1}(1) \Big) = \Psi(\zeta_\mu, \lambda),$$

and by cancelling and grouping other terms (recall that $\theta(u) = u\xi'(u) - \xi(u)$),

$$\lim_{N \to \infty} F_N = \sup_{u \in [d,D]} \inf_{\substack{\zeta \in \mathcal{M}_{0,u} \\ \lambda \in \mathbb{R}}} \Big[-\lambda u + \Psi(\zeta_\mu, \lambda) + \frac{1}{2}\theta(u) - \frac{1}{2} \int_0^u \zeta(s) \, \mathrm{d}\theta(s) \Big],$$

where $F_N = F_N(1, \mu)$. An integration by parts gives that

$$\lim_{N \to \infty} F_N = \sup_{u \in [d,D]} \inf_{\substack{\zeta \in \mathcal{M}_{0,u} \\ \lambda \in \mathbb{R}}} \left[-\lambda u + \Psi(\zeta_\mu, \lambda) + \frac{1}{2} \int_0^u \theta(s) \, d\zeta(s) \right].$$

Since $\xi^*(\xi'(s)) = \theta(s)$ for $s \geqslant 0$, if U is a uniform random variable over $[0, 1]$, we can write (again, recall (2.4))

$$\int_0^u \theta(s) \, d\zeta(s) = \int_0^u \xi^*(\xi'(s)) \, d\zeta(s) = \mathbb{E}\xi^*(\xi'(\zeta^{-1}(U))) = \mathbb{E}\xi^*(\zeta_\mu^{-1}(U) - \mu^{-1}(U)),$$

and thus

$$\lim_{N \to \infty} F_N = \sup_{u \in [d,D]} \inf_{\substack{\zeta \in \mathcal{M}_{0,u} \\ \lambda \in \mathbb{R}}} \left[-\lambda u + \Psi(\zeta_\mu, \lambda) + \frac{1}{2} \mathbb{E}\xi^*(\zeta_\mu^{-1}(U) - \mu^{-1}(U)) \right].$$

Notice that as ζ varies in $\in \mathcal{M}_{0,u}$, the measures ζ_μ defined in (2.5) span the set

$$\mathcal{D}_u = \left\{ \nu \in \mathcal{M}_b(\mathbb{R}_+) : \forall s, \nu(s) \leqslant \mu(s); \operatorname{supp}(\nu) \subseteq [0, \xi'(u) + \mu^{-1}(1)] \right\},$$

which means that, recalling that we write $X_\nu = \nu^{-1}(U)$,

(3.1)
$$\lim_{N \to \infty} F_N = \sup_{u \in [d,D]} \inf_{\substack{\nu \in \mathcal{D}_u \\ \lambda \in \mathbb{R}}} \left[-\lambda u + \Psi(\nu, \lambda) + \frac{1}{2} \mathbb{E}\xi^*(X_\nu - X_\mu) \right].$$

3.2. **Removing the constraint on the support.** Let us first show that we can remove the constraint $\operatorname{supp}(\nu) \subseteq [0, \xi'(u) + \mu^{-1}(1)]$ in $\mathcal{D}_u$. without changing the value of the right side of (3.1). For $\nu \in \mathcal{M}_b(\mathbb{R}_+)$, let $\widetilde{\nu}(s) = 1$ for $s \geqslant \xi'(u) + \mu^{-1}(1)$ and $\widetilde{\nu}(s) = \nu(s)$ otherwise. This corresponds to the truncation

$$X_{\widetilde{\nu}} = \min(X_\nu, \xi'(u) + \mu^{-1}(1)).$$

Let us show that

(3.2)
$$\xi^*(X_\nu - X_\mu) \geqslant \xi^*(X_{\widetilde{\nu}} - X_\mu) + u|X_\nu - X_{\widetilde{\nu}}|.$$

If $X_{\widetilde{\nu}} = X_\nu$ then the two sides are equal. Otherwise, $X_\nu > X_{\widetilde{\nu}} = \xi'(u) + \mu^{-1}(1)$. Since $\mu^{-1}(1) \geqslant X_\mu$, this implies that $X_\nu - X_\mu > X_{\widetilde{\nu}} - X_\mu \geqslant \xi'(u)$. It remains to observe that $\xi^{*\prime}(s) = \xi'^{-1}(s) \geqslant u$ if $s \geqslant \xi'(u)$, so (3.2) holds and

(3.3)
$$\mathbb{E}\xi^*(X_\nu - X_\mu) \geqslant \mathbb{E}\xi^*(X_{\widetilde{\nu}} - X_\mu) + u\,\mathbb{E}|X_\nu - X_{\widetilde{\nu}}|.$$

On the other hand, if we define

(3.4)
$$\Gamma_u(\nu) := \inf_{\lambda \in \mathbb{R}} \left(-\lambda u + \Psi(\nu, \lambda) \right),$$

we will show below in Subsection 3.5 below that

(3.5)
$$|\Gamma_u(\nu) - \Gamma_u(\widetilde{\nu})| \leqslant \frac{u}{2} \mathbb{E}|X_\nu - X_{\widetilde{\nu}}|.$$

The last two inequalities imply that

$$\Gamma_u(\widetilde{\nu}) + \frac{1}{2} \mathbb{E}\xi^*(X_{\widetilde{\nu}} - X_\mu) \leqslant \Gamma_u(\nu) + \frac{1}{2} \mathbb{E}\xi^*(X_\nu - X_\mu),$$

which means that the constraint $\operatorname{supp}(\nu) \subseteq [0, \xi'(u) + \mu^{-1}(1)]$ can be removed and

(3.6)
$$\lim_{N \to \infty} F_N = \sup_{u \in [d,D]} \inf_{\substack{\nu \in \mathcal{D}_\mu \\ \lambda \in \mathbb{R}}} \left[-\lambda u + \Psi(\nu, \lambda) + \frac{1}{2} \mathbb{E}\xi^*(X_\nu - X_\mu) \right],$$

where $\mathcal{D}_\mu := \left\{ \nu \in \mathcal{M}_b(\mathbb{R}_+) : \forall s, \nu(s) \leqslant \mu(s) \right\}$.

3.3. **Using convexity.** Since $X_\nu - X_\mu \geqslant 0$ for $\nu \in \mathcal{D}_\mu$, for such ν we have that $\mathbb{E}\xi^*(X_\nu - X_\mu) = \mathbb{E}\xi^*(|X_\nu - X_\mu|)$. Since $\xi^*(|s|)$ is convex and symmetric, $\mathbb{E}\xi^*(|X_\nu - X_\mu|)$ is a Wasserstein distance between μ and ν with the cost function $\xi^*(|x - y|)$ (see e.g. [32, Theorem 2.18 and Remark 2.19(ii)]) and, therefore, $\mathbb{E}\xi^*(X_\nu - X_\mu)$ is convex in ν on $\mathcal{D}_\mu$. Also, the Auffinger-Chen representation [2] for the solution of equation (2.1) (see also [10]) implies that $\Psi(\nu, \lambda)$ is convex in (ν, λ). Therefore, by Sion's minimax theorem, see [26, Corollary 3.3], we have

$$(3.7) \qquad \lim_{N \to \infty} F_N = \inf_{\substack{\nu \in \mathcal{D}_\mu \\ \lambda \in \mathbb{R}}} \sup_{u \in [d,D]} \Big[-\lambda u + \Psi(\nu, \lambda) + \frac{1}{2}\mathbb{E}\xi^*(X_\nu - X_\mu) \Big].$$

Using that the boundary condition in (2.1) satisfies

$$\log \int e^{\sigma x + \lambda \sigma^2} \, \mathrm{d}P_1(\sigma) \geqslant \log \int e^{\sigma x} \, \mathrm{d}P_1(\sigma) + \lambda d \mathbb{1}_{\{\lambda \geqslant 0\}} + \lambda D \mathbb{1}_{\{\lambda \leqslant 0\}},$$

we see that

$$\sup_{u \in [d,D]} \Big[-\lambda u + \Psi(\nu, \lambda) \Big] = -\lambda d \mathbb{1}_{\{\lambda \geqslant 0\}} - \lambda D \mathbb{1}_{\{\lambda \leqslant 0\}} + \Psi(\nu, \lambda) \geqslant \Psi(\nu, 0).$$

This implies that infimum over λ is achieved at $\lambda = 0$ and, recalling that $\psi(\nu) = \Psi(\nu, 0)$,

$$(3.8) \qquad \lim_{N \to \infty} F_N = \inf_{\nu \in \mathcal{D}_\mu} \Big[\psi(\nu) + \frac{1}{2}\mathbb{E}\xi^*(X_\nu - X_\mu) \Big].$$

3.4. **Removing the stochastic constraint.** It remains to remove the constraint in $\mathcal{D}_\mu$, namely, $\nu(s) \leqslant \mu(s)$. The reason we can do this is because, for arbitrary $\nu \in \mathcal{M}_b(\mathbb{R}_+)$ and $\widetilde{\nu}$ with the c.d.f. $\widetilde{\nu}(s) = \min(\nu(s), \mu(s))$, we have

$$\psi(\widetilde{\nu}) \leqslant \psi(\nu) \text{ and } \mathbb{E}\xi^*(X_{\widetilde{\nu}} - X_\mu) = \mathbb{E}\xi^*(X_\nu - X_\mu).$$

The second equality holds because $\xi^*(s) = 0$ for $s \leqslant 0$ and $X_{\widetilde{\nu}} = \max(X_\nu, X_\mu)$. The first inequality follows from the monotonicity of ψ in ν, which can be seen as follows. First of all, in the definition of $\Psi(\nu, h)$ in (2.3), if we take any $a \geqslant \nu^{-1}(1)$ and let $\Phi^a(t, x)$ be the solution of the equation

$$(3.9) \qquad \begin{cases} \partial_t \Phi^a = -\dfrac{1}{2}\big(\partial_x^2 \Phi^a + \mu(t)(\partial_x \Phi^a)^2 \big) & \text{on } [0, a] \times \mathbb{R}, \\[2mm] \Phi^a(a, x) = \log \displaystyle\int_{\mathbb{R}} \exp\big(\sigma x + \lambda \sigma^2 \big) \, \mathrm{d}P_1(\sigma) & \text{for } x \in \mathbb{R}, \end{cases}$$

and define $\mathcal{P}^a(\nu, \lambda) := \Phi^a(0, 0)$, then

$$(3.10) \qquad \Psi(\nu, h) = \mathcal{P}\Big(\nu, h - \frac{1}{2}\nu^{-1}(1) \Big) = \mathcal{P}^a\Big(\nu, h - \frac{1}{2}a \Big).$$

This means that when we compare $\Psi(\nu, \lambda)$ and $\Psi(\nu', \lambda)$, we can solve the PDE on the same interval that includes the support of both measures. Since the solution is monotone in the c.d.f. $\nu(s)$, monotonicity of the mapping $\nu \mapsto \Psi(\nu, \lambda)$ follows. This proves

$$(3.11) \qquad \lim_{N \to \infty} F_N = \inf_{\nu \in \mathcal{M}_b(\mathbb{R}_+)} \Big[\psi(\nu) + \frac{1}{2}\mathbb{E}\xi^*(X_\nu - X_\mu) \Big].$$

Finally, rescaling $\xi \to t\xi$ and recalling the definition of ψ in (2.3), we get

$$(3.12) \qquad \lim_{N \to \infty} F_N(t, \mu) = \inf_{\nu \in \mathcal{M}_b(\mathbb{R}_+)} \Big[\psi(\nu) + \frac{t}{2}\mathbb{E}\xi^*\Big(\frac{X_\nu - X_\mu}{t} \Big) \Big].$$

This finishes the proof that Proposition 2.1 implies Theorem 1.1, up to the verification of (3.5).

3.5. **Lipschitz continuity.** We now show (3.5). Let us recall the definition of the set in (2.7) and define

$$f_N(\nu, \varepsilon) = \frac{1}{N} \mathbb{E} \log \int\!\!\!\int_{\Omega_N^\varepsilon(u)} \exp\Big(\sum_{i=1}^N \sigma_i z_i^\nu(\alpha) - \frac{1}{2}\nu^{-1}(1)|\sigma|^2\Big) \mathrm{d}P_N(\sigma) \mathrm{d}\mathfrak{R}(\alpha).$$

Let us first suppose that ν is discrete. If we recall (3.4), the results of [22, Section 7] (specialized to the one-dimensional case) show that, for discrete ν,

(3.13) $$\Gamma_u(\nu) = \lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} f_N(\nu, \varepsilon).$$

Given $\nu, \widetilde{\nu} \in \mathcal{M}_b(\mathbb{R}_+)$, we can interpolate between $f_N(\nu, \varepsilon)$ and $f_N(\widetilde{\nu}, \varepsilon)$ by replacing $z_i^\nu(\alpha)$ by $\sqrt{t}z_i^\nu(\alpha) + \sqrt{1-t}z_i^{\widetilde{\nu}}(\alpha)$ with covariance $(t\nu^{-1} + (1-t)\widetilde{\nu}^{-1})(\alpha^1 \cdot \alpha^2)$ and replacing $\nu^{-1}(1)$ by $t\nu^{-1}(1) + (1-t)\widetilde{\nu}^{-1}(1)$. Then the derivative of $f_N(\nu, \varepsilon)$ in t along this interpolation path equals

$$-\frac{1}{2}\mathbb{E}\big\langle R_{1,2}(\nu^{-1}(\alpha^1 \cdot \alpha^2) - \widetilde{\nu}^{-1}(\alpha^1 \cdot \alpha^2))\big\rangle,$$

where $\langle \cdot \rangle$ is the average with respect to the Gibbs measure

$$\mathrm{d}G_N(\sigma, \alpha) \sim \exp\Big(\sum_{i=1}^N \sigma_i\big(\sqrt{t}z_i^\nu(\alpha) + \sqrt{1-t}z_i^{\widetilde{\nu}}(\alpha)\big)$$
$$-\frac{1}{2}\big(t\nu^{-1}(1) + (1-t)\widetilde{\nu}^{-1}(1)\big)|\sigma|^2\Big) \mathrm{d}P_N(\sigma)\mathrm{d}\mathfrak{R}(\alpha)$$

on $\Omega_N^\varepsilon(u) \times \mathrm{supp}(\mathfrak{R})$. Since, by Cauchy's inequality, $|R_{1,2}| \leqslant u + \varepsilon$ for $\sigma^1, \sigma^2 \in \Omega_N^\varepsilon(u)$, the above derivative is bounded by

$$\frac{u+\varepsilon}{2}\mathbb{E}\big\langle|\nu^{-1}(\alpha^1 \cdot \alpha^2) - \widetilde{\nu}^{-1}(\alpha^1 \cdot \alpha^2)|\big\rangle = \frac{u+\varepsilon}{2}\mathbb{E}|X_\nu - X_{\widetilde{\nu}}|,$$

where we also used the fact that the distribution of $\alpha^1 \cdot \alpha^2 \sim U[0,1]$ under $\mathbb{E}G_N^{\otimes 2}$ is the same as under $\mathbb{E}\mathfrak{R}^{\otimes 2}$ by the properties of the Ruelle probability cascades (see e.g. [19, Theorem 4.4]). This and (3.13) imply (3.5) for discrete ν, and by extension for all $\nu \in \mathcal{M}_b(\mathbb{R}_+)$.

4. Proof of the Parisi formula

The goal of this section is to prove Theorem 2.2, which we recall implies Proposition 2.1. We first prove the upper and then the lower bound.

4.1. **Upper bound.** The upper bound is proved by the standard Guerra replica-symmetry-breaking interpolation [8]. By Lipschitz continuity (1.7), it is enough to consider discrete μ and suppose that the infimum in Theorem 2.2 is taken also over discrete distributions $\zeta \in \mathcal{M}_{0,u}$ such that $\zeta^{-1}(1) = u$. Given such ζ, let $z(\alpha)$ and $y(\alpha)$ be independent Gaussian processes (conditionally on $\mathfrak{R}$) indexed by $\alpha \in \mathrm{supp}(\mathfrak{R})$ with covariances

$$\mathbb{E}z(\alpha^1)z(\alpha^2) = \xi'(\zeta^{-1}(\alpha^1 \cdot \alpha^2)), \ \mathbb{E}y(\alpha^1)y(\alpha^2) = \theta(\zeta^{-1}(\alpha^1 \cdot \alpha^2)),$$

and let $z_i(\alpha)$ be independent copies of $z(\alpha)$ for $i \geqslant 1$. We assume these processes to also be independent of H_N and $z_i^\mu(\alpha)$, conditionally on $\mathfrak{R}$. Consider an interpolating free energy, for $t \in [0,1]$,

(4.1) $$\varphi(t) := \frac{1}{N}\mathbb{E}\log \int\!\!\!\int_{\Omega_N^\varepsilon(u)} \exp(H_{N,t}(\sigma, \alpha)) \, \mathrm{d}P_N(\sigma) \, \mathrm{d}\mathfrak{R}(\alpha).$$

where the interpolation Hamiltonian is defined by

$$H_{N,t}(\sigma,\alpha) := \sqrt{t}H_N(\sigma) + \sqrt{1-t}\sum_{i=1}^{N}\sigma_i z_i(\alpha) + \sqrt{t}\sqrt{N}y(\alpha) + \sum_{i=1}^{N}\sigma_i z_i^{\mu}(\alpha).$$

One can see that

$$\frac{2}{N}\mathbb{E}\,\partial_t H_{N,t}(\sigma^1,\alpha^1)\,H_{N,t}(\sigma^2,\alpha^2) = \Delta(R_{1,2},\alpha^1\cdot\alpha^2)$$

$$:= \xi(R_{1,2}) - R_{1,2}\xi'(\zeta^{-1}(\alpha^1\cdot\alpha^2)) + \theta(\zeta^{-1}(\alpha^1\cdot\alpha^2)).$$

Since $R_{1,1} = N^{-1}|\sigma^1|^2 \in (u-\varepsilon, u+\varepsilon)$ whenever $\sigma^1 \in \Omega_N^{\varepsilon}(u)$ and we also assumed that $\zeta^{-1}(\alpha^1\cdot\alpha^1) = \zeta^{-1}(1) = u$, we have $|\Delta(R_{1,1},\alpha^1\cdot\alpha^1)| = \mathcal{O}(\varepsilon)$. Therefore, by the usual Gaussian integration by parts,

$$\varphi'(t) = -\frac{1}{2}\mathbb{E}\big\langle\Delta(R_{1,2},\alpha^1\cdot\alpha^2)\big\rangle + \mathcal{O}(\varepsilon),$$

where $\langle\cdot\rangle$ is the average with respect to the Gibbs measure

$$dG_N(\sigma,\alpha) \sim \exp(H_{N,t}(\sigma,\alpha))\,dP_N(\sigma)d\mathfrak{R}(\alpha)$$

on $\Omega_N^{\varepsilon}(u) \times \mathrm{supp}(\mathfrak{R})$. When ξ is convex, $\xi(a) - a\xi'(b) + \theta(b) \geqslant 0$, which proves that $\varphi(1) \leqslant \varphi(0) + \mathcal{O}(\varepsilon)$. When ξ is only convex on $\mathbb{R}_+$, one can add a small perturbation that enforces the Ghirlanda-Guerra identities and, as a result, enforces asymptotic positivity of $R_{1,2}$ (see [28] or [19, Chapter 3]).

We can bound $\varphi(0)$ from above by adding $-N\lambda u + \lambda|\sigma|^2$ to the Hamiltonian, which on $\Omega_N^{\varepsilon}(u)$ is bounded in absolute value by $N|\lambda|\varepsilon$ and, as a result,

$$\varphi(0) \leqslant |\lambda|\varepsilon - \lambda u + \frac{1}{N}\mathbb{E}\log\iint\exp\Big(\sum_{i=1}^{N}\sigma_i\big(z_i(\alpha) + z_i^{\mu}(\alpha)\big) + \lambda|\sigma|^2\Big)\,dP_N(\sigma)d\mathfrak{R}(\alpha)$$

$$= |\lambda|\varepsilon - \lambda u + \mathbb{E}\log\iint\exp\Big(\sigma_1\big(z_1(\alpha) + z_1^{\mu}(\alpha)\big) + \lambda\sigma_1^2\Big)\,dP_1(\sigma_1)d\mathfrak{R}(\alpha)$$

$$= |\lambda|\varepsilon - \lambda u + \mathcal{P}\big(\zeta_\nu,\lambda\big),$$

by the standard properties of the Ruelle probability cascades and the fact that $z_1(\alpha) + z_1^{\mu}(\alpha)$ has covariance $\zeta_\mu^{-1}(\alpha^1\cdot\alpha^2)$, where $\zeta_\mu(s)$ was defined in (2.4), (2.5). On the other hand, again, by the standard properties of the Ruelle probability cascades (recall the notation in (2.8)),

$$\varphi(1) = p_N^{\varepsilon}(u) + \frac{1}{N}\mathbb{E}\log\int\exp(\sqrt{N}y(\alpha))\,d\mathfrak{R}(\alpha)$$

$$= p_N^{\varepsilon}(u) + \mathbb{E}\log\int\exp(y(\alpha))\,d\mathfrak{R}(\alpha) = p_N^{\varepsilon}(u) + \frac{1}{2}\int_0^u\zeta(s)\,d\theta(s).$$

Putting everything together shows that

$$(4.2)\qquad \lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}p_N^{\varepsilon}(u) \leqslant -\lambda u + \mathcal{P}(\zeta_\mu,\lambda) - \frac{1}{2}\int_0^u\zeta(s)\,d\theta(s),$$

for all discrete distributions $\zeta \in \mathcal{M}_{0,u}$ such that $\zeta^{-1}(1) = u$. Since continuous extension of $\mathcal{P}(\zeta_\mu,\lambda)$ to all $\zeta \in \mathcal{M}_{0,u}$ not necessarily satisfying $\zeta^{-1}(1) = u$ is exactly

$$(4.3)\qquad \mathcal{P}\Big(\zeta_\mu,\lambda + \frac{\xi'(u) - \xi'(\zeta^{-1}(1))}{2}\Big)$$

(this is analogous to why the term $-\frac{1}{2}\mu^{-1}(1)|\sigma|^2$ was included in the definition of $F_N(t,\mu)$), this finishes the proof of the upper bound.

4.2. **Lower bound.** The proof of the lower bound is identical to the one-dimensional case of [22], with some simplifications due to the one-dimensional nature of our problem and one minor modification to account for the presence of the term $\sum_{i=1}^{N} \sigma_i z_i^{\mu}(\alpha)$ that we will now explain.

The main effect of this term is that the cavity fields (in the first term) of the Aizenman-Sims-Starr representation will be of the form $c_i(\sigma, \alpha) := z_i(\sigma) + z_i^{\mu}(\alpha)$ for $(\sigma, \alpha) \in \Omega_N^{\varepsilon}(u) \times \operatorname{supp}(\mathfrak{R})$ with covariance

$$(4.4) \qquad C_{\ell, \ell'} := \mathbb{E} c_i(\sigma^{\ell}, \alpha^{\ell}) c_i(\sigma^{\ell'}, \alpha^{\ell'}) = \xi'(R_{\ell, \ell'}) + \mu^{-1}(\alpha^{\ell} \cdot \alpha^{\ell'}).$$

To understand the distribution of the array $(C_{\ell, \ell'})_{\ell, \ell' \geqslant 1}$ under the Gibbs measure that arises in the cavity computation, we can use the synchronization mechanism from [20] to synchronize the overlaps $R_{1,2}$ and $q_{1,2} := \mu^{-1}(\alpha^1 \cdot \alpha^2)$. This can be done by including terms in the perturbation Hamiltonian with covariances given by monomials $R_{1,2}^n q_{1,2}^m$ and then use Theorem 4 in [20] to show that both $R_{1,2}$ and $q_{1,2}$ are non-decreasing 1-Lipschitz functions of their sum in the thermodynamic limit.

If we think of the sum $S_{1,2} := R_{1,2} + q_{1,2}$ as the quantile transform $\nu^{-1}(U)$ of $\nu = \mathcal{L}(S_{1,2})$ and uniform $U \sim U[0,1]$, then both $R_{1,2}$ and $q_{1,2}$ are non-decreasing functions of U, which means they must be quantile transforms of their distributions. The distribution of $q_{1,2}$ is μ for all N by the properties of the Ruelle probability cascades ([19, Theorem 4.4]) and, thus, in the limit. If the limiting distribution of $R_{1,2}$ (as usual, along some subsequence) is ζ then (recalling (2.4))

$$\xi'(R_{\ell, \ell'}) + \mu^{-1}(\alpha^{\ell} \cdot \alpha^{\ell'}) \overset{d}{=} \xi'(\zeta^{-1}(U)) + \mu^{-1}(U) = \zeta_{\mu}^{-1}(U).$$

This means that $C_{\ell, \ell'} \sim \zeta_{\mu}$. Similarly, the cavity fields $y(\sigma)$ coming from the Onsager correction in the second term in the Aizenman-Sims-Starr scheme will have covariance

$$\theta(R_{\ell, \ell'}) \overset{d}{=} \theta(\zeta^{-1}(U)),$$

in the thermodynamic limit. If $\zeta^{-1}(1) = u$ then the lower bound one obtains by the cavity computation is equal to

$$(4.5) \qquad \inf_{\lambda \in \mathbb{R}} \Big[-\lambda u + \mathcal{P}(\zeta_{\mu}, \lambda) - \frac{1}{2} \int_0^u \zeta(s) \, \mathrm{d}\theta(s) \Big].$$

For general $\zeta \in \mathcal{M}_{0,u}$, we again appeal to the fact that (4.3) is a continuous extension from general ζ of $\mathcal{P}(\zeta_{\mu}, \lambda)$ for ζ satisfying $\zeta^{-1}(1) = u$.

5. Proof of Theorem 1.2

The goal of this section is to prove Theorem 1.2. We obtain this result by combining Theorem 1.1 with the observation in (1.14) that F_N satisfies a Hamilton-Jacobi equation, up to a small error term. Denote by $\langle \cdot \rangle$ the Gibbs measure

$$\mathrm{d}G_N(\sigma, \alpha) \sim \exp\Big(\sqrt{t} H_N(\sigma) + \sum_{i=1}^{N} \sigma_i z_i^{\mu}(\alpha)$$

$$- \frac{1}{2} N(t-s)\xi(N^{-1}|\sigma|^2) - \frac{1}{2}\mu^{-1}(1)|\sigma|^2 + h|\sigma|^2\Big) \, \mathrm{d}P_N(\sigma) \mathrm{d}\mathfrak{R}(\alpha).$$

Similarly to the observations in [13, Section 1] concerning the Curie-Weiss model (with ξ replaced by the square function there), we have

$$\partial_s \mathsf{F}_N = \frac{1}{2} \mathbb{E} \big\langle \xi \big(N^{-1}|\sigma|^2 \big) \big\rangle,$$

while

$$\partial_h \mathsf{F}_N = \mathbb{E} \left\langle N^{-1}|\sigma|^2 \right\rangle, \tag{5.1}$$

and

$$\partial_h^2 \mathsf{F}_N = N \, \mathbb{E} \left\langle \left(N^{-1}|\sigma|^2 - \mathbb{E} \left\langle N^{-1}|\sigma|^2 \right\rangle \right)^2 \right\rangle. \tag{5.2}$$

Notice in particular that $\partial_h \mathsf{F}_N \geqslant 0$, $\partial_h^2 \mathsf{F}_N \geqslant 0$, and since the support of the measure P_1 is assumed to be bounded, the derivatives $\partial_s \mathsf{F}_N$ and $\partial_h \mathsf{F}_N$ are bounded uniformly in N. Moreover, since ξ is locally Lispschitz continuous, there exists a constant $C < \infty$ such that, for every N,

$$\begin{aligned} |2\partial_s \mathsf{F}_N - \xi(\partial_h \mathsf{F}_N)| &= \left| \mathbb{E} \left\langle \xi \left(N^{-1}|\sigma|^2 \right) \right\rangle - \xi \left(\mathbb{E} \left\langle N^{-1}|\sigma|^2 \right\rangle \right) \right| \\ &\leqslant C \mathbb{E} \left\langle \left| (N^{-1}|\sigma|^2 - \mathbb{E} \left\langle N^{-1}|\sigma|^2 \right\rangle \right| \right\rangle \\ &\leqslant C(N^{-1}\partial_h^2 \mathsf{F}_N)^{\frac{1}{2}}. \end{aligned} \tag{5.3}$$

We fix $t \geqslant 0$ and $\mu \in \mathcal{M}_b(\mathbb{R}_+)$, and denote by $f = f(s,h) : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ the candidate limit for F_N, namely

$$f(s,h) := \sup_{h' \in \mathbb{R}} \left(\widehat{\Psi}(h') - \frac{s}{2} \xi^* \left(\frac{2(h'-h)}{s} \right) \right), \tag{5.4}$$

where we set

$$\widehat{\Psi}(h') := \inf_{\nu \in \mathcal{M}_b(\mathbb{R}_+)} \left(\Psi(\nu, h') + \frac{t}{2} \mathbb{E} \xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right).$$

Notice that we do not display the dependency of f and $\widehat{\Psi}$ on $t \geqslant 0$ and $\mu \in \mathcal{M}_b(\mathbb{R}_+)$; we allow ourselves to do this since these parameters will be kept fixed throughout the section. For the same reason, from now on, *we write $\mathsf{F}_N(s,h)$ in place of $\mathsf{F}_N(s,t,\mu,h)$*.

Recalling that, by Theorem 1.1, the quantity $\widehat{\Psi}(h')$ is the limit of $\mathsf{F}_N(0,h')$, and using (5.1) and (5.2), it is clear that $\widehat{\Psi}$ is uniformly Lipschitz continuous, nondecreasing, and convex. One can check that these properties transfer to the function f: it is uniformly Lipschitz continuous over $\mathbb{R}_+ \times \mathbb{R}$, and for each fixed $s \geqslant 0$, the mapping $h \mapsto f(s,h)$ is nondecreasing and convex (see for instance [5, Lemmas I.3.3.2 and I.3.3.3]). In particular, by the Rademacher theorem, the function f is differentiable almost everywhere. Moreover, the expression for f in (5.4) is a Hopf-Lax formula; as a consequence, see [5, Theorem I.3.3.5], for every $(s,h) \in (0,\infty) \times \mathbb{R}$, if f is differentiable at (s,h), then

$$2\partial_s f(s,h) - \xi(\partial_h f(s,h)) = 0. \tag{5.5}$$

Our goal is to show that $\mathsf{F}_N(s,h)$ converges to $f(s,h)$. While we refrain from writing down a general statement, we list here all the properties of these functions that will be used below:

(1) the functions are uniformly Lipschitz, with a common Lipschitz constant;
(2) the functions are nondecreasing and convex in h;
(3) for each h, we have $\lim_{N \to \infty} \mathsf{F}_N(0,h) = f(0,h)$;
(4) the function f satisfies the equation (5.5) almost everywhere, while the function F_N satisfies the same equation, up to an error that we will show to be small after integration in h, uniformly over s.

Proof of Theorem 1.2. We split the proof into two steps.

Step 1. We write down an equation for the difference between F_N and f and state some elementary bounds. We denote

$$w_N := \mathsf{F}_N - f \quad \text{and} \quad \mathrm{err}_N := 2\partial_s \mathsf{F}_N - \xi(\partial_h \mathsf{F}_N),$$

so that, almost everywhere in $\mathbb{R}_+ \times \mathbb{R}$,

$$
\begin{aligned}
2\partial_s w_N &= \xi(\partial_h \mathsf{F}_N) - \xi(\partial_h f) + \mathrm{err}_N \\
&= \int_0^1 \partial_u \left(\xi(u\partial_h \mathsf{F}_N + (1-u)\partial_h f) \right) \mathrm{d}u + \mathrm{err}_N \\
&= \mathbf{b}_N \partial_h w_N + \mathrm{err}_N,
\end{aligned}
$$

where we have set

$$\mathbf{b}_N := \int_0^1 \xi'(u\partial_h \mathsf{F}_N + (1-u)\partial_h f) \, \mathrm{d}u.$$

Let $\phi \in C^\infty(\mathbb{R})$ be a smooth function such that $\phi(0) = 0$ and $|\phi'| \leqslant 1$, and define $v_N := \phi(w_N)$. By the chain rule, we have

$$(5.6) \qquad 2\partial_s v_N - \mathbf{b}_N \partial_h v_N = \phi'(w_N)\mathrm{err}_N \qquad \text{a.e. in } \mathbb{R}_+ \times \mathbb{R}.$$

It will be convenient to be allowed to differentiate $\mathbf{b}_N$ in h. In order to make this rigorous, we regularize $\mathbf{b}_N$ a bit, by convolution with a smooth kernel. Let $\zeta \in C_c^\infty(\mathbb{R})$ be a smooth function with compact support such that $\int_\mathbb{R} \zeta = 1$, and for each $\varepsilon > 0$, $s \geqslant 0$ and $h \in \mathbb{R}$, denote

$$f_\varepsilon(s,h) := \varepsilon^{-1} \int_\mathbb{R} f(s, h - h')\zeta(\varepsilon^{-1}h') \, \mathrm{d}h',$$

and

$$\mathbf{b}_{N,\varepsilon} := \int_0^1 \xi'(u\partial_h \mathsf{F}_N + (1-u)\partial_h f_\varepsilon) \, \mathrm{d}u.$$

One can check that for each fixed N and $s \geqslant 0$, the function $\mathbf{b}_{N,\varepsilon}(s,\cdot)$ converges to $\mathbf{b}_N(s,\cdot)$ almost everywhere in $\mathbb{R}$ as ε tends to zero (see for instance [5, Theorem C.5.7]). Moreover,

$$\partial_h \mathbf{b}_{N,\varepsilon} = \int_0^1 (u\partial_h^2 \mathsf{F}_N + (1-u)\partial_h^2 f_\varepsilon)\xi''(u\partial_h \mathsf{F}_N + (1-u)\partial_h f_\varepsilon) \, \mathrm{d}u,$$

and since $\partial_h \mathsf{F}_N$, $\partial_h f_\varepsilon$, $\partial_h^2 \mathsf{F}_N$, and $\partial_h^2 f_\varepsilon$, are all nonnegative, and ξ'' maps $\mathbb{R}_+$ to $\mathbb{R}_+$, we deduce that

$$(5.7) \qquad\qquad\qquad\qquad \partial_h \mathbf{b}_{N,\varepsilon} \geqslant 0.$$

Notice also that, since F_N and f are Lipschitz with a common Lipschitz constant, we have that $\|\mathbf{b}_{N,\varepsilon}\|_{L^\infty}$ is bounded uniformly over N and ε. We write

$$R := 1 + \sup_{N,\varepsilon} \|\mathbf{b}_{N,\varepsilon}\|_{L^\infty} < \infty.$$

Step 2. We fix $S \geqslant 1$ for the remainder of the proof, and study the quantity

$$J_N(s) := \int_{-R(S-s)}^{R(S-s)} v_N(s, h) \, \mathrm{d}h, \qquad s \in [0, S].$$

The function J_N is Lipschitz continuous, and for almost every $s \in [0, S]$,

$$\partial_s J_N(s) = \int_{-R(S-s)}^{R(S-s)} \partial_s v_N(s, h) \, \mathrm{d}h - Rv_N(s, R(S-s)) - Rv_N(s, -R(S-s)).$$

By (5.6), we also have

$$2\int_{-R(S-s)}^{R(S-s)} \partial_s v_N \, \mathrm{d}h = \int_{-R(S-s)}^{R(S-s)} \left(\mathbf{b}_{N,\varepsilon} \partial_h v_N + (\mathbf{b}_N - \mathbf{b}_{N,\varepsilon})\partial_h v_N + \phi'(w_N)\mathrm{err}_N \right) \mathrm{d}h,$$

where we kept it implicit in the notation that the functions in the integrands are evaluated at (s, h). We now estimate the contribution of each term on the right side in turn. By the definition of R and an integration by parts, we have

$$\left| \int_{-R(S-s)}^{R(S-s)} \left(\mathbf{b}_{N,\varepsilon} \partial_h v_N + v_N \partial_h \mathbf{b}_{N,\varepsilon} \right) \, \mathrm{d}h \right| \le R v_N(s, R(S-s)) + R v_N(s, -R(S-s))$$

(recall that $v_N \ge 0$). Using also (5.7), we deduce that

$$\int_{-R(S-s)}^{R(S-s)} \mathbf{b}_{N,\varepsilon} \partial_h v_N \, \mathrm{d}h - R v_N(s, R(S-s)) - R v_N(s, -R(S-s)) \le 0.$$

Recalling also that for each fixed N, we have that $\mathbf{b}_{N,\varepsilon}(s\cdot)$ converges to $\mathbf{b}_N(s,\cdot)$ almost everywhere, and using the dominated convergence theorem, we see that

$$\lim_{\varepsilon \to 0} \int_{-R(S-s)}^{R(S-s)} (\mathbf{b}_N - \mathbf{b}_{N,\varepsilon}) \partial_h v_N \, \mathrm{d}h = 0.$$

Summarizing, we have shown that, for almost every $s \in [0, S]$,

$$\partial_s J_N(s) \le \frac{1}{2} \int_{-R(S-s)}^{R(S-s)} \phi'(w_N) \mathsf{err}_N \, \mathrm{d}h.$$

Recalling that $|\phi'| \le 1$ and using (5.3), we deduce that

$$\partial_s J_N(s) \le C N^{-\frac{1}{2}} \int_{-R(S-s)}^{R(S-s)} \left(\partial_h^2 \mathsf{F}_N \right)^{\frac{1}{2}} \, \mathrm{d}h.$$

Allowing the constant C to depend on R and S, we can use Jensen's inequality to deduce that, for almost every $s \in [0, S]$,

$$\partial_s J_N(s) \le C N^{-\frac{1}{2}} \left(\int_{-R(S-s)}^{R(S-s)} \partial_h^2 \mathsf{F}_N \, \mathrm{d}h \right)^{\frac{1}{2}},$$

and since F_N is Lipschitz uniformly over N, the integral on the right side is bounded. To sum up, we have thus shown that for every $s \in [0, S]$,

$$J_N(s) = J_N(0) + \int_0^s \partial_s J_N(r) \, \mathrm{d}r \le J_N(0) + C s N^{-\frac{1}{2}}.$$

Recalling the definition of v_N, fixing $s = \frac{S}{2}$, this implies in particular, up to a redefinition of $C < \infty$,

$$\int_{-\frac{RS}{2}}^{\frac{RS}{2}} \phi(\mathsf{F}_N - f)(s, h) \, \mathrm{d}h \le \int_{-\frac{RS}{2}}^{\frac{RS}{2}} \phi(\mathsf{F}_N - f)(0, h) \, \mathrm{d}h + C N^{-\frac{1}{2}}.$$

Notice also that the constant $C < \infty$ does not depend on our choice of function ϕ such that $|\phi'| \le 1$. We thus deduce that

$$\int_{-\frac{RS}{2}}^{\frac{RS}{2}} |\mathsf{F}_N - f|(s, h) \, \mathrm{d}h \le \int_{-\frac{RS}{2}}^{\frac{RS}{2}} |\mathsf{F}_N - f|(0, h) \, \mathrm{d}h + C N^{-\frac{1}{2}}.$$

Finally, by the dominated convergence theorem, the integral on the right side converges to 0 as N tends to infinity. We have thus shown that

$$\lim_{N \to \infty} \int_{-\frac{RS}{2}}^{\frac{RS}{2}} |\mathsf{F}_N - f|(s, h) \, \mathrm{d}h = 0.$$

Recall that $R > 0$ and that our choice of $S \ge 1$ was arbitrary. To conclude for the pointwise convergence of F_N to f, it suffices to use the Lipschitz regularity of F_N. Explicitly, for every $\varepsilon > 0$, we can write

$$\mathsf{F}_N(s, h) = \mathsf{F}_N(s, h) - \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathsf{F}_N(s, h') \, \mathrm{d}h' + \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathsf{F}_N(s, h') \, \mathrm{d}h',$$

and we have seen above that the last integral converges to the corresponding integral with F_N replaced by f as N tends to infinity. Moreover, by the Lipschitz continuity of F_N,

$$\left| \mathsf{F}_N(s, h) - \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathsf{F}_N(s, h') \, \mathrm{d}h' \right| \leqslant C\varepsilon.$$

Hence, sending N to infinity first and then ε to zero allows us to conclude that for each $s \geqslant 0$ and $h \in \mathbb{R}$, we have indeed $\lim_{N \to \infty} \mathsf{F}_N(s, h) = f(s, h)$. $\qquad\square$

References

[1] Agliari, E., Barra, A., Burioni, R., Di Biasio, A.: Notes on the p-spin glass studied via Hamilton-Jacobi and smooth-cavity techniques. J. Math. Phys. 53, no. 6, 063304 (2012)

[2] Auffinger, A., Chen, W.-K.: The Parisi formula has a unique minimizer. Comm. Math. Phys. 335, no. 3, 1429–1444 (2015)

[3] Barra, A., Di Biasio, A., Guerra, F.: Replica symmetry breaking in mean-field spin glasses through the HamiltonJacobi technique. J. Stat. Mech. Theory E, 2010.09 (2010): P09006.

[4] Contucci, P., Mingione, E.: A multi-scale spin-glass mean-field model. Comm. Math. Phys. 268, no. 3, 1323–1344 (2018)

[5] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[6] Gikhman I.I., Skorokhod, A.V.: *The theory of stochastic processes I*. Classics in Mathematics. Springer-Verlag, New York (2004)

[7] Guerra, F.: Sum rules for the free energy in the mean field spin glass model, Mathematical physics in mathematics and physics (Siena, 2000). Fields Inst. Commun. 30, 161–170 (2001)

[8] Guerra, F.: Broken replica symmetry bounds in the mean field spin glass model. Comm. Math. Phys. 233, no. 1, 1–12 (2003)

[9] Jagannath, A., Ko, J., Sen, S.: A connection between MAX κ-CUT and the inhomogeneous Potts spin glass in the large degree limit. Ann. of Applied Probab. 28, no. 3, 1536–1572 (2018)

[10] Jagannath, A., Tobasco, I.: A dynamic programming approach to the Parisi functional. Proc. Amer. Math. Soc. 144, no. 7, 3135–3150 (2016)

[11] Ko, J.: Free energy of multiple systems of spherical spin glasses with constrained overlaps. arXiv:1806.09772 (2018)

[12] M. Mézard, G. Parisi, and M. Virasoro. *Spin glass theory and beyond: an introduction to the replica method and its applications*, volume 9. World Scientific Publishing Company, 1987

[13] Mourrat, J.-C.: Hamilton-Jacobi equations for mean-field disordered systems. arXiv:1811.01432 (2018)

[14] Mourrat, J.-C.: Hamilton-Jacobi equations for finite-rank matrix inference. arXiv:1904.05294 (2019)

[15] Mourrat, J.-C.: Parisi's formula is a Hamilton-Jacobi equation in Wasserstein space. arXiv: 1906.08471 (2019)

[16] Panchenko, D.: Free energy in the generalized Sherrington–Kirkpatrick mean field model. Rev. Math. Phys. 17, no. 7, 793–857 (2005)

[17] Panchenko, D.: The Parisi ultrametricity conjecture. Ann. of Math. (2) 177, no. 1, 383–393 (2013)

[18] Panchenko, D.: The Parisi formula for mixed p-spin models. Ann. Probab. 42, no. 3, 946–958 (2014)

[19] Panchenko, D.: *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer-Verlag, New York (2013)

[20] Panchenko, D.: Free energy in the multi-species Sherrington-Kirkpatrick model, Ann. Probab., 43, no. 6, 3494–3513 (2015)

[21] Panchenko, D.: Free energy in the Potts spin glass. Ann. Probab. 46, no. 2, 829–864 (2018)

[22] Panchenko, D.: Free energy in the mixed p-spin models with vector spins. Ann. Probab. 46, no. 2, 865–896 (2018)

[23] Parisi, G.: Infinite number of order parameters for spin-glasses. Phys. Rev. Lett. 43, 1754–1756 (1979)

[24] Parisi, G.: A sequence of approximate solutions to the S-K model for spin glasses. J. Phys. A 13, L-115 (1980)
[25] Ruelle, D.: A mathematical reformulation of Derrida's REM and GREM. Comm. Math. Phys. 108, no. 2, 225–239 (1987)
[26] Sion, M.: On general minimax theorems. Pacific Journal of Mathematics 8, no.1, 171–176 (1958)
[27] Sherrington, D., Kirkpatrick, S.: Solvable model of a spin glass. Phys. Rev. Lett. 35, 1792–1796 (1975)
[28] Talagrand, M.: On Guerra's broken replica-symmetry bound. C. R. Math. Acad. Sci. Paris 337, no. 7, 477–480 (2003)
[29] Talagrand, M.: Parisi measures. J. Funct. Anal. 231, no. 2, 269–286 (2006)
[30] Talagrand, M.: The Parisi formula. Ann. of Math. (2) 163, no. 1, 221–263 (2006)
[31] Talagrand, M.: *Mean-Field Models for Spin Glasses*. Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, vol. 54, 55. Springer-Verlag (2011)
[32] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

(J.-C. Mourrat) Courant Institute of Mathematical Sciences, New York University, New York, New York, USA

Email address: jcm777@nyu.edu

(D. Panchenko) Department of Mathematics, University of Toronto, Toronto, Ontario, Canada

Email address: panchenk@math.toronto.edu