

# Parisi formula for Ising and Spherical Perceptrons

October 7, 2026

## Abstract

We prove a Parisi formula for Ising and spherical perceptrons with Gaussian disorder and i.i.d. random potentials having a common Lipschitz bound and integrable values at zero. We assume  $M/N \rightarrow \alpha \in (0, \infty)$ , where  $M$  is the number of Gaussian vectors and  $N$  is the number of spins. The proof combines cavity in  $M$  and  $N$  with the Aizenman–Sims–Starr scheme for the lower bound and Guerra’s interpolation for the upper bound. As consequences, we answer several problems posed by Talagrand (2011) and establish a quenched large deviation principle for the empirical distribution of the Gaussian projections, resolving a conjecture of Bolthausen and Kistler (2010). We also resolve the sharp-threshold conjecture of Aubin, Perkins, and Zdeborová (2019) for the  $u$ -function binary perceptron at every positive width, and establish the predicted variational formulas for the entropy and capacity of the negative spherical perceptron. Finally, the zero-temperature limit yields the scalar projection-pursuit formulas conjectured by Montanari and Zhou (2025, 2026), in both the supervised and unlabelled settings.

## Contents

|          |                                                      |           |
|----------|------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                  | <b>2</b>  |
| <b>2</b> | <b>Proof of the free energy formula</b>              | <b>11</b> |
| <b>3</b> | <b>Cavity in <math>M</math></b>                      | <b>22</b> |
| <b>4</b> | <b>Scalar transforms and supporting inequalities</b> | <b>31</b> |
| <b>5</b> | <b>Cavity in <math>N</math></b>                      | <b>38</b> |
| <b>6</b> | <b>Perturbation, positivity, and synchronization</b> | <b>43</b> |
| <b>7</b> | <b>Quenched large deviations</b>                     | <b>49</b> |
| <b>8</b> | <b>Applications and singular limits</b>              | <b>52</b> |

## Disclaimer

This document and its arguments were generated through discussions with ChatGPT 5.6 Pro. An initial draft for the capacity of the negative spherical perceptron was generated on July 15, 2026, and we extended the argument to i.i.d. lipschitz potentials, both Ising and spherical, on September 17, 2026. While distilling the main ideas, carefully checking the proof, and polishing the draft, we learned on October 6, 2026, that OpenAI had released overlapping results across three manuscripts [23, 24, 25]. Given these circumstances, we decided to post this working draft on Hexagon.

# 1 Introduction

We study Ising and spherical perceptrons with Gaussian disorder and i.i.d. random Lipschitz potentials. Let  $\nu$  be a Borel probability measure on  $\mathbb{R}$ , and let  $(\theta, x) \mapsto u_\theta(x)$  be Borel measurable on  $\mathbb{R}^2$ . We assume that  $\text{Lip}(u_\theta) \leq L < \infty$  for  $\nu$ -almost every  $\theta$  and that  $\int_{\mathbb{R}} |u_\theta(0)| \nu(d\theta) < \infty$ .

Let  $\Theta_1, \dots, \Theta_M$  be i.i.d. with law  $\nu$ , and let  $g_1, \dots, g_M \in \mathbb{R}^N$  be independent standard Gaussian vectors, independent of  $(\Theta_k)_{k \leq M}$ . The Hamiltonian of the perceptron model is defined by

$$H_{N,M}^\nu(\sigma) = \sum_{k=1}^M u_{\Theta_k}(h_k(\sigma)), \quad \text{where} \quad h_k(\sigma) = \frac{g_k \cdot \sigma}{\sqrt{N}}$$

is the Gaussian projection. We consider the spherical and Ising configuration spaces

$$\Sigma_N^S = \{\sigma \in \mathbb{R}^N : \|\sigma\|_2^2 = N\}, \quad \Sigma_N^I = \{-1, 1\}^N.$$

Let  $\mu_N^S$  and  $\mu_N^I$  be the uniform probability measures on  $\Sigma_N^S$  and  $\Sigma_N^I$ , respectively. For  $\kappa \in \{S, I\}$ , the partition function is

$$Z_{N,M}^{\kappa,\nu} = \int_{\Sigma_N^\kappa} e^{H_{N,M}^\nu(\sigma)} \mu_N^\kappa(d\sigma).$$

Our main object is the normalized free energy  $N^{-1} \log Z_{N,M}^{\kappa,\nu}$ . Throughout, we take  $M = M_N$  with  $M_N/N \rightarrow \alpha \in (0, \infty)$ .

The study of perceptron model in statistical physics was initiated by Gardner and Gardner–Derrida [14, 15]. For constraints  $h_k(\sigma) \geq b$ , they studied the volume of configurations satisfying all  $M$  inequalities and the density  $M/N$  at which solutions disappear. The entropy records the logarithmic volume or number of solutions; the storage capacity is the critical density for their existence. Krauth and Mézard subsequently predicted the capacity for Ising spins [18].

For spherical spins and nonnegative margin  $b$ , Shcherbina and Tirozzi proved Gardner’s replica-symmetric entropy formula [32]. For Ising spins, Talagrand developed cavity arguments, comparing systems with slightly different numbers of spins or Gaussian vectors, and proved replica-symmetric formulas under restrictions on the potential and density [34]. Bolthausen, Nakajima, Sun, and Xu later obtained small-density formulas for a broad class of activations [7]. Nakajima and Sun proved concentration and universality results, together with sharp-threshold sequences for broad classes of Ising constraints [22]. Their results control random fluctuations and changes in the disorder distribution, and locate the satisfiability transition around a sequence of densities. Proving that this sequence converges requires an additional argument.

Beyond replica symmetry, Györgyi and Reimann developed formulas involving a full overlap distribution for the spherical perceptron at positive temperature [16]. Franz and Parisi studied the constraints  $h_k(\sigma) \geq b$  with  $b < 0$  and their jamming transition, where feasible configurations disappear as  $M/N$  increases [13]. Their predictions concern both the entropy and the geometry near this transition. For Ising spins, Aubin, Perkins, and Zdeborová studied the symmetric  $U$ -function constraints  $|h_k(\sigma)| \geq K$ . They identified a large- $K$  regime in which the first-moment estimate, based on the expected number of solutions, no longer gives the predicted capacity [3].

Our main theorem determines the limiting free energy under the assumptions above. We also prove a quenched large-deviation principle for the empirical distribution of the Gaussian projections  $h_k(\sigma)$ . For a typical fixed realization of the Gaussian vectors, this describes the exponential probabilities of atypical projection distributions when  $\sigma$  is sampled from  $\mu_N^\kappa$ . Finally, the zero-temperature limit determines optimal empirical averages over one-dimensional projections, in both supervised and unlabelled settings.

Define the overlap-path space by

$$\mathcal{Q} = \{q : [0, 1] \rightarrow [0, 1] : q \text{ is right-continuous and nondecreasing, } q(1) = 1\}.$$

We identify paths that agree almost everywhere on  $[0, 1)$  and equip  $\mathcal{Q}$  with the  $L^1$  metric. The spin cost  $\mathcal{A}_\kappa$  is defined in (3) for spherical spins and (5) for Ising spins. The channel functional  $\Phi_u$  is defined in (2) through the scalar Parisi PDE (1). These definitions follow [theorem 1.1](#).

**Theorem 1.1** (Free energy for i.i.d. Lipschitz channels). *For either prior,*

$$\frac{1}{N} \log Z_{N,M_N}^{\kappa,\nu} \longrightarrow \inf_{q \in \mathcal{Q}} \left\{ \mathcal{A}_\kappa(q) + \alpha \int_{\mathbb{R}} \Phi_{u_\theta}(q) \nu(d\theta) \right\} \quad \text{in } L^1.$$

*Convergence holds in  $L^2$  if  $\int_{\mathbb{R}} |u_\theta(0)|^2 \nu(d\theta) < \infty$ , and almost surely and in every finite  $L^p$  if there is  $B < \infty$  such that  $|u_\theta(x)| \leq B$  for every  $x$  and  $\nu$ -almost every  $\theta$ .*

Note that a deterministic potential  $u$  corresponds to a point-mass channel law. In this case, write  $H_{N,M}^u$ ,  $Z_{N,M}^{\kappa,u}$ , and  $F_{N,\kappa}^u = N^{-1} \log Z_{N,M_N}^{\kappa,u}$ , and put

$$\mathcal{P}_{\kappa,u}(\alpha) = \inf_{q \in \mathcal{Q}} \{ \mathcal{A}_\kappa(q) + \alpha \Phi_u(q) \}.$$

The theorem gives convergence in  $L^2$ , since  $u(0)$  is finite.

We next define the functionals  $\Phi_u$  and  $\mathcal{A}_\kappa$  on the overlap paths  $q \in \mathcal{Q}$ . Recall that these paths are right-continuous and nondecreasing, take values in  $[0, 1]$ , and satisfy  $q(1) = 1$ . Let

$$\mathcal{P}_b = \{a : [0, 1) \rightarrow [0, \infty) : a \text{ is bounded, right-continuous and nondecreasing}\}.$$

These paths will enter the variational definition of the Ising spin cost. For  $c \in \mathcal{P}_b$ , write  $c_* = c(1-)$  and  $\gamma_c(t) = \text{Leb}\{r \in [0, 1) : c(r) \leq t\}$ . Thus  $\gamma_c$  is the distribution function of  $c(U)$  for  $U$  uniform on  $[0, 1)$ . We use the same notation for  $q \in \mathcal{Q}$ , restricting  $q$  to  $[0, 1)$ .

For  $T > 0$ , a nonnegative Lebesgue-integrable function  $\gamma : [0, T] \rightarrow [0, \infty)$ , and a globally Lipschitz function  $h : \mathbb{R} \rightarrow \mathbb{R}$ , let  $f_{\gamma,h}$  solve the backward Parisi PDE

$$\partial_t f + \frac{1}{2} (f_{xx} + \gamma(t) f_x^2) = 0, \quad f(T, x) = h(x). \quad (1)$$

We use the stochastic-control solution; its existence and stability are proved in [theorem 8.1](#). The channel functional is

$$\Phi_u(q) = f_{\gamma_q, u}(0, 0), \quad T = 1. \quad (2)$$

The two spin costs are as follows. For spherical spins,

$$\mathcal{A}_S(q) = \frac{1}{2} \int_0^1 \left\{ \frac{1}{\int_t^1 \gamma_q(s) ds} - \frac{1}{1-t} \right\} dt. \quad (3)$$

The outer integrand is nonnegative, and the value may be infinite. For Ising spins and  $a \in \mathcal{P}_b$ , use the same PDE with terminal function  $\log \cosh$  and terminal time  $a_*$ :

$$\mathcal{S}_I(a) = f_{\gamma_a, \log \cosh}(0, 0) - \frac{a_*}{2}, \quad T = a_*. \quad (4)$$

At  $a = 0$ , define  $\mathcal{S}_I(0) = 0$ . The Ising cost is

$$\mathcal{A}_I(q) = \sup_{a \in \mathcal{P}_b} \left\{ \mathcal{S}_I(a) + \frac{1}{2} \int_0^1 a(r) q(r) dr \right\}. \quad (5)$$

Below, we write  $\langle a, q \rangle = \int_0^1 a(r) q(r) dr$ .

## 1.1 Direct consequences of the free energy formula

In this section, we prove direct consequences of Theorem 1.1, which may be of independent interest.

**Gaussian external fields.** For a standard Gaussian vector  $g^0 \in \mathbb{R}^N$ , independent of all other randomness, let

$$Z_{N,M}^{\kappa,u,b} = \int e^{\sum_{k \leq M} u(h_k(\sigma)) + b g^0 \cdot \sigma} \mu_N^\kappa(d\sigma), \quad F_{N,\kappa}^{u,b} = \frac{1}{N} \log Z_{N,M_N}^{\kappa,u,b}, \quad (6)$$

and denote the marked version by  $Z_{N,M}^{\kappa,\nu,b}$ . Write

$$\mathbb{L}(q) = \int_0^1 (1 - q(r)) dr, \quad \mathcal{P}_{\kappa,u}^{\text{fld}}(\alpha, b) = \inf_{q \in \mathcal{Q}} \left\{ \mathcal{A}_\kappa(q) + \alpha \Phi_u(q) + \frac{b^2}{2} \mathbb{L}(q) \right\}. \quad (7)$$

For a deterministic potential, the Gaussian-field extension of theorem 1.1 gives  $F_{N,\kappa}^{u,b} \rightarrow \mathcal{P}_{\kappa,u}^{\text{fld}}(\alpha, b)$  in  $L^2$ . For i.i.d. channels, replace  $\Phi_u$  by  $\mathbb{E}_\nu \Phi_{u_\Theta}$ ; the same  $L^1$  and  $L^2$  conclusions as in theorem 1.1 hold. Mixing in linear channels proves this extension; see section 8.1.

**Spherical i.i.d. fields.** Let  $\pi$  be a Borel probability measure on  $\mathbb{R}$  satisfying

$$v_\pi := \int_{\mathbb{R}} h^2 \pi(dh) < \infty.$$

Let  $h_1, h_2, \dots$  be i.i.d. with law  $\pi$ , independent of the Gaussian constraint vectors  $(g_k)_{k \leq M_N}$ . Define

$$Z_{N,S}^{u,\pi} := \int_{S_N} \exp \left\{ \sum_{k=1}^{M_N} u \left( \frac{g_k \cdot \sigma}{\sqrt{N}} \right) + \sum_{i=1}^N h_i \sigma_i \right\} \mu_N^S(d\sigma),$$

$$F_{N,S}^{u,\pi} := \frac{1}{N} \log Z_{N,S}^{u,\pi}.$$

Here  $\mathbb{E}$  and  $\mathbb{P}$  denote expectation and probability with respect to the joint law of the constraint vectors  $(g_k)_{k \leq M_N}$  and external fields  $(h_i)_{i \leq N}$ .

**Corollary 1.2** (Spherical i.i.d. external fields). *Let  $u : \mathbb{R} \rightarrow \mathbb{R}$  be real-valued and globally Lipschitz, let  $M_N/N \rightarrow \alpha \in (0, \infty)$ , and let  $\pi$  satisfy  $\int h^2 \pi(dh) < \infty$  as above. Then*

$$F_{N,S}^{u,\pi} \longrightarrow \mathcal{P}_{S,u}^{\text{fld}}(\alpha, \sqrt{v_\pi}) \quad \text{in } L^2.$$

*Proof.* For constraint vectors  $\mathbf{g} = (g_k)_{k=1}^{M_N}$  and a deterministic field  $\mathbf{h} \in \mathbb{R}^N$ , define

$$\widehat{f}_N(\mathbf{g}, \mathbf{h}) := \frac{1}{N} \log \int_{S_N} \exp \left\{ \sum_{k=1}^{M_N} u \left( \frac{g_k \cdot \sigma}{\sqrt{N}} \right) + \mathbf{h} \cdot \sigma \right\} \mu_N^S(d\sigma).$$

For  $\mathbf{h}, \tilde{\mathbf{h}} \in \mathbb{R}^N$ , Cauchy–Schwarz gives  $|(\mathbf{h} - \tilde{\mathbf{h}}) \cdot \sigma| \leq \sqrt{N} \|\mathbf{h} - \tilde{\mathbf{h}}\|_2$ . Comparing the two exponential weights therefore yields

$$\left| \widehat{f}_N(\mathbf{g}, \mathbf{h}) - \widehat{f}_N(\mathbf{g}, \tilde{\mathbf{h}}) \right| \leq \frac{\|\mathbf{h} - \tilde{\mathbf{h}}\|_2}{\sqrt{N}}. \quad (8)$$

For an orthogonal matrix  $O \in \mathbb{R}^{N \times N}$ , write  $O\mathbf{g} = (Og_k)_{k=1}^{M_N}$ . The change of variables  $\sigma = O^\top \tau$ , together with rotational invariance of  $\mu_N^S$ , gives  $\widehat{f}_N(\mathbf{g}, \mathbf{h}) = \widehat{f}_N(O\mathbf{g}, O\mathbf{h})$ . Under the Gaussian row

law,  $O\mathbf{g}$  has the same distribution as  $\mathbf{g}$ . Writing  $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^N$  and choosing  $O$  such that  $O\mathbf{h} = \|\mathbf{h}\|_2 e_1$ , we obtain

$$\widehat{f}_N(\mathbf{g}, \mathbf{h}) \stackrel{d}{=} \widehat{f}_N(\mathbf{g}, \|\mathbf{h}\|_2 e_1) \quad (\mathbf{h} \in \mathbb{R}^N),$$

where the equality in distribution is with respect to the Gaussian constraint vectors. Let  $z_1, z_2, \dots$  be independent standard Gaussian random variables, independent of both the external fields and the constraint vectors, and put  $\mathbf{h}^{(N)} = (h_1, \dots, h_N)$  and  $z^{(N)} = (z_1, \dots, z_N)$ . In the remainder of the proof,  $\mathbb{E}$  also averages over this auxiliary Gaussian sequence. Define, using the same constraint vectors in both expressions,

$$X_N := \widehat{f}_N(\mathbf{g}, \|\mathbf{h}^{(N)}\|_2 e_1), \quad Y_N := \widehat{f}_N(\mathbf{g}, \sqrt{v_\pi} \|z^{(N)}\|_2 e_1).$$

Conditioning on  $\mathbf{h}^{(N)}$  or  $z^{(N)}$  and applying the preceding rotational identity gives

$$\begin{aligned} X_N &\stackrel{d}{=} \widehat{f}_N(\mathbf{g}, \mathbf{h}^{(N)}) = F_{N,S}^{u,\pi}, \\ Y_N &\stackrel{d}{=} \widehat{f}_N(\mathbf{g}, \sqrt{v_\pi} z^{(N)}) \stackrel{d}{=} F_{N,S}^{u,\sqrt{v_\pi}}, \end{aligned}$$

where  $F_{N,S}^{u,\sqrt{v_\pi}}$  is the Gaussian-field free energy from (6) with  $b = \sqrt{v_\pi}$ . By (8),

$$|X_N - Y_N| \leq \left| \frac{\|\mathbf{h}^{(N)}\|_2}{\sqrt{N}} - \sqrt{v_\pi} \frac{\|z^{(N)}\|_2}{\sqrt{N}} \right|.$$

Using  $(\sqrt{x} - \sqrt{y})^2 \leq |x - y|$  for  $x, y \geq 0$ , we conclude that

$$\begin{aligned} \mathbb{E}|X_N - Y_N|^2 &\leq \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N h_i^2 - \frac{v_\pi}{N} \sum_{i=1}^N z_i^2 \right|^2 \\ &\leq \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N h_i^2 - v_\pi \right|^2 + v_\pi \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N z_i^2 - 1 \right|^2 \longrightarrow 0. \end{aligned}$$

Indeed, both terms tend to zero by the law of large numbers, since  $\mathbb{E}h_1^2 = v_\pi < \infty$  and  $\mathbb{E}z_1^2 = 1$ . Set

$$P_\pi := \mathcal{P}_{S,u}^{\text{fld}}(\alpha, \sqrt{v_\pi}).$$

The Gaussian-field extension of Theorem 1.1 and the distributional identity for  $Y_N$  imply that  $\mathbb{E}|Y_N - P_\pi|^2 \rightarrow 0$ . Consequently,

$$\mathbb{E} \left| F_{N,S}^{u,\pi} - P_\pi \right|^2 = \mathbb{E}|X_N - P_\pi|^2 \leq 2\mathbb{E}|X_N - Y_N|^2 + 2\mathbb{E}|Y_N - P_\pi|^2 \longrightarrow 0,$$

which proves the asserted  $L^2$  convergence.  $\square$

**Bipartite transposition.** Summing over one party converts a bipartite Ising model with a homogeneous deterministic field on only that party into a perceptron. For  $\beta, b \in \mathbb{R}$ , define

$$Z_{N,M}^{\text{bip}}(\beta, b) = 2^{-(N+M)} \sum_{\substack{\sigma \in \{-1, +1\}^N \\ \tau \in \{-1, +1\}^M}} \exp \left\{ \frac{\beta}{\sqrt{N}} \sum_{k=1}^M \sum_{i=1}^N g_{ki} \tau_k \sigma_i + b \sum_{i=1}^N \sigma_i \right\},$$

and put

$$u_{\alpha,\beta,b}^{\text{bip}}(x) = \log \cosh(b + \beta \sqrt{\alpha} x).$$

**Corollary 1.3** (Bipartite Ising formula with a field on one party). *For every  $\beta, b \in \mathbb{R}$ , if  $M_N/N \rightarrow \alpha \in (0, \infty)$ , then*

$$\frac{1}{N} \log Z_{N, M_N}^{\text{bip}}(\beta, b) \longrightarrow \alpha \mathcal{P}_{\mathbf{I}, u_{\alpha, \beta, b}^{\text{bip}}}(1/\alpha)$$

*in  $L^2$  and hence in  $L^1$  and probability. Moreover,*

$$\text{Var} \left( \frac{1}{N} \log Z_{N, M_N}^{\text{bip}}(\beta, b) \right) \leq \frac{\beta^2 M_N}{N^2}. \quad (9)$$

*Proof.* Put  $\alpha_N = M_N/N$  and  $u_N^{\text{bip}}(x) = \log \cosh(b + \beta \sqrt{\alpha_N} x)$ . Summing first over  $\sigma$  gives the exact identity

$$Z_{N, M_N}^{\text{bip}}(\beta, b) = Z_{M_N, N}^{\mathbf{I}, u_N^{\text{bip}}}.$$

Indeed, the columns of the  $M_N$ -by- $N$  disorder matrix are independent standard Gaussian rows in  $\mathbb{R}^{M_N}$ . Write  $G = (g_{ki})_{k \leq M_N, i \leq N}$  for this matrix. The one-Lipschitz property of  $\log \cosh$ , Cauchy–Schwarz, and the standard Gaussian operator-norm estimate give

$$\left| \frac{1}{M_N} \log Z_{M_N, N}^{\mathbf{I}, u_N^{\text{bip}}} - \frac{1}{M_N} \log Z_{M_N, N}^{\mathbf{I}, u_{\alpha, \beta, b}^{\text{bip}}} \right| \leq |\beta| |\sqrt{\alpha_N} - \sqrt{\alpha}| \frac{\sqrt{N} \|G\|_{\text{op}}}{M_N} = o_{L^2}(1).$$

Apply [theorem 1.1](#) with spin dimension  $M_N$ , row count  $N$ , and limiting density  $1/\alpha$ , and multiply by  $M_N/N \rightarrow \alpha$ . Finally,  $\partial_{g_{ki}}[N^{-1} \log Z^{\text{bip}}] = \beta(N\sqrt{N})^{-1} \langle \tau_k \sigma_i \rangle$ , so Gaussian Poincaré gives (9).  $\square$

**Gaussian spins.** The spherical field formula gives the Euclidean Shcherbina–Tirozzi model [\[32\]](#). Let  $g^0, g_1, g_2, \dots$  be independent standard Gaussian vectors in  $\mathbb{R}^N$ , and define, for  $c > 0$ ,

$$\mathcal{Z}_{N, M}^{\text{ST}}(u, h, c) = \int_{\mathbb{R}^N} \exp \left\{ \sum_{k=1}^M u \left( \frac{g_k \cdot x}{\sqrt{N}} \right) + h g^0 \cdot x - c \|x\|_2^2 \right\} dx.$$

For  $\rho > 0$ , put

$$u_\rho(z) = u(\sqrt{\rho} z)$$

and

$$\mathcal{P}_{\text{ST}}(\alpha; u, h, c) := \frac{1}{2} \log(2\pi e) + \sup_{\rho > 0} \left\{ \frac{1}{2} \log \rho - c\rho + \inf_{q \in \mathcal{Q}} \left[ \mathcal{A}_{\text{S}}(q) + \alpha \Phi_{u_\rho}(q) + \frac{h^2 \rho}{2} \mathcal{L}(q) \right] \right\}. \quad (10)$$

**Theorem 1.4** (Gaussian-spin free energy). *Let  $u : \mathbb{R} \rightarrow \mathbb{R}$  be real-valued and globally Lipschitz, let  $c > 0$ ,  $h \in \mathbb{R}$ , and  $M_N/N \rightarrow \alpha \in (0, \infty)$ . Then*

$$\frac{1}{N} \log \mathcal{Z}_{N, M_N}^{\text{ST}}(u, h, c) \longrightarrow \mathcal{P}_{\text{ST}}(\alpha; u, h, c) \quad \text{in } L^1.$$

In particular, under the hypotheses of Talagrand’s Research Problem 3.1.2, [theorem 1.4](#) computes the limit [\[34, Research Problem 3.1.2, p. 193\]](#). It does not resolve the separate Research Problem 3.2.13 on uniqueness of the scalar replica-symmetric equations [\[34, Research Problem 3.2.13, p. 222\]](#); the optimizer in (10) is not assumed to be a constant path.

*Proof.* Polar disintegration with  $x = \sqrt{\rho} \sigma$ ,  $\sigma \in S_N$ , gives

$$\mathcal{Z}_{N, M_N}^{\text{ST}}(u, h, c) = C_N \int_0^\infty \rho^{N/2-1} e^{-cN\rho} Z_{N, M_N}^{\text{S}, u_\rho, h\sqrt{\rho}} d\rho, \quad C_N = \frac{(\pi N)^{N/2}}{\Gamma(N/2)}. \quad (11)$$

Stirling's formula yields

$$\frac{1}{N} \log C_N \longrightarrow \frac{1}{2} \log(2\pi e). \quad (12)$$

For  $s > 0$ , define, using the same disorder for every  $s$ ,

$$Y_N(s) := \frac{1}{N} \log Z_{N,M_N}^{S,u_{s^2},hs}, \quad Y(s) := \mathcal{P}_{S,u_{s^2}}^{\text{fld}}(\alpha, hs).$$

Since  $\rho = s^2$  gives  $\rho^{N/2-1} d\rho = 2s^{N-1} ds$ , we obtain

$$\frac{1}{N} \log \mathcal{Z}_{N,M_N}^{\text{ST}}(u, h, c) = \frac{1}{N} \log(2C_N) + \frac{1}{N} \log \int_0^\infty s^{N-1} e^{N[Y_N(s) - cs^2]} ds.$$

For these functions  $Y_N$  and  $Y$ , [theorem 8.4](#) gives

$$\frac{1}{N} \log \int_0^\infty s^{N-1} e^{N[Y_N(s) - cs^2]} ds \xrightarrow{L^1} \sup_{s>0} \{Y(s) + \log s - cs^2\}.$$

Adding (12), with  $N^{-1} \log 2 \rightarrow 0$ , proves  $L^1$  convergence of the full free energy to

$$\frac{1}{2} \log(2\pi e) + \sup_{\rho>0} \left\{ \frac{1}{2} \log \rho - c\rho + \mathcal{P}_{S,u_\rho}^{\text{fld}}(\alpha, h\sqrt{\rho}) \right\},$$

where we have returned to  $\rho = s^2$ . Substituting the definition of  $\mathcal{P}_{S,u_\rho}^{\text{fld}}$  from (7) identifies this expression with  $\mathcal{P}_{\text{ST}}(\alpha; u, h, c)$  in (10).  $\square$

## 1.2 Quenched large deviations

The free energy formula also determines nonlinear weights depending on the empirical distribution of the constraint fields. With the Gaussian rows  $(g_k)_k$  held fixed, define

$$L_{N,\sigma} = \frac{1}{M_N} \sum_{k \leq M_N} \delta_{h_k(\sigma)}, \quad \rho_N^\kappa = \mu_N^\kappa \circ L_{N,\cdot}^{-1}.$$

On the space  $\mathcal{P}(\mathbb{R})$  of probability measures, equipped with its weak topology, put

$$\mathcal{J}_{\kappa,\alpha}(\mu) = \sup_{u \in \text{BL}(\mathbb{R})} \left\{ \int u d\mu - \frac{\mathcal{P}_{\kappa,u}(\alpha)}{\alpha} \right\}, \quad (13)$$

where  $\text{BL}(\mathbb{R})$  denotes bounded Lipschitz functions.

**Theorem 1.5** (Quenched empirical measure large deviations). *For either prior, on one event of probability one,  $\rho_N^\kappa$  satisfies a large-deviation principle with speed  $M_N$  and good convex rate function  $\mathcal{J}_{\kappa,\alpha}$ . On the same event, for every bounded weakly continuous  $\Psi : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}$ ,*

$$\frac{1}{N} \log \int e^{M_N \Psi(L_{N,\sigma})} \mu_N^\kappa(d\sigma) \longrightarrow \alpha \sup_\mu \{ \Psi(\mu) - \mathcal{J}_{\kappa,\alpha}(\mu) \}. \quad (14)$$

Note that no convexity of  $\Psi$  is assumed in Theorem 1.5. For unnormalized Ising counting, add  $\log 2$  to the limit in Eq. (14). At  $M_N = N$ , this proves the quenched Laplace principle in [6, Definition 2.1 and Conjecture 2.2] for the Gaussian Ising perceptron. The result concerns constraint fields under the spin prior; it does not assert replica symmetry for a conditioned spin measure. The proof in [section 7](#) uses differentiability of the pressure in bounded channel directions.

### 1.3 Scalar projection pursuit and supervised learning

We next consider the problem of maximizing an empirical objective over a single projection direction. Let  $(X_i)_i$  be independent standard Gaussian vectors in  $\mathbb{R}^d$ , where  $n/d \rightarrow \alpha \in (0, \infty)$ . Fix a teacher dimension  $k \geq 1$ , deterministic matrices  $W_{*,d} \in \mathbb{R}^{d \times k}$  with orthonormal columns, and labels

$$Y_i = \varphi(W_{*,d}^\top X_i, \varepsilon_i),$$

where  $\varphi$  is measurable and real-valued and the  $\varepsilon_i$  are i.i.d. with a fixed law, independently of the covariates. The objective is  $n^{-1} \sum_i h(Y_i, w \cdot X_i)$ , with  $\|w\|_2 = 1$  and bounded continuous  $h$ .

The limiting value uses the following scalar functional. Write  $C_b$  for bounded continuous functions and let  $\mathcal{U}$  be the nonnegative, nondecreasing integrable functions on  $[0, 1]$ . For  $H \in C_b(\mathbb{R})$  and  $c > 0$ , set

$$\mathbf{H}_c H(x) = \sup_z \{H(x+z) - z^2/(2c)\}.$$

Let  $f_{\mu,c}^H$  solve (1) on  $[0, 1]$  with coefficient  $\mu$  and terminal function  $h = \mathbf{H}_c H$ . The function  $h$  is bounded and Lipschitz, so the definition applies to every  $\mu \in \mathcal{U}$ .

**Theorem 1.6** (Scalar Montanari–Zhou formula). *Under the preceding assumptions, for every fixed  $h \in C_b(\mathbb{R}^2)$ ,*

$$\max_{\|w\|_2=1} \frac{1}{n} \sum_{i=1}^n h(Y_i, w \cdot X_i) \longrightarrow \sup_{r \in B_k(1)} \inf_{\mu \in \mathcal{U}, c>0} \mathbf{F}_{\alpha,h}(\mu, c, r)$$

*almost surely and in every  $L^p$  with  $p < \infty$ , where  $B_k(1)$  is the closed unit ball and*

$$\mathbf{F}_{\alpha,h}(\mu, c, r) = \mathbb{E} \left[ f_{\mu,c}^{h(Y,\cdot)}(|r|^2, r \cdot G) \right] + \frac{1}{2\alpha} \int_{|r|^2}^1 \frac{dt}{c + \int_t^1 \mu(s) ds}. \quad (15)$$

*Here  $G \sim N(0, I_k)$  is independent of  $\varepsilon$ , and  $Y = \varphi(G, \varepsilon)$ . At  $|r| = 1$ , the inner infimum equals  $\mathbb{E}h(Y, r \cdot G)$ .*

*Moreover, for every fixed  $H \in C_b(\mathbb{R})$ ,*

$$\max_{\|w\|_2=1} \frac{1}{n} \sum_{i=1}^n H(w \cdot X_i) \longrightarrow \inf_{\mu \in \mathcal{U}, c>0} \left\{ f_{\mu,c}^H(0, 0) + \frac{1}{2\alpha} \int_0^1 \frac{dt}{c + \int_t^1 \mu(s) ds} \right\}$$

*almost surely and in every  $L^p$  with  $p < \infty$ .*

These formulas establish the supervised scalar conjecture [20, Conjecture 2.2, equations (14)–(15)] and the  $m = 1$  projection-pursuit conjecture [19, Conjecture 2.1 and Remark 1]. The first applies, in particular, under their sub-Gaussian label assumption. The limits also identify their feasible-distribution duals, which equal the corresponding probability limit inferiors [20, equation (9)][19, equation (3)].

### 1.4 Hard constraints and capacities

We use the free-energy formula for Lipschitz potentials to derive formulas for the hard-constraint free energy at margin  $\theta \in \mathbb{R}$ . Let  $C \subsetneq \mathbb{R}$  be a nonempty closed set containing either  $[\theta, \infty)$  or  $(-\infty, \theta]$ . By reflection, we may assume that

$$[\theta_C, \infty) \subset C$$

for some finite  $\theta_C$ ; we call such a set *tail-accessible*. For  $Z \sim N(0, 1)$ , put

$$d_C(x) = \text{dist}(x, C), \quad u_{\beta, C}(x) = -\beta d_C(x), \quad \pi_C = \mathbb{P}(Z \in C) \in (0, 1).$$

The feasible set and its normalized entropy are

$$\begin{aligned} \mathcal{F}_{M, N}^\kappa(C) &= \{\sigma \in \Sigma_N^\kappa : h_k(\sigma) \in C \text{ for all } k \leq M\}, \\ \text{Ent}_{N, \kappa}(C) &= \frac{1}{N} \log \mu_N^\kappa(\mathcal{F}_{M, N}^\kappa(C)), \quad \log 0 = -\infty. \end{aligned}$$

For Ising spins, the unnormalized counting entropy is  $\log 2 + \text{Ent}_{N, \text{I}}(C)$ . Define the hard channel and variational value by

$$\mathcal{P}_{\kappa, \infty, C}(\alpha) = \inf_{q \in \mathcal{Q}: \mathcal{A}_\kappa(q) < \infty} \{\mathcal{A}_\kappa(q) + \alpha \Phi_{\infty, C}(q)\}, \quad \Phi_{\infty, C}(q) = \lim_{\beta \rightarrow \infty} \Phi_{u_{\beta, C}}(q),$$

both taking values in  $[-\infty, 0]$ , with  $\mathcal{P}_{\kappa, \infty, C}(0) = 0$ .

**Theorem 1.7** (Closed-set hardening). *Let  $C$  be tail-accessible and  $M_N/N \rightarrow \alpha > 0$ . For either prior,*

$$\mathcal{P}_{\kappa, u_{\beta, C}}(\alpha) \downarrow \mathcal{P}_{\kappa, \infty, C}(\alpha) \quad (\beta \rightarrow \infty). \quad (16)$$

*For spherical spins,*

$$\text{Ent}_{N, \text{S}}(C) \xrightarrow{\mathbb{P}} \mathcal{P}_{\text{S}, \infty, C}(\alpha) \quad (17)$$

*in the extended real line. For Ising spins, if  $\mathcal{P}_{\text{I}, \infty, C}(\alpha) > -\log 2$ , then*

$$\begin{aligned} \text{Ent}_{N, \text{I}}(C) &\xrightarrow{\mathbb{P}} \mathcal{P}_{\text{I}, \infty, C}(\alpha), \\ \frac{1}{N} \log \# \mathcal{F}_{M_N, N}^{\text{I}}(C) &\xrightarrow{\mathbb{P}} \log 2 + \mathcal{P}_{\text{I}, \infty, C}(\alpha) > 0. \end{aligned}$$

*If  $\mathcal{P}_{\text{I}, \infty, C}(\alpha) < -\log 2$ , then*

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(\mathcal{F}_{M_N, N}^{\text{I}}(C) \neq \emptyset) < 0. \quad (18)$$

Define the candidate Ising capacity by the variational ratio

$$\alpha_c^{\text{I}}(C) = \inf_{q \in \mathcal{Q}: \mathcal{A}_{\text{I}}(q) < \infty} \frac{\log 2 + \mathcal{A}_{\text{I}}(q)}{-\Phi_{\infty, C}(q)},$$

where the ratio is zero if  $\Phi_{\infty, C}(q) = -\infty$ . The denominator is positive by (131). Under a common infinite sequence of rows, define for either prior

$$\mathfrak{M}_N^\kappa(C) = \sup\{M \in \mathbb{N}_0 : \mathcal{F}_{M, N}^\kappa(C) \neq \emptyset\}.$$

**Theorem 1.8** (Ising capacity). *For every tail-accessible  $C$ , the critical constraint density satisfies*

$$0 < \alpha_c^{\text{I}}(C) \leq \frac{\log 2}{-\log \pi_C} < \infty. \quad (19)$$

*It is characterized by positivity of the variational counting entropy:*

$$\alpha_c^{\text{I}}(C) = \sup\{\alpha \geq 0 : \log 2 + \mathcal{P}_{\text{I}, \infty, C}(\alpha) > 0\}. \quad (20)$$

If  $\alpha \neq \alpha_c^I(C)$ , then the probability that all  $M_N$  constraints admit a common feasible Ising configuration satisfies

$$\mathbb{P}(\mathcal{F}_{M_N, N}^I(C) \neq \emptyset) \longrightarrow \begin{cases} 1, & 0 < \alpha < \alpha_c^I(C), \\ 0, & \alpha > \alpha_c^I(C). \end{cases} \quad (21)$$

If  $0 < \alpha < \alpha_c^I(C)$ , then

$$\frac{1}{N} \log \# \mathcal{F}_{M_N, N}^I(C) \xrightarrow{\mathbb{P}} \log 2 + \mathcal{P}_{1, \infty, C}(\alpha) > 0,$$

with the convention  $\log 0 = -\infty$ . If  $\alpha > \alpha_c^I(C)$ , then

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(\mathcal{F}_{M_N, N}^I(C) \neq \emptyset) < 0.$$

Moreover,  $\mathfrak{M}_N^I(C) < \infty$  almost surely for each  $N$ , and

$$\frac{\mathfrak{M}_N^I(C)}{N} \xrightarrow{\mathbb{P}} \alpha_c^I(C). \quad (22)$$

**Nonmonotone constraints.** The  $U$ -function constraint corresponds to

$$C_K^U = (-\infty, -K] \cup [K, \infty), \quad K > 0.$$

Its soft channel is  $-\beta(K - |x|)_+$ . The two theorems give its spherical hard volume and Ising entropy and capacity, with  $\pi_{C_K^U} = 2\mathbb{P}(Z \geq K)$ . For  $K > K_* \simeq 0.817$ , this gives the full-path variational characterization called for by the wide-regime analysis of [3]; it does not numerically evaluate the capacity or identify an optimizing path. More examples and the limits of the tail-accessibility assumption are discussed in [section 8.5](#). For general  $C$ , a spherical volume formula does not imply an emptiness threshold.

**Spherical half-spaces.** For  $C_\theta = [\theta, \infty)$ , geometry turns the hard-volume formula into a satisfiability threshold. Define

$$\alpha_c^S(\theta) = \sup\{\alpha \geq 0 : \mathcal{P}_{S, \infty, C_\theta}(\alpha) > -\infty\}.$$

**Theorem 1.9** (Spherical half-space capacity). *For every  $\theta \in \mathbb{R}$  and positive  $\alpha \neq \alpha_c^S(\theta)$ ,*

$$\mathbb{P}(\mathcal{F}_{\lfloor \alpha N \rfloor, N}^S(C_\theta) \neq \emptyset) \longrightarrow \mathbf{1}_{\{\alpha < \alpha_c^S(\theta)\}}.$$

*For every  $\alpha \in (0, \alpha_c^S)$ , the hard-volume limit in (17) is finite. Moreover,  $\mathfrak{M}_N^S(C_\theta) < \infty$  almost surely and*

$$\frac{\mathfrak{M}_N^S(C_\theta)}{N} \xrightarrow{\mathbb{P}} \alpha_c^S(\theta). \quad (23)$$

*No assertion is made when  $M_N/N \rightarrow \alpha_c^S(C_\theta)$ .*

For negative margins, the entropy formula and the capacity characterization above establish the variational predictions of Franz and Parisi [13]. Their further predictions concerning the structure of the minimizing overlap distribution and the jamming critical exponents are not addressed here.

**Relation to three problems of Talagrand.** At zero Ising margin, [theorems 1.7](#) and [1.8](#) give a positive deterministic capacity, convergence of the normalized random capacity, the full subcritical variational entropy, and exponentially small supercritical satisfiability probability. These provide variational formulas for the threshold and entropy requested in Research Problem 2.1.1 [[34](#), Research Problem 2.1.1, p. 152]. The formulas optimize over all overlap paths; we do not identify a constant-path optimizer or evaluate the capacity numerically.

The Gaussian-spin formula in [theorem 1.4](#) answers Research Problem 3.1.2. For every negative margin, with Talagrand’s  $\tau$  equal to our  $\theta$ , [theorems 1.7](#) and [1.9](#) give the full-path hard-volume limit at every density and the SAT–UNSAT transition away from the critical density, answering the free energy and capacity aspects of Research Problem 8.4.3 [[35](#), Research Problem 8.4.3, p. 25]. We make no claim at the critical load and do not determine optimizer structure or critical exponents.

## 2 Proof of the free energy formula

In this section, we outline the proof of the Parisi formula for the soft free energy, [theorem 1.1](#), and state the main ingredients. We establish matching lower and upper bounds for the limiting free energy. The lower bound follows the Aizenman–Sims–Starr scheme, using increments obtained by adding finitely many constraints and coordinates. The upper bound uses a Guerra interpolation.

For both bounds, we add a Gaussian perturbation, independent of the constraint vectors, that changes the expected normalized log partition function by  $o(1)$ . This perturbation enforces the joint Ghirlanda–Guerra identities for subsequential limits of the disorder-averaged replica laws of the spin overlaps and cascade ancestry labels. Together with synchronization, these identities yield the limiting-array representation used in the cavity calculations.

Cavity in  $M$  determines the asymptotic effect of adding a fixed number of constraints. Cavity in  $N$  bounds the increment produced by adding a fixed block of coordinates. The overlap equation supplies the supporting equality needed to turn these calculations into the two variational bounds. We state these results below and deduce [theorem 1.1](#) from them at the end of this section. Sections [3–6](#) provide the proofs of these results and verify their limiting-array hypotheses.

Let us first work with centered smooth channels; this assumption will be removed later. For the cavity arguments, assume

$$u_\theta(0) = 0, \quad u_\theta \in C^3(\mathbb{R}), \quad L_j := \sup_{\theta \in \mathbb{R}} \|u_\theta^{(j)}\|_\infty < \infty \quad (1 \leq j \leq 3). \quad (24)$$

Put  $L = L_1$ , fix  $\kappa \in \{\text{I}, \text{S}\}$ , and write  $u_k = u_{\Theta_k}$ , where the marks  $\Theta_k \sim \nu$  are i.i.d. and independent of the Gaussian vectors and all further randomness introduced below. We work at zero external field (meaning  $b = 0$ ) and abbreviate  $\mathcal{A} = \mathcal{A}_\kappa$  and  $\mathcal{S} = \mathcal{S}_\kappa$ .

### 2.1 Conditional constraint insertion and reconstruction

We first isolate the effect of adding a fixed number of independent constraints  $u(h_k)$  to the Hamiltonian. Throughout, the channels and marks satisfy the centered smooth assumptions stated above, and  $L = \sup_\theta \|u'_\theta\|_\infty$ .

The key result in this section is proposition [2.1](#), which has two conclusions. First, adding a fixed number of independent constraints preserves the limiting joint distribution of the spin overlaps and ancestry labels  $(R_{\ell m}, r_{\ell m})_{\ell, m \geq 1}$ , after averaging over the disorder. Second, under the reconstruction hypotheses (stated in the proposition), the proposition identifies the limiting distribution of the score-overlap array  $Q$  ([36](#)) jointly with  $(R, r)$ .

**The finite-volume measure and its overlap array.** Fix  $d \geq 2$ . Let  $\mathfrak{R}_d$  be the RPC with masses  $j/d$ ,  $1 \leq j < d$ , viewed as a random probability measure on the leaf set  $\mathcal{L}_d = \mathbb{N}^{d-1}$ . For two leaves  $\eta, \eta'$ , let  $|\eta \wedge \eta'|$  be the number of initial coordinates they have in common. Define

$$r(\eta, \eta') := \begin{cases} |\eta \wedge \eta'|/d, & \eta \neq \eta', \\ 1, & \eta = \eta'. \end{cases}$$

We also write  $\eta \cdot \eta' = r(\eta, \eta')$ . Let  $G_N$  be a random probability measure on

$$\mathcal{S}_N := \Sigma_N^\kappa \times \mathcal{L}_d \times \mathbf{E}_N, \quad s = (\sigma, \eta, \xi),$$

where  $\mathbf{E}_N$  is any measurable space, allowing for possibly more integrated coordinates; it is a singleton when none are present. Let  $\mathcal{F}_N$  be a sigma-algebra determining  $G_N$ . In the applications, it contains the retained constraint vectors and channel marks, the cascade weights, and any Gaussian source fields and perturbations used to define  $G_N$ . It excludes the marks and Gaussian noises of constraints to be inserted.

Conditional on  $\mathcal{F}_N$ , sample replicas  $s^\ell = (\sigma^\ell, \eta^\ell, \xi^\ell)$  independently from  $G_N$ . For a function  $A$  of  $n$  replicas, write

$$\langle A \rangle_N := \int A(s^1, \dots, s^n) \prod_{\ell=1}^n G_N(ds^\ell).$$

Define

$$R_{\ell m} := \frac{\sigma^\ell \cdot \sigma^m}{N}, \quad r_{\ell m} := r(\eta^\ell, \eta^m). \quad (25)$$

Thus  $R_{\ell\ell} = r_{\ell\ell} = 1$ . The array  $(r_{\ell m})$  records the shared ancestry of the sampled leaves; this is their genealogy. The pair-ancestry distribution of  $\mathfrak{R}_d$  has quantile

$$\rho_d(v) = \begin{cases} j/d, & v \in [j/d, (j+1)/d), \quad 0 \leq j \leq d-2, \\ 1, & v \in [(d-1)/d, 1). \end{cases} \quad (26)$$

Let  $U = (U_{\ell m})$  be the limiting ancestry array of finite RPCs with ancestry grids of vanishing mesh, with  $U_{\ell\ell} = 1$ . Its law is uniquely determined by [26, Theorems 2.13, 2.17 and Remark 2.1], and  $U_{12}$  is uniform on  $[0, 1)$ . Write  $\mathbb{E}_U$  for expectation under this array law; its entries are not independent.

We work along a subsequence on which the disorder-averaged replica laws of  $(R, r)$  converge. Denoting the limiting array by  $(R^\infty, r^\infty)$ , the required representation is

$$(R_{\ell m}^\infty, r_{\ell m}^\infty)_{\ell < m} \stackrel{d}{=} (q(U_{\ell m}), \rho_d(U_{\ell m}))_{\ell < m}, \quad q \in \mathcal{Q}, \quad (27)$$

with diagonal  $(1, 1)$ . Here  $\mathcal{Q}$  consists of right-continuous nondecreasing paths  $[0, 1] \rightarrow [0, 1]$ , with value one at 1, identified almost everywhere on  $[0, 1)$ . The path  $q$  is deterministic at the fixed parameters. The array  $U$  represents the limiting law; it is not a coordinate of  $G_N$  and need not be recoverable from  $(R^\infty, r^\infty)$ . Section 6 verifies this representation using joint Ghirlanda–Guerra identities and synchronization.

**The scalar response.** We now define the scalar quantities associated with a covariance path  $c \in \mathcal{Q}$ . Although  $\rho_d$  is a fixed step function,  $c$  need not be constant on its intervals. We therefore use auxiliary cascades on grids adapted to  $c$ . These cascades are used only for the scalar calculation; they do not add coordinates to  $G_N$  or change  $\mathfrak{R}_d$ .

First suppose that  $c$  is constant on each interval  $[m_j, m_{j+1})$  of a grid

$$\pi = \{0 = m_0 < \dots < m_{J+1} = 1\}.$$

Let  $\mathfrak{R}_\pi$  be the RPC with masses  $m_1, \dots, m_J$ , with leaves indexed by  $\mathbb{N}^J$ , and put

$$r^\pi(\eta, \eta') := m_{|\eta \wedge \eta'|}.$$

In particular,  $r^\pi(\eta, \eta) = m_J$ . For distinct replicas set  $r_{\ell m}^\pi = r^\pi(\eta^\ell, \eta^m)$ , and set  $r_{\ell \ell}^\pi = 1$  separately. Thus, unlike the finite-volume convention, the auxiliary label of two distinct replicas on the same leaf is  $m_J$ , not one. Write  $\mathbb{E}_\pi$  for the averaged law of these ancestry labels when the leaves are sampled from the unweighted cascade  $\mathfrak{R}_\pi$ .

Independently of the cascade weights, take a centered Gaussian field  $(\bar{X}^c(\eta))_{\eta \in \mathbb{N}^J}$  with covariance

$$\mathbb{E}[\bar{X}^c(\eta) \bar{X}^c(\eta')] = c(r^\pi(\eta, \eta')).$$

Its variance is  $c_* := c(1-) = c(m_J)$ . Fix a channel  $u$ , and write

$$\gamma(\mathrm{d}z) := (2\pi)^{-1/2} e^{-z^2/2} \mathrm{d}z.$$

Define

$$Z_{\pi, c, u} := \int_{\mathbb{N}^J} \int_{\mathbb{R}} e^{u(\bar{X}^c(\eta) + \sqrt{1-c_*} z)} \gamma(\mathrm{d}z) \mathfrak{R}_\pi(\mathrm{d}\eta).$$

Set

$$X^c(\eta, z) := \bar{X}^c(\eta) + \sqrt{1-c_*} z, \quad G_{\pi, c, u}(\mathrm{d}\eta, \mathrm{d}z) := \frac{e^{u(X^c(\eta, z))}}{Z_{\pi, c, u}} \mathfrak{R}_\pi(\mathrm{d}\eta) \gamma(\mathrm{d}z). \quad (28)$$

Write  $\langle \cdot \rangle_{\pi, c, u}$  for independent-replica expectation under  $G_{\pi, c, u}$ , and put  $X^{c, \ell} := X^c(\eta^\ell, z^\ell)$ . The scalar calculation in [theorem 3.2](#) gives

$$\mathbb{E} \log Z_{\pi, c, u} = \Phi_u(c),$$

where  $\Phi_u$  is the channel functional from [\(2\)](#), and defines a step response  $p_{c, u}$  by

$$\mathbb{E} \left\langle f(r_{12}^\pi) u'(X^{c, 1}) u'(X^{c, 2}) \right\rangle_{\pi, c, u} = \sum_{j=0}^J (m_{j+1} - m_j) f(m_j) p_{c, u}(m_j) \quad (29)$$

for every bounded Borel  $f : [0, 1] \rightarrow \mathbb{R}$ , with  $p_{c, u}$  constant on each  $[m_j, m_{j+1})$ . The response  $p_{c, u}$  is nonnegative, nondecreasing, bounded by  $\|u'\|_\infty^2$ , and constant on every interval on which  $c$  is constant. For step paths  $c, \tilde{c}$  on a common grid, Gaussian covariance differentiation gives

$$\left. \frac{\mathrm{d}}{\mathrm{d}s} \Phi_u((1-s)c + s\tilde{c}) \right|_{s=0+} = -\frac{1}{2} \int_0^1 p_{c, u}(r) (\tilde{c}(r) - c(r)) \mathrm{d}r.$$

Thus the two-replica derivative correlation also determines the first variation of  $\Phi_u$  in the covariance path.

For general  $c \in \mathcal{Q}$ , choose step paths  $c_n \rightarrow c$  in  $L^1([0, 1])$  on grids  $\pi_n$  of vanishing mesh. The averaged finite-replica laws of the ancestry labels and completed inputs, and the responses  $p_{c_n, u}$ , converge independently of the approximation. Their limits define  $\mathbb{E}\langle \cdot \rangle_{c, u}$  and  $p_{c, u}$ , respectively; response convergence is in  $L^1$ . For example,

$$\mathbb{E}\langle \chi(X^c) \rangle_{c, u} := \lim_{n \rightarrow \infty} \mathbb{E}\langle \chi(X^{c_n}) \rangle_{\pi_n, c_n, u}$$

for bounded continuous  $\chi : \mathbb{R} \rightarrow \mathbb{R}$ . Section 3 proves existence, independence of approximation, and continuity.

For the random channels, define

$$\Phi(c) := \int_{\mathbb{R}} \Phi_{u_\theta}(c) \nu(d\theta), \quad p_c(r) := \int_{\mathbb{R}} p_{c,u_\theta}(r) \nu(d\theta). \quad (30)$$

Each channel is normalized separately before averaging over its mark. Averaging the fixed-channel derivative identity gives, for step paths on a common grid,

$$\left. \frac{d}{ds} \Phi((1-s)c + s\tilde{c}) \right|_{s=0+} = -\frac{1}{2} \langle p_c, \tilde{c} - c \rangle, \quad \langle a, b \rangle := \int_0^1 a(r)b(r) dr. \quad (31)$$

The map  $c \mapsto p_c$  is  $L^1$ -continuous, and for all  $c, \tilde{c} \in \mathcal{Q}$ ,

$$|\Phi(c) - \Phi(\tilde{c})| \leq \frac{1}{2} L^2 \|c - \tilde{c}\|_{L^1}. \quad (32)$$

**The Gaussian fields to be inserted.** Return to the finite-volume measure  $G_N$ . Let  $\mathbf{C}_N$  be an  $\mathcal{F}_N$ -measurable covariance kernel on  $\mathcal{S}_N$ , with constant diagonal  $\mathbf{C}_N(s, s) = v_N$ , where

$$0 \leq v_N \leq 1 + \varepsilon_N, \quad v_N \rightarrow v \in [0, 1], \quad \varepsilon_N \rightarrow 0.$$

Conditional on  $\mathcal{F}_N$ , construct a centered measurable Gaussian process  $\bar{X}_N$  with covariance

$$\mathbb{E}[\bar{X}_N(s)\bar{X}_N(s') \mid \mathcal{F}_N] = \mathbf{C}_N(s, s'),$$

using Gaussian randomness independent of replica sampling and of the new channel marks. Define

$$X_N(s, z) := \bar{X}_N(s) + \sqrt{1 + \varepsilon_N - v_N} z, \quad z \text{ integrated against } \gamma.$$

For a deterministic  $c \in \mathcal{Q}$  with  $c_* \leq v$ , assume that, for every fixed  $n \geq 2$ ,

$$(R_{\ell m}, r_{\ell m}, \mathbf{C}_N(s^\ell, s^m))_{1 \leq \ell < m \leq n} \implies (q(U_{\ell m}), \rho_d(U_{\ell m}), c(U_{\ell m}))_{1 \leq \ell < m \leq n}. \quad (33)$$

This convergence averages over  $\mathcal{F}_N$  and replicas sampled from  $G_N$ , before insertion. The diagonals of the displayed triples are  $(1, 1, v_N) \rightarrow (1, 1, v)$ . Different constraints use conditionally independent copies of  $\bar{X}_N$  and separate Gaussian integration coordinates.

For  $n \geq 2$ , set

$$\mathbf{D}_n := ([-1, 1] \times [0, 1])^{n \times n}, \quad A_N^{(n)}(s^1, \dots, s^n) := ((R_{\ell m}, r_{\ell m}))_{1 \leq \ell, m \leq n}.$$

Thus  $A_N^{(n)}$  depends only on the spin and leaf coordinates of the  $n$  states. Define the limiting array block

$$A_\infty^{(n)} := (a_{\ell m})_{1 \leq \ell, m \leq n}, \quad a_{\ell m} := \begin{cases} (q(U_{\ell m}), \rho_d(U_{\ell m})), & \ell \neq m, \\ (1, 1), & \ell = m, \end{cases}$$

where  $U$  has the array law specified in (27), and  $\mathbb{E}_U$  denotes expectation under that law. For a deterministic bounded continuous function  $F : \mathbf{D}_n \rightarrow \mathbb{R}$ , write

$$F_N(s^1, \dots, s^n) := F(A_N^{(n)}(s^1, \dots, s^n)).$$

The function  $F$  is fixed as  $N$  varies and does not depend on the disorder. The assumed convergence of the replica array under  $G_N$  means that

$$\mathbb{E}\langle F_N \rangle_N \longrightarrow \mathbb{E}_U F(A_\infty^{(n)})$$

for every  $n \geq 2$  and every  $F$  as above.

**Proposition 2.1** (Independent constraint insertion and reconstruction). *Assume the preceding hypotheses on  $G_N$ , its limiting array, and the new Gaussian fields, with deterministic limiting paths  $q, c$ .*

*Insertion. Fix  $k \geq 1$ . Give the new constraints independent marks  $\Theta_b \sim \nu$ , independent of  $\mathcal{F}_N$  and the Gaussian noises, and write  $u_b = u_{\Theta_b}$ . Let  $\bar{X}_{N,b}$ ,  $1 \leq b \leq k$ , be the conditionally independent Gaussian fields specified above, and put*

$$X_b(s, z_b^0) := \bar{X}_{N,b}(s) + \sqrt{1 + \varepsilon_N - v_N} z_b^0.$$

*Write  $z^0 = (z_1^0, \dots, z_k^0)$ , and let  $\gamma$  be standard Gaussian probability measure on  $\mathbb{R}$ . With all marks and Gaussian fields held fixed, define*

$$Z_{N,k}^+ := \int_{\mathcal{S}_N \times \mathbb{R}^k} \exp \left\{ \sum_{b=1}^k u_b(X_b(s, z_b^0)) \right\} G_N(ds) \prod_{b=1}^k \gamma(dz_b^0),$$

$$G_{N,+k}(ds, dz^0) := \frac{\exp\{\sum_{b=1}^k u_b(X_b(s, z_b^0))\}}{Z_{N,k}^+} G_N(ds) \prod_{b=1}^k \gamma(dz_b^0).$$

*For a bounded measurable function  $B$  of  $n \geq 1$  augmented states, define*

$$\langle B \rangle_{+k} := \int B((s^\ell, z^{0,\ell})_{\ell=1}^n) \prod_{\ell=1}^n G_{N,+k}(ds^\ell, dz^{0,\ell}),$$

*and write  $X_b^\ell := X_b(s^\ell, z_b^{0,\ell})$ .*

*For every fixed  $n \geq 2$  and every deterministic bounded continuous  $F : \mathcal{D}_n \rightarrow \mathbb{R}$ ,*

$$\mathbb{E} \langle F_N \rangle_{+k} - \mathbb{E} \langle F_N \rangle_N \longrightarrow 0.$$

*Consequently,*

$$\mathbb{E} \langle F_N \rangle_{+k} \longrightarrow \mathbb{E}_U F_\infty.$$

*Moreover,*

$$\mathbb{E} \log Z_{N,k}^+ \longrightarrow k\Phi(c). \quad (34)$$

*For each  $k \in \{1, 2\}$ , every  $n \geq 2$ , and every  $F$  as above,*

$$\mathbb{E} \left\langle F_N \prod_{b=1}^k u'_b(X_b^1) u'_b(X_b^2) \right\rangle_{+k} \longrightarrow \mathbb{E}_U [F_\infty p_c(U_{12})^k]. \quad (35)$$

**Reconstruction.** *Consider full measures  $G_{N,M}$ , with  $M = M_N \rightarrow \infty$ , on states with the same spin and leaf coordinates as above. Suppose that their disorder-averaged replica laws, including the channel marks and constraint inputs, are invariant under permutations of the  $M$  constraints.*

*For  $j = 1, 2$ , let  $G_{N,M}^{(-j)}$  denote the normalized measure obtained by deleting constraints  $1, \dots, j$  and omitting their added integration coordinates. Suppose that this measure is independent of the deleted marks and Gaussian noises, and that restoring these constraints by the preceding insertion gives  $G_{N,M}$ .*

*Assume that, along a common subsequence, the two deleted measures satisfy the insertion hypotheses, with deterministic paths  $q^{(-j)}, c^{(-j)}$ ,  $j = 1, 2$ , not assumed equal. All limits in this reconstruction statement are taken along that subsequence. In each deleted system, suppose that every restored constraint has the same off-diagonal covariance rule*

$$C_N(s^\ell, s^m) = K(R_{\ell m}, r_{\ell m}), \quad \ell \neq m,$$

where  $K : [-1, 1] \times [0, 1] \rightarrow \mathbb{R}$  is a fixed Borel function, and that

$$c^{(-j)}(v) = K(q^{(-j)}(v), \rho_d(v)) \quad \text{for almost every } v \in [0, 1), \quad j = 1, 2.$$

Then the two deleted measures and the full measure have the same limiting spin-overlap path:

$$q^{(-1)} = q^{(-2)} =: q \quad \text{almost everywhere on } [0, 1).$$

Their limiting covariance paths also agree:

$$c^{(-1)} = c^{(-2)} =: c, \quad c(v) = K(q(v), \rho_d(v)) \quad \text{almost everywhere on } [0, 1).$$

In particular, the full measure has limiting overlap array (27) with this path  $q$ .

Finally, let  $X_b^\ell = \bar{X}_{N,b}(s_{-b}^\ell) + \sqrt{1 + \varepsilon_N - v_N} z_b^{0,\ell}$  be the input of constraint  $b$  at replica  $\ell$ , and put  $u_b = u_{\Theta_b}$ . Define

$$Q_{\ell m} := \frac{1}{M} \sum_{b=1}^M u'_b(X_b^\ell) u'_b(X_b^m). \quad (36)$$

Every joint subsequential limit  $(R^\infty, r^\infty, Q^\infty)$  of the disorder-averaged replica laws under  $G_{N,M}$  satisfies

$$Q_{12}^\infty = \bar{p}_c(R_{12}^\infty, r_{12}^\infty) = p_c(U_{12}) \quad \text{almost surely,} \quad (37)$$

where  $\bar{p}_c : [0, 1]^2 \rightarrow \mathbb{R}$  is a bounded Borel function satisfying

$$\bar{p}_c(q(v), \rho_d(v)) = p_c(v) \quad \text{for almost every } v \in [0, 1).$$

The final equality in (37) is understood in the limiting representation (27). For every bounded Borel function  $\chi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  that is continuous in its second argument,

$$\mathbb{E} \left\langle \left| \frac{1}{M} \sum_{b=1}^M \chi(\Theta_b, X_b^1) - m_c(\chi) \right|^2 \right\rangle \rightarrow 0, \quad (38)$$

$$m_c(\chi) := \int_{\mathbb{R}} \mathbb{E} \langle \chi(\theta, X^c) \rangle_{c, u_\theta} \nu(d\theta).$$

The scalar expectation is the normalized finite-cascade limit defined above, with the channel mark averaged afterward.

In particular, define the diagonal derivative correlation by

$$p_{\text{diag}}(c) := \int_{\mathbb{R}} \mathbb{E} \langle u'_\theta(X^c)^2 \rangle_{c, u_\theta} \nu(d\theta). \quad (39)$$

Applying (38) with  $\chi(\theta, x) = u'_\theta(x)^2$  gives  $\mathbb{E} \langle |Q_{11} - p_{\text{diag}}(c)|^2 \rangle \rightarrow 0$ . This one-replica quantity need not equal  $p_c(1-)$ , which is a two-replica correlation. Even when two replicas share a scalar cascade leaf, their terminal coordinates are sampled separately. Conditional Jensen in the scalar transition law gives

$$0 \leq p_c(r) \leq p_{\text{diag}}(c) \leq L^2 \quad \text{for a.e. } r \in [0, 1), \quad p_c(1-) \leq p_{\text{diag}}(c).$$

To explain the reconstruction, put  $S_b := u'_b(X_b^1) u'_b(X_b^2)$ , so that  $Q_{12} = M^{-1} \sum_b S_b$ . Constraint exchangeability gives

$$\begin{aligned} \mathbb{E} \langle F Q_{12} \rangle &= \mathbb{E} \langle F S_1 \rangle, \\ \mathbb{E} \langle F Q_{12}^2 \rangle &= \frac{M-1}{M} \mathbb{E} \langle F S_1 S_2 \rangle + \frac{1}{M} \mathbb{E} \langle F S_1^2 \rangle. \end{aligned} \quad (40)$$

The last term is bounded in absolute value by  $\|F\|_\infty L^4/M$ . Deleting and restoring one constraint in the first identity and two in the second therefore gives

$$\mathbb{E}\langle FQ_{12} \rangle \longrightarrow \mathbb{E}_U[F_\infty p_c(U_{12})], \quad (41)$$

$$\mathbb{E}\langle FQ_{12}^2 \rangle \longrightarrow \mathbb{E}_U[F_\infty p_c(U_{12})^2]. \quad (42)$$

The common covariance rule and the plateau property make  $p_c(U_{12})$  a Borel function of  $(R_{12}^\infty, r_{12}^\infty)$ . These identities determine the conditional first and second moments of  $Q_{12}^\infty$  given the limiting overlap-ancestry array. Its conditional variance is zero, proving (37). When several constraints are inserted together, their weights share the random normalizer  $Z_{N,k}^+$ . These conclusions therefore use the normalized RPC calculation, not independence of the unnormalized constraint weights alone. Section 3.2 proves the insertion and reconstruction statements, including agreement of the deleted and full limiting paths.

## 2.2 Identifying the overlap equation

In this section we identify a self-consistency equation the order parameter  $q$ , corresponding to spin overlap  $R_{12}$ , must satisfy. Since this equation is used in the upper bound as well, we specify the interpolation and the perturbation before stating the overlap equation. Fix  $d \geq 2$ ,  $t \in [0, 1]$ , a trial path  $q^0 \in \mathcal{Q}_d$ , and a bounded nonnegative nondecreasing path  $a^0$  on the same grid. Here  $\mathcal{Q}_d$  consists of paths constant on  $[j/d, (j+1)/d)$ . Write  $b_* = b(1-)$  for a path's last off-diagonal value and define its lookup on the finite ancestry labels by

$$b^\#(\rho_d(u)) = b(u) \quad \text{for a.e. } u \in [0, 1].$$

In particular  $b^\#(1) = b_*$ . Let  $\bar{Y}_k$  be independent Gaussian fields on the cascade, with covariance  $q^{0,\#}(r_{\eta\eta'})$  and variance  $q_*^0$ , and put

$$Y_k(\eta, z_k^0) = \bar{Y}_k(\eta) + \sqrt{1 - q_*^0} z_k^0, \quad X_{k,t} = \sqrt{t} h_k(\sigma) + \sqrt{1 - t} Y_k(\eta, z_k^0).$$

The variables  $z_k^0$  are integrated against standard Gaussian measure. Independently, let  $W^{a^0} = (W_i^{a^0})_{i \leq N}$  have independent spatial coordinates of covariance  $a^{0,\#}(r_{\eta\eta'})$  and variance  $a_*^0$ , and set  $W_t = \sqrt{1 - t} W^{a^0}$ . All Gaussian ancestor marks are independent of the cascade weights.

Use a centered Gaussian perturbation  $V_N(\sigma, \eta)$ , independent of the channel marks, constraint vectors, and source fields, whose law does not depend on  $M$  or  $t$ . It has amplitudes  $\mathbf{x} \in [1, 2]^{\mathbb{N}}$ ;  $\mathbb{E}_{\mathbf{x}}$  averages their independent uniform laws, while  $\mathbb{E}$  holds them fixed. Use the same realization for every constraint count. Section 6 constructs this perturbation and proves the properties stated below.

Define

$$\mathcal{H}_{N,M,t} = \sum_{k \leq M} u_k(X_{k,t}) + W_t \cdot \sigma - \frac{N}{2}(1 - t)a_*^0 + V_N, \quad (43)$$

and let  $Z_{N,M,t}$  normalize

$$G_{N,M,t}(d\sigma, d\eta, dz^0) = \frac{e^{\mathcal{H}_{N,M,t}}}{Z_{N,M,t}} \mu_N^\kappa(d\sigma) \mathfrak{R}_d(d\eta) \gamma^{\otimes M}(dz^0). \quad (44)$$

Adding  $V_N$  changes normalized expected pressure by  $o(1)$ , uniformly in the amplitudes and time. At  $t = 1$ , removing it gives the original zero-field model. For every bounded Borel  $f$ , the finite-volume ancestry marginal is exactly

$$\mathbb{E}\langle f(r_{12}) \rangle_{N,M,t} = \int_0^1 f(\rho_d(s)) ds. \quad (45)$$

Deleting a constraint leaves its Gaussian noises independent of the remaining measure. Its covariance on restoration is

$$tR_{\ell m} + (1-t)q^{0,\sharp}(r_{\ell m}) \quad (\ell \neq m), \quad \text{Var}(X_{k,t}^\ell) = 1. \quad (46)$$

Thus, if the spin path is  $q$ , constraint reconstruction uses

$$c = (1-t)q^0 + tq.$$

Here are the two scalar responses. For a step source  $a$ , let  $w^a$  be a Gaussian field on  $\mathfrak{R}_\pi$  with covariance  $a(r^\pi)$  and variance  $a_*$ . For Ising spins, set

$$\mathbb{P}_{\pi,a}^{\text{I}}(\text{d}\eta, \text{d}\tau) = \frac{e^{\tau w^a(\eta) - a_*/2}}{Z_{\pi,a}^{\text{I}}} \mathfrak{R}_\pi(\text{d}\eta) \frac{\delta_{-1} + \delta_1}{2}(\text{d}\tau).$$

Its two-replica correlation defines  $\hat{q}_a^{\text{I}}$ ; for general  $a$ , take the finite-cascade limit established in Section 4:

$$\mathbb{E}\langle f(U_{12})\tau^1\tau^2 \rangle_a^{\text{I}} = \int_0^1 f(r)\hat{q}_a^{\text{I}}(r) \text{d}r.$$

For the sphere, the response has an explicit integral formula. Put  $x_a(s) = \inf\{r \in [0, 1) : a(r) \geq s\}$ , with  $\inf \emptyset = 1$ , and set

$$\beta_-(a) = a_* - \int_0^1 a(r) \text{d}r, \quad \mathfrak{D}_a = (\beta_-(a), \infty).$$

For  $\beta \in \mathfrak{D}_a$ , define

$$\begin{aligned} \pi_{a,\beta}(s) &= \beta - \int_s^{a_*} x_a(v) \text{d}v, \\ q_{a,\beta}^{\text{co}}(r) &= \int_0^{a(r)} \frac{\text{d}s}{\pi_{a,\beta}(s)^2}, \quad V_a(\beta) = \frac{1}{\beta} + \int_0^{a_*} \frac{\text{d}s}{\pi_{a,\beta}(s)^2}. \end{aligned} \quad (47)$$

Section 4 proves that  $V_a(\beta) = 1$  has a unique solution  $\bar{\beta}(a) \in \mathfrak{D}_a$ . Define

$$\hat{q}_a^{\text{S}}(r) = q_{a,\bar{\beta}(a)}^{\text{co}}(r) \quad (r < 1), \quad \hat{q}_a^{\text{S}}(1) = 1.$$

The spherical spin transform is

$$\mathcal{S}_{\text{S}}(a) = \frac{1}{2} \left\{ \bar{\beta}(a) - 1 - a_* - \log \bar{\beta}(a) + \int_0^{a_*} \frac{\text{d}s}{\pi_{a,\bar{\beta}(a)}(s)} \right\}.$$

Both responses are nonnegative, nondecreasing, bounded by one, and constant on plateaus of  $a$ .

**Proposition 2.2** (Scalar spin identities). *For either prior, every  $q \in \mathcal{Q}$ , and every  $a \in \mathcal{P}_{\text{b}}$ ,*

$$\begin{aligned} \mathcal{A}_\kappa(q) &= \sup_{b \in \mathcal{P}_{\text{b}}} \{\mathcal{S}_\kappa(b) + \tfrac{1}{2}\langle b, q \rangle\}, \\ \mathcal{A}_\kappa(\hat{q}_a^\kappa) &= \mathcal{S}_\kappa(a) + \tfrac{1}{2}\langle a, \hat{q}_a^\kappa \rangle < \infty. \end{aligned}$$

For a fixed step source  $a$  on  $\pi$ , let  $W^a$  have  $N$  independent coordinates distributed as  $w^a$ . Then

$$\frac{1}{N} \mathbb{E} \log \int \int e^{W^a(\eta) \cdot \sigma - Na_*/2} \mu_N^\kappa(\text{d}\sigma) \mathfrak{R}_\pi(\text{d}\eta) \longrightarrow \mathcal{S}_\kappa(a). \quad (48)$$

For Ising spins, equality holds for every  $N$ .

Section 4 proves these identities; the Ising conjugate formula is already the definition of  $\mathcal{A}_I$ .

Call  $(q, a)$  a supporting pair when  $\mathcal{A}(q) = \mathcal{S}(a) + \langle a, q \rangle / 2$ . Applying the conjugate inequality to each of two supporting pairs and adding gives

$$\langle q_1 - q_0, a_1 - a_0 \rangle \geq 0.$$

This is the sign used in the upper bound. The Ising supporting equality uses the Chen–Issa–Mourrat concavity theorem [9]; the spherical one follows from a Gaussian integral. Neither requires convexity of the channel functional  $\Phi$ .

For the next two propositions, retain fixed amplitudes and a joint array-convergent subsequence for the full measures and their one- and two-constraint deletions. For Ising spins also include the system obtained by deleting one coordinate from dimension  $D$ , retaining the denominator  $\sqrt{D}$  and using  $V_{D-1}$ , and its constraint deletions.

Section 6 supplies the required limits as follows. From any dimension sequence, one may select a common subsequence for any fixed finite collection of constraint-count changes and the preceding deletions. For almost every amplitude vector at  $t = 1$ , and for almost every amplitude–time pair in the interpolation, every further joint array limit has the form (27). Compactness supplies such refinements at each retained parameter. The paths of different systems are not assumed to agree.

**Proposition 2.3** (The overlap equation). *Let  $M/N \rightarrow \alpha \in (0, \infty)$ , and let  $q$  be the limiting spin path of  $G_{N,M,t}$ . Set*

$$c = (1 - t)q^0 + tq, \quad a = (1 - t)a^0 + t\alpha p_c.$$

*Then*

$$q = \hat{q}_a \quad \text{almost everywhere}, \quad \mathcal{A}(q) = \mathcal{S}(a) + \frac{1}{2}\langle q, a \rangle.$$

*For spherical spins, the precision is identified by*

$$\bar{\beta}(a) = 1 + a_* - \langle q, a \rangle = \frac{1}{1 - q(1-)}.$$

The two priors lead to different short proofs. For Ising spins, retaining  $\sqrt{D}$  when deleting a coordinate leaves a Gaussian linear term and a quadratic term independent of that coordinate. Constraint restoration concentrates this term, and coordinate exchangeability identifies the response. On the sphere, integration by parts in the tangent direction  $\sigma^2 - R_{12}\sigma^1$  cancels the second-derivative and diagonal terms exactly. The remaining identity involves only  $Q_{12}$ , so constraint reconstruction and the three-replica RPC law determine  $q$ . Section 5 gives both proofs.

### 2.3 The logarithmic increment from a fixed block

At  $t = 1$ , with the amplitudes fixed, write  $A_N(M) = \mathbb{E} \log Z_{N,M,1}$ . The next proposition computes the partition-function ratio needed for the lower bound.

**Proposition 2.4** (Inserting a fixed block of coordinates). *Fix  $K \geq 1$ . If  $M_N/N \rightarrow \alpha$  and  $q$  is the limiting spin path in dimension  $N$  under the preceding hypotheses, then*

$$\liminf_N \{A_{N+K}(M_N) - A_N(M_N)\} \geq K\mathcal{A}(q).$$

*No array limit at dimension  $N + K$  is required. For each compact  $I \subset (0, \infty)$ , there is  $C_{K,I} < \infty$  such that, uniformly in the amplitudes and sufficiently large  $N$ ,*

$$A_{N+K}(M) - A_N(M) \geq -C_{K,I} \quad (M/N \in I). \quad (49)$$

All  $K$  coordinates are inserted together. The expansion uses  $Q$  and the single coefficient  $M^{-1} \sum_k \{u_k''(h_k) - h_k u_k'(h_k)\}$ , reconstructed before insertion. Their qualitative limits suffice because their multipliers are bounded for a fixed coordinate cutoff. The overlap equation has already fixed the scalar precision. Consequently the spherical proof passes only expected logarithms to the limit; it does not need a limit theorem for the inserted coordinates or their second moments. The lower bound below combines this proposition with constraint insertion at the same dimension. The upper bound uses only the supporting equality in [theorem 2.3](#).

## 2.4 Proof of the theorem

Two elementary bounds complete the stated inputs. Independent row insertion and Jensen's inequality give, uniformly in the amplitudes,

$$|A_N(M+1) - A_N(M)| \leq C_{\text{row}}, \quad |\Phi(q)| \leq C_{\text{row}}, \quad C_{\text{row}} := \mathbb{E} e^{L|Z|}, \quad Z \sim N(0,1). \quad (50)$$

The same centered-channel bounds, Gaussian Poincaré, and resampling one mark give, for the original model and  $M/N$  in a fixed compact subset of  $(0, \infty)$ ,

$$\text{Var}\left(N^{-1} \log Z_{N,M}^{\kappa,\nu}\right) \leq C/N. \quad (51)$$

The estimates are proved in [Sections 3](#) and [6](#), respectively.

*Proof of [theorem 1.1](#).* First assume the centered smooth hypotheses [\(24\)](#), and abbreviate  $\Phi = \mathbb{E}_\nu \Phi_{u_\Theta}$ ,  $\mathcal{A} = \mathcal{A}_\kappa$ , and  $\mathcal{S} = \mathcal{S}_\kappa$ . Write

$$\mathcal{P}_\kappa(\alpha) := \inf_{q \in \mathcal{Q}} \{\mathcal{A}(q) + \alpha \Phi(q)\}, \quad F_{N,\kappa} := N^{-1} \log Z_{N,M}^{\kappa,\nu}.$$

*Lower bound.* For the Aizenman–Sims–Starr lower bound [\[2\]](#), fix  $d = 2$ , use the interpolation at  $t = 1$ , and set

$$m_N = 2 + \lfloor \alpha N \rfloor, \quad A_N(M) = \mathbb{E} \log Z_{N,M,1}, \quad \bar{A}_N = \mathbb{E}_{\mathbf{x}} A_N(m_N),$$

where  $A_N(M)$  keeps the amplitudes fixed. Fix  $K \geq 1$ . The number  $k_{N,K} = m_{N+K} - m_N$  satisfies

$$|k_{N,K} - \alpha K| \leq 1.$$

Insert these constraints first and then all  $K$  coordinates:

$$\begin{aligned} \Delta_{N,K}(\mathbf{x}) &:= A_{N+K}(m_{N+K}) - A_N(m_N) \\ &= A_N(m_{N+K}) - A_N(m_N) + A_{N+K}(m_{N+K}) - A_N(m_{N+K}). \end{aligned}$$

From any subsequence, select one on which  $k_{N,K} = k$  is constant and apply the common selection stated above to both constraint counts and all deletions required by the cavity propositions. At almost every amplitude vector, take a joint array-convergent refinement. [Proposition 2.1](#) identifies the same spin path  $q$  at the two counts and gives  $A_N(m_{N+K}) - A_N(m_N) \rightarrow k\Phi(q)$ . [Proposition 2.4](#), applied at the second count, gives a lower limit  $K\mathcal{A}(q)$  for the remaining term. Hence

$$\liminf_N \Delta_{N,K}(\mathbf{x}) \geq K\{\mathcal{A}(q) + \alpha\Phi(q)\} + (k - \alpha K)\Phi(q) \geq K\mathcal{P}_\kappa(\alpha) - C_{\text{row}}.$$

The bound is independent of  $q$ , so a refinement of any subsequence realizing a smaller lower limit would give a contradiction. The row bound [\(50\)](#) and the uniform block bound [\(49\)](#) provide a lower envelope for  $\Delta_{N,K}$  independent of  $N, \mathbf{x}$ . Fatou's lemma gives

$$\liminf_N (\bar{A}_{N+K} - \bar{A}_N) \geq K\mathcal{P}_\kappa(\alpha) - C_{\text{row}}.$$

Telescope at fixed  $K$  in each residue class modulo  $K$ , divide by the dimension, and let  $K \rightarrow \infty$ . This yields  $\liminf_N \bar{A}_N/N \geq \mathcal{P}_\kappa(\alpha)$ . Removing the perturbation costs  $o(1)$  in normalized pressure. Since  $|M_N - m_N| = o(N)$ , the row bound transfers the result to every  $M_N/N \rightarrow \alpha$ , proving  $\liminf_N \mathbb{E} F_{N,\kappa} \geq \mathcal{P}_\kappa(\alpha)$ .

*Upper bound.* Fix  $d \geq 2$ . The conjugate formula in [theorem 2.2](#) makes  $\mathcal{A}$  convex and lower semicontinuous. Compactness of  $\mathcal{Q}_d$ , continuity of  $\Phi$ , and the finite competitor  $\hat{q}_0 \in \mathcal{Q}_d$  give a minimizer  $q^0$  of  $\mathcal{A} + \alpha\Phi$  on  $\mathcal{Q}_d$ . Put

$$a^0 = \alpha p_{q^0}.$$

Minimality along  $(1-s)q^0 + sq$ , the channel derivative [\(31\)](#), and the convex secant inequality give

$$\mathcal{A}(q) - \frac{1}{2}\langle a^0, q \rangle \geq \mathcal{A}(q^0) - \frac{1}{2}\langle a^0, q^0 \rangle, \quad q \in \mathcal{Q}_d. \quad (52)$$

For infinite  $\mathcal{A}(q)$  this is automatic. The plateau properties put  $\bar{q} := \hat{q}_{a^0}$  in  $\mathcal{Q}_d$ . The conjugate inequality, [\(52\)](#), and scalar supporting equality imply

$$\mathcal{S}(a^0) \leq \mathcal{A}(q^0) - \frac{1}{2}\langle a^0, q^0 \rangle \leq \mathcal{A}(\bar{q}) - \frac{1}{2}\langle a^0, \bar{q} \rangle = \mathcal{S}(a^0).$$

Thus  $(q^0, a^0)$  is a supporting pair, including at boundary minimizers; no identity  $q^0 = \hat{q}_{a^0}$  is needed.

Use [\(43\)](#) with these trial paths and set  $\alpha_N = M_N/N$ ,  $\varphi_N^\mathbf{x}(t) = N^{-1}\mathbb{E} \log Z_{N,M_N,t}$ , and  $\bar{\varphi}_N(t) = \mathbb{E}_\mathbf{x} \varphi_N^\mathbf{x}(t)$ . Write  $Q_{t,\ell m}$  for [\(36\)](#) with inputs  $X_{k,t}$ , and let brackets below denote replicas of  $G_{N,M_N,t}$ . Define

$$\mathcal{R}_N^\mathbf{x}(t) := \mathbb{E} \langle (R_{12} - q^{0,\sharp}(r_{12}))(\alpha_N Q_{t,12} - a^{0,\sharp}(r_{12})) \rangle_{t,\mathbf{x}}.$$

Gaussian integration by parts, conditional on the channel marks, gives the two contributions

$$-\frac{\alpha_N}{2} \mathbb{E} \langle (R_{12} - q^{0,\sharp}(r_{12})) Q_{t,12} \rangle_{t,\mathbf{x}}, \quad \frac{1}{2} \mathbb{E} \langle a^{0,\sharp}(r_{12}) R_{12} \rangle_{t,\mathbf{x}}.$$

The row diagonals cancel because both fields have variance one; the source diagonal cancels the derivative of  $-N(1-t)a_*^0/2$ . The exact ancestry marginal [\(45\)](#) then gives

$$(\varphi_N^\mathbf{x})'(t) = \frac{1}{2}\langle a^0, q^0 \rangle - \frac{1}{2}\mathcal{R}_N^\mathbf{x}(t). \quad (53)$$

Apply the common selection on the amplitude–time product space. For almost every  $(\mathbf{x}, t)$ ,  $0 < t < 1$ , any further subsequence has a joint array-convergent refinement for the full and required deleted systems. If its spin path is  $q_t$ , the cavity propositions give

$$c_t = (1-t)q^0 + tq_t, \quad a_t = (1-t)a^0 + t\alpha p_{c_t}, \quad q_t = \hat{q}_{a_t}.$$

Joint reconstruction of  $(R, r, Q)$  passes the bounded remainder to  $\mathcal{R}_t = \langle q_t - q^0, \alpha p_{c_t} - a^0 \rangle$ ; the trial lookups are continuous on the finite ancestry set. Both  $(q_t, a_t)$  and  $(q^0, a^0)$  are supporting pairs, so

$$t\mathcal{R}_t = \langle q_t - q^0, a_t - a^0 \rangle \geq 0.$$

Every such refinement has a nonnegative limit. Thus the selected subsequence has  $\liminf_N \mathcal{R}_N^\mathbf{x}(t) \geq 0$  almost everywhere. The bound

$$|\mathcal{R}_N^\mathbf{x}(t)| \leq 2\{\alpha_N L^2 + \|a^0\|_\infty\}$$

permits Fatou's lemma; since the starting subsequence was arbitrary,  $\liminf_N \mathbb{E}_\mathbf{x} \int_0^1 \mathcal{R}_N^\mathbf{x}(t) dt \geq 0$ . It remains to evaluate the endpoints. Removing  $V_N$  costs  $o(1)$  in normalized expected pressure,

uniformly in the amplitudes. At  $t = 0$ , the independent row and spin recursions add. The row part contributes  $\alpha_N \Phi(q^0)$ , and the spin-block limit in [theorem 2.2](#) evaluates the remaining part. Thus

$$\bar{\varphi}_N(0) \longrightarrow \mathcal{S}(a^0) + \alpha \Phi(q^0).$$

At  $t = 1$ , the unperturbed model is the original perceptron, so

$$\bar{\varphi}_N(1) = \mathbb{E} F_{N,\kappa} + o(1).$$

Integrating [\(53\)](#) now yields

$$\limsup_N \mathbb{E} F_{N,\kappa} \leq \mathcal{S}(a^0) + \alpha \Phi(q^0) + \frac{1}{2} \langle a^0, q^0 \rangle = \min_{q \in \mathcal{Q}_d} \{ \mathcal{A}(q) + \alpha \Phi(q) \}.$$

For  $d = 2^n$ , approximate any  $q \in \mathcal{Q}$  from below by  $q_d(r) = q(j/d)$  on  $[j/d, (j+1)/d)$ , with  $q_d(1) = 1$ . Then  $q_d \rightarrow q$  in  $L^1$  and the conjugate formula gives  $\mathcal{A}(q_d) \leq \mathcal{A}(q)$ . Continuity of  $\Phi$  shows that the grid minima decrease to  $\mathcal{P}_\kappa(\alpha)$ . Together with the lower bound and the original centered-model variance estimate [\(51\)](#), this proves convergence in  $L^2$ .

*General Lipschitz channels and the convergence assertions.* For a centered uniformly Lipschitz family, convolve with one common mollifier and subtract the mollified value at zero. The resulting jointly Borel family  $u_{\theta,\delta}$  satisfies [\(24\)](#) and  $\varepsilon_\delta := \sup_\theta \|u_{\theta,\delta} - u_\theta\|_\infty = O(L\delta)$ . Comparison of exponential weights bounds the normalized pressure difference pointwise by  $(M_N/N)\varepsilon_\delta$ . The scalar recursions likewise bound the variational-value difference by  $\alpha\varepsilon_\delta$ . Send  $N \rightarrow \infty$ , then  $\delta \downarrow 0$ , to obtain the centered formula in  $L^2$ . For general channels set  $\tilde{u}_\theta = u_\theta - u_\theta(0)$ . Exactly,

$$\log Z_{N,M}^{\kappa,\nu}(u) = \sum_{k \leq M} u_{\Theta_k}(0) + \log Z_{N,M}^{\kappa,\nu}(\tilde{u}), \quad \Phi_{u_\theta} = u_\theta(0) + \Phi_{\tilde{u}_\theta}.$$

The  $L^1$  law of large numbers, or its  $L^2$  version under the stated second moment, proves the corresponding limits. Finally, if  $|u_\theta| \leq B$  almost surely, resampling one complete row changes  $\log Z$  by at most  $2B$ . Bounded differences give summable deviation probabilities for each fixed positive normalized deviation; Borel–Cantelli and convergence of the means yield almost-sure convergence. The deterministic bound  $|F_{N,\kappa}| \leq M_N B/N$  then gives convergence in every finite  $L^p$ .  $\square$

**Organization** Section [3](#) proves [theorem 2.1](#) and the channel derivative formula. Section [4](#) proves [theorem 2.2](#). Section [5](#) proves [theorems 2.3](#) and [2.4](#). Section [6](#) constructs the perturbation and establishes the ancestry identity, common subsequence selection, and concentration estimates. Section [7](#) proves the quenched large-deviation principle; Section [8](#) proves the Gaussian-field extension and the remaining application results.

### 3 Cavity in $M$

One- and two-row insertions determine the conditional first and second moments of  $Q_{12}$ . We compute these moments on finite cascades and transfer them to the model.

We use the centered smooth family [\(24\)](#) and the averaged responses [\(30\)](#) throughout Sections [3–6](#). Write  $\mathbf{N}_k = \gamma^{\otimes k}$ , where we recall  $\gamma$  denotes the standard Gaussian law. Deletion and restoration always concern a complete marked row; all estimates below are uniform in its mark.

### 3.1 Independent constraint insertion

Begin with a step path  $c$  on a finite grid  $\pi$ . Take independent completed fields  $X_b^c$  from (28), with Gaussian mark trees independent of the entire cascade Poisson tree. Set

$$G_{\pi,c,+k}(\mathrm{d}\eta, \mathrm{d}\mathbf{z}) = Z_{\pi,c,k}^{-1} e^{\sum_{b \leq k} u_b(X_b^c(\eta, z_b))} \mathfrak{R}_\pi(\mathrm{d}\eta) \mathbf{N}_k(\mathrm{d}\mathbf{z}).$$

Here  $Z_{\pi,c,k}$  normalizes the measure, and  $X_b^{c,\ell} = X_b^c(\eta^\ell, z_b^\ell)$ . For general paths, the brackets  $\mathbb{E}\langle \cdot \rangle_{c,+k}$  denote the limits of these averaged finite-replica laws, constructed below. The following conditional specialization of marked-RPC invariance [26, Theorem 4.4] tracks both the inserted inputs and their genealogy.

**Lemma 3.1** (Marked-RPC reweighting). *Consider a finite RPC with masses  $0 < m_1 < \dots < m_d < 1$ . Conditional on root data  $\mathcal{F}_0$ , independent of the Poisson tree, attach a standard Borel mark tree independent of that entire tree. Given ancestral marks, child subtrees are independent with a common law. Apply one common positive finite measurable function  $L$  to each marked branch and the root data, giving weights  $L_\eta$ . Define*

$$S_d = L, \quad S_{j-1} = (\mathbb{E}_j S_j^{m_j})^{1/m_j}, \quad j = d, \dots, 1,$$

where  $\mathbb{E}_j$  integrates the child mark given its ancestors. Assume these conditional moments are finite and positive. Then  $\mathcal{Z} = \sum_\eta v_\eta L_\eta$  is finite and positive. After a genealogy-preserving relabelling, the weights  $v_\eta L_\eta / \mathcal{Z}$  have the original RPC law and are independent of the mark tree with transitions

$$\nu_j^+(\mathrm{d}y \mid \text{ancestors}) = (S_j / S_{j-1})^{m_j} \nu_j(\mathrm{d}y \mid \text{ancestors}). \quad (54)$$

In particular, for every bounded finite-genealogy test  $F$ ,

$$\mathbb{E}[\langle F \rangle_+ \mid \mathcal{F}_0] = \mathbb{E}_{\text{RPC}}^{\mathbf{m}} F, \quad \mathbf{m} = (m_1, \dots, m_d),$$

and

$$\mathbb{E}[\log \mathcal{Z} - \log S_0 \mid \mathcal{F}_0] = 0.$$

Thus  $\mathbb{E} \log \mathcal{Z} = \mathbb{E} \log S_0$  if  $\mathbb{E} |\log S_0| < \infty$ . All mark and weight assertions are conditional on  $\mathcal{F}_0$ .

*Proof.* Conditional on the root data, apply [26, Theorem 4.4] with  $X_d = \log L$  and  $X_j = \log S_j$ . Its transitions (4.43) are (54); the conditional moment assumption makes the recursion finite. The theorem gives independence of the full reweighted cascade and tilted mark tree, including genealogy. For unnormalized weights, [26, Theorem 4.3 and (4.53)] also gives

$$T' := \sum_\eta w_\eta L_\eta / S_0 \stackrel{\text{d}}{=} T := \sum_\eta w_\eta.$$

By [26, Lemmas 2.2 and 2.4],  $T$  is a deterministic scale times a stable Poisson total with integrable logarithm. Hence  $\log \mathcal{Z} - \log S_0 = \log T' - \log T$  has conditional mean zero and an absolute first moment bounded independently of the root data. This proves the logarithmic assertions and positivity and finiteness of  $\mathcal{Z} = S_0 T' / T$ . Depth zero is immediate.  $\square$

For the finite-step calculation, define

$$T_{m,s} h(x) = m^{-1} \log \mathbb{E} e^{mh(x + \sqrt{s}Z)} \quad (m > 0), \quad T_{0,s} h(x) = \mathbb{E} h(x + \sqrt{s}Z), \quad Z \sim N(0, 1). \quad (55)$$

Write  $c = c_j$  on  $[m_j, m_{j+1})$ , where  $0 = m_0 < \dots < m_{J+1} = 1$ , and set  $c_{J+1} = 1$ . The Cole–Hopf transform on each interval where  $\gamma_c$  is constant gives the PDE value (2) as

$$X_{J+1} = u, \quad X_j = T_{m_{j+1}, c_{j+1} - c_j} X_{j+1} \quad (j = J, \dots, 0), \quad \Phi_u(c) = [T_{0, c_0} X_0](0).$$

Indeed,  $\gamma_c = m_{j+1}$  on  $(c_j, c_{j+1})$ , and it is zero below  $c_0$ . Thus the root is averaged without tilting and the terminal interval has mass one. For general paths, control stability and the quantile identity

$$\int_0^1 |\gamma_c(t) - \gamma_{\tilde{c}}(t)| \, dt = \|c - \tilde{c}\|_{L^1}$$

identify the PDE value with the continuous extension of this recursion. The control lemma applies on any finite horizon  $T > 0$ : under  $t = Ts$ , its diffusion coefficient is  $T$  and its nonlinear coefficient is  $T\gamma(Ts)$ . Additive constants give  $\Phi_u = u(0) + \Phi_{u-u(0)}$ , while centering leaves normalized laws unchanged.

**Corollary 3.2** (Independent constraint insertions). *For a step path  $c$  on  $\pi$ , every fixed  $k \geq 1$ , and every bounded Borel test  $F$  of finitely many replica ancestries,*

$$\mathbb{E}\langle F \rangle_{\pi, c, +k} = \mathbb{E}_\pi F, \quad \mathbb{E} \log Z_{\pi, c, k} = k\Phi(c). \quad (56)$$

If these replicas include 1, 2, then

$$\mathbb{E} \left\langle F \prod_{b=1}^k u'_b(X_b^{c,1}) u'_b(X_b^{c,2}) \right\rangle_{\pi, c, +k} = \mathbb{E}_\pi [F p_c(r_{12}^\pi)^k]. \quad (57)$$

For bounded Borel measurable  $\chi_b(\theta, x)$ , continuous in  $x$ ,

$$\mathbb{E} \left\langle \prod_{b=1}^k \chi_b(\Theta_b, X_b^{c,1}) \right\rangle_{\pi, c, +k} = \prod_{b=1}^k \mathbb{E} \langle \chi_b(\Theta, X^{c,1}) \rangle_{\pi, c}. \quad (58)$$

For arbitrary  $c \in \mathcal{Q}$ , these identities hold for the limiting replica laws, with  $r^\pi, \mathbb{E}_\pi$  replaced by  $U, \mathbb{E}_U$ . The limiting expected log normalizer is  $k\Phi(c)$ . The response  $p_c$  is nondecreasing, takes values in  $[0, L^2]$ , and is constant on plateaus of  $c$ . Both  $\Phi(c)$  and  $p_c$  are continuous for  $L^1$  convergence of  $c$ . The derivative formula (31) also holds.

*Finite-step proof.* Condition on the fresh marks. Represent each row by independent ancestral Gaussian increments of variances  $c_0, c_1 - c_0, \dots, c_J - c_{J-1}$ , integrate its terminal Gaussian, and apply theorem 3.1 conditional on the root Gaussians. The bound  $|u_b(x)| \leq L|x|$  makes the recursion finite and its root log value integrable. Independent rows have factorized exponential moments, so the recursion adds and the tilted transitions factor even for different terminals [26, Theorem 2.9 and (2.59)–(2.60)]. The expected log normalizer is therefore  $\sum_b \Phi_{u_b}(c)$ . Conditional on any sampled genealogy, the entire tilted mark trees are independent across rows; their one-replica marginals do not depend on the genealogy. Averaging the marks proves (56) and (58).

For one terminal  $u$ , start with  $g_0 \sim N(0, c_0)$  and tilt each transition  $N(g_j, c_{j+1} - c_j)$  by  $e^{m_{j+1}(X_{j+1}(g_{j+1}) - X_j(g_j))}$ , including the terminal step. Differentiating the recursion shows that  $M_j = X'_j(g_j)$  is a bounded martingale. Replicas branching after level  $j$  have conditionally independent continuations and terminal scores of conditional mean  $M_j$ . Hence

$$p_{c,u}(r) = \mathbb{E} M_j^2 \quad \text{on } [m_j, m_{j+1}).$$

Other branches integrate to one, so this identity holds conditional on the full genealogy. Independent marks give

$$\mathbb{E}_{\Theta_1, \dots, \Theta_k} \prod_{b=1}^k p_{c, u_{\Theta_b}}(r) = p_c(r)^k,$$

proving (57). Conditional Jensen gives monotonicity and  $0 \leq p_c(r) \leq p_{\text{diag}}(c) \leq L^2$ ; the inequality passes to general paths by their  $L^1$  limits. Zero-variance transitions give the plateau property.

Gaussian covariance differentiation, with completed diagonal variance fixed at one, leaves only the two-replica score term and gives (31). The recursion and its derivatives are measurable in the marks, and the uniform derivative bounds justify mark averaging. Integrating the derivative gives (32).  $\square$

To pass through the random normalizer, we need positive and inverse weight moments. Under any base  $\mu$  independent of fresh row noises, the completed inputs at one sampled state are independent standard normals. Thus, for  $W = e^{\sum_{b \leq k} u_b(X_b)}$  and every  $p > 0$ ,

$$\mathbb{E}\mu(W^p + W^{-p}) \leq 2(\mathbb{E}e^{pL|Z|})^k < \infty. \quad (59)$$

For one fresh row under an independent probability measure  $\Gamma$ , Jensen also gives, with  $A = \log \Gamma(e^{L|X|}) \geq 0$ ,

$$|\log \Gamma(e^{u_{\Theta}(X)})| \leq A, \quad \mathbb{E}A^p \leq \mathbb{E}e^{pL|Z|} \quad (p \geq 1). \quad (60)$$

These bounds also control mark fluctuations and row-count changes.

**Lemma 3.3** (Convergence after changing density). *Let  $\mu_n$  be random probability measures,  $W_n > 0$  measurable weights, and  $A_n(s, s')$  maps into a fixed Polish space; single-sample marks may be included. For conditional independent samples, put*

$$Z_n = \mu_n(W_n), \quad \mu_n^+(ds) = Z_n^{-1} W_n(s) \mu_n(ds).$$

*Suppose that every finite sampled array, including its diagonal, converges:*

$$\left( (A_n(s^\ell, s^m))_{\ell, m \leq k}, (W_n(s^\ell))_{\ell \leq k} \right) \Longrightarrow ((A_{\ell m})_{\ell, m \leq k}, (W^\ell)_{\ell \leq k}), \quad W^\ell > 0,$$

*and, for every  $p \geq 1$ ,*

$$\sup_n \mathbb{E}\mu_n(W_n^p + W_n^{-p}) < \infty.$$

*Then the sampled arrays under  $\mu_n^+$  converge to a unique consistent replica law determined by the array on the right. There is also a positive random variable  $Z$ , determined by that array, such that*

$$Z_n \Longrightarrow Z, \quad \mathbb{E} \log Z_n \longrightarrow \mathbb{E} \log Z.$$

*No random-measure representation of the limiting array is required. If one is given, the limits agree with its ordinary change of density.*

*Proof.* Fix a bounded continuous test  $F$  of a  $k$ -replica array, with  $|F| \leq 1$ , and take  $P$  conditional independent samples  $s^1, \dots, s^P$  from  $\mu_n$ . Set

$$\begin{aligned} H_n(s^1, \dots, s^k) &= F((A_n(s^\ell, s^m))_{\ell, m \leq k}) \prod_{\ell=1}^k W_n(s^\ell), \\ B_n &= \mu_n^{\otimes k}(H_n), \quad B_{n,P} = P^{-k} \sum_{i_1, \dots, i_k \leq P} H_n(s^{i_1}, \dots, s^{i_k}), \\ Z_{n,P} &= P^{-1} \sum_{i \leq P} W_n(s^i), \quad M_p = 1 + \sup_n \mathbb{E}\mu_n(W_n^p + W_n^{-p}). \end{aligned}$$

Single-sample marks are included in the argument of  $F$  when present. Repeated sample indices are allowed in  $B_{n,P}$ ; this is why the array-convergence hypothesis includes the diagonal.

Conditional independence gives  $\mathbb{E}|Z_{n,P} - Z_n|^2 \leq M_2/P$ . To estimate the numerator, expand  $\mathbb{E}|B_{n,P} - B_n|^2$  as a sum over two ordered  $k$ -tuples of sample indices. If all  $2k$  indices are different, the corresponding centered product has conditional expectation zero. There are at most  $\binom{2k}{2}P^{2k-1}$  remaining pairs of tuples. For a tuple with distinct indices of multiplicities  $m_1, \dots, m_j$ , where  $\sum_i m_i = k$ , Hölder with exponents  $k/m_i$  and  $k/(k - m_i)$  (or equality when  $m_i = k$ ) gives

$$\mathbb{E}[|H_n(s^{i_1}, \dots, s^{i_k})|^2 \mid \mu_n, W_n, A_n] \leq \prod_{i=1}^j \mu_n(W_n^{2m_i}) \leq \mu_n(W_n^{2k}).$$

Also  $|B_n|^2 \leq \mu_n(W_n^{2k})$ . Cauchy–Schwarz therefore bounds each remaining centered product in absolute expectation by  $4M_{2k}$ , and hence

$$\mathbb{E}|B_{n,P} - B_n|^2 \leq 4 \binom{2k}{2} M_{2k}/P.$$

Jensen bounds all positive and inverse moments of  $Z_n, Z_{n,P}$ . Since  $|B_n| \leq Z_n^k$  and  $|B_{n,P}| \leq Z_{n,P}^k$ , expanding the difference using the larger denominator gives

$$\left| \frac{B_{n,P}}{Z_{n,P}^k} - \frac{B_n}{Z_n^k} \right| \leq \frac{|B_{n,P} - B_n|}{\max(Z_n, Z_{n,P})^k} + k \frac{|Z_{n,P} - Z_n|}{\max(Z_n, Z_{n,P})}.$$

Cauchy–Schwarz, with exponents 2, 2, now yields

$$\sup_n \mathbb{E} \left| B_{n,P}/Z_{n,P}^k - B_n/Z_n^k \right| \leq C_k P^{-1/2},$$

uniformly over  $|F| \leq 1$ .

Place the consistent limiting input arrays on one probability space. Given their first  $P$  samples, choose indices  $I_1, \dots, I_k$  independently with probabilities  $W^i / \sum_{j \leq P} W^j$ , and let  $\Lambda_{P,k}$  be the averaged law of  $(A_{I_\ell I_m})_{\ell, m \leq k}$ , together with any sampled marks. For fixed  $P$ , its expectation of  $F$  is the limit of  $\mathbb{E}B_{n,P}/Z_{n,P}^k$ : the normalized empirical integral is a bounded continuous function of the finite array and its positive weights. The preceding estimate shows, for all bounded continuous  $|F| \leq 1$ ,

$$|\Lambda_{P,k}(F) - \Lambda_{Q,k}(F)| \leq C_k(P^{-1/2} + Q^{-1/2}).$$

For probability measures on a Polish space, the supremum over these tests equals the supremum over bounded measurable tests. Thus  $\Lambda_{P,k}$  is Cauchy in total variation and has a probability limit  $\Lambda_k$ . Consistency in  $k$  holds for each  $P$  and passes to the limit. Comparing the exact tilted law with its empirical law, then taking  $n \rightarrow \infty$  and  $P \rightarrow \infty$ , proves convergence to  $\Lambda_k$  and its uniqueness.

On the limiting input array put  $Z_P = P^{-1} \sum_{i \leq P} W^i$ . Finite-array convergence and the Portmanteau theorem transfer the bound

$$\mathbb{E}|Z_{n,P} - Z_{n,Q}|^2 \leq 2M_2(P^{-1} + Q^{-1})$$

to  $Z_P, Z_Q$ . Consequently  $Z_P \rightarrow Z$  in  $L^2$ . Jensen and Fatou give  $\mathbb{E}(Z^p + Z^{-p}) \leq M_p$ , so  $Z > 0$ . For any bounded Lipschitz function, comparison successively through  $Z_{n,P}$  and  $Z_P$ , first sending  $n \rightarrow \infty$  and then  $P \rightarrow \infty$ , proves  $Z_n \Rightarrow Z$ . Finally  $|\log z|^2 \leq C(z + z^{-1})$  makes  $\log Z_n$  uniformly integrable; continuous mapping gives convergence of its expectation. If a random-measure representation of the limiting array is supplied, the same empirical estimates identify these limits with its ordinary change of density and normalizer.  $\square$

*General paths in theorem 3.2. Covariance convergence.* Choose step paths  $c_n \rightarrow c$  in  $L^1$  on grids  $\pi_n$  of mesh  $h_n \rightarrow 0$ , adding zero-variance levels if necessary. Couple their finite genealogy arrays by  $r_{\ell m}^n = \kappa_{\pi_n}(U_{\ell m})$  off the diagonal, where  $\kappa_\pi$  rounds down to the left grid endpoint. This exact finite-RPC law follows from monotone rounding [26, (2.118)–(2.122)] and the GG characterization [26, Theorems 2.10 and 2.13]. Since  $U_{12}$  is uniform and  $c_n$  is constant on grid cells,

$$\mathbb{E}|c_n(r_{12}^n) - c(U_{12})| = \|c_n - c\|_1.$$

Every completed Gaussian covariance matrix therefore converges in probability, with diagonal fixed at one, regardless of  $c_n(1-)$ . When true diagonals differ from same-leaf pair marks, adjoin an independent uniform label to each sampled state and assign diagonal values only to identical complete states. Distinct replicas then have distinct complete states almost surely; all weights and Gaussian fields ignore the extra labels.

*Normalized limits.* Condition on the fresh marks and couple finite Gaussian vectors by their matrix square roots, using independent copies across rows. This gives joint convergence of genealogy, completed inputs, and weights. The bound (59) and theorem 3.3 define their unique normalized replica limits and expected log normalizers. They depend only on  $c$ , not its approximation. Finite-RPC invariance passes to continuous tests and then to Borel tests by equality of finite measures, so the limiting genealogy still has law  $U$ . Uniform moments permit mark averaging.

*Response continuity.* For continuous  $f$ , define

$$\nu_c(f) = \mathbb{E}\langle f(U_{12})u'_\Theta(X^{c,1})u'_\Theta(X^{c,2}) \rangle_c.$$

Helly selection and  $0 \leq p_{c_n} \leq L^2$  give an  $L^1$ -convergent further subsequence of every subsequence. The finite-step score identity identifies every such limit  $p$  by

$$\int_0^1 f(u)p(u) \, du = \nu_c(f),$$

since replacing  $f(r_{12}^n)$  by  $f(U_{12})$  costs at most  $L^2\omega_f(h_n)$ . Continuous tests determine finite measures, so  $p$  is unique almost everywhere. This defines  $p_c$  and proves full-sequence  $L^1$  convergence. Boundedness passes  $p_{c_n}^k$  to  $p_c^k$ , giving all three insertion identities. The logarithmic bound (32) passes to the limit as well.

For an arbitrary sequence  $c_n \rightarrow c$ , choose successively finer step approximations to each  $c_n$ . Their completed arrays converge to the same array with covariance  $c(U)$ ; the uniform empirical bounds in theorem 3.3 give  $\nu_{c_n}(f) \rightarrow \nu_c(f)$ . The same Helly argument proves  $p_{c_n} \rightarrow p_c$  in  $L^1$ . To preserve a plateau of  $c$ , use  $2^{-n}\lfloor 2^n c \rfloor$  off the diagonal and refine its grid. These step approximations show that  $p_c$  is constant there. One-replica convergence gives  $p_{\text{diag}}(c_n) \rightarrow p_{\text{diag}}(c)$ . Passing the finite-grid inequality  $p_{c_n} \leq p_{\text{diag}}(c_n)$  almost everywhere, then taking the essential supremum, gives

$$p_{\text{diag}}(c) \geq p_c(1-).$$

Convergence of  $p_{c_n}(1-)$  is neither needed nor asserted.  $\square$

### 3.2 Array limits and reconstruction

We transfer the insertion identities to the finite-volume measures of Section 2. Their limiting observed array is  $(q(U_{\ell m}), \rho_d(U_{\ell m}))$  off the diagonal, with diagonal  $(1, 1)$ .

*Proof of theorem 2.1. Insertion.* For every fixed  $k$ , apply theorem 3.3 with

$$\mu_N = G_N \otimes \mathbf{N}_k, \quad W_N = \exp\left\{\sum_{b=1}^k u_b(X_b)\right\}.$$

Conditional on the old environment, sampled old states, and fresh marks, the completed Gaussian input vectors are centered and independent across rows, with off-diagonal covariance  $\mathbf{C}_N$  and true diagonal  $1 + \varepsilon_N \rightarrow 1$ , including  $\varepsilon_N = 0$ . Thus (33) gives joint convergence of the old array, completed inputs, and weights. The bound (59), with its Gaussian variance replaced by the bounded  $1 + \varepsilon_N$ , supplies all positive and inverse weight moments, uniformly in the marks. The lemma gives convergence of the normalized laws and expected logarithms; these conclusions pass through the mark average. Explicitly,

$$\begin{aligned} \mathbb{E} \log Z_{N,k}^+ &\longrightarrow k\Phi(c), \\ \mathbb{E}\langle F\Psi \rangle_{+k} &\longrightarrow \mathbb{E}\langle F_\infty \Psi \rangle_{c,+k}, \end{aligned}$$

for bounded continuous old-array tests  $F$  and bounded  $\Psi$  measurable in the fresh marks and continuous in the inserted inputs. Here  $F_\infty$  evaluates  $F$  on the observed limiting array. Corollary 3.2 preserves that array's law and, with  $\Psi = \prod_b u'_b(X_b^1)u'_b(X_b^2)$ , gives (35).

Restoring deleted rows gives the full measure exactly and preserves the limiting observed array. Its spin marginal, and hence its nondecreasing quantile  $q$ , agree across the full and deleted systems. The common limiting rule  $c = K(q, \rho_d)$  identifies  $c$ ; a uniform error  $\|K_N - K\|_\infty \rightarrow 0$  does not change the covariance limit. The same argument applies to any fixed collection of row counts linked by finitely many insertions.

*Reconstruction.* Put  $S_k = u'_k(X_k^1)u'_k(X_k^2)$ , so  $|S_k| \leq L^2$  and  $Q_{12} = M^{-1} \sum_k S_k$ . For a bounded continuous old-array test  $F$ , exchangeability gives

$$\mathbb{E}\langle FQ_{12} \rangle = \mathbb{E}\langle FS_1 \rangle, \quad \mathbb{E}\langle FQ_{12}^2 \rangle = \frac{M-1}{M} \mathbb{E}\langle FS_1 S_2 \rangle + O(M^{-1}).$$

The repeated-index error is at most  $\|F\|_\infty L^4/M$ . One- and two-row restoration therefore give

$$\begin{aligned} \mathbb{E}\langle FQ_{12} \rangle &\longrightarrow \mathbb{E}_U[F_\infty p_c(U_{12})], \\ \mathbb{E}\langle FQ_{12}^2 \rangle &\longrightarrow \mathbb{E}_U[F_\infty p_c(U_{12})^2]. \end{aligned}$$

The common covariance rule  $c = K(q, \rho_d)$  and the plateau property give  $p_c(u) = \bar{p}_c(q(u), \rho_d(u))$  almost everywhere, with  $\bar{p}_c$  bounded Borel; see theorem 6.3. For any joint subsequential limit  $(A, Q)$ , where  $A$  retains the finite observed overlap-ancestry array containing  $(R_{12}, r_{12})$ , the first- and second-moment identities against continuous tests identify finite signed Borel measures. Hence

$$\mathbb{E}[Q \mid A] = \bar{p}_c(R_{12}, r_{12}), \quad \mathbb{E}[Q^2 \mid A] = \bar{p}_c(R_{12}, r_{12})^2.$$

The conditional variance vanishes, so  $Q = \bar{p}_c(R_{12}, r_{12})$  almost surely; in the canonical representation this equals  $p_c(U_{12})$ . Boundedness gives compactness, and every joint limit has this form.

For  $T_N = M^{-1} \sum_k \chi(\Theta_k, X_k^1)$ , restore one row to compute  $\mathbb{E}\langle T_N \rangle$  and two rows to compute  $\mathbb{E}\langle T_N^2 \rangle$ . The product identity (58) gives the limits  $m_c(\chi)$  and  $m_c(\chi)^2$ ; repeated indices contribute  $O(M^{-1})$ . This proves (38). The scalar limits exist by the same covariance and weight-moment passage, so no separate approximation argument is needed.

For the family of Section 2.2, deletion leaves the row noises independent, and their covariance is (46). The lookup  $q^{0,\sharp}$  is a function on the fixed finite ancestry alphabet, so convergence of the observed  $(R, r)$ -array passes its covariance to the limit. Since  $q^{0,\sharp}(\rho_d(u)) = q^0(u)$ , this verifies (33) with  $c = (1-t)q^0 + tq$ .  $\square$

The following comparison will control changes in the insertion field during coordinate lifting.

**Lemma 3.4** (Comparison of tilted averages). *Let  $\mu$  be a random probability measure on a measurable space. For  $Z_H = \mu(e^H) \in (0, \infty)$ , write  $\langle F \rangle_H = (e^H \mu / Z_H)^{\otimes n}(F)$ . Let  $F$  be a bounded  $n$ -replica test, with deterministic bound  $\|F\|_\infty$ .*

(i) General comparison. *For real measurable fields  $X, Y$ , put*

$$d^2 = \mathbb{E}\mu((Y - X)^2), \quad B = \max_{\substack{H \in \{X, Y\} \\ \epsilon = \pm 1}} \mathbb{E}\mu(e^{4\epsilon H}).$$

*If  $d, B < \infty$ , then*

$$|\mathbb{E} \log Z_Y - \mathbb{E} \log Z_X| \leq \sqrt{B} d, \\ |\mathbb{E} \langle F \rangle_Y - \mathbb{E} \langle F \rangle_X| \leq 2n \|F\|_\infty \sqrt{B} d.$$

*No independence among  $\mu, X, Y, F$  is required. If  $|Y - X| \leq \varepsilon$ , both bounds hold before averaging with  $\sqrt{B} d$  replaced by  $\varepsilon$ , requiring only finite positive normalizers; the averaged log bounds then hold whenever the logarithms are integrable.*

(ii) Gaussian comparison. *Condition on an environment determining  $\mu$  and  $F$ . Let  $X, Y$  be conditionally centered Gaussian fields with covariance kernels  $C_X, C_Y$ , whose diagonal values are bounded by a finite deterministic constant. If  $\|C_Y - C_X\|_\infty \leq \delta$ , then*

$$|\mathbb{E} \log Z_Y - \mathbb{E} \log Z_X| \leq \delta, \\ |\mathbb{E} \langle F \rangle_Y - \mathbb{E} \langle F \rangle_X| \leq n(2n + 1) \|F\|_\infty \delta.$$

*Proof.* For (i), set  $D = Y - X$ ,  $H_t = X + tD$ , and  $Z_t = \mu(e^{H_t})$ . Holding  $F$  fixed, differentiation gives

$$\frac{d}{dt} \log Z_t = \langle D \rangle_t, \quad \frac{d}{dt} \langle F \rangle_t = \left\langle F \left( \sum_{i=1}^n D(s^i) - nD(s^{n+1}) \right) \right\rangle_t.$$

Convexity bounds  $\mathbb{E}\mu(e^{\pm 4H_t})$  by  $B$ . Jensen gives  $Z_t^{-4} \leq \mu(e^{-4H_t})$ , so Hölder yields

$$\mathbb{E} \langle |D| \rangle_t \leq d (\mathbb{E}\mu(e^{4H_t}))^{1/4} (\mathbb{E}\mu(e^{-4H_t}))^{1/4} \leq \sqrt{B} d.$$

Integration proves (i); truncation and this bound justify differentiation. For  $|D| \leq \varepsilon$ , use  $\langle |D| \rangle_t \leq \varepsilon$  directly, before averaging.

For (ii), take conditionally independent endpoint copies and set  $V_t = \sqrt{1-t}X + \sqrt{t}Y$ ,  $Z_t = \mu(e^{V_t})$ , and  $\Delta_{ij} = (C_Y - C_X)(s^i, s^j)$ . Gaussian integration by parts gives

$$\frac{d}{dt} \mathbb{E} \log Z_t = \frac{1}{2} \mathbb{E} \langle \Delta_{11} - \Delta_{12} \rangle_t \tag{61}$$

and

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \langle F \rangle_t &= \frac{1}{2} \mathbb{E} \left\langle F \left[ \sum_{i,j \leq n} \Delta_{ij} - 2n \sum_{i \leq n} \Delta_{i,n+1} \right. \right. \\ &\quad \left. \left. - n\Delta_{n+1,n+1} + n(n+1)\Delta_{n+1,n+2} \right] \right\rangle_t. \end{aligned}$$

These identities hold first for finite state spaces. If  $B_0$  bounds the field variances, then  $\mathbb{E}\mu(e^{pV_t} + e^{-pV_t}) \leq 2e^{p^2 B_0/2}$ . The empirical measure argument of [theorem 3.3](#), using conditional samples independent of the Gaussian noises, passes the integrated identities to general  $\mu$ . Bounding each covariance difference by  $\delta$  and summing the absolute coefficients proves (ii).  $\square$

### 3.3 The quadratic coefficient for block insertion

The block calculation at  $t = 1$  uses the coefficient

$$B_N = \frac{1}{M} \sum_{k=1}^M \{u_k''(h_k) - h_k u_k'(h_k)\}.$$

The diagonal correlation  $Q_{11}$  is already covered by (38). We recover  $B_N$  by Gaussian integration by parts and the limiting GG identity. The spectral bound below controls its unbounded factor  $h_k$ .

**Lemma 3.5** (Uniform constraint bounds). *Let  $G$  be the  $M \times N$  Gaussian constraint matrix and put  $\Lambda_N = \|G\|_{\text{op}}^2/M$ . If  $M/N$  stays in a compact subset of  $(0, \infty)$ , then, for every  $r, c \geq 0$ ,*

$$\sup_N \mathbb{E} \Lambda_N^r < \infty, \quad \sup_N \mathbb{E} e^{c\sqrt{\Lambda_N}} < \infty.$$

On either spin space,

$$\frac{1}{M} \sum_k h_k(\sigma)^2 \leq \Lambda_N, \quad \frac{1}{M} \sum_k |h_k(\sigma)| \leq \sqrt{\Lambda_N}. \quad (62)$$

*Proof.* Use  $\|\sigma\|^2 = N$ , Cauchy–Schwarz, and the Gaussian matrix tail bound  $\mathbb{P}\{\|G\|_{\text{op}} \geq \sqrt{M} + \sqrt{N} + s\} \leq e^{-s^2/2}$ .  $\square$

In particular, the common bounds on  $u_k', u_k''$  give the uniform envelope

$$|B_N| \leq C_u(1 + \sqrt{\Lambda_N}). \quad (63)$$

**Lemma 3.6** (The quadratic coefficient). *At  $t = 1$ , assume the full measure and its one- and two-constraint deletions satisfy the array-limit hypotheses of theorem 2.1, with common spin path  $q$ . Put*

$$\overline{B} = \langle q, p_q \rangle - p_{\text{diag}}(q).$$

Then

$$\begin{aligned} \mathbb{E} \langle |Q_{11} - p_{\text{diag}}(q)|^2 \rangle &\longrightarrow 0, \\ \mathbb{E} \langle |B_N - \overline{B}|^2 \rangle &\longrightarrow 0. \end{aligned}$$

*Proof.* The first limit is (38) with  $\chi(\theta, x) = u_\theta'(x)^2$ .

To recover  $B_N$ , first add back its diagonal derivative term. Put

$$C_N = B_N + Q_{11}, \quad K_{\ell m} = R_{\ell m} Q_{\ell m}, \quad \overline{K}_N = \mathbb{E} \langle K_{12} \rangle.$$

For an  $n$ -replica test  $F$ , let

$$D_n(F) = \frac{1}{M\sqrt{N}} \sum_{k,i} \mathbb{E} \langle \sigma_i^1 u_k'(h_k^1) \partial_{g_{ki}} F \rangle,$$

where the derivative holds the replica states and marks fixed. Gaussian integration by parts conditional on the marks, using independence of the perturbation from the constraint vectors, gives

$$\mathbb{E} \langle F C_N^1 \rangle = \mathbb{E} \left\langle F \left( n K_{1,n+1} - \sum_{\ell=2}^n K_{1\ell} \right) \right\rangle - D_n(F). \quad (64)$$

The derivatives of  $u'_k(h_k^1)$  and of the first replica's density produce  $u''_k(h_k^1)$  and  $u'_k(h_k^1)^2$ ; the other density derivatives produce the displayed overlaps. With  $n = 1, F = 1$ , this yields  $\mathbb{E}\langle C_N \rangle = \overline{K}_N$ .

The uniform derivative bounds and  $\|\sigma^\ell\| = \sqrt{N}$  give

$$\|\nabla_{g_k} C_N^1\| \leq \frac{C_u}{M}(1 + |h_k^1|), \quad \|\nabla_{g_k} K_{12}\| \leq \frac{C_u}{M}.$$

Contracting with  $\sigma^1$  and using (62) yields

$$|D_1(C_N^1)| + |D_2(K_{12})| \leq \frac{C_u}{M}(1 + \mathbb{E}\sqrt{\Lambda_N}) = O(M^{-1}).$$

The same bound and (63) justify these integrations by parts by truncation. Applying (64) to these two tests gives

$$\mathbb{E}\langle (C_N - \overline{K}_N)^2 \rangle = \{2\mathbb{E}\langle K_{12}K_{13} \rangle - \mathbb{E}\langle K_{12}^2 \rangle - \overline{K}_N^2\} - D_2(K_{12}) - D_1(C_N^1).$$

Row reconstruction identifies the limiting  $K_{12} = q(U_{12})p_q(U_{12})$ . The GG identity for  $U$ , applied to the bounded Borel function  $u \mapsto q(u)p_q(u)$ , makes the expression in braces vanish. This is the identity of [26, Theorem 2.10], extended by [26, Theorem 2.17] and equality of the resulting finite dimensional laws to bounded Borel tests. Thus  $C_N \rightarrow \langle q, p_q \rangle$  in averaged  $L^2$ . Subtract the established limit of  $Q_{11}$  to obtain the result.  $\square$

## 4 Scalar transforms and supporting inequalities

The Ising response is a two-spin correlation. The spherical response is the explicit function obtained from a Gaussian quadratic integral. We prove that both responses give supporting pairs for the spin cost. The same Gaussian integral supplies the spherical block increment. Write  $\mathcal{S}_I, \mathcal{S}_S, \mathcal{A}_I, \mathcal{A}_S$ , and  $\hat{q}_a^I, \hat{q}_a^S$  for the two priors. For a partition  $\pi = (x_0, \dots, x_n)$ , write a step source as

$$0 = x_0 < x_1 < \dots < x_n = 1, \quad a(r) = a_j \quad (x_{j-1} \leq r < x_j), \quad 0 \leq a_1 \leq \dots \leq a_n, \quad (65)$$

and put  $a_* = a_n$ . The cascade  $\mathfrak{R}_\pi$  has masses  $x_1, \dots, x_{n-1}$ . Its off-diagonal ancestry is the left endpoint of the common ancestry interval, including  $x_{n-1}$  for two replicas on the same leaf; the true replica diagonal is one. Write  $\mathbb{E}_\pi$  for its averaged genealogy law. Gaussian ancestor marks are independent of the entire Poisson tree; root increments form the common environment in theorem 3.1. Limiting Ising replica laws use the uniform genealogy  $U$ , with expectation  $\mathbb{E}_U$ , and are defined by finite partitions of vanishing mesh. Zero-variance levels may always be added.

### 4.1 The spherical Gaussian integral

Use the domain  $\mathfrak{D}_a$ , precision  $\pi_{a,\beta}$ , and functions  $V_a, q_{a,\beta}^{\text{co}}$  defined in Section 2. Put

$$K_a(\beta) = -\frac{1}{2} \log \beta + \frac{1}{2} \int_0^{a_*} \frac{dt}{\pi_{a,\beta}(t)}. \quad (66)$$

For a step source, let  $w^a$  have covariance path  $a$  and variance  $a_*$ , and set

$$K_{a,R}(\beta) = \mathbb{E} \log \int \mathfrak{R}_\pi(d\eta) \int_{-R}^R \frac{d\tau}{\sqrt{2\pi}} e^{\tau w^a(\eta) - \beta \tau^2/2}.$$

The next lemma also defines this quantity for general sources.

**Lemma 4.1** (Gaussian recursion, precision, and logarithmic cutoff). *For a step source and  $\beta \in \mathfrak{D}_a$ ,*

$$\mathbb{E} \log \int \mathfrak{R}_\pi(d\eta) \int_{\mathbb{R}} \frac{d\tau}{\sqrt{2\pi}} e^{\tau w^a(\eta) - \beta \tau^2/2} = K_a(\beta).$$

*The function  $V_a$  is continuous and strictly decreasing from infinity to zero on  $\mathfrak{D}_a$ . Its unique solution  $V_a(\beta(a)) = 1$  satisfies*

$$1 + \beta_-(a) \leq \bar{\beta}(a) \leq 1 + a_*, \quad \inf_t \pi_{a, \bar{\beta}(a)}(t) \geq 1.$$

*If step paths  $a_n$  on refining partitions satisfy  $a_n(1-) = a_*$ ,  $\|a_n - a\|_{L^1} \rightarrow 0$ ,  $\beta_n \rightarrow \beta$ , and  $\inf_{n,t} \pi_{a_n, \beta_n}(t) > 0$ , then*

$$K_{a_n}(\beta_n) \rightarrow K_a(\beta), \quad \|q_{a_n, \beta_n}^{\text{co}} - q_{a, \beta}^{\text{co}}\|_{L^1} \rightarrow 0, \quad V_{a_n}(\beta_n) \rightarrow V_a(\beta).$$

*For each fixed  $R > 0$ , the quantities  $K_{a_n, R}(\beta_n)$  also converge to a limit  $K_{a, R}(\beta)$  independent of the approximation, and*

$$K_{a, R}(\beta) \uparrow K_a(\beta) \quad (R \rightarrow \infty). \quad (67)$$

*Proof. Gaussian recursion.* For a step source, set

$$d_n = \beta, \quad d_j = d_{j+1} - x_j(a_{j+1} - a_j), \quad 1 \leq j < n.$$

These numbers are positive precisely when  $\beta > \beta_-(a)$ . The leaf integral is  $\beta^{-1/2} e^{w^2/(2\beta)}$ . Square completion at level  $j$ , with  $\Delta_j = a_{j+1} - a_j$ , gives

$$\frac{1}{x_j} \log \mathbb{E}_Z e^{x_j[C + (w + \sqrt{\Delta_j} Z)^2/(2d_{j+1})]} = C + \frac{1}{2x_j} \log \frac{d_{j+1}}{d_j} + \frac{w^2}{2d_j}.$$

This is the Gaussian recursion of [36, Lemma 3.5]; positivity of  $d_j$  checks exactly the fractional moments in [theorem 3.1](#). Averaging the unchanged root Gaussian of variance  $a_1$  gives

$$K_a(\beta) = -\frac{1}{2} \log \beta + \frac{a_1}{2d_1} + \frac{1}{2} \sum_{j=1}^{n-1} \frac{1}{x_j} \log \frac{d_{j+1}}{d_j}. \quad (68)$$

We need only the one-coordinate marginal of the normalized scalar integral, to control the logarithmic cutoff below. Under the tilted transition (54), the next field, given its current value  $w$ , is Gaussian with mean  $d_{j+1}w/d_j$  and variance  $\Delta_j d_{j+1}/d_j$ . At the leaf, the coordinate has conditional law  $\mathcal{N}(w/\beta, \beta^{-1})$ . Consequently its averaged law is

$$\tau \stackrel{d}{=} \frac{Z_n}{\sqrt{\beta}} + \frac{\sqrt{a_1}}{d_1} Z_0 + \sum_{j=1}^{n-1} \sqrt{\frac{a_{j+1} - a_j}{d_j d_{j+1}}} Z_j, \quad (69)$$

with independent standard Gaussians. Its variance is

$$V_a(\beta) = \frac{1}{\beta} + \frac{a_1}{d_1^2} + \sum_{j=1}^{n-1} \frac{a_{j+1} - a_j}{d_j d_{j+1}}.$$

On  $(0, a_1]$ ,  $\pi_{a, \beta} = d_1$ ; on  $(a_j, a_{j+1}]$ ,  $\pi_{a, \beta}(t) = d_j + x_j(t - a_j)$ . Integrating  $\pi^{-1}$  and  $\pi^{-2}$  gives (66) and (47). In particular,

$$q_{a, \beta}^{\text{co}}(r) = \frac{a_1}{d_1^2} + \sum_{i=1}^{j-1} \frac{a_{i+1} - a_i}{d_i d_{i+1}}, \quad x_{j-1} \leq r < x_j. \quad (70)$$

*Precision and continuity.* For general sources the explicit integrals give

$$V'_a(\beta) = -\frac{1}{\beta^2} - 2 \int_0^{a_*} \frac{dt}{\pi_{a,\beta}(t)^3} < 0.$$

With  $\delta = \beta - \beta_-(a)$ , layer cake gives  $\pi_{a,\beta}(t) = \delta + \int_0^t x_a(s) ds \leq \delta + t$  and  $\beta \leq \delta + a_*$ . Thus  $V_a(\beta) \geq \delta^{-1}$ ; also  $V_a(\beta) \rightarrow 0$  as  $\beta \rightarrow \infty$ . At  $\beta = 1 + a_*$ ,  $\pi_{a,\beta}(t) \geq 1 + t$ , so  $V_a(1 + a_*) \leq 1$ . These facts prove the existence, uniqueness, and bounds for  $\bar{\beta}(a)$ . Endpoint-preserving step approximations exist by rounding source values and setting a shrinking terminal interval equal to  $a_*$ . The quantile identities

$$\int_0^{a_*} |x_{a_n} - x_a| = \|a_n - a\|_{L^1}, \quad \int_0^{a_*} x_a(t) dt = a_* - \int_0^1 a(r) dr$$

imply  $\|\pi_{a_n,\beta_n} - \pi_{a,\beta}\|_\infty \leq |\beta_n - \beta| + \|a_n - a\|_{L^1}$ . The positive lower bound on precision permits dominated convergence in  $K_a, V_a$ . If that lower bound is  $\underline{\pi}$ , then

$$\|q_{a_n,\beta_n}^{\text{co}} - q_{a,\beta}^{\text{co}}\|_{L^1} \leq \underline{\pi}^{-2} \|a_n - a\|_{L^1} + a_* \|\pi_{a_n,\beta_n}^{-2} - \pi_{a,\beta}^{-2}\|_\infty \rightarrow 0.$$

*Logarithmic cutoff.* At fixed  $R$ , use the probability base  $\mathfrak{R}_{\pi_n} \otimes \gamma(\cdot \mid [-R, R])$  and weight  $e^{\tau w^{a_n} - (\beta_n - 1)\tau^2/2}$ . Its positive and inverse moments of order  $p$  are bounded by  $\exp\{pR^2 \sup_n |\beta_n - 1|/2 + p^2 R^2 a_*/2\}$ . The sampled covariance arrays converge, including their common diagonal  $a_*$ . Lemma 3.3 therefore passes the expected logarithmic normalizer to the limit, including the factor  $\gamma([-R, R])$ , and defines  $K_{a,R}(\beta)$ .

For a finite step approximation, let  $L_R(w)$  denote the scalar coordinate integral and put

$$P_{\beta,R}(w) = \frac{L_R(w)}{L_\infty(w)} = \mathbb{P}\{|w/\beta + Z/\sqrt{\beta}| \leq R\}.$$

Jensen's inequality under the normalized full scalar integral gives

$$0 \leq K_{a_n}(\beta_n) - K_{a_n,R}(\beta_n) \leq \mathbb{E}\langle -\log P_{\beta_n,R}(w^{a_n}) \rangle.$$

By (69), the averaged coordinate is  $\mathcal{N}(0, V_{a_n}(\beta_n))$ . Its conditional distribution given the field is  $\mathcal{N}(w/\beta_n, \beta_n^{-1})$ , so characteristic functions show that the averaged field under this normalized integral is  $\mathcal{N}(0, \beta_n^2 V_{a_n}(\beta_n) - \beta_n)$ . These variances are uniformly bounded. For  $R \geq 1$ , integration over  $[-1, 1]$  gives  $0 \leq -\log P_{\beta_n,R}(w) \leq C(1 + w^2)$ , uniformly in  $n$ . The left side tends to zero uniformly on bounded field intervals; Gaussian tails therefore make the logarithmic loss tend to zero uniformly in  $n$ . Passing to the limit at fixed  $R$ , and then letting  $R \rightarrow \infty$ , proves (67).  $\square$

## 4.2 The scalar spin transforms

**Block transforms and terminal completion.** For a finite-step source as in (65), let  $(v_\eta)$  be the weights of  $\mathfrak{R}_\pi$ , and let  $W^a = (W_1^a, \dots, W_K^a)$  consist of independent copies of the Gaussian field  $w^a$ , with variance  $a_*$ . Use normalized counting measure  $\mu_K^I$  on  $\Sigma_K^I = \{-1, 1\}^K$  and normalized surface measure  $\mu_K^S$  on  $\Sigma_K^S = S^{K-1}(\sqrt{K})$ . Set

$$\mathcal{S}_{\kappa,K}(a) = \frac{1}{K} \mathbb{E} \log \sum_\eta v_\eta \int_{\Sigma_K^\kappa} e^{W^a(\eta) \cdot \tau - K a_*/2} \mu_K^\kappa(d\tau). \quad (71)$$

Adding zero-variance levels leaves this finite recursion unchanged. If the Gaussian diagonal is  $d \geq a_*$ , complete the field by  $\sqrt{d - a_*} z$ , integrate  $z$  inside the spin integral, and replace the compensation by  $-Kd/2$ . The resulting transform satisfies

$$\mathcal{S}_{\kappa,K}(a; d) = \mathcal{S}_{\kappa,K}(a),$$

because  $\|\tau\|^2 = K$  and

$$\int e^{\sqrt{c}z \cdot \tau - c\|\tau\|^2/2} \mathbf{N}_K(dz) = 1 \quad (c \geq 0).$$

These transforms satisfy the uniform continuity bound

$$|\mathcal{S}_{\kappa,K}(a) - \mathcal{S}_{\kappa,K}(\tilde{a})| \leq \frac{1}{2}\|a - \tilde{a}\|_{L^1}. \quad (72)$$

Indeed, for  $a_t = (1-t)a + t\tilde{a}$  on a common grid, the leaf factor lies between  $e^{-Ka_{t,*}/2 - \sqrt{K}\|W^{a_t}\|}$  and  $e^{-Ka_{t,*}/2 + \sqrt{K}\|W^{a_t}\|}$ . Its recursive conditional moments are finite and positive, so [theorem 3.1](#) preserves ancestry. Apply (61) to  $W^{a_t}(\eta) \cdot \tau$ . Its diagonal derivative cancels the compensation  $-Ka_{t,*}/2$ , giving

$$\left| \frac{d}{dt} \mathcal{S}_{\kappa,K}(a_t) \right| \leq \frac{1}{2} \mathbb{E}_t \left\langle \frac{|\tau^1 \cdot \tau^2|}{K} |\tilde{a}(r_{12}) - a(r_{12})| \right\rangle \leq \frac{1}{2} \|\tilde{a} - a\|_{L^1}.$$

Integration proves the bound; its unique  $L^1$ -continuous extension defines  $\mathcal{S}_{\kappa,K}$  on  $\mathcal{P}_b$ .

**The spherical transform.** The variance constraint enters the block transform through the multiplier functional

$$F_a(\beta) = K_a(\beta) + \frac{1}{2}(\beta - 1 - a_*), \quad \beta \in \mathfrak{D}_a.$$

The explicit formula for  $K_a$  gives

$$F'_a(\beta) = \frac{1}{2}\{1 - V_a(\beta)\}.$$

Thus  $F_a$  is strictly convex, with unique minimizer  $\bar{\beta}(a)$ . The spherical block transform is

$$\mathcal{S}_S(a) = \min_{\beta \in \mathfrak{D}_a} F_a(\beta) = F_a(\bar{\beta}(a)), \quad (73)$$

and

$$\lim_{K \rightarrow \infty} \mathcal{S}_{S,K}(a) = \mathcal{S}_S(a). \quad (74)$$

For a finite-step source with  $a_* > 0$ , apply [36, Proposition 3.1 and (3.34)] with

$$\xi(s) = a_* s^2/2, \quad h = 0, \quad b = \beta, \quad x(s) = \text{Leb}\{r \in [0, 1) : a(r) \leq a_* s\}.$$

The overlap levels are  $a_j/a_*$ , with masses  $x_j$ ; coinciding terminal levels are allowed in [36, (3.2)]. Both surface measures are normalized, so his spin pressure is  $\varphi_K(0) = \mathcal{S}_{S,K}(a) + a_*/2$ . His  $d(s) = a_* \int_s^1 x(u) du$  satisfies  $d(0) = \beta_-(a)$  and  $\beta - d(s) = \pi_{a,\beta}(a_* s)$ . Changing variables in (3.34) gives

$$W(x, \beta) - a_*/2 = K_a(\beta) + (\beta - 1 - a_*)/2 = F_a(\beta).$$

Proposition 3.1 therefore proves (74). Its restriction  $b > 1$  loses no minimizer, since  $V_a(\bar{\beta}(a)) = 1$  implies  $\bar{\beta}(a) > 1$  for  $a_* > 0$ . For  $a = 0$ , both sides are zero and  $\bar{\beta}(0) = 1$ .

For general  $a$ , choose endpoint-preserving step approximations  $a_m \rightarrow a$ . Their minimizing precisions lie in  $[1, 1 + a_*]$  and satisfy  $\pi \geq 1$ . The stability in [theorem 4.1](#) shows that every subsequential limit solves  $V_a(\beta) = 1$ , hence  $\bar{\beta}(a_m) \rightarrow \bar{\beta}(a)$  and  $F_{a_m}(\bar{\beta}(a_m)) \rightarrow F_a(\bar{\beta}(a))$ . The uniform bound (72) then passes both (73) and (74) to  $a$ .

For  $q_* = q(1-) < 1$ ,  $\gamma_q = 1$  above  $q_*$ , so (3) becomes

$$\mathcal{A}_S(q) = \frac{1}{2} \int_0^{q_*} \frac{dt}{\int_t^1 \gamma_q(s) ds} + \frac{1}{2} \log(1 - q_*). \quad (75)$$

For its finite-grid expression, put  $\mathcal{Q}_{<1} = \{q \in \mathcal{Q} : q(1-) < 1\}$  and  $\lambda_{\gamma_q}(t) = \int_t^1 \gamma_q(s) ds$ . If  $q = q_j$  on the grid intervals and  $q_n < 1$ , its susceptibilities are

$$\lambda_n = 1 - q_n, \quad \lambda_j = \lambda_{\gamma_q}(q_j) = \lambda_{j+1} + x_j(q_{j+1} - q_j), \quad j < n.$$

The Crisanti–Sommers entropy [36, (4.4)–(4.6)] has the following expression in these quantile coordinates:

$$2\mathcal{A}_S(q) = \frac{q_1}{\lambda_1} + \sum_{j=1}^{n-1} \frac{1}{x_j} \log \frac{\lambda_j}{\lambda_{j+1}} + \log \lambda_n. \quad (76)$$

The quantile identity and dominated convergence also give

$$\mathcal{A}_S(q_m) \rightarrow \mathcal{A}_S(q) \quad (77)$$

whenever  $q_m \rightarrow q$  in  $L^1$  with a common endpoint bound below one.

**Ising spins as constraint insertions.** For a step source  $a$  on  $\pi$ , choose  $A > \max\{1, \|a\|_\infty\}$ , and put

$$u_A(x) = \log \cosh(\sqrt{A}x) - A/2, \quad c_a(r) = a(r)/A \quad (r < 1), \quad c_a(1) = 1.$$

For  $X = (w^a + \sqrt{A - a_*}z)/\sqrt{A}$ , terminal Gaussian integration gives

$$\begin{aligned} \int e^{u_A(X)} \gamma(dz) &= e^{-a_*/2} \cosh w^a, \\ \frac{\int u'_A(X) e^{u_A(X)} \gamma(dz)}{\int e^{u_A(X)} \gamma(dz)} &= \sqrt{A} \tanh w^a. \end{aligned}$$

The first identity identifies  $K$  spin coordinates with  $K$  independent constraint insertions; the second identifies their response. The recursion for this spin insertion is the step-coefficient solution of (1) with terminal  $\log \cosh$  and horizon  $a_*$ , followed by subtraction of  $a_*/2$ . Thus it is exactly (4). For general sources, use endpoint-preserving step approximations  $a_n$  on partitions  $\pi_n$  of vanishing mesh. The finite bases  $\mathfrak{R}_{\pi_n} \otimes (\delta_{-1} + \delta_1)/2$  and weights  $e^{\tau w^{a_n} - a_*/2}$  have uniformly bounded positive and inverse moments of every order. Their sampled Gaussian arrays converge, with off-diagonal path  $a$  and fixed diagonal  $a_*$ . Lemma 3.3 therefore defines the limiting averaged Ising replica laws  $\mathbb{E}\langle \cdot \rangle_a^I$ . Control stability and the quantile identity pass the transform identification to the limit. The constraint identities in theorem 3.2 give ancestry preservation and

$$\mathcal{S}_{I,K}(a) = \mathcal{S}_{I,1}(a) = \mathcal{S}_I(a), \quad K \geq 1, \quad (78)$$

$$\mathcal{S}_I(a) = \Phi_{u_A}(c_a), \quad \hat{q}_a^I = A^{-1} p_{c_a, u_A}.$$

In particular,  $\hat{q}_a^I$  is nonnegative, nondecreasing, bounded by one, and constant on the plateaus of  $a$ . Choose one  $A$  for both endpoints of a segment  $a_s = (1-s)a + s\tilde{a}$ . The constraint derivative and response continuity give

$$\left. \frac{d}{ds} \right|_{s=0+} \mathcal{S}_I(a_s) = -\frac{1}{2} \langle \hat{q}_a^I, \tilde{a} - a \rangle. \quad (79)$$

For general sources, apply the integrated finite-grid identity and pass to the limit using the  $L^1$ -continuity of the response and  $0 \leq \hat{q}^I \leq 1$ . The fixed  $A$  keeps every completed constraint variance equal to one.

**The Chen–Issa–Mourrat concavity input.** Concavity follows by identifying the spin transform with their symmetric-Ising functional:

$$\mathcal{S}_I(a) = -\psi^\circ(a/2).$$

Indeed,  $\sqrt{2}w^{a/2}$  has covariance  $a$ , the diagonal subtraction is  $a_*/2$ , and the Ising prior is normalized as in (71). The identity extends from finite grids by (72). Convexity of  $\psi^\circ$  in [9, Proposition 2.2] gives concavity of  $\mathcal{S}_I$ .

**Corollary 4.2** (Responses and supporting equalities). *For either prior,  $\hat{q}_a^\kappa$ , with true diagonal one, belongs to  $\mathcal{Q}$  and is constant on every plateau of  $a$ . The finite and limiting Ising scalar laws preserve genealogy:*

$$\mathbb{E}\langle F \rangle_{\pi;a}^I = \mathbb{E}_\pi F, \quad \mathbb{E}\langle F \rangle_a^I = \mathbb{E}_U F$$

for every bounded Borel test of finitely many replica ancestries. Their supporting equality is

$$\mathcal{A}_I(\hat{q}_a^I) = \mathcal{S}_I(a) + \frac{1}{2}\langle a, \hat{q}_a^I \rangle < \infty. \quad (80)$$

For the spherical response,  $\hat{q}_a^S(1-) = 1 - \bar{\beta}(a)^{-1} \leq a_*/(1 + a_*) < 1$ , and

$$\begin{aligned} K_a(\bar{\beta}(a)) &= \mathcal{S}_S(a) + \frac{1}{2}\{1 + a_* - \bar{\beta}(a)\}, \\ \mathcal{A}_S(\hat{q}_a^S) &= \mathcal{S}_S(a) + \frac{1}{2}\langle a, \hat{q}_a^S \rangle. \end{aligned}$$

*Proof of the Ising assertions.* The constraint specialization gives genealogy preservation and the response properties. Concavity and (79) give

$$\mathcal{S}_I(\tilde{a}) \leq \mathcal{S}_I(a) - \frac{1}{2}\langle \hat{q}_a^I, \tilde{a} - a \rangle.$$

The supremum over  $\tilde{a} \in \mathcal{P}_b$ , after adding  $\langle \hat{q}_a^I, \tilde{a} \rangle/2$ , is attained at  $a$ , proving (80).  $\square$

### 4.3 The spherical supporting equalities

*Proof of the spherical assertions of theorem 4.2.* The definition (73) gives the normalizer identity. Also  $\hat{q}_a^S(1-) = 1 - \bar{\beta}(a)^{-1} \leq a_*/(1 + a_*)$ ; monotonicity and the plateau property follow from (47).

One finite-grid calculation proves the supporting equality and the Fenchel inequality needed below. Given  $q, a$  on a common grid with  $q_n < 1$ , set  $d_1 = \lambda_1^{-1}$  and  $d_{j+1} = d_j + x_j(a_{j+1} - a_j)$ . Define  $h(z) = z - 1 - \log z \geq 0$ , and recall  $x_n = 1$ . Substituting (68) and (76), then collecting terms, gives

$$2F_a(d_n) + \langle q, a \rangle = 2\mathcal{A}_S(q) - \sum_{j=2}^n \left( \frac{1}{x_{j-1}} - \frac{1}{x_j} \right) h(d_j \lambda_j). \quad (81)$$

Here  $\lambda_1 = 1 - \int_0^1 q$  cancels the root term. For the remaining terms,  $\lambda_{j-1} = \lambda_j + x_{j-1}(q_j - q_{j-1})$  makes the coefficient of  $d_j$  equal to  $-\lambda_j(1/x_{j-1} - 1/x_j)$ , including  $j = n$ . Since  $\mathcal{S}_S(a) \leq F_a(d_n)$ , this yields

$$\mathcal{S}_S(a) + \frac{1}{2}\langle q, a \rangle \leq \mathcal{A}_S(q). \quad (82)$$

For  $q = \hat{q}_a^S$ , the explicit response formula (70) and  $V_a(\bar{\beta}) = 1$  give, with the minimizing precisions,

$$q_1 = a_1 d_1^{-2}, \quad q_n = 1 - d_n^{-1}, \quad q_{j+1} - q_j = \frac{d_j^{-1} - d_{j+1}^{-1}}{x_j}.$$

Thus  $\lambda_j = d_j^{-1}$ , the matching relation in [36, Lemma 4.1, (4.8)–(4.11)]; all gaps vanish and  $d_n$  is the minimizer. Consequently

$$\mathcal{S}_S(a) = \mathcal{A}_S(q) - \frac{1}{2}\langle q, a \rangle.$$

Endpoint-preserving step approximations of  $a$ , the stability in [theorem 4.1](#), and the common response bound  $q(1-) \leq a_*/(1 + a_*)$  pass this equality to general sources by (77). The same continuity and (72) extend (82) to all  $q \in \mathcal{Q}_{<1}$ ,  $a \in \mathcal{P}_b$ .  $\square$

For later use, let  $\bar{q}_x$  be the constant off-diagonal path:

$$\bar{q}_x(r) = x \quad (r < 1), \quad \bar{q}_x(1) = 1, \quad 0 \leq x < 1.$$

**Theorem 4.3** (Spherical conjugate formula). *For every  $q \in \mathcal{Q}$ , the cost defined in (3) satisfies*

$$\mathcal{A}_S(q) = \sup_{a \in \mathcal{P}_b} \{ \mathcal{S}_S(a) + \frac{1}{2}\langle a, q \rangle \}. \quad (83)$$

For every  $L^1$ -continuous  $\Psi : \mathcal{Q} \rightarrow \mathbb{R}$ ,

$$\inf_{q \in \mathcal{Q}} \{ \mathcal{A}_S(q) + \Psi(q) \} = \inf_{q \in \mathcal{Q}_{<1}} \{ \mathcal{A}_S(q) + \Psi(q) \}. \quad (84)$$

*Proof.* Denote the supremum in (83) by  $C(q)$ . The Fenchel inequality gives  $C(q) \leq \mathcal{A}_S(q)$  when  $q(1-) < 1$ . For step  $q$ , construct its inverse source by

$$a_1 = \frac{q_1}{\lambda_1^2}, \quad a_{j+1} - a_j = \frac{1}{x_j} \left( \frac{1}{\lambda_{j+1}} - \frac{1}{\lambda_j} \right).$$

This is a bounded nonnegative nondecreasing source. With  $d_j = \lambda_j^{-1}$ , its scalar response is  $q$  and  $V_a(d_n) = q_n + d_n^{-1} = 1$ , so the trial precision is the minimizer. All gaps in (81) vanish, giving equality, also when  $n = 1$  and the sum is empty. For general  $q(1-) \leq \bar{q} < 1$ , take step approximations  $q_m$  with this same endpoint bound. Their inverse sources obey  $\|a_m\|_\infty \leq \bar{q}/(1 - \bar{q})^2$ , and hence

$$C(q) \geq \mathcal{S}_S(a_m) + \frac{1}{2}\langle q, a_m \rangle = \mathcal{A}_S(q_m) + \frac{1}{2}\langle q - q_m, a_m \rangle \longrightarrow \mathcal{A}_S(q).$$

For any  $q \in \mathcal{Q}$ , put  $q_\delta = (1 - \delta)q$  off the diagonal and  $q_\delta(1) = 1$ . Nonnegativity of the sources and interchange of the two suprema give

$$C(q_\delta) = \sup_{a \in \mathcal{P}_b} \{ \mathcal{S}_S(a) + \frac{1}{2}(1 - \delta)\langle a, q \rangle \} \uparrow C(q).$$

Since  $\gamma_{q_\delta} \downarrow \gamma_q$ , monotone convergence in (3) gives  $\mathcal{A}_S(q_\delta) \uparrow \mathcal{A}_S(q)$ . The identity  $C(q_\delta) = \mathcal{A}_S(q_\delta)$  therefore proves  $C(q) = \mathcal{A}_S(q)$ , including infinite values. In particular, the spherical supporting equality is

$$\mathcal{A}_S(\hat{q}_a^S) = \mathcal{S}_S(a) + \frac{1}{2}\langle a, \hat{q}_a^S \rangle.$$

Finally,  $q_\delta \rightarrow q$  in  $L^1$  and  $\mathcal{A}_S(q_\delta) \rightarrow \mathcal{A}_S(q)$ . Applying this approximation to every finite-cost competitor proves (84); the reverse inequality follows from inclusion of domains.  $\square$

*Proof of theorem 2.2.* The Ising conjugate formula is (5); the spherical formula is [theorem 4.3](#). Corollary 4.2 gives the supporting equalities. The left-hand side of (48) is  $\mathcal{S}_{\kappa, N}(a)$  from (71). Its asserted value and limit follow from (78) and (74), respectively.  $\square$

## 5 Cavity in $N$

We first prove the overlap equation in [theorem 2.3](#). For Ising spins, separating one coordinate determines its normalized law. On the sphere, integration by parts gives the equation directly. We then compute the logarithmic increment from inserting a fixed block of coordinates at  $t = 1$ . This last calculation is the only one that uses the quadratic coefficient  $B_N$ .

### 5.1 The Ising overlap equation

*Proof of the Ising assertion of [theorem 2.3](#).* Put  $D = N + 1$ , write  $\sigma^D = (\sigma, \tau)$ , and separate the last entries  $z_k$  of the constraint vectors. In the measure with the last coordinate deleted, retain the denominator  $\sqrt{D}$ :

$$X_k^- = \sqrt{t/D} g_k \cdot \sigma + \sqrt{1-t} Y_k, \quad X_k^D = X_k^- + \sqrt{t/D} z_k \tau.$$

Use  $V_N$  in this deleted measure. Its fresh constraint variance is  $1 - t/D$ , and its off-diagonal covariance is  $t(N/D)R + (1-t)q^{0,\sharp}(r)$ . The variable-diagonal version of [theorem 2.1](#) therefore applies. Select a common refinement for this measure and its one- and two-constraint deletions, and write  $q^-$  for their common limiting spin path. Set

$$c^- = (1-t)q^0 + tq^-, \quad a^- = (1-t)a^0 + t\alpha p_{c^-}.$$

No comparison with the usual  $N$ -dimensional denominator is made.

The lifted and full overlaps differ by at most  $2/D$ , while their ancestries agree. The covariance of  $V_N$  is  $N^{3/4}$  times a uniformly bounded Lipschitz kernel. Consequently replacing  $V_D(\sigma, \tau, \eta)$  by  $V_N(\sigma, \eta)$  changes its covariance by  $O(N^{-1/4})$ , uniformly in the states and amplitudes. Apply [theorem 3.4\(ii\)](#) with the perturbations removed from the common density. The replacement changes bounded replicated tests by  $o(1)$ .

Since  $\tau^2 = 1$ , Taylor expansion gives

$$\begin{aligned} \sum_{k \leq M} \{u_k(X_k^D) - u_k(X_k^-)\} &= \tau Z_N + \frac{tM}{2D} C_N + \varepsilon_N, \\ Z_N &= \sqrt{t/D} \sum_{k \leq M} z_k u'_k(X_k^-), \quad C_N = M^{-1} \sum_{k \leq M} u''_k(X_k^-). \end{aligned}$$

Before insertion, the centered quadratic term has conditional second moment at most  $2t^2 ML_2^2/D^2$ . The cubic remainder has averaged second moment  $O(N^{-1})$ , by the uniform third-derivative bound and Gaussian moments. Thus  $\mathbb{E}\mu_N(\varepsilon_N^2) = O(N^{-1})$ , where  $\mu_N$  is the deleted measure times the symmetric Ising prior of  $\tau$ . The exact increment is uniformly Lipschitz in  $(z_k)$ , with squared constant at most  $tML^2/D$ , and its conditional Gaussian mean has absolute value at most  $tML_2/(2D)$ . The approximate increment has the same positive and inverse exponential moment bounds, since  $C_N$  is bounded and  $Z_N$  has bounded conditional variance. The fresh source term  $\tau W_{t,D}$  also has bounded Gaussian variance. Lemma [3.4\(i\)](#) therefore transfers the Taylor estimate through the normalizer.

Apply [\(38\)](#) with  $\chi(\theta, x) = u''_\theta(x)$  and with  $\chi(\theta, x) = u'_\theta(x)^2$ . It gives

$$C_N \longrightarrow c'', \quad Q_{11}^- \longrightarrow p_{\text{diag}}(c^-) \quad \text{in averaged } L^2,$$

where  $c''$  is deterministic on the selected refinement. Row reconstruction also gives  $Q_{12}^- \Rightarrow p_{c^-}(U_{12})$ , jointly with the retained array. Conditional on the old replicas,  $Z_N + W_{t,D}$  is Gaussian, with off-diagonal covariance converging to  $a^-(U_{12})$  and diagonal converging to  $(1-t)a_*^0 + t\alpha p_{\text{diag}}(c^-)$ .

The covariance and coefficient limits, together with the preceding moment bounds, verify [theorem 3.3](#). Thus the normalized insertion law passes to its scalar limit. Represent the excess diagonal variance by an independent terminal Gaussian of variance  $\delta = t\alpha\{p_{\text{diag}}(c^-) - p_{c^-}(1-)\}$ . Its integral is  $e^{\delta\tau^2/2} = e^{\delta/2}$ . The factor from  $C_N$  tends to  $e^{t\alpha c''/2}$ . Both factors are constants and cancel from normalization. The remaining law is exactly the scalar Ising law with weight  $e^{\tau w^{a^-} - a^*/2}$ . This argument uses concentration of  $C_N$ ; the equality  $\tau^2 = 1$  alone would not remove a coefficient depending on the old configuration.

The scalar Ising law preserves the genealogy, by [theorem 4.2](#). The old array after insertion therefore still has law  $(q^-(U), \rho_d(U))$ . Since full and old overlaps differ by  $O(N^{-1})$ , the limiting path  $q$  of the full system equals  $q^-$ . Coordinate exchangeability gives

$$\mathbb{E}\langle \tau^1 \tau^2 f(R_{12}^D, r_{12}) \rangle_D = \mathbb{E}\langle R_{12}^D f(R_{12}^D, r_{12}) \rangle_D.$$

Passing this bounded identity to the scalar law yields

$$\int_0^1 f(q(r), \rho_d(r)) \{ \hat{q}_a^1(r) - q(r) \} dr = 0, \quad a = (1-t)a^0 + t\alpha p_{(1-t)q^0+tq}.$$

Both paths factor through  $(q, \rho_d)$ , by their plateau properties. Continuous tests determine the resulting signed measure, so  $q = \hat{q}_a^1$  almost everywhere. The supporting equality follows from [theorem 4.2](#).  $\square$

## 5.2 The spherical overlap equation

*Proof of the spherical assertion of [theorem 2.3](#).* Let  $F = f(R_{12}, r_{12})$ , with  $f$  continuously differentiable in its first variable; the second variable lies in the finite ancestry alphabet. First retain only the first  $J$  terms of  $V_N$ , and write  $H^{[J]}$  and  $\langle \cdot \rangle_J$  for this Hamiltonian and its replica averages. Put

$$v = \sigma^2 - R_{12}\sigma^1, \quad d_\ell = N^{-1}v \cdot \sigma^\ell; \quad d_1 = 0, \quad d_2 = 1 - R_{12}^2, \quad d_3 = R_{23} - R_{12}R_{13}.$$

The vector  $v$  is tangent to the sphere at  $\sigma^1$ . Its spherical divergence and its action on the test are

$$\text{div}_S(v/N) = -(1 - N^{-1})R_{12}, \quad (v/N) \cdot \nabla F = N^{-1}d_2\partial_R f.$$

Integrating the divergence of  $(v/N)Fe^{H^{[J]}(\sigma^1)}$  in the first replica therefore gives

$$(1 - N^{-1})\mathbb{E}\langle R_{12}F \rangle_J = N^{-1}\mathbb{E}\langle d_2\partial_R f \rangle_J + N^{-1}\mathbb{E}\langle Fv \cdot \nabla H^{[J]}(\sigma^1) \rangle_J.$$

For the Gaussian integrations, condition on the underlying cascade weights and channel marks, and hold each replica state fixed inside its integral. Differentiating a two-replica average with respect to a Gaussian coordinate  $g$  gives

$$\partial_g \langle A \rangle_J = \left\langle \partial_g A + A \{ \partial_g H^{[J],1} + \partial_g H^{[J],2} - 2\partial_g H^{[J],3} \} \right\rangle_J.$$

The two numerator densities give the coefficients 1, 1, and the square of the normalizer gives  $-2$ . Neither  $F$  nor  $v$  has an explicit Gaussian derivative; their dependence through the sampling measure is included in these density terms.

Write  $u_k^\ell = u'_k(X_k^\ell)$  and  $H_{\text{con}} = \sum_k u_k(X_k)$ . Since  $\partial_{g_{ki}} X_k^\ell = \sqrt{t/N} \sigma_i^\ell$ , Gaussian integration by parts gives the full constraint contribution:

$$\begin{aligned} N^{-1}\mathbb{E}\langle Fv \cdot \nabla H_{\text{con}}(\sigma^1) \rangle_J &= \frac{t}{N} \sum_k \mathbb{E} \left\langle F \left[ d_1 u''_k(X_k^1) + u'_k(d_1 u_k^1 + d_2 u_k^2 - 2d_3 u_k^3) \right] \right\rangle_J \\ &= \frac{tM}{N} \mathbb{E}\langle F(d_2 Q_{12} - 2d_3 Q_{13}) \rangle_J. \end{aligned}$$

The  $u''$  term and the diagonal  $(u')^2$  term both multiply  $d_1 = 0$ , which is the required cancellation.

For the source, set  $b_{11} = (1-t)a_*^0$  and  $b_{1\ell} = (1-t)a^{0,\sharp}(r_{1\ell})$  for  $\ell = 2, 3$ . Its covariance and the same density derivative give

$$N^{-1}\mathbb{E}\langle Fv \cdot W_t(\eta^1) \rangle_J = \mathbb{E}\langle F(d_1b_{11} + d_2b_{12} - 2d_3b_{13}) \rangle_J.$$

Again the diagonal term vanishes. Thus the constraints and source together contribute  $\mathbb{E}\langle F(d_2A_{12}^N - 2d_3A_{13}^N) \rangle_J$ , where

$$A_{\ell m}^N = t(M/N)Q_{\ell m} + (1-t)a^{0,\sharp}(r_{\ell m}).$$

The truncated perturbation has covariance  $N^{3/4}k_J(R, r)$ , where  $k_J(R, r) = \sum_{j \leq J} 4^{-j}x_j^2 K_j((1+R)/2, r)$ . Differentiating this covariance in its first spin argument shows that its contribution is exactly

$$N^{-1/4}\mathbb{E}\langle F[d_2\partial_R k_J(R_{12}, r_{12}) - 2d_3\partial_R k_J(R_{13}, r_{13})] \rangle_J.$$

Here the self term also vanishes because  $d_1 = 0$ . Since  $|d_2|, |d_3| \leq 1$  and  $\|\partial_R k_J\|_\infty \leq \frac{1}{2} \sum_j 4^{-j}x_j^2 \leq 2/3$ , this is bounded by  $2\|F\|_\infty N^{-1/4}$ , uniformly in  $J$ .

Now keep  $N$  fixed and let  $J \rightarrow \infty$ . The covariance tail is  $O(N^{3/4}4^{-J})$ , so [theorem 3.4\(ii\)](#) passes all the bounded replica averages in the combined identity to the full perturbation. The covariance-derivative tail is  $O(4^{-J})$ , which also passes the displayed perturbation term to the limit. No convergence of sample gradients is needed. Dropping the truncation subscript, we obtain

$$(1 - N^{-1})\mathbb{E}\langle R_{12}F \rangle = N^{-1}\mathbb{E}\langle d_2\partial_R f \rangle + \mathbb{E}\langle F(d_2A_{12}^N - 2d_3A_{13}^N) \rangle + \varepsilon_N, \quad (85)$$

$$|\varepsilon_N| \leq 2\|F\|_\infty N^{-1/4}.$$

Row reconstruction gives  $A_{\ell m}^N \Rightarrow a(U_{\ell m})$ , jointly with the overlap-ancestry array. We may now let  $N \rightarrow \infty$  in (85): the test-derivative term and  $\varepsilon_N$  vanish. To evaluate the limit, recall the conditional triple law of the uniform RPC: for bounded  $H$ ,

$$2\mathbb{E}[H(U_{13}, U_{23}) \mid U_{12} = r] = \int_0^r H(s, s) \, ds + \int_r^1 \{H(s, r) + H(r, s)\} \, ds + rH(r, r). \quad (86)$$

The GG identity gives the joint law of  $(U_{12}, U_{13})$  as one half product measure plus one half diagonal measure; ultrametricity and replica symmetry then give this formula.

Set

$$B = \int_0^1 qa, \quad C(r) = 1 - B + \int_r^1 a, \quad I(r) = \int_0^r qa,$$

$$D(r) = 1 - \int_r^1 q, \quad \Delta(r) = D(r) - rq(r).$$

Apply (86) to  $H(s, v) = \{q(v) - q(r)q(s)\}a(s)$ . Collecting terms yields

$$\int_0^1 f(q(r), \rho_d(r)) \{q(r)C(r) + I(r) - a(r)\Delta(r)\} \, dr = 0.$$

On every common fiber of  $q$  and  $\rho_d$ , the path  $a$  is constant. The expression in braces is also constant there: the derivatives of its integral terms cancel. Thus it factors through  $(q, \rho_d)$ , by the factorization argument in [theorem 6.3](#). Smooth tests in the first variable and arbitrary tests on the finite ancestry alphabet are dense in continuous tests. Equality of the resulting finite measures therefore implies

$$q(r)C(r) + I(r) = a(r)\Delta(r) \quad \text{for a.e. } r. \quad (87)$$

This equation can be solved without assuming differentiability of the paths. The function  $CD + rI$  is absolutely continuous, and (87) gives

$$(CD + rI)' = -aD + Cq + I + rqa = 0 \quad \text{a.e.}$$

Its value at  $r = 1$  is one. Expanding the product and using (87) proves

$$P(r)\Delta(r) = 1, \quad P(r) := C(r) + ra(r). \quad (88)$$

Right-continuous versions give this identity at every  $r < 1$ . At the terminal point it yields

$$(1 + a_* - B)(1 - q_*) = 1.$$

Thus  $q_* < 1$ , and  $\beta := 1 + a_* - B = (1 - q_*)^{-1}$ . Layer cake gives  $P(r) = \pi_{a,\beta}(a(r))$ , using the precision function defined in Section 2. Moreover  $P(0) = 1/\Delta(0) > 0$ , so  $\beta \in \mathfrak{D}_a$ .

In the sense of Stieltjes measures,  $dP = r da$  and  $d\Delta = -r dq$ . The continuous parts of (88) give  $dq = da/P^2$  away from zero. At a jump at  $r > 0$ , the same identity gives exactly

$$q(r) - q(r-) = \frac{a(r) - a(r-)}{P(r-)P(r)} = \int_{a(r-)}^{a(r)} \frac{dv}{[P(r-) + r\{v - a(r-)\}]^2}.$$

At zero, (87) gives  $q(0) = a(0)/P(0)^2$ . Together these identities prove

$$q(r) = \int_0^{a(r)} \pi_{a,\beta}(v)^{-2} dv = q_{a,\beta}^{\text{co}}(r).$$

The terminal identity says  $V_a(\beta) = q_* + \beta^{-1} = 1$ . Uniqueness of this precision gives  $\beta = \bar{\beta}(a)$ , hence  $q = \hat{q}_a^S$ . The supporting equality follows from [theorem 4.2](#).  $\square$

### 5.3 The logarithmic increment of a fixed block

*Proof of theorem 2.4.* Fix  $K \geq 1$ , put  $D = N + K$ , and use  $t = 1$ . The old measure has  $M$  constraints in dimension  $N$ . For  $\tau \in \mathbb{R}^K$ , use

$$\iota_\tau^I \sigma = (\sigma, \tau), \quad \iota_\tau^S \sigma = \left( \sqrt{\frac{D - \|\tau\|^2}{N}} \sigma, \tau \right).$$

In the Ising case  $\tau \in \{-1, 1\}^K$ . On every fixed cube  $[-R, R]^K$ , the old and full overlaps satisfy

$$\max_{\ell, m} |R_{\ell m}^D - R_{\ell m}^N| \leq C_{K,R}/N,$$

and their ancestry labels agree. The covariance comparison used above therefore replaces  $V_D \circ \iota$  by  $V_N$  with error  $O_{K,R}(N^{-1/4})$  in expected logarithms. In this comparison the common density contains neither perturbation; any old partition function in the insertion ratio cancels from the difference.

For spherical spins, the density of the last  $K$  coordinates is

$$f_{D,K}(\tau) = \frac{\Gamma(D/2)}{(\pi D)^{K/2} \Gamma((D-K)/2)} \left( 1 - \frac{\|\tau\|^2}{D} \right)_+^{(D-K-2)/2}.$$

Conditionally on  $\tau$ , the first coordinates have the spherical law specified by  $\iota_\tau^S$ . Stirling's formula gives

$$f_{D,K}(\tau) = (2\pi)^{-K/2} e^{-\|\tau\|^2/2} \{1 + O_{K,R}(N^{-1})\} \quad \text{uniformly on } [-R, R]^K.$$

Let  $\nu_{N,R}$  be this prior conditioned on the cube, and  $\zeta_{N,R}$  its mass. For Ising spins take  $R \geq 1$ ,  $\nu_{N,R}$  the product Ising prior and  $\zeta_{N,R} = 1$ . Write  $\mu_N = G_{N,M,1} \otimes \nu_{N,R}$ .

Separate  $g_k^D = (g_k, z_k)$ , where  $z_k \in \mathbb{R}^K$ , and put

$$c_D(\tau) = \sqrt{1 - \|\tau\|^2/D}, \quad \xi_k = D^{-1/2} z_k \cdot \tau, \quad Z_i = D^{-1/2} \sum_k z_{ki} u'_k(h_k).$$

The exact new constraint input is  $c_D(\tau)h_k + \xi_k$ , for both priors. Taylor expansion gives

$$\begin{aligned} U_N &:= \sum_k \{u_k(c_D h_k + \xi_k) - u_k(h_k)\} = E_N + \frac{1}{2} T_N + \mathcal{R}_N, \\ E_N &= \tau \cdot Z + \frac{M\|\tau\|^2}{2D} B_N, \quad T_N = \sum_k u''_k(h_k) \{\xi_k^2 - \|\tau\|^2/D\}. \end{aligned}$$

Here  $B_N = M^{-1} \sum_k \{u''_k(h_k) - h_k u'_k(h_k)\}$ . Before insertion the  $z_k$ 's are independent of  $\mu_N$ , so

$$\mathbb{E}_z T_N^2 = \frac{2\|\tau\|^4}{D^2} \sum_k u''_k(h_k)^2 \leq C_{K,R}/N.$$

Taylor's theorem also gives

$$|\mathcal{R}_N| \leq C_{K,R} \left\{ D^{-2} \sum_k (|h_k| + h_k^2) + D^{-1} \sum_k |h_k \xi_k| + \sum_k |\xi_k|^3 + D^{-3} \sum_k |h_k|^3 \right\}.$$

Using  $\sum_k h_k^2 \leq M\Lambda_N$ ,  $\sum_k |h_k|^3 \leq (M\Lambda_N)^{3/2}$ , and Gaussian moments, [theorem 3.5](#) gives

$$\mathbb{E} \mu_N (T_N^2 + |\mathcal{R}_N|^2) \leq C_{K,R}/N.$$

For fixed old state and  $\tau$ , the squared Lipschitz constant of  $U_N$  in  $(z_k)$  is at most  $ML^2\|\tau\|^2/D$ , and

$$|\mathbb{E}_z U_N| \leq L|c_D - 1| \sum_k |h_k| + \frac{ML_2\|\tau\|^2}{2D} \leq C_{K,R}(1 + \sqrt{\Lambda_N}).$$

Gaussian concentration bounds its positive and negative exponential moments. The same bound holds for  $E_N$ , since its conditional Gaussian variance is bounded and  $|B_N| \leq C_u(1 + \sqrt{\Lambda_N})$ . Hölder and the spectral exponential bound therefore give, for each fixed  $p < \infty$ ,

$$\sup_{0 \leq s \leq 1} \mathbb{E} \mu_N \left( e^{p\{(1-s)E_N + sU_N\}} + e^{-p\{(1-s)E_N + sU_N\}} \right) \leq C_{p,K,R}. \quad (89)$$

Only the exact and quadratic endpoint exponents are used here; no exponential moment of the cubic Taylor bound is required. Lemma [3.4\(i\)](#) now changes the expected log normalizer by  $O_{K,R}(N^{-1/2})$  when  $U_N$  is replaced by  $E_N$ .

On the selected refinement, row reconstruction and [theorem 3.6](#) give

$$Q_{12} \Rightarrow p_q(U_{12}), \quad Q_{11} \longrightarrow p_{\text{diag}}(q), \quad B_N \longrightarrow \langle q, p_q \rangle - p_{\text{diag}}(q),$$

jointly with the old array; the last two limits hold in averaged  $L^2$ . Conditional on the old measure and sampled states, the inserted Gaussian vector has covariance

$$\mathbb{E}_z Z_i^\ell Z_j^m = \delta_{ij}(M/D)Q_{\ell m}.$$

Put  $a = \alpha p_q$ . Its limit consists of  $K$  independent Gaussian trees of covariance path  $a$ , each completed by independent terminal variance  $\alpha\{p_{\text{diag}}(q) - p_q(1-)\}$ . The covariance limits and (89) verify [theorem 3.3](#) at fixed  $K, R$ . Equivalently, approximate the paths by finite cascades, preserving  $a_*$ ; sampled covariances, coefficients, and logarithmic normalizers have the same limit. The prior mass  $\zeta_{N,R}$  is restored after using the probability base  $\mu_N$ .

Integrating the terminal Gaussians cancels the  $p_{\text{diag}}(q)$  term. For Ising spins the resulting block weight is

$$e^{K\langle q, a \rangle / 2} \prod_{i=1}^K e^{\tau_i w_i^a - a_*/2} \left( \frac{\delta_{-1} + \delta_1}{2} \right)^{\otimes K} (d\tau).$$

The independent Gaussian trees have product normalized transitions; their logarithmic RPC recursion is additive. Its expected logarithm is  $K\{\mathcal{S}_I(a) + \langle q, a \rangle / 2\}$ . The Ising overlap equation already proved identifies this as  $K\mathcal{A}_I(q)$ .

For spherical spins the limiting cutoff weight is

$$\mathbf{1}_{[-R, R]^K}(\tau) \exp \left\{ \sum_{i=1}^K \tau_i w_i^a - \frac{\beta}{2} \|\tau\|^2 \right\} \frac{d\tau}{(2\pi)^{K/2}}, \quad \beta = 1 + a_* - \langle q, a \rangle.$$

The same product recursion gives expected logarithm  $KK_{a,R}(\beta)$ . Restriction to the cube can only decrease the finite dimensional normalizer. The spherical overlap equation gives  $\beta = \bar{\beta}(a)$ , and the logarithmic cutoff limit (67) consequently yields

$$\liminf_N \{A_{N+K}(M) - A_N(M)\} \geq KK_a(\bar{\beta}(a)) = K\mathcal{A}_S(q).$$

The last equality uses the supporting identity and  $1 + a_* - \bar{\beta}(a) = \langle q, a \rangle$ . This proves the claimed increment for both priors. Only the old array was used; no limiting array in dimension  $N + K$  is needed.

Finally, take  $R = 1$  to obtain the uniform lower bound, without any array-limit hypothesis. The preceding lifting and Taylor comparisons are uniform for  $M/N$  in a compact subset of  $(0, \infty)$  and for all amplitudes. The Gaussian part of  $E_N$  has mean zero under  $\mu_N$ . Gaussian integration by parts gives

$$\mathbb{E}\langle B_N \rangle = \mathbb{E}\langle R_{12}Q_{12} - Q_{11} \rangle \geq -2L^2.$$

The coordinate prior is independent of the old measure and  $\nu_{N,1}(\|\tau\|^2) \leq K$ . Jensen therefore gives

$$A_{N+K}(M) - A_N(M) \geq \log \zeta_{N,1} + \mathbb{E} \log \mu_N(e^{E_N}) + o(1) \geq \log \zeta_{N,1} - MKL^2/D + o(1) \geq -C_K.$$

Here  $\zeta_{N,1} = 1$  for Ising spins and  $\zeta_{N,1} \rightarrow \gamma([-1, 1])^K > 0$  on the sphere.  $\square$

## 6 Perturbation, positivity, and synchronization

The free energy bounds require the conditional cavity hypotheses to hold simultaneously for the interpolating measures and their fixed row and dimension shifts. We construct a perturbation that enforces the joint GG identities at vanishing pressure cost. Pressure concentration permits a common subsequence selection; synchronization then represents the limiting joint array through a limiting genealogy array. Throughout, assume (24) and use the common bound  $L = L_1$ .

## 6.1 The perturbation and its covariance

Fix  $\kappa \in \{I, S\}$ ,  $d \geq 2$ , and the trial paths  $q^0 \in \mathcal{Q}_d$  and  $a^0$  used in Section 2.2. Use the finite RPC  $\mathfrak{R}_d$  with masses  $j/d$ ,  $1 \leq j < d$ , and ancestry quantile  $\rho_d$ , together with the independent constraint vectors and Gaussian fields specified there. Trial covariances use the lookup functions  $q^{0,\sharp}, a^{0,\sharp}$ ; in particular their values at the terminal label  $r = 1$  are  $q_*^0, a_*^0$ . We keep  $d$  fixed while taking the dimension limit.

For the interpolation (43)–(44), define the perturbation as follows.

Put  $R^+ = (1 + R)/2$ , enumerate  $(\theta_j, k_j) \in (\mathbb{Q} \cap [0, 1]) \times \mathbb{N}_+$ , and define

$$K_j(x, r) = \frac{[\theta_j x + (1 - \theta_j)r]^{k_j}}{k_j}.$$

Let  $G_{j,N}$  have independent centered Gaussian noises, independent of the constraint vectors, source fields, and cascade weights, with

$$\mathbb{E}G_{j,N}(s^1)G_{j,N}(s^2) = NK_j(R_{12}^+, r_{12}), \quad s^\ell = (\sigma^\ell, \eta^\ell).$$

For  $\mathbf{x} = (x_j)_{j \geq 1} \in [1, 2]^\mathbb{N}$ , set

$$V_N = N^{-1/8} \sum_{j \geq 1} 2^{-j} x_j G_{j,N}. \quad (90)$$

Write  $b_h = h/d$  for  $0 \leq h \leq d - 2$  and  $b_{d-1} = 1$ . The finite ancestral construction is

$$K_j(x, r) = K_j(x, 0) + \sum_{h=1}^{d-1} \mathbf{1}_{\{r \geq b_h\}} \{K_j(x, b_h) - K_j(x, b_{h-1})\}.$$

Every polynomial in braces, and  $K_j(x, 0)$ , has nonnegative coefficients in  $x$ . Since  $R^+$  and the ancestry-cluster indicators are Gram kernels, the Schur product theorem proves positive semidefiniteness. Finite tensor representations in the spin variables and independent Gaussian marks at each ancestor give measurable fields. Choose these marks independently of the entire Poisson tree, the source-field marks, the constraint vectors, and the channel marks. The source fields have the same construction using their trial lookup covariances. On  $[0, 1]^2$ ,  $0 \leq K_j \leq 1$  and  $\text{Lip}(K_j) \leq 1$ , so

$$\sum_{j \geq 1} 4^{-j} (\|K_j\|_\infty + \text{Lip}(K_j)) \leq \frac{2}{3}. \quad (91)$$

For each fixed  $N$ ,  $V_N$  is therefore an  $L^2$  Gaussian limit of finite mixtures, uniformly over bounded amplitudes. The amplitudes are suppressed in  $V_N$ . Choose these processes once for each  $N$ ; inserting or deleting constraints leaves  $V_N$  unchanged.

The lower-bound model is  $t = 1$ ; without  $V_N$  it is the original zero-field model. For that endpoint we fix  $d = 2$ .

**Proposition 6.1** (Perturbation, positivity, and synchronization). *The series (90) is well defined. Adding  $V_N$  changes  $N^{-1} \mathbb{E} \log Z$  by  $o(1)$ , uniformly in the amplitudes and in  $t$ . Let  $M/N \rightarrow \alpha$ , and fix  $t \in [0, 1]$ . For every  $n \geq 2$ , bounded continuous  $F$  of  $(R_{\ell m}, r_{\ell m})_{\ell, m \leq n}$ , and bounded continuous  $f$  on  $[-1, 1] \times [0, 1]$ ,*

$$\mathbb{E}_{\mathbf{x}} \left| \mathbb{E} \langle F f(R_{1,n+1}, r_{1,n+1}) \rangle - \frac{1}{n} \mathbb{E} \langle F \rangle \mathbb{E} \langle f(R_{12}, r_{12}) \rangle - \frac{1}{n} \sum_{\ell=2}^n \mathbb{E} \langle F f(R_{1\ell}, r_{1\ell}) \rangle \right| \longrightarrow 0. \quad (92)$$

The inner expectations keep the amplitudes fixed. The same assertions hold for the Ising coordinate-deleted measures specified in Section 2.2, whose constraints retain the denominator of the larger dimension. Every averaged replicated-array limit at fixed amplitudes and time that satisfies the joint  $(R, r)$  GG identities has  $R_{12} \geq 0$  and the RPC representation (27) for a deterministic  $q \in \mathcal{Q}$ , with diagonal  $(1, 1)$  and the ancestry law in (93).

The scalar covariance terms give positivity of the limiting spin overlap; the mixed terms give joint identities for spin overlap and ancestry. The common subsequence selection below applies simultaneously to the row deletions and the coordinate-deleted Ising families required by the overlap equation.

## 6.2 Finite ancestry and pressure concentration

Integrating spins and terminal Gaussians gives the marked-RPC integral

$$L_N(\eta) = \int e^{\mathcal{H}_{N,M,t}(\sigma, \eta, z^0)} \mu_N^\kappa(d\sigma) \gamma^{\otimes M}(dz^0), \quad Z_N = \int L_N(\eta) \mathfrak{R}_d(d\eta).$$

Conditional on the channel marks, constraint vectors, and common Gaussian increments, the ancestor marks are independent of the Poisson tree. The bound  $|u_k(x)| \leq L|x|$  and Gaussian exponential moments give every positive and inverse moment of the unnormalized weight, averaged against the spin, cascade, and terminal-Gaussian product measure. For fixed  $N, M$ , these bounds are uniform in finite truncations of  $V_N$ , and verify the conditional moment hypotheses of theorem 3.1. Its genealogy conclusion gives

$$\mathbb{E} \langle F((r_{\ell m})_{\ell, m \leq n}) \rangle_{N,M,t} = \mathbb{E} \mathfrak{R}_d F((r_{\ell m})_{\ell, m \leq n}), \quad (93)$$

for every bounded Borel genealogy test  $F$ , with fixed amplitudes. The covariance tail of the perturbation tends uniformly to zero at fixed  $N$ , so theorem 3.4(ii) passes this identity from finite perturbation sums to the full series. The same argument applies after fixed dimension or constraint shifts, constraint deletion, and coordinate restriction.

**Cost and concentration of the perturbation.** Write  $a_N = N^{-1/8}$ , and let  $Z_N^{\text{pert}}$  and  $Z_N^0$  be the normalizers with and without  $V_N$ . At fixed  $d$ , uniformly over the amplitudes,  $t \in [0, 1]$ , and constraint densities in a fixed compact subset of  $(0, \infty)$ , we claim

$$N^{-1} \left| \mathbb{E} \log Z_N^{\text{pert}} - \mathbb{E} \log Z_N^0 \right| \leq C a_N^2 = o(1), \\ \text{Var}(N^{-1} \log Z_N^{\text{pert}}) \leq C_d / N. \quad (94)$$

For the first bound, apply theorem 3.4(ii) to 0 and  $V_N$ , using the normalized unperturbed measure as base. By (91),  $\delta \leq C N a_N^2$ ; dividing by  $N$  gives the claim. The second bound follows from a concentration estimate for marked cascades.

**Lemma 6.2** (Pressure concentration on a finite RPC). *Let a finite RPC have first mass  $m > 0$ , and let  $\mathcal{Z}(g) = \sum_\eta v_\eta L(g, Y_\eta)$  satisfy the conditional moment hypotheses of theorem 3.1. Here  $g$  is a standard Gaussian vector independent of the Poisson and branch-mark trees, whose laws do not depend on  $g$ . If  $\log L(g, y)$  is  $D$ -Lipschitz in  $g$ , uniformly in the branch marks  $y$ , then*

$$\text{Var}(\log \mathcal{Z}) \leq D^2 + C m^{-2}, \quad (95)$$

where  $C$  is universal, independent of the later masses and the weight.

*Proof.* We separate the Gaussian fluctuations of the conditional mean from the cascade fluctuations at fixed  $g$ . Put  $b(g) = \log S_0(g)$ , with  $S_0$  from the marked recursion. For unnormalized cascade weights  $w_\eta$ , set  $T = \sum_\eta w_\eta$  and  $T' = \sum_\eta w_\eta L(g, Y_\eta)/S_0(g)$ . Conditional only on  $g$ , [theorem 3.1](#) gives

$$T' \stackrel{d}{=} T, \quad \log \mathcal{Z} - b(g) = \log T' - \log T. \quad (96)$$

The total-mass calculation in [\[26, Lemma 2.4\]](#) identifies  $T$  as a deterministic scale times a positive  $m$ -stable variable. Later levels enter only through finite fractional moments, since their masses are strictly increasing.

For any  $Q > 0$  with  $\mathbb{E}e^{-sQ} = e^{-cs^m}$ , take  $E \sim \text{Exp}(1)$  independent of  $Q$ . Then

$$\mathbb{P}(E/Q > s) = e^{-cs^m}, \quad \log(E/Q) \stackrel{d}{=} m^{-1}(\log E' - \log c), \quad E' \sim \text{Exp}(1).$$

Since  $\log E$  and  $\log(E/Q)$  have finite second moments, so does  $\log Q$ . Independence of  $E$  and  $Q$  now yields

$$\text{Var}(\log Q) = (m^{-2} - 1) \text{Var}(\log E) \leq Cm^{-2}.$$

Centering the two identically distributed logarithms in [\(96\)](#) therefore gives

$$\mathbb{E}[\log \mathcal{Z} \mid g] = b(g), \quad \text{Var}(\log \mathcal{Z} \mid g) \leq Cm^{-2}.$$

No independence of  $T, T'$  is needed. Finally, the bound  $f(g, y) \leq f(\tilde{g}, y) + D\|g - \tilde{g}\|$  passes through every logarithmic moment transform in the backward recursion by order preservation and translation equivariance. Thus  $b$  is  $D$ -Lipschitz. Gaussian Poincaré and total variance prove [\(95\)](#).  $\square$

Condition on all channel marks; the following constants are uniform in those marks. The common Gaussian vector consists of the constraint vectors, common trial-row and spin-source increments, and the root noises of  $V_N$ . All higher ancestral increments are independent branch marks; terminal Gaussians are integrated inside  $L_N$ . Using  $\|\sigma\|^2 = N$ ,  $|u'_k| \leq L$ , and [\(91\)](#), the squared Lipschitz constant of the Hamiltonian in the common Gaussian vector is at most

$$D_N^2 \leq tML^2 + (1-t)q^0(0)ML^2 + (1-t)Na^0(0) + CN a_N^2 \leq C'N.$$

Logarithmic spin and terminal integration preserves this bound. The first cascade mass is  $m = 1/d$ , so [theorem 6.2](#) gives

$$\text{Var}(\log Z_N^{\text{pert}} \mid \text{channel marks}) \leq CN + Cd^2.$$

The bound is uniform over finite perturbation sums and bounded amplitudes. To pass to the full sum, let  $V_N^{(J)}$  retain its first  $J$  fields and let  $\langle \cdot \rangle_J$  be the associated Gibbs average. Independence and the vanishing diagonal variance of the Gaussian tail give, at fixed  $N, M$ ,

$$\mathbb{E} \left| \log Z_N^{\text{pert}} - \log Z_N^{(J)} \right|^2 \leq \mathbb{E} \left\langle e^{2|V_N - V_N^{(J)}|} - 1 \right\rangle_J \longrightarrow 0.$$

Here  $|\log \langle e^V \rangle| \leq \log \langle e^{|V|} \rangle$  and Jensen's inequality give the displayed bound. Thus it also holds for the full perturbation, after fixed constraint or dimension shifts.

To include mark fluctuations, resample one mark and delete its complete row. Under the deleted law  $\Gamma$ , the change in  $\log Z$  is at most  $2A$ , where  $A = \log \Gamma(e^{L|X|})$  and  $X$  is the fresh completed row input. By [\(60\)](#),  $\mathbb{E}A^2 \leq \mathbb{E}e^{2L|Z|}$  for a standard Gaussian  $Z$ . Conditional Jensen and Efron–Stein therefore bound the variance of the conditional mean given the marks by  $CM$ . Returning to the full disorder law, total variance gives

$$\text{Var}(\log Z_{N,M,t}) \leq CN + Cd^2 + CM = O_d(N), \quad (97)$$

which proves (94) and supplies concentration for the joint GG identities below. For the original centered model, omit the cascade, source fields, and perturbation. The same Gaussian gradient bound and mark-resampling argument give  $\text{Var}(\log Z_{N,M}^{\kappa,\nu}) \leq ML^2 + CM$ . Dividing by  $N^2$  proves (51).

### 6.3 Positivity, Ghirlanda–Guerra identities, and synchronization

The following representation identifies the array limits once the joint GG identities hold.

**Lemma 6.3** (Uniform refinement of a finite ancestry array). *Let  $(R, r)$  be jointly weakly exchangeable positive semidefinite arrays with diagonal  $(1, 1)$ , satisfying the joint GG identities, and suppose that  $r$  has the ancestry law of  $\mathfrak{R}_d$ . Then  $R_{12} \geq 0$  almost surely and the joint array has the representation (27) for a deterministic  $q \in \mathcal{Q}$ . This path is uniquely determined almost everywhere by the law of  $R_{12}$ .*

*Every bounded Borel function  $h$  that is constant almost everywhere on each common fiber of  $q$  and  $\rho_d$  admits a Borel factorization  $h(u) = \bar{h}(q(u), \rho_d(u))$  almost everywhere, with  $\bar{h} : [0, 1]^2 \rightarrow \mathbb{R}$ . If a nondecreasing covariance path  $c$  factors through  $(q, \rho_d)$ , then so does  $p_c$ . In particular, for the trial paths of Section 2.2,  $t \in [0, 1]$ , and  $\alpha > 0$ ,*

$$c = (1 - t)q^0 + tq, \quad p_c, \quad a = (1 - t)a^0 + t\alpha p_c, \quad \hat{q}_a$$

*all have such factorizations. Their evaluations at  $U_{12}$  are therefore measurable functions of the retained pair  $(R_{12}, r_{12})$ , although the uniform refinement  $U$  itself need not be observable.*

*Proof.* The scalar GG identities for  $R$  and the positivity principle [26, Theorem 2.16] give

$$\mathbb{P}(R_{12} < 0) = 0.$$

Put  $S = (R + r)/2$ . Panchenko’s synchronization theorem [27, Theorem 4], applied with weights  $1/2, 1/2$ , gives deterministic nondecreasing 2-Lipschitz maps  $L_1, L_2$  such that

$$R_{\ell m} = L_1(S_{\ell m}), \quad r_{\ell m} = L_2(S_{\ell m}).$$

The scalar array  $S$  also satisfies GG. The characterization and existence results [26, Theorems 2.13, 2.17 and Remark 2.1] represent its off-diagonal law as  $s_0(U_{\ell m})$ , where  $s_0$  is the nondecreasing quantile of  $S_{12}$  and  $U$  is the uniform limiting genealogy array from finite RPCs. These results apply to the off-diagonal Gram–de Finetti array via the Dovbysh–Sudakov representation, as in the proof of Theorem 2.17; the specified true diagonal is kept separately. Thus  $q = L_1 \circ s_0$  and  $\rho = L_2 \circ s_0$  give the joint representation. Since  $U_{12}$  is uniform, the nondecreasing  $\rho$  is the quantile of  $r_{12}$ , hence equals  $\rho_d$  almost everywhere. Likewise  $q$  is the quantile of  $R_{12}$ . Choose right-continuous versions and set  $q(1) = 1$ ; then

$$((R_{\ell m}, r_{\ell m}))_{\ell < m} \stackrel{d}{=} ((q(U_{\ell m}), \rho_d(U_{\ell m})))_{\ell < m}, \quad q \in \mathcal{Q}.$$

The diagonals on both sides are  $(1, 1)$ , independently of the off-diagonal endpoint values. The relation  $s_0 = (q + \rho_d)/2$  holds almost everywhere.

For the factorization, put  $s = (q + \rho_d)/2$ . Monotonicity makes its fibers exactly the common fibers of  $q$  and  $\rho_d$ . Only countably many have positive length. A generalized inverse of  $s$ , with the corresponding almost-everywhere constants assigned at those fiber values, gives a Borel  $H$  with  $h = H \circ s$  almost everywhere. Set  $\bar{h}(x, y) = H((x + y)/2)$ . If  $c$  factors through  $(q, \rho_d)$ , it is constant almost everywhere on each common fiber; the plateau property of  $p_c$  gives the second assertion.

On such a fiber the lookup identities  $q^{0,\sharp} \circ \rho_d = q^0$  and  $a^{0,\sharp} \circ \rho_d = a^0$  show that the trial paths are constant there as well. Thus  $c$  is constant there, and the plateau property of  $p_c$  in [theorem 3.2](#) implies the same for  $p_c$  and  $a$ . [Theorem 4.2](#) gives the same for  $\hat{q}_a$ .  $\square$

*Proof of Proposition 6.1.* Apply the argument proving [\[26, Theorems 3.2–3.3\]](#) to one perturbation component at a time. For fixed  $j$ , the field  $g_{N,j} = G_{j,N}/\sqrt{N}$  has covariance  $K_j$ , bounded by one with deterministic diagonal  $1/k_j$ , and coefficient  $s_{N,j}x_j$ , where  $s_{N,j} = 2^{-j}N^{3/8} \rightarrow \infty$ . Fix all other amplitudes and absorb the remaining Hamiltonian into the random reference measure. On the fixed finite-depth label tree, the selected Gaussian noise is independent of this reference environment. These are precisely the covariance and independence properties used in the cited proof: Gaussian integration by parts is conditional on the reference environment, while its convexity estimate uses pressure concentration over all disorder.

By [\(97\)](#), uniformly over amplitudes in  $[0, 3]$ ,

$$\mathbf{v}_N := \sup_{\mathbf{x}} \mathbb{E} |\log Z_{N,M,t} - \mathbb{E} \log Z_{N,M,t}| = O_d(\sqrt{N}), \quad \frac{\mathbf{v}_N}{s_{N,j}^2} = O_{d,j}(N^{-1/4}) \rightarrow 0.$$

The cited argument therefore makes the GG defect with pair test  $K_j(R^+, r)$  tend to zero after averaging  $x_j$ , for every bounded test of the full overlap–ancestry array. Averaging the remaining amplitudes preserves this conclusion. The rational mixed powers defining  $K_j$ , together with constants, span all bivariate polynomials, so uniform approximation gives [\(92\)](#).

For the Ising coordinate-deleted measures, the constraint coefficient is  $\sqrt{t/D}$  in dimension  $D - 1$ , and the completed variance is  $1 - t/D$ . The Gaussian gradient bound, the cascade moment bounds, and the independence of the perturbing noises used above are unchanged. Thus their pressure concentration and GG defect estimates hold with the same orders, uniformly in  $t$ . Their fresh-constraint covariance rules converge uniformly to  $tR + (1 - t)q^{0,\sharp}(r)$ .

**Common subsequence selection.** Fix a starting dimension sequence and finitely many changes in the constraint count, including the coordinate-deleted Ising families when they are needed. Enumerate countable dense families of GG tests for these systems, with nonnegative defects  $e_{N,j}$ . Each defect has vanishing parameter average: the parameter law  $\mu$  is the amplitude law for the lower specialization, or its product with Lebesgue measure on  $t \in (0, 1)$  for the interpolation. The latter assertion follows by dominated convergence; the lower endpoint is selected separately. Choose  $N_k$  so that

$$\sum_{j=1}^k \int e_{N_k,j} d\mu \leq 2^{-k}.$$

Tonelli’s theorem gives  $e_{N_k,j} \rightarrow 0$  for every  $j$  outside one  $\mu$ -null set. Uniform approximation gives all continuous GG tests. The same selection accommodates countably many fixed constraint-count changes or coordinate-deleted families, without a rate uniform over them. At each retained parameter, compactness gives a further, possibly parameter-dependent, subsequence on which any fixed finite collection of arrays converges jointly. Their spin paths are not yet asserted to agree.

**Positivity and synchronization of the limits.** The joint GG identities of any limiting array extend to bounded Borel tests: for each replica count they express equality of finite Borel measures on a compact space. Positive semidefiniteness, joint weak exchangeability, and the diagonals pass to the limit, and [\(93\)](#) preserves the finite ancestry law. [Lemma 6.3](#) therefore gives positivity and the representation [\(27\)](#).

No constraint derivative enters the perturbation or the selection.  $\square$

**The perturbation under coordinate insertion.** The processes  $V_N$  are independent of the constraints and spatial source fields, and their covariance depends on the spins only through their overlap. The averaged Ising measures are therefore invariant under coordinate permutations. For insertion of  $K$  coordinates, write  $D = N + K$ . The embeddings are

$$\iota_\tau^I \sigma = (\sigma, \tau), \quad \iota_\tau^S \sigma = \left( \sqrt{\frac{D - \|\tau\|^2}{N}} \sigma, \tau \right).$$

For fixed  $K$  and  $\tau, \tau' \in [-R, R]^K$ , they change spin overlaps by  $O_{K,R}(N^{-1})$  and leave ancestry unchanged. The covariance scale  $N^{3/4}$  and (91) give

$$\sup_{s, s', \tau, \tau'} |\text{Cov}(V_D(\iota_\tau s), V_D(\iota_{\tau'} s')) - \text{Cov}(V_N(s), V_N(s'))| \leq C_{K,R} N^{-1/4}.$$

This is the comparison used in Section 5. Also,  $\sum_j 4^{-j} x_j^2 \|\partial_R K_j((1+R)/2, r)\|_\infty \leq C$ . After spherical and Gaussian integration by parts, this bound makes the perturbation contribution to the tangent identity  $O(N^{-1/4})$ . One may first truncate the series: covariance comparison passes its left side to the limit, and the derivative tail bounds its integrated right side. The common selection above supplies all array hypotheses needed in Sections 2 and 5.

## 7 Quenched large deviations

We prove [theorem 1.5](#). The new step is a uniform curvature bound for the pressure as a function of finitely many channel coefficients. It gives the differentiability needed for the Gärtner–Ellis theorem. The spectral bound (62) then permits passage from finitely many empirical averages to the full empirical measure.

### 7.1 Differentiability of the channel pressure

For a bounded function  $f$ , put  $\text{osc}(f) = \sup f - \inf f$ .

**Proposition 7.1** (Differentiability in bounded channel directions). *Fix  $\alpha > 0$ ,  $\kappa \in \{\text{S}, \text{I}\}$ , and  $f_1, \dots, f_k \in \text{BL}(\mathbb{R})$ . Define*

$$\Lambda_f(\lambda) = \frac{1}{\alpha} \mathcal{P}_{\kappa, u_\lambda}(\alpha), \quad u_\lambda = \sum_{i=1}^k \lambda_i f_i, \quad C_f = \frac{1}{4} \sum_{i=1}^k \text{osc}(f_i)^2.$$

*Then  $\Lambda_f$  is convex and belongs to  $C^{1,1}(\mathbb{R}^k)$ , with*

$$\|\nabla \Lambda_f(\lambda) - \nabla \Lambda_f(\lambda')\|_2 \leq C_f \|\lambda - \lambda'\|_2.$$

*Proof.* First let  $q \in \mathcal{Q}$  be a step path, and choose a finite grid  $\pi$  containing its jump points. On the finite cascade  $\mathfrak{R}_\pi$ , use the completed Gaussian field  $X^q$  from (28), whose construction does not depend on  $\lambda$ , and define

$$G_{q,\lambda}^\pi(d\eta, dz) = \frac{e^{u_\lambda(X^q(\eta, z))}}{Z_{q,\lambda}^\pi} \mathfrak{R}_\pi(d\eta) \gamma(dz), \quad Z_{q,\lambda}^\pi = \int e^{u_\lambda(X^q(\eta, z))} \mathfrak{R}_\pi(d\eta) \gamma(dz).$$

The finite-step one-constraint identity (56) gives  $\Phi_{u_\lambda}(q) = \mathbb{E} \log Z_{q,\lambda}^\pi$ ; for bounded Lipschitz channels the identity follows also by uniform mollification. Differentiation in  $\lambda$  is justified by boundedness of the  $f_i$ . For  $h \in \mathbb{R}^k$ , writing  $h \cdot f = \sum_i h_i f_i$ , it gives

$$D_\lambda^2 \Phi_{u_\lambda}(q)[h, h] = \mathbb{E} \text{Var}_{G_{q,\lambda}^\pi}(h \cdot f(X^q)) \leq \frac{1}{4} \text{osc}(h \cdot f)^2 \leq C_f \|h\|_2^2. \quad (98)$$

The first inequality is the variance bound for a random variable taking values in an interval of length  $\text{osc}(h \cdot f)$ . Thus  $\lambda \mapsto \Phi_{u_\lambda}(q) - C_f \|\lambda\|_2^2/2$  is concave for every step path  $q$ .

For an arbitrary  $q \in \mathcal{Q}$ , choose step paths  $q_n \rightarrow q$  in  $L^1$ . On each compact set  $K \subset \mathbb{R}^k$ ,  $L_K := \sup_{\lambda \in K} \text{Lip}(u_\lambda) < \infty$ , and scalar path continuity (32) gives

$$\sup_{\lambda \in K} |\Phi_{u_\lambda}(q_n) - \Phi_{u_\lambda}(q)| \leq \frac{1}{2} L_K^2 \|q_n - q\|_{L^1} \rightarrow 0.$$

Concavity passes to this limit. In particular, for every  $q \in \mathcal{Q}$ ,

$$\Phi_{u_{\lambda+h}}(q) + \Phi_{u_{\lambda-h}}(q) - 2\Phi_{u_\lambda}(q) \leq C_f \|h\|_2^2.$$

Only this semiconcavity bound is needed for general paths.

The main variational formula now shows that

$$\Lambda_f(\lambda) - \frac{C_f}{2} \|\lambda\|_2^2 = \inf_{\substack{q \in \mathcal{Q} \\ \mathcal{A}_\kappa(q) < \infty}} \left\{ \frac{\mathcal{A}_\kappa(q)}{\alpha} + \Phi_{u_\lambda}(q) - \frac{C_f}{2} \|\lambda\|_2^2 \right\} \quad (99)$$

is concave: each function inside the infimum is concave by the finite-step bound (98) and the preceding limit argument. The set of finite-cost paths is nonempty and independent of  $\lambda$ . On the other hand,  $\Lambda_f$  is finite and convex, since it is the pointwise limit of the convex functions  $M_N^{-1} \mathbb{E} \log Z_{N, M_N}^{\kappa, u_\lambda}$ .

For completeness, if a finite convex function  $g$  satisfies that  $g - C \|\cdot\|_2^2/2$  is concave, then its central differences are at most  $C \|h\|_2^2$ . For any two subgradients  $p, p'$  at  $x$ , the subgradient inequalities therefore give  $t(p - p') \cdot h \leq C t^2 \|h\|_2^2$ . Letting  $t \downarrow 0$  shows that the subgradient is unique, so  $g$  is differentiable. Convolution with a compactly supported smooth probability density preserves convexity and the same upper curvature bound; hence  $0 \leq D^2 g_\varepsilon \leq C I_k$ . The gradients converge locally to  $\nabla g$ , proving that  $\nabla g$  is  $C$ -Lipschitz. Apply this fact to (99). No selection or uniqueness of a minimizing path is used.  $\square$

## 7.2 finite dimensional large deviations

Fix  $\kappa, \alpha$ , and abbreviate

$$\Lambda(u) = \frac{\mathcal{P}_{\kappa, u}(\alpha)}{\alpha}, \quad \Lambda_N(u) = \frac{1}{M_N} \log Z_{N, M_N}^{\kappa, u}.$$

All finite dimensional large-deviation principles must hold on one probability-one event. We construct this event for a countable class of channels, then extend the pressure limits to every bounded Lipschitz channel using the spectral bound.

Let  $\mathbf{G}_N$  be the  $M_N \times N$  matrix with rows  $g_k^\top$ , and fix  $C > (1 + \alpha^{-1/2})^2$ . The Gaussian matrix tail estimate in theorem 3.5 and Borel–Cantelli imply, almost surely,

$$\sup_{\sigma \in \Sigma_N^\kappa} \int x^2 L_{N, \sigma}(dx) \leq \frac{\|\mathbf{G}_N\|_{\text{op}}^2}{M_N} \leq C \quad \text{for all sufficiently large } N. \quad (100)$$

Thus  $\rho_N^\kappa$  is eventually supported on

$$K_C = \left\{ \mu \in \mathcal{P}(\mathbb{R}) : \int x^2 \mu(dx) \leq C \right\}.$$

This set is weakly compact, by tightness and lower semicontinuity of the second moment.

For  $u, v \in \text{BL}(\mathbb{R})$ , set  $d_C(u, v) = \sup_{\mu \in K_C} \int |u - v| \, d\mu$ . When  $\rho_N^\kappa(K_C) = 1$ , direct comparison of the integrands gives

$$|\Lambda_N(u) - \Lambda_N(v)| \leq d_C(u, v), \quad |\Lambda(u) - \Lambda(v)| \leq d_C(u, v). \quad (101)$$

To justify the deterministic second inequality, first fix  $u, v$ . The first inequality holds with probability tending to one, while  $|\Lambda_N(u) - \Lambda_N(v)| \leq \|u - v\|_\infty$  always. Take expectations and use [theorem 1.1](#).

Let  $\mathcal{D}$  be the countable rational vector space generated by the constant function one and compactly supported piecewise linear functions with rational breakpoints and values. It is dense for  $d_C$  in  $\text{BL}(\mathbb{R})$ : cutting off  $u$  outside  $[-R - 1, R + 1]$ , while leaving it unchanged on  $[-R, R]$ , costs at most  $C\|u\|_\infty/R^2$ ; the resulting compactly supported function can be approximated uniformly by rational piecewise linear functions. Bounded-channel almost-sure convergence in [theorem 1.1](#), applied to a deterministic channel, gives one event of probability one on which  $\Lambda_N(v) \rightarrow \Lambda(v)$  for all  $v \in \mathcal{D}$ , and (100) holds. On this event, (101) and approximation give

$$\Lambda_N(u) \rightarrow \Lambda(u) \quad \text{for every } u \in \text{BL}(\mathbb{R}). \quad (102)$$

Indeed, the limit superior of the absolute difference is at most  $2d_C(u, v)$  for each  $v \in \mathcal{D}$ .

Fix an environment in this event. For any finite family  $f = (f_1, \dots, f_k)$  of bounded Lipschitz functions, define

$$T_f(\mu) = \left( \int f_1 \, d\mu, \dots, \int f_k \, d\mu \right), \quad Y_N(\sigma) = T_f(L_{N,\sigma}).$$

Its logarithmic moment generating function satisfies the exact identity

$$\frac{1}{M_N} \log \int e^{M_N \lambda \cdot Y_N(\sigma)} \mu_N^\kappa(d\sigma) = \Lambda_N(\lambda \cdot f) \rightarrow \Lambda_f(\lambda).$$

The limit is finite and differentiable on all of  $\mathbb{R}^k$ , by [theorem 7.1](#); there is no boundary of its effective domain at which a steepness condition is needed. The vectors  $Y_N$  lie in a fixed compact rectangle. The Gärtner–Ellis theorem [[12](#), Theorem 2.3.6] therefore gives a full large-deviation principle at speed  $M_N$ , with good rate

$$J_f(z) = \sup_{\lambda \in \mathbb{R}^k} \{\lambda \cdot z - \Lambda_f(\lambda)\}. \quad (103)$$

This conclusion holds for every such finite family on the same event.

### 7.3 The empirical measure principle and nonlinear free energies

We continue on this event and write  $\mathcal{J} = \mathcal{J}_{\kappa,\alpha}$ . Its definition (13) is a supremum of continuous affine functions of  $\mu$ , so  $\mathcal{J}$  is convex and lower semicontinuous. It is nonnegative because  $\Lambda(0) = 0$ . It is infinite outside  $K_C$ : if  $\mu(x^2) > C$ , choose  $R < \infty$  with  $\mu(f_R) > C$ , where  $f_R(x) = \min\{x^2, R\}$  is bounded and Lipschitz. Since  $\nu(f_R) \leq C$  for  $\nu \in K_C$ , (102) gives  $\Lambda(tf_R) \leq tC$  for every  $t > 0$ . Consequently

$$\mathcal{J}(\mu) \geq t\{\mu(f_R) - C\} \rightarrow \infty.$$

Thus all sublevel sets of  $\mathcal{J}$  are compact.

*Passage to empirical measures.* Enumerate  $\mathcal{D} = \{f_j : j \geq 1\}$ . The continuous coordinate map

$$T : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}^\mathbb{N}, \quad T(\mu) = (\mu(f_j))_{j \geq 1},$$

restricts to a homeomorphism of  $K_C$  onto its compact image in the product topology, since the functions in  $\mathcal{D}$  separate measures. Every finite projection of  $T_\# \rho_N^\kappa$  satisfies the good LDP (103) on

the same quenched event. The Dawson–Gärtner theorem [12, Theorem 4.6.1] gives a good LDP in the product space, with rate  $z \mapsto \sup_k J_{(f_1, \dots, f_k)}(z_1, \dots, z_k)$ . The laws are eventually supported on the closed set  $T(K_C)$ . Restricting to that set and transporting by  $T^{-1}$  gives the LDP on  $K_C$ , with rate

$$\sup_{u \in \text{span}_{\mathbb{R}} \mathcal{D}} \{\mu(u) - \Lambda(u)\} = \mathcal{J}(\mu).$$

The equality follows from  $d_C$ -density and (101). Eventual support in  $K_C$  and  $\mathcal{J} = \infty$  outside it extend this to the weak topology on  $\mathcal{P}(\mathbb{R})$ , proving the asserted large-deviation principle.

*Nonlinear free energies.* For bounded weakly continuous  $\Psi$ , Varadhan’s lemma [12, Theorem 4.3.1] now gives

$$\frac{1}{M_N} \log \int e^{M_N \Psi(\mu)} \rho_N^\kappa(\mathrm{d}\mu) \longrightarrow \sup_{\mu} \{\Psi(\mu) - \mathcal{J}(\mu)\}.$$

The integral on the left equals the spin integral in (14). Multiplication by  $M_N/N \rightarrow \alpha$  proves that formula; replacing normalized Ising counting measure by the sum adds  $\log 2$ . This completes the proof of theorem 1.5.

## 8 Applications and singular limits

We first prove the Gaussian-field extension, then pass to zero temperature for scalar projection pursuit and justify the radial passage used in theorem 1.4. Removing quadratic truncations and hardening distance penalties give the ground-state and capacity results. Throughout this section,  $\mathbf{G}_N$  denotes the  $M_N \times N$  matrix with rows  $g_k^\top$ .

### 8.1 Gaussian external fields

We now prove the Gaussian-field extension, using the notation of (6)–(7).

*Proof of the Gaussian-field extension.* Put  $\ell_b(x) = bx$ . On a step path, a Gaussian recursion step of mass  $m$  and variance  $\Delta$  contributes  $b^2 m \Delta / 2$ . Summing and then using path continuity gives

$$\Phi_{\ell_b}(q) = \frac{b^2}{2} \int_0^1 \gamma_q(s) \, \mathrm{d}s = \frac{b^2}{2} \mathsf{L}(q).$$

Apply theorem 1.1 at zero field to  $T_N = \lfloor (\alpha + 1)N \rfloor$  rows, each having channel  $u_\Theta$  with probability  $\alpha/(\alpha + 1)$  and  $\ell_b$  otherwise. This is a Borel family indexed by  $\mathbb{R}$ : assign label  $e^\theta > 0$  to  $u_\theta$ , label 0 to  $\ell_b$ , and the zero channel to negative labels. This mixture has a common Lipschitz bound and the required origin moment. Its variational value is

$$\inf_{q \in \mathcal{Q}} \left\{ \mathcal{A}_\kappa(q) + \alpha \mathbb{E}_\nu \Phi_{u_\Theta}(q) + \frac{b^2}{2} \mathsf{L}(q) \right\}.$$

Conditional on the counts  $M'_N, K'_N = T_N - M'_N$ , the linear rows sum to  $b\sqrt{K'_N/N} g^0 \cdot \sigma$ . Thus the mixture partition  $Z_N^{\text{mix}}$  can be coupled to the desired partition  $Z_N(b)$  using one original-row stream and one  $g^0$ , both independent of the counts and of each other;  $g^0$  is standard Gaussian even when  $K'_N = 0$ , when its coefficient vanishes.

For a fresh original row, centering the channel in (60) bounds its logarithmic increment in  $L^p$  by  $C_p + \|u_\Theta(0)\|_p$ ,  $p \in \{1, 2\}$ , uniformly over the independent deleted law. Conditional-count Minkowski

and  $N^{-1}|(b' - b)g^0 \cdot \sigma| \leq |b' - b|\|g^0\|_2/\sqrt{N}$  give

$$\left\| \frac{\log Z_N^{\text{mix}} - \log Z_N(b)}{N} \right\|_p \leq \frac{C_p}{N} \|M'_N - M_N\|_p + |b| \left\| \sqrt{K'_N/N} - 1 \right\|_p = o(1).$$

The last step uses the binomial count moments and  $M_N/N \rightarrow \alpha$ . This proves the asserted  $L^1$  and  $L^2$  limits; the spherical entropy representation is inherited from the mixture theorem. For a deterministic channel, the squared gradient norms of the normalized pressure in the rows and in  $g^0$  are bounded by  $M_N L^2/N^2$  and  $b^2/N$ , respectively. Gaussian Poincaré gives

$$\text{Var}(F_{N,\kappa}^{u,b}) \leq \frac{M_N \text{Lip}(u)^2}{N^2} + \frac{b^2}{N}. \quad (104)$$

□

## 8.2 Zero temperature and the Montanari–Zhou formula

The zero-temperature passage uses control values for integrable coefficients and measurable marked terminals. We prove their stability and integrability by extending the standard step-coefficient Parisi control argument; see [17, Lemma 18].

**Lemma 8.1** (Scalar control with integrable coefficients). *Let  $\lambda \in L_+^1([0, 1])$ ,  $a \geq 0$ , and let  $K : \mathbb{R} \rightarrow \mathbb{R}$  be  $L$ -Lipschitz. Define*

$$\mathcal{T}_\lambda^a K(t, x) = \sup_v \mathbb{E} \left[ K \left( x + \sqrt{a}(B_1 - B_t) + \int_t^1 \lambda(s) v_s \, ds \right) - \frac{1}{2} \int_t^1 \lambda(s) v_s^2 \, ds \right].$$

*Here  $B$  is standard Brownian motion. The controls are progressively measurable for its increments after  $t$ , also when  $a = 0$ , and satisfy  $\mathbb{E} \int_t^1 \lambda(s) v_s^2 \, ds < \infty$ . The value is finite and  $L$ -Lipschitz in  $x$ , with*

$$|\mathcal{T}_\lambda^a K(t, x) - K(x)| \leq L \sqrt{\frac{2a(1-t)}{\pi}} + \frac{L^2}{2} \int_t^1 \lambda(s) \, ds. \quad (105)$$

*For bounded  $K$ , it is also bounded in absolute value by  $\|K\|_\infty$ . The supremum is unchanged by restricting to  $|v| \leq L$ . The terminal map preserves order and is a contraction in the sup norm. For another such triple  $(\tilde{\lambda}, \tilde{a}, \tilde{K})$  with the same Lipschitz bound  $L$ ,*

$$\begin{aligned} \|\mathcal{T}_\lambda^a K(t, \cdot) - \mathcal{T}_{\tilde{\lambda}}^{\tilde{a}} \tilde{K}(t, \cdot)\|_\infty &\leq \|K - \tilde{K}\|_\infty + \frac{3}{2} L^2 \|\lambda - \tilde{\lambda}\|_{L^1(t, 1)} \\ &\quad + L \sqrt{\frac{2(1-t)}{\pi}} |\sqrt{a} - \sqrt{\tilde{a}}|. \end{aligned} \quad (106)$$

*For step  $\lambda$  and  $a > 0$ , the value is the backward Cole–Hopf/heat recursion solving, between the step endpoints,*

$$\partial_t f + \frac{a}{2} f_{xx} + \frac{\lambda(t)}{2} f_x^2 = 0, \quad f(1, \cdot) = K.$$

*For general  $\lambda$ , it is the uniform limit on  $[0, 1] \times \mathbb{R}$  of these values for any nonnegative step approximation in  $L^1$ . This approximation-independent extension is the stochastic-control solution. It satisfies dynamic programming at every deterministic intermediate time. At zero diffusion,*

$$\mathcal{T}_\lambda^0 K(t, x) = \mathbb{H}_{\int_t^1 \lambda(s) \, ds} K(x), \quad \mathbb{H}_0 K = K. \quad (107)$$

If  $(\theta, x) \mapsto K_\theta(x)$  is jointly Borel measurable on  $\mathbb{R}^2$ , with common Lipschitz bound  $L$ , then  $(\theta, t, x) \mapsto \mathcal{T}_\lambda^a K_\theta(t, x)$  is jointly Borel measurable. For any Borel probability measure  $\nu$  on  $\mathbb{R}$  satisfying  $\int_{\mathbb{R}} |K_\theta(0)| \nu(d\theta) < \infty$ , these values are integrable at every  $(t, x)$ , and the preceding comparison bounds may be integrated over the mark.

*Proof. Clipping and comparison.* Weighted Cauchy–Schwarz makes the drift of every admissible control integrable. Since  $|K(x)| \leq |K(0)| + L|x|$ , its terminal payoff and cost are integrable. For  $w = (-L) \vee (v \wedge L)$ , the terminal payoff lost by clipping is at most  $L \int_t^1 \lambda |v - w|$ , whereas

$$\frac{1}{2} \int_t^1 \lambda (v^2 - w^2) \geq L \int_t^1 \lambda |v - w|.$$

Thus clipping improves the payoff pathwise. The zero control gives the lower growth bound. The upper bound follows from  $L|v| - v^2/2 \leq L^2/2$ , proving (105). For bounded terminals the zero control and nonnegative cost give the sup-norm bound. Comparing translations and terminals gives the Lipschitz bound, order preservation, and contraction. For the same clipped control, changing  $\lambda$  changes the terminal expectation by at most  $L^2 \|\lambda - \tilde{\lambda}\|_1$ , and the cost by at most half this amount. Changing  $a$  changes only the Brownian displacement, with expected magnitude  $\sqrt{2(1-t)/\pi} |\sqrt{a} - \sqrt{\tilde{a}}|$ . Taking suprema proves (106). All these arguments also hold on any subinterval.

*Step coefficients.* For  $a > 0$ , on an interval of length  $\Delta$  where  $\lambda = m > 0$ , the backward operator is

$$F(x) \mapsto \frac{a}{m} \log \mathbb{E} \exp \left\{ \frac{m}{a} F(x + \sqrt{a\Delta} G) \right\}, \quad G \sim N(0, 1).$$

For  $m = 0$ , use  $\mathbb{E}F(x + \sqrt{a\Delta} G)$ . Mollify  $K$  uniformly, preserving its Lipschitz bound. The resulting recursion solves the PDE, has  $|f_x| \leq L$ , and has bounded  $f_{xx}$  for each fixed mollification. For  $X_s = x + \sqrt{a}(B_s - B_t) + \int_t^s \lambda(u) v_u du$ , apply Itô's formula on the successive intervals, stopping first when  $|X_s|$  exceeds a finite level. For clipped controls, linear growth of  $f$ , bounded  $f_x$ , and the integrable Brownian maximum justify removing the stopping. This gives

$$\mathbb{E} \left[ K(X_1) - \frac{1}{2} \int_t^1 \lambda(s) v_s^2 ds \right] = f(t, x) - \frac{1}{2} \mathbb{E} \int_t^1 \lambda(s) (v_s - f_x(s, X_s))^2 ds.$$

Here  $K$  denotes the mollified terminal. The feedback  $v_s = f_x(s, X_s)$  has a bounded Lipschitz drift and gives equality. Thus the recursion equals the control value. Applying the same verification on subintervals gives dynamic programming. Terminal contraction removes the mollification.

At  $a = 0$ , put  $A = \int_t^1 \lambda$ . If  $A > 0$ , the displacement  $z = \int_t^1 \lambda v$  satisfies  $\int_t^1 \lambda v^2 \geq z^2/A$  pathwise; constant controls  $v = z/A$  give the reverse inequality after taking the supremum over deterministic  $z$ . This proves (107). If  $A = 0$ , the value is  $K$ . Splitting the quadratic cost between two intervals gives the Hopf–Lax semigroup identity and hence dynamic programming.

*Integrable coefficients and marks.* Take nonnegative step functions  $\lambda_n \rightarrow \lambda$  in  $L^1$ ; interval averages preserve monotonicity when  $\lambda$  is nondecreasing. Estimate (106) gives uniform convergence to the control value, independently of the approximation. The same estimate on subintervals and terminal contraction pass dynamic programming to the limit. For marked terminals, the finite-step Gaussian recursions are jointly Borel measurable in  $(\theta, t, x)$ . At zero diffusion the Hopf–Lax supremum can be taken over rational displacements. The uniform limits therefore establish measurability for integrable coefficients. Clipped payoffs are bounded in absolute value by

$$|K_\theta(0)| + L|x| + L\sqrt{a}|B_1 - B_t| + \frac{3}{2}L^2\|\lambda\|_1.$$

The assumed first moment makes this bound integrable over marks and Brownian paths. Fubini and dominated convergence therefore apply; no measurable choice of an optimizing control is needed.  $\square$

Write  $\mathcal{T}_\lambda = \mathcal{T}_\lambda^1$ . For  $H \in C_b(\mathbb{R})$  with  $\|H\|_\infty \leq B$ , maximizing displacements in  $H_c H$  have magnitude at most  $2\sqrt{Bc}$ . Comparison using an optimizer gives a spatial increment bound  $2\sqrt{B/c}|h| + h^2/(2c)$ ; subdivision removes the quadratic term. Thus  $H_c H$  is bounded by  $B$  and  $2\sqrt{B/c}$ -Lipschitz. More generally, if  $H$  is  $L$ -Lipschitz, its Hopf–Lax transform is finite,  $L$ -Lipschitz, and  $H \leq H_c H \leq H + cL^2/2$ . If  $(\theta, x) \mapsto H_\theta(x)$  is jointly Borel measurable on  $\mathbb{R}^2$ , with each  $H_\theta$  continuous, the transform is jointly Borel measurable by its rational-displacement supremum. Consequently [theorem 8.1](#) defines  $f_{\mu,c}^{H_\theta} = \mathcal{T}_\mu(H_c H_\theta)$ , with the mark expectation taken after this scalar evolution.

**Lemma 8.2** (Zero-temperature variational comparison). *Fix  $\alpha > 0$ , a nonempty parameter set  $\Lambda$ ,  $b : \Lambda \rightarrow \mathbb{R}$ , and a Borel probability measure  $\nu$  on  $\mathbb{R}$ . For each  $\lambda \in \Lambda$ , let  $(\theta, z) \mapsto H_{\theta,\lambda}(z)$  be jointly Borel measurable on  $\mathbb{R}^2$ , and assume*

$$\sup_{\theta \in \mathbb{R}, \lambda \in \Lambda} \text{Lip}(H_{\theta,\lambda}) \leq L, \quad \int_{\mathbb{R}} |H_{\theta,\lambda}(0)| \nu(d\theta) < \infty \quad \text{for every } \lambda.$$

Let  $\mathcal{C}$  be the nondecreasing right-continuous functions  $\gamma : [0, 1] \rightarrow [0, 1]$  equal to one on some terminal interval, and write  $D_\gamma(t) = \int_t^1 \gamma(s) ds$ . Define

$$\begin{aligned} \mathcal{J}_\beta(\lambda, \gamma) &= \mathbb{E} \mathcal{T}_{\beta\gamma} H_{\Theta,\lambda}(0, 0) + \frac{1}{2\alpha\beta} \int_0^1 \left\{ \frac{1}{D_\gamma(t)} - \frac{1}{1-t} \right\} dt + b(\lambda), \\ \mathcal{F}(\lambda, \mu, c) &= \mathbb{E} \mathcal{T}_\mu(H_c H_{\Theta,\lambda})(0, 0) + \frac{1}{2\alpha} \int_0^1 \frac{dt}{c + \int_t^1 \mu(s) ds} + b(\lambda). \end{aligned} \tag{108}$$

For every fixed  $(\lambda, \mu, c) \in \Lambda \times \mathcal{U} \times (0, \infty)$ , there are  $\gamma_\beta \in \mathcal{C}$  such that  $\mathcal{J}_\beta(\lambda, \gamma_\beta) \rightarrow \mathcal{F}(\lambda, \mu, c)$ . Conversely, for every  $\beta \geq 2$  and  $\gamma \in \mathcal{C}$ ,  $\mathcal{F}(\lambda, \beta\gamma, \beta^{-1}) \leq \mathcal{J}_\beta(\lambda, \gamma) + C_{\alpha,L}(1 + \log \beta)/\beta$ , uniformly in  $\lambda$  and the mark law. Consequently, in  $\mathbb{R} \cup \{-\infty\}$ ,

$$\lim_{\beta \rightarrow \infty} \inf_{\lambda \in \Lambda, \gamma \in \mathcal{C}} \mathcal{J}_\beta(\lambda, \gamma) = \inf_{\lambda \in \Lambda, \mu \in \mathcal{U}, c > 0} \mathcal{F}(\lambda, \mu, c).$$

*Proof. Uniform reverse comparison.* For  $c > 0$ , use the Hopf–Lax bound  $H \leq H_c H \leq H + cL^2/2$  and the elementary inequality

$$\frac{1}{c+x} - \frac{1}{x} + \frac{1}{y} \leq \frac{1}{c+y} \quad (0 < x \leq y).$$

Apply the latter with  $x = \beta D_\gamma(t) \leq y = \beta(1-t)$  and integrate. Order preservation and translation equivariance of  $\mathcal{T}_{\beta\gamma}$  then give

$$\mathcal{F}(\lambda, \beta\gamma, c) \leq \mathcal{J}_\beta(\lambda, \gamma) + \frac{cL^2}{2} + \frac{1}{2\alpha\beta} \log \left( 1 + \frac{\beta}{c} \right).$$

Take  $c = \beta^{-1}$ . The error is independent of  $\lambda$ , terminal offsets, and the mark law, so taking infima proves the lower limit.

*Recovery.* Fix  $(\lambda, \mu, c)$ , taking  $\mu$  right-continuous. For  $\beta > c$ , set  $T = 1 - c/\beta$ ,  $\mu_\beta = \mu \wedge \beta$ , and  $\gamma_\beta = \mu_\beta/\beta$  below  $T$ , and one above. This is admissible, and its entropy is

$$\frac{1}{2\alpha\beta} \int_0^1 \left\{ \frac{1}{D_{\gamma_\beta}(t)} - \frac{1}{1-t} \right\} dt = \frac{1}{2\alpha} \int_0^T \frac{dt}{c + \int_t^T \mu_\beta(s) ds} + \frac{\log(c/\beta)}{2\alpha\beta}.$$

It converges to the entropy in  $\mathcal{F}$ , since  $\mu_\beta \rightarrow \mu$  in  $L^1$  and the denominators are at least  $c$ . For the channel, split at  $T$  and apply [theorem 8.1](#): remove diffusion from the terminal interval, whose drift mass is  $c$ ; replace  $\mu_\beta$  by  $\mu$  below  $T$ ; and restore the omitted  $\mu$ -evolution on  $[T, 1]$ . Terminal contraction bounds the total error by

$$2L\mathbb{E}|G|\sqrt{c/\beta} + \frac{3}{2}L^2\|\mu_\beta - \mu\|_1 + \frac{L^2}{2}\int_T^1 \mu(s) ds \rightarrow 0.$$

This proves recovery uniformly in the mark. Taking the infimum over fixed candidates proves the upper limit, also when it is  $-\infty$ .  $\square$

**Proposition 8.3** (Marked spherical ground states). *Let  $\nu$  be a Borel probability measure on  $\mathbb{R}$ , and let  $(\theta, x) \mapsto H_\theta(x)$  be jointly Borel measurable on  $\mathbb{R}^2$ , with  $\sup_{\theta \in \mathbb{R}} \|H_\theta\|_\infty \leq B$  and  $\sup_{\theta \in \mathbb{R}} \text{Lip}(H_\theta) \leq L$ . For i.i.d. marks  $\Theta_k \sim \nu$ , independent of the Gaussian rows and  $M_N/N \rightarrow \alpha \in (0, \infty)$ , almost surely,*

$$E_N := \max_{\|v\|_2=1} \frac{1}{M_N} \sum_{k \leq M_N} H_{\Theta_k}(g_k \cdot v) \rightarrow \inf_{\mu \in \mathcal{U}, c > 0} \mathcal{F}_\nu(\mu, c),$$

$$\mathcal{F}_\nu(\mu, c) := \mathbb{E}_\nu f_{\mu, c}^{H_\Theta}(0, 0) + \frac{1}{2\alpha} \int_0^1 \frac{dt}{c + \int_t^1 \mu(s) ds}.$$

Convergence also holds in every finite  $L^p$ .

*Proof. Finite temperature.* The class  $\mathcal{C}$  in [theorem 8.2](#) consists precisely of the distribution functions  $\gamma_q$  with  $q(1-) < 1$ , up to endpoint conventions. This restriction does not change the finite-temperature infimum: replacing  $q$  by  $q \wedge (1 - \varepsilon)$  on  $[0, 1]$ , retaining  $q(1) = 1$ , decreases the spherical entropy, while (32), applied to  $\beta H_\Theta$ , changes its  $\beta^{-1}$ -scaled channel term by at most  $\beta L^2 \varepsilon / 2$ . Let  $\varepsilon \downarrow 0$ . The recursion (55) and [theorem 8.1](#) give  $\beta^{-1} \Phi_{\beta H}(q) = \mathcal{T}_{\beta \gamma_q} H(0, 0)$ , first on finite grids and then by approximation. Consequently [theorem 1.1](#) and (75) give, for the partition function with channels  $\beta H_{\Theta_k}$ ,

$$\lim_N \frac{\log Z_N(\beta)}{\beta M_N} = \inf_{\gamma \in \mathcal{C}} \mathcal{J}_\beta(\gamma), \quad (109)$$

where  $\mathcal{J}_\beta$  is (108) with a singleton parameter set and zero penalty. The limit holds almost surely, simultaneously for integer  $\beta$ . By [theorem 8.2](#),  $\lim_{\beta \rightarrow \infty} \inf_\gamma \mathcal{J}_\beta(\gamma) = \inf_{\mu, c} \mathcal{F}_\nu(\mu, c)$ .

*From pressure to the maximum.* The objective defining  $E_N$  has Lipschitz constant at most  $L\|\mathbf{G}_N\|_{\text{op}}/\sqrt{M_N}$ . The cap estimate used in (120) therefore gives, for  $0 < \eta \leq 1$ ,

$$0 \leq E_N - \frac{\log Z_N(\beta)}{\beta M_N} \leq \frac{L\|\mathbf{G}_N\|_{\text{op}}}{\sqrt{M_N}} \eta + \frac{N}{\beta M_N} \log \frac{8}{\eta}.$$

Gaussian operator-norm tails imply that the first coefficient is eventually bounded almost surely. Take  $\eta = \beta^{-1}$ , let  $N \rightarrow \infty$  on the common event in (109), then let integer  $\beta \rightarrow \infty$ . This proves the almost-sure limit. Boundedness gives every finite  $L^p$  convergence.  $\square$

*Proof of theorem 1.6.* First suppose  $h$  is bounded, jointly Borel measurable, and uniformly  $L$ -Lipschitz in its prediction coordinate. Rotate the teacher coordinates to write  $X_i = (G_i, Z_i)$ , where  $G_i \sim N(0, I_k)$  and  $Z_i \sim N(0, I_{d-k})$  are independent, and  $Y_i = \varphi(G_i, \varepsilon_i)$ . For  $r \in B_k(1)$ , set  $s_r = \sqrt{1 - |r|^2}$  and

$$H_{r, (G, Y)}(z) = h(Y, r \cdot G + s_r z), \quad M_{n, d}(r) = \max_{\|v\|_2=1} \frac{1}{n} \sum_i H_{r, (G_i, Y_i)}(Z_i \cdot v).$$

These channels are uniformly bounded and  $L$ -Lipschitz, and their marks are independent of the residual rows. Identify  $(G, Y) \in \mathbb{R}^{k+1}$  with a real label through a fixed Borel isomorphism, the same for every  $r$ ; the resulting channels are jointly Borel measurable with the same bounds. Since  $n/(d-k) \rightarrow \alpha$ , [theorem 8.3](#) gives  $M_{n,d}(r) \rightarrow V(r)$  almost surely at each fixed  $r$ , with  $V(r)$  given by its marked functional. To maximize over the teacher overlap, we need uniform convergence. Comparing the objectives on the same disorder gives

$$|M_{n,d}(r) - M_{n,d}(r')| \leq L|r - r'| \left( \frac{1}{n} \sum_i |G_i|^2 \right)^{1/2} + L|s_r - s_{r'}| \frac{\|Z\|_{\text{op}}}{\sqrt{n}}.$$

The two random coefficients are eventually bounded almost surely, and  $|s_r - s_{r'}| \leq \sqrt{2|r - r'|}$ . The deterministic limit inherits this modulus. Intersect this probability-one event with the convergence events for a countable dense set of overlaps. Finite nets then give uniform convergence on  $B_k(1)$ . The original maximum is exactly  $\max_r M_{n,d}(r)$ , so its limit is  $\sup_r V(r)$ .

For  $a = |r|^2 < 1$ , change variables in the marked functional by

$$t_{\text{new}} = a + (1-a)t, \quad x_{\text{new}} = r \cdot G + \sqrt{1-a}x, \quad c_{\text{new}} = (1-a)c, \quad \mu_{\text{new}}(t) = \mu\left(\frac{t-a}{1-a}\right).$$

The last formula applies on  $[a, 1)$ , with zero extension below  $a$ . For step coefficients it transforms both the scalar evolution and its reciprocal-integral term into [\(15\)](#); [\(106\)](#) extends the identity to integrable coefficients. Conversely every admissible coefficient restricted to  $[a, 1)$  has an admissible inverse under this change, so the infima agree. At  $|r| = 1$  the channel is constant in its residual argument and its value is  $\mathbb{E}h(Y, r \cdot G)$ . In [\(15\)](#) the entropy interval is then empty; letting  $c \downarrow 0$  gives the same value. Applying [theorem 8.3](#) directly to a deterministic channel gives the unlabelled assertion.

Finally let  $h \in C_b(\mathbb{R}^2)$ ,  $\|h\|_{\infty} \leq B$ , and regularize only the prediction coordinate:

$$h_L^-(y, z) = \inf_w \{h(y, w) + L|z - w|\}, \quad h_L^+(y, z) = \sup_w \{h(y, w) - L|z - w|\}.$$

These functions satisfy  $-B \leq h_L^- \leq h \leq h_L^+ \leq B$  and are uniformly  $L$ -Lipschitz in  $z$ . Taking the extrema over rational  $w$  proves joint measurability. For each fixed  $R, y$ , continuity of  $h(y, \cdot)$  gives

$$\eta_L(R, y) := \sup_{|z| \leq R} (h_L^+ - h_L^-)(y, z) \longrightarrow 0, \quad 0 \leq \eta_L(R, y) \leq 2B.$$

The supremum can again be taken over rational points, so  $\eta_L$  is measurable. Writing  $X$  for the full covariate matrix,

$$\sup_{\|w\|_2=1} \frac{1}{n} \sum_i (h_L^+ - h_L^-)(Y_i, w \cdot X_i) \leq \frac{1}{n} \sum_i \eta_L(R, Y_i) + \frac{2B\|X\|_{\text{op}}^2}{nR^2}.$$

For each integer  $L, R$ , bounded-variable concentration and Borel–Cantelli give convergence of the empirical first term to  $\mathbb{E}\eta_L(R, Y)$ , also for triangular arrays. Dominated convergence makes this expectation tend to zero as  $L \rightarrow \infty$ ; Gaussian operator-norm control bounds the other coefficient almost surely. Thus the difference of the two envelope optimization limits tends to zero, first as  $L \rightarrow \infty$ , then  $R \rightarrow \infty$ . The functional for  $h$  is well defined by [theorem 8.1](#) and the Hopf–Lax terminal bounds. Its terminal transform, scalar evolution, infimum, and outer supremum preserve order, so its value lies between the two envelope values. The same squeeze proves convergence of the empirical maximum to this value. The one-variable envelopes give the unlabelled case. Boundedness gives convergence in every finite  $L^p$ .  $\square$

### 8.3 Radial passage for Gaussian spins

The following lemma justifies the radial optimization used in [theorem 1.4](#).

**Lemma 8.4** (Radial Laplace passage). *Let  $u, h, c$  and  $M_N$  satisfy the assumptions of [theorem 1.4](#). On the same disorder, define*

$$Y_N(s) := \frac{1}{N} \log Z_{N, M_N}^{S, u_{s^2}, hs},$$

$$Y(s) := \inf_{q \in \mathcal{Q}} \left\{ \mathcal{A}_S(q) + \alpha \Phi_{u_{s^2}}(q) + \frac{h^2 s^2}{2} \mathcal{L}(q) \right\}.$$

Then

$$\frac{1}{N} \log \int_0^\infty s^{N-1} e^{N[Y_N(s) - cs^2]} ds \xrightarrow{L^1} \sup_{s>0} \{Y(s) + \log s - cs^2\}. \quad (110)$$

The supremum is finite.

*Proof. Shell convergence.* For each  $s > 0$ , the [Gaussian-field extension](#) and (104) give  $Y_N(s) \rightarrow Y(s)$  in  $L^2$ . Writing  $L = \text{Lip}(u)$ ,  $u_0 = u(0)$ , and  $\alpha_N = M_N/N$ , comparison of the Gibbs weights gives

$$|Y_N(s) - Y_N(t)| \leq D_N |s - t|, \quad D_N = L \sqrt{\alpha_N} \frac{\|\mathbf{G}_N\|_{\text{op}}}{\sqrt{N}} + |h| \frac{\|g^0\|_2}{\sqrt{N}},$$

$$|Y_N(s) - \alpha_N u_0| \leq s D_N, \quad s, t \geq 0. \quad (111)$$

Here  $Y_N(0) = \alpha_N u_0$ , and [theorem 3.5](#) and Gaussian norm moments give  $\sup_N \mathbb{E} D_N^p < \infty$  for every finite  $p$ . The expected first bound makes  $Y$  locally Lipschitz; finite nets then give

$$\sup_{s \in [a, A]} |Y_N(s) - Y(s)| \rightarrow 0 \quad \text{in probability and in } L^1, \quad 0 < a < A < \infty. \quad (112)$$

*Radial localization.* Putting  $\rho = s^2$  in (11) gives

$$\mathcal{Z}_{N, M_N}^{\text{ST}}(u, h, c) = 2C_N \int_0^\infty s^{N-1} e^{N[Y_N(s) - cs^2]} ds. \quad (113)$$

Choose  $\bar{\alpha} \geq \sup_N \alpha_N$ . On  $\{D_N \leq B\}$ , (111) implies, for  $0 < a < 1$ ,

$$\int_0^a s^{N-1} e^{N[Y_N(s) - cs^2]} ds \leq \frac{a^N}{N} e^{N(\bar{\alpha}|u_0| + Ba)}.$$

For  $A$  large enough that  $\log s + \bar{\alpha}|u_0| + Bs \leq cs^2/2$  on  $[A, \infty)$ ,

$$\int_A^\infty s^{N-1} e^{N[Y_N(s) - cs^2]} ds \leq \int_A^\infty s^{-1} e^{-Ncs^2/2} ds \leq \frac{e^{-NcA^2/2}}{NcA^2}.$$

These tails have arbitrarily negative exponential rates as  $a \downarrow 0$  and  $A \uparrow \infty$ . Apply the Laplace principle on  $[a, A]$  using (112), then remove the cutoffs and let  $B \uparrow \infty$ , using tightness of  $D_N$ . This proves (110) in probability. The limiting envelope  $|Y(s) - \alpha u_0| \leq Cs$  makes the supremum finite and sends its objective to  $-\infty$  at both radial ends; a fixed positive  $s$  supplies the lower bound.

*Uniform integrability.* The nonquadratic part differs from  $M_N u_0$  by at most  $\sqrt{N} D_N \|x\|_2$ . Completing the square and applying Jensen under  $N(0, (2c)^{-1} I_N)$  give, respectively,

$$\frac{1}{N} \log \mathcal{Z}_{N, M_N}^{\text{ST}}(u, h, c) \leq \frac{1}{2} \log \frac{2\pi}{c} + \alpha_N u_0 + \frac{D_N^2}{2c},$$

$$\frac{1}{N} \log \mathcal{Z}_{N, M_N}^{\text{ST}}(u, h, c) \geq \frac{1}{2} \log \frac{\pi}{c} + \alpha_N u_0 - \frac{D_N}{\sqrt{2c}}.$$

The moment bounds on  $D_N$  give uniform integrability of  $N^{-1} \log \mathcal{Z}_{N,M_N}^{\text{ST}}$ . By (113) and (12), the normalized log radial integral differs from it by the bounded deterministic sequence  $N^{-1} \log(2C_N)$ . It is therefore uniformly integrable as well, which upgrades the limit to  $L^1$ .  $\square$

**Constant-path restriction.** For  $q = \bar{q}_{x/\rho}$ ,  $0 \leq x < \rho$ , the objective in (10), apart from  $\frac{1}{2} \log(2\pi e)$ , becomes

$$\alpha \mathbb{E}_Z \log \mathbb{E}_\xi \exp\{u(\sqrt{x} Z + \sqrt{\rho - x} \xi)\} + \frac{1}{2} \left\{ \frac{x}{\rho - x} + \log(\rho - x) \right\} - c\rho + \frac{h^2}{2}(\rho - x).$$

where  $Z, \xi$  are independent standard Gaussians. For each fixed  $\rho$ , its infimum over  $0 \leq x < \rho$  is the constant-path inner problem; the replica-symmetric restriction of  $\mathcal{P}_{\text{ST}}$  is obtained after taking the outer supremum over  $\rho$ . Theorem 1.4 does not assert constant-path optimality or uniqueness of its stationary equations.

## 8.4 Quadratic penalties and hard constraints

This subsection proves two singular limits of the Lipschitz-channel formula. The first gives quadratic free energies and their zero-temperature ground states. The second hardens distance penalties into exact closed constraints.

### 8.4.1 Quadratic penalties and ground states

Fix  $\theta \in \mathbb{R}$  and  $\beta > 0$ , and put

$$\ell_\theta(x) = (\theta - x)_+, \quad u_{\beta,\theta}^{\text{har}}(x) = -\frac{\beta}{2} \ell_\theta(x)^2.$$

For  $R > 0$ , its capped Lipschitz approximation is

$$u_{\beta,\theta,R}(x) = -\frac{\beta}{2} (\ell_\theta(x) \wedge R)^2.$$

Define

$$\begin{aligned} \Phi_{\beta,\theta}^{\text{har}}(q) &= \lim_{R \rightarrow \infty} \Phi_{u_{\beta,\theta,R}}(q), \\ \mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha) &= \inf_{q \in \mathcal{Q}} \left\{ \mathcal{A}_\kappa(q) + \alpha \Phi_{\beta,\theta}^{\text{har}}(q) \right\}, \\ F_{N,\kappa}^{\text{har}}(\beta, \theta) &= \frac{1}{N} \log \int_{\Sigma_N^\kappa} \exp \left\{ -\frac{\beta}{2} \sum_{k=1}^{M_N} (\theta - h_k(\sigma))_+^2 \right\} \mu_N^\kappa(d\sigma), \\ e_{N,\kappa}(\theta) &= \frac{1}{2N} \min_{\sigma \in \Sigma_N^\kappa} \sum_{k=1}^{M_N} (\theta - h_k(\sigma))_+^2. \end{aligned}$$

**Theorem 8.5** (Quadratic free energy and ground state). *Let  $M_N/N \rightarrow \alpha \in (0, \infty)$ ,  $\theta \in \mathbb{R}$ , and  $\kappa \in \{\text{S}, \text{I}\}$ . For every  $\beta > 0$ ,*

$$F_{N,\kappa}^{\text{har}}(\beta, \theta) \longrightarrow \mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha) \quad \text{in } L^1, \quad (114)$$

and

$$\mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha) = \lim_{R \rightarrow \infty} \mathcal{P}_{\kappa,u_{\beta,\theta,R}}(\alpha). \quad (115)$$

There is a deterministic  $e_{\text{gs},\kappa}(\alpha, \theta) \geq 0$  such that

$$e_{N,\kappa}(\theta) \longrightarrow e_{\text{gs},\kappa}(\alpha, \theta) \quad \text{in } L^1, \quad (116)$$

and

$$e_{\text{gs},\kappa}(\alpha, \theta) = \lim_{\beta \rightarrow \infty} -\frac{1}{\beta} \mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha). \quad (117)$$

The proof first sends  $N \rightarrow \infty$  at fixed  $(\beta, R)$ , then  $R \rightarrow \infty$  at fixed  $\beta$ , and finally  $\beta \rightarrow \infty$ .

*Proof of Theorem 8.5. Finite temperature.* Put

$$d_R(x) = u_{\beta,\theta,R}(x) - u_{\beta,\theta}^{\text{har}}(x) = \frac{\beta}{2}(\ell_\theta(x)^2 - R^2)_+, \quad \varepsilon_R = \mathbb{E}d_R(Z) \longrightarrow 0, \quad Z \sim N(0, 1).$$

We first compare one row. Let  $\Gamma$  be a possibly random probability measure under which  $X$  has annealed standard-Gaussian marginal. Write  $\mathcal{N}_R = \Gamma(e^{u_{\beta,\theta,R}(X)})$ ,  $\mathcal{N}_\infty = \Gamma(e^{u_{\beta,\theta}^{\text{har}}(X)})$ , and  $\Gamma_R = \mathcal{N}_R^{-1} e^{u_{\beta,\theta,R}(X)} \Gamma$ . Jensen gives

$$0 \leq \log \mathcal{N}_R - \log \mathcal{N}_\infty = -\log \Gamma_R(e^{-d_R(X)}) \leq \Gamma_R(d_R(X)) \leq \Gamma(d_R(X)).$$

The last inequality follows because  $d_R$  increases and  $e^{u_{\beta,\theta,R}}$  decreases as functions of  $\ell_\theta^2$ . Consequently

$$0 \leq \mathbb{E}(\log \mathcal{N}_R - \log \mathcal{N}_\infty) \leq \varepsilon_R. \quad (118)$$

For the finite-volume capped pressure  $F_{N,\kappa}^R$ , replace the rows successively. The normalized measure with the current row deleted is independent of that row, whose marginal at each spin is standard normal. Thus (118) gives

$$0 \leq \mathbb{E}[F_{N,\kappa}^R(\beta, \theta) - F_{N,\kappa}^{\text{har}}(\beta, \theta)] \leq \frac{M_N}{N} \varepsilon_R, \quad (119)$$

with a pointwise nonnegative difference.

For a step path  $q$ , the completed field on its finite cascade has the same standard-Gaussian marginal under the averaged unweighted measure. Jensen therefore gives

$$-\frac{\beta}{2} \mathbb{E}(\theta - Z)_+^2 \leq \Phi_{u_{\beta,\theta,R}}(q) \leq 0.$$

At fixed  $\beta, R$ , the capped channel is Lipschitz, so (32) extends these bounds to every  $q \in \mathcal{Q}$ . They then persist in the decreasing limit  $\Phi_{\beta,\theta}^{\text{har}}(q)$ . Commuting  $\inf_R$  and  $\inf_q$  on  $\{q : \mathcal{A}_\kappa(q) < \infty\}$  therefore gives

$$\mathcal{P}_{\kappa,u_{\beta,\theta,R}}(\alpha) \downarrow \inf_{q: \mathcal{A}_\kappa(q) < \infty} \{\mathcal{A}_\kappa(q) + \alpha \Phi_{\beta,\theta}^{\text{har}}(q)\} = \mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha).$$

The values are finite by the Jensen bound,  $\mathcal{A}_\kappa \geq 0$ , and the trial path  $\bar{q}_0$ . This proves (115). At fixed  $R$ , theorem 1.1 gives convergence of  $F_{N,\kappa}^R$  to its variational value in  $L^2$ . The triangle inequality and (119), followed by  $R \rightarrow \infty$ , prove (114).

*Zero temperature.* Set  $X_{N,\kappa}(\beta) = -F_{N,\kappa}^{\text{har}}(\beta, \theta)/\beta$ . Since the prior is normalized, Hölder's inequality makes this nonincreasing in  $\beta$ , with  $X_{N,\kappa}(\beta) \geq e_{N,\kappa}(\theta) \geq 0$ . A minimizing Ising atom has mass  $2^{-N}$ , so

$$0 \leq X_{N,\text{I}}(\beta) - e_{N,\text{I}}(\theta) \leq \frac{\log 2}{\beta}.$$

For the spherical prior write  $x = \sigma/\sqrt{N}$ , put  $\theta_+ = \max\{\theta, 0\}$ , and set

$$\mathcal{L}_{N,\theta}(x) = \frac{1}{2N} \|(\theta \mathbf{1} - \mathbf{G}_N x)_+\|_2^2, \quad L_{N,\theta} = \frac{\|\mathbf{G}_N\|_{\text{op}}}{\sqrt{N}} \left( \theta_+ \sqrt{\frac{M_N}{N}} + \frac{\|\mathbf{G}_N\|_{\text{op}}}{\sqrt{N}} \right).$$

Contraction of the positive-part map and  $\|(\theta \mathbf{1} - \mathbf{G}_N x)_+\|_2 \leq \theta_+ \sqrt{M_N} + \|\mathbf{G}_N\|_{\text{op}}$  show that  $\mathcal{L}_{N,\theta}$  is  $L_{N,\theta}$ -Lipschitz on the unit sphere. For  $N \geq 2$  and  $0 < r \leq 1$ , its Euclidean radius- $r$  cap about a minimizer has normalized measure at least  $(r/8)^{N-1}$ , by integration in the tangent-plane chart of the hemisphere. Restricting the partition function to this cap yields

$$0 \leq X_{N,S}(\beta) - e_{N,S}(\theta) \leq L_{N,\theta} r + \frac{1}{\beta} \log \frac{8}{r}. \quad (120)$$

By [theorem 3.5](#),  $\sup_N \mathbb{E} L_{N,\theta} < \infty$ . Taking  $r = \beta^{-1}$  gives, for both priors and  $\beta \geq 1$ ,

$$\limsup_{N \rightarrow \infty} \mathbb{E} |X_{N,\kappa}(\beta) - e_{N,\kappa}(\theta)| \leq \frac{C_{\alpha,\theta} + \log(8\beta)}{\beta} \rightarrow 0. \quad (121)$$

The finite-temperature limit  $e_{\kappa}^{\text{soft}}(\beta) = -\mathcal{P}_{\kappa,\beta,\theta}^{\text{har}}(\alpha)/\beta$  inherits nonnegativity and monotonicity; denote its limit by  $e_{\text{gs},\kappa}(\alpha, \theta)$ . Using (114) in (121), first let  $N \rightarrow \infty$ , then  $\beta \rightarrow \infty$ . The triangle inequality proves (116) and (117).  $\square$

#### 8.4.2 Closed-set hardening

For the proof, abbreviate the soft pressure and its limiting value by

$$F_{N,\kappa}^{\beta,C} = \frac{1}{N} \log \int_{\Sigma_N^{\kappa}} \exp \left\{ -\beta \sum_{k=1}^{M_N} d_C(h_k(\sigma)) \right\} \mu_N^{\kappa}(\mathrm{d}\sigma),$$

$$\mathcal{P}_{\kappa,\beta,C}(\alpha) = \mathcal{P}_{\kappa,u_{\beta,C}}(\alpha).$$

We now prove [theorem 1.7](#). The single-constraint estimate in the proof is formulated for weighted measures because, after some constraints are hardened, the remaining measure is neither uniform nor a product measure.

**Proposition 8.6** (Sequential hardening preserves the limiting entropy). *Fix a tail-accessible closed set  $C$  and  $\kappa \in \{\text{S}, \text{I}\}$ , and suppose that, for some finite  $\bar{\alpha} > 0$ ,  $M_N \leq \bar{\alpha}N$  for all large  $N$ . Assume that, for some  $c_0 > 0$  with  $c_0 < \log 2$  when  $\kappa = \text{I}$  and for every fixed  $\beta > 0$ ,*

$$\mathbb{P}(F_{N,\kappa}^{\beta,C} < -c_0) \rightarrow 0. \quad (122)$$

Then, for every  $\varepsilon > 0$ ,

$$\lim_{\beta \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P} \left( F_{N,\kappa}^{\beta,C} - \text{Ent}_{N,\kappa}(C) > \varepsilon \right) = 0.$$

*Proof. Single-constraint claim.* Fix  $\rho, \zeta > 0$  and a tail-accessible closed set  $C$ , and choose  $\theta_C \in \mathbb{R}$  such that  $[\theta_C, \infty) \subset C$ . There are constants  $b_{\text{row}}, c, A, w_0 > 0$  and  $N_0 < \infty$ , depending only on  $(\rho, \zeta, C, \theta_C)$ , such that the following holds for every  $N \geq N_0$ . Let  $\nu$  be any deterministic probability measure on a subset  $\Omega_N \subset S_N$  satisfying

$$\sup_{\sigma_0 \in \Omega_N} \nu \left\{ \sigma : \|\sigma - \sigma_0\|_2 \leq \rho \sqrt{N} \right\} \leq e^{-\zeta N}. \quad (123)$$

For  $g \sim \mathbf{N}_N$ , set

$$\begin{aligned} X_g(\sigma) &= N^{-1/2} g \cdot \sigma, \\ p_{\nu,C}(g) &= \nu\{X_g \in C\}, \\ s_{\nu,\beta,C}(g) &= \int e^{-\beta d_C(X_g(\sigma))} \nu(d\sigma), \\ D_{\nu,\beta,C}(g) &= \log \frac{s_{\nu,\beta,C}(g)}{p_{\nu,C}(g)}, \end{aligned}$$

We suppress the common argument  $g$  in the remainder of this claim and its proof. Set  $D_{\nu,\beta,C} = +\infty$  when  $p_{\nu,C} = 0$ . Uniformly for  $w_0 \leq w \leq b_{\text{row}}N$ ,

$$\mathbb{P}_g(p_{\nu,C}(g) \leq e^{-w}) \leq Ae^{-cw} + Ae^{-cN}. \quad (124)$$

Moreover, there is a deterministic function  $\omega_{\rho,\zeta,C}(\beta) \downarrow 0$  as  $\beta \rightarrow \infty$  such that, for each fixed  $\beta$  and all sufficiently large  $N$ ,

$$\mathbb{E}_g(D_{\nu,\beta,C} \wedge b_{\text{row}}N) \leq \omega_{\rho,\zeta,C}(\beta) + ANe^{-cN}. \quad (125)$$

All constants are uniform over  $\nu$  satisfying (123). The same conclusions hold conditionally if  $\nu$  is random, independent of  $g$ , and (123) holds almost surely.

*Proof of the claim.* Decrease  $b_{\text{row}}$  so that  $2b_{\text{row}} < \zeta$ . Given  $w \in [w_0, b_{\text{row}}N]$ , let  $K = \lfloor e^w/4 \rfloor$  and sample  $\sigma^1, \dots, \sigma^K$  independently from  $\nu$ . A union bound and (123) show that the  $\nu^{\otimes K}$ -probability of a pair within distance  $\rho\sqrt{N}$  is at most  $K^2 e^{-\zeta N} \leq e^{-cN}$ . Conditional on pairwise separation, the centered Gaussian variables  $X_g(\sigma^i)$  have variance one and canonical distances at least  $\rho$ . Sudakov minoration gives  $\mathbb{E}_g \max_{i \leq K} X_g(\sigma^i) \geq c_\rho \sqrt{\log K}$ , and the Gaussian concentration inequality for the maximum then gives, after increasing  $w_0$ ,

$$\mathbb{P}_g \left( \max_{i \leq K} X_g(\sigma^i) < \theta_C \mid \sigma^1, \dots, \sigma^K \right) \leq e^{-cw} \quad (126)$$

for every pairwise separated realization  $(\sigma^1, \dots, \sigma^K)$ . On the other hand, on  $\{p_{\nu,C}(g) \leq e^{-w}\}$ , conditional independence of the samples gives

$$\mathbb{P} \left( X_g(\sigma^i) \notin C \text{ for every } i \leq K \mid g \right) = (1 - p_{\nu,C}(g))^K \geq \frac{3}{4}$$

once  $w_0$  is large. Averaging first in the samples and then in  $g$ , and noting that  $x \notin C$  implies  $x < \theta_C$ , combining this inequality with (126) proves (124).

Write  $\Delta_{\nu,\beta,C}(g) = s_{\nu,\beta,C}(g) - p_{\nu,C}(g)$  and  $Y_{\nu,C}(g) = -\log p_{\nu,C}(g)$ , with  $Y_{\nu,C} = +\infty$  when  $p_{\nu,C} = 0$ . Since  $X_g(\sigma)$  is standard Gaussian for every  $\sigma \in S_N$ , Fubini gives the uniform estimate

$$\mathbb{E}_g \Delta_{\nu,\beta,C} = \varepsilon_C(\beta) := \mathbb{E} \left[ e^{-\beta d_C(Z)} \mathbf{1}_{\{Z \notin C\}} \right] \longrightarrow 0. \quad (127)$$

Here dominated convergence applies because closedness gives  $d_C(Z) > 0$  on  $\{Z \notin C\}$ . For  $w \in [w_0, b_{\text{row}}N]$ ,

$$D_{\nu,\beta,C} \wedge b_{\text{row}}N \leq e^w \Delta_{\nu,\beta,C} + (Y_{\nu,C} \wedge b_{\text{row}}N) \mathbf{1}_{\{Y_{\nu,C} > w\}}. \quad (128)$$

Indeed, on  $\{Y_{\nu,C} \leq w\}$  use  $\log(1 + \Delta_{\nu,\beta,C}/p_{\nu,C}) \leq \Delta_{\nu,\beta,C}/p_{\nu,C} \leq e^w \Delta_{\nu,\beta,C}$ , while on  $\{Y_{\nu,C} > w\}$  use  $s_{\nu,\beta,C} \leq 1$ , hence  $D_{\nu,\beta,C} \leq Y_{\nu,C}$ . Integrating (124) yields

$$\mathbb{E}_g[(Y_{\nu,C} \wedge b_{\text{row}}N) \mathbf{1}_{\{Y_{\nu,C} > w\}}] \leq A(1+w)e^{-cw} + ANe^{-cN}.$$

For large  $\beta$ , take

$$w_\beta = w_0 \vee \frac{1}{2} \log \frac{1}{\varepsilon_C(\beta)};$$

for fixed  $\beta$  this lies below  $b_{\text{row}}N$  when  $N$  is large. Then  $e^{w_\beta \varepsilon_C(\beta)} \rightarrow 0$  and  $(1 + w_\beta)e^{-cw_\beta} \rightarrow 0$ . Equations (127) and (128) prove (125) with some  $\omega_{\rho, \zeta, C}(\beta) \downarrow 0$ .

*Sequential hardening.* Choose  $c_1 > c_0$ , with  $c_1 < \log 2$  when  $\kappa = \text{I}$ . We first verify the spread hypothesis in the single-constraint claim. If a probability measure  $\nu$  on  $S_N$  satisfies  $d\nu/d\mu_N^S \leq e^{c_1 N}$ , then a spherical cap estimate gives, for fixed  $0 < \rho < 1$ ,

$$\sup_{\sigma_0 \in S_N} \nu\{\|\sigma - \sigma_0\|_2 \leq \rho\sqrt{N}\} \leq e^{c_1 N} C_\rho \sqrt{N} \rho^{N-3}.$$

In the spherical case, choose  $\rho$  sufficiently small and then  $\zeta > 0$  so that this bound is at most  $e^{-\zeta N}$  for all large  $N$ . On the cube, if  $\max_{\sigma \in \Sigma_N^1} \nu(\sigma) \leq e^{-sN}$  for some  $s > 0$ , a ball of radius  $\rho\sqrt{N}$  contains at most

$$\exp\{N\mathbf{h}(\rho^2/4) + o(N)\}$$

vertices, where  $\mathbf{h}(x) = -x \log x - (1-x) \log(1-x)$  is binary entropy. Choosing  $\rho$  so that  $\mathbf{h}(\rho^2/4) < s/2$  and taking any  $0 < \zeta < s/2$  gives (123) for all large  $N$ . In the application below,  $s = \log 2 - c_1 > 0$  in the Ising case. Fix the resulting pair  $(\rho, \zeta)$  for the chosen prior.

Write

$$U_k(\sigma) = e^{-\beta d_C(h_k(\sigma))}, \quad I_k(\sigma) = \mathbf{1}_{\{h_k(\sigma) \in C\}}.$$

For  $0 \leq m \leq M_N$ , introduce the soft-hard hybrids

$$Z_m = \int \left( \prod_{k \leq m} I_k(\sigma) \right) \left( \prod_{m < k \leq M_N} U_k(\sigma) \right) \mu_N^\kappa(d\sigma).$$

Thus  $N^{-1} \log Z_0 = F_{N, \kappa}^{\beta, C}$  and  $N^{-1} \log Z_{M_N} = \text{Ent}_{N, \kappa}(C)$ . Delete row  $m$  and let

$$\begin{aligned} \hat{Z}_m &= \int \left( \prod_{k < m} I_k \right) \left( \prod_{m < k \leq M_N} U_k \right) d\mu_N^\kappa, \\ \nu_m(d\sigma) &= \hat{Z}_m^{-1} \left( \prod_{k < m} I_k(\sigma) \right) \left( \prod_{m < k \leq M_N} U_k(\sigma) \right) \mu_N^\kappa(d\sigma) \end{aligned}$$

when  $\hat{Z}_m > 0$ ; otherwise set  $\nu_m = \mu_N^\kappa$ . Conditional on all rows except  $g_m$ ,  $\nu_m$  is independent of  $g_m$ . When  $\hat{Z}_m > 0$ ,

$$Z_{m-1} = \hat{Z}_m s_{\nu_m, \beta, C}(g_m), \quad Z_m = \hat{Z}_m p_{\nu_m, C}(g_m).$$

Let  $D_m = D_{\nu_m, \beta, C}(g_m)$ . Whenever  $Z_{m-1} > 0$ ,  $D_m = \log Z_{m-1} - \log Z_m \in [0, \infty]$ .

On  $\{\hat{Z}_m \geq e^{-c_1 N}\}$  the unnormalized density defining  $\nu_m$  is at most one. Hence

$$\begin{aligned} \frac{d\nu_m}{d\mu_N^S} &\leq e^{c_1 N} & (\kappa = \text{S}), \\ \max_\sigma \nu_m(\sigma) &\leq 2^{-N} e^{c_1 N} = e^{-(\log 2 - c_1)N} & (\kappa = \text{I}). \end{aligned}$$

The preceding spread estimates and the single-constraint claim, conditionally on all constraints except  $g_m$ , give constants  $b_{\text{row}}, c, A > 0$  and  $\omega_{\rho, \zeta, C}(\beta) \downarrow 0$ , uniform in  $m$ , such that

$$\mathbb{E} \left[ (D_m \wedge b_{\text{row}}N) \mathbf{1}_{\{\hat{Z}_m \geq e^{-c_1 N}\}} \right] \leq \omega_{\rho, \zeta, C}(\beta) + ANe^{-cN}. \quad (129)$$

To telescope even when a hard constraint makes a partition function zero, truncate logarithms at a common floor. For  $z \geq 0$ , put  $\ell_N(z) = \log(z \vee e^{-c_1 N})$  and  $L_m = \ell_N(Z_{m-1}) - \ell_N(Z_m) \geq 0$ . Both hybrids lie below the floor when  $\hat{Z}_m < e^{-c_1 N}$ ; otherwise  $L_m \leq D_m \wedge c_1 N$ , since  $Z_{m-1} \leq 1$ . Thus

$$L_m \leq C_1(D_m \wedge b_{\text{row}} N) \mathbf{1}_{\{\hat{Z}_m \geq e^{-c_1 N}\}}, \quad C_1 = \max\{1, c_1/b_{\text{row}}\}.$$

The event on the right is measurable without row  $m$ . Sum (129) and telescope to obtain

$$\frac{1}{N} \mathbb{E}[\ell_N(Z_0) - \ell_N(Z_{M_N})] \leq C_1 \bar{\alpha} \omega_{\rho, \zeta, C}(\beta) + o_N(1).$$

Fix  $0 < \eta < \min\{\varepsilon, c_1 - c_0\}$ . If  $Z_0 \geq e^{-c_0 N}$  and the floored loss is at most  $\eta N$ , then  $\ell_N(Z_{M_N}) \geq -(c_0 + \eta)N > -c_1 N$ ; hence both endpoint logs are unclipped and their difference is at most  $\eta N$ . Markov's inequality gives

$$\mathbb{P}\left(F_{N, \kappa}^{\beta, C} - \text{Ent}_{N, \kappa}(C) > \varepsilon\right) \leq \mathbb{P}(Z_0 < e^{-c_0 N}) + \frac{C_1 \bar{\alpha} \omega_{\rho, \zeta, C}(\beta) + o_N(1)}{\eta}.$$

Apply (122), first letting  $N \rightarrow \infty$  at fixed  $\beta$ , then  $\beta \rightarrow \infty$ .  $\square$

*Proof of theorem 1.7.* At fixed  $q$ , monotonicity of the scalar recursion gives  $\Phi_{u_{\beta, C}}(q) \downarrow \Phi_{\infty, C}(q)$ . Hence the soft variational values decrease to a limit no smaller than the hard value. Conversely, for every  $q$  with  $\mathcal{A}_{\kappa}(q) < \infty$ ,

$$\mathcal{P}_{\kappa, \beta, C}(\alpha) \leq \mathcal{A}_{\kappa}(q) + \alpha \Phi_{u_{\beta, C}}(q).$$

Letting  $\beta \rightarrow \infty$  and then taking the infimum proves (16), including the value  $-\infty$ .

Since a hard indicator is bounded above by its soft distance weight,

$$\text{Ent}_{N, \kappa}(C) \leq F_{N, \kappa}^{\beta, C}. \quad (130)$$

Write  $P_{\infty} = \mathcal{P}_{\kappa, \infty, C}(\alpha)$  and  $P_{\beta} = \mathcal{P}_{\kappa, \beta, C}(\alpha)$ . Suppose first that  $P_{\infty} > -\infty$  for spherical spins, or that  $P_{\infty} > -\log 2$  for Ising spins. Choose  $c_0 > -P_{\infty}$ , with  $c_0 < \log 2$  in the Ising case. Because  $P_{\beta} \geq P_{\infty}$ , the main free energy theorem at the fixed finite Lipschitz channel  $u_{\beta, C}$  implies (122). Apply theorem 8.6. Choose first  $\beta$  with  $P_{\beta} - P_{\infty}$  small and then send  $N \rightarrow \infty$ . The hybrid controls the lower tail of  $\text{Ent}_{N, \kappa}(C)$ , while (130) controls the upper tail. Thus  $\text{Ent}_{N, \kappa}(C) \rightarrow P_{\infty}$  in probability, proving the finite spherical case and the Ising entropy assertions.

If  $P_{\infty} = -\infty$  for spherical spins, fix  $L > 0$  and choose a finite  $\beta$  with  $P_{\beta} < -L - 1$ . The fixed- $\beta$  soft formula and (130) give  $\mathbb{P}(\text{Ent}_{N, \kappa}(C) > -L) \rightarrow 0$ ; then let  $L \rightarrow \infty$ .

Finally suppose  $P_{\infty} < -\log 2$  for Ising spins. Choose  $\eta > 0$  and then a finite  $\beta$  so that  $P_{\beta} < -\log 2 - 3\eta$ . For all large  $N$ ,

$$\mathbb{E} F_{N, \text{I}}^{\beta, C} < -\log 2 - 2\eta.$$

On the satisfiable event, one configuration has zero penalty and hence  $F_{N, \text{I}}^{\beta, C} \geq -\log 2$ . Since  $\text{Lip}(u_{\beta, C}) = \beta$ , Gaussian concentration with row-matrix Lipschitz constant  $\beta \sqrt{M_N}/N$  gives

$$\mathbb{P}(\mathcal{F}_{M_N, N}^{\text{I}}(C) \neq \emptyset) \leq 2 \exp \left\{ -\frac{\eta^2 N^2}{2 M_N \beta^2} \right\}.$$

As  $M_N/N \rightarrow \alpha$ , this proves (18).  $\square$

## 8.5 Capacity consequences

This subsection proves the general Ising capacity theorem, the spherical half-space threshold, and the explicit spherical trial bound summarized in [section 1](#).

### 8.5.1 Ising capacities for tail-accessible sets

*Proof of [theorem 1.8](#).* Fix a tail-accessible closed set  $C$  and  $\beta > 0$ . For a step path  $q$ , choose a finite grid  $\pi$  containing its jump points. Jensen's inequality on the finite cascade gives

$$\Phi_{u_{\beta,C}}(q) \leq \log \mathbb{E} \int e^{-\beta d_C(X^q(\eta,z))} \mathfrak{N}_{\pi}(d\eta) \mathbf{N}_1(dz) = \log \mathbb{E} e^{-\beta d_C(Z)}.$$

The equality uses the standard-Gaussian marginal of the completed field under the averaged unweighted measure. At this fixed  $\beta$ ,  $\text{Lip}(u_{\beta,C}) = \beta$ , so [\(32\)](#) extends the inequality to every  $q \in \mathcal{Q}$  by step-path approximation. Only then let  $\beta \rightarrow \infty$ ; dominated convergence yields

$$\Phi_{\infty,C}(q) \leq \log \pi_C < 0. \quad (131)$$

Put  $B_C(q) = -\Phi_{\infty,C}(q)$ ,  $r_C = \alpha_c^I(C)$ , and

$$s_C(\alpha) = \log 2 + \mathcal{P}_{I,\infty,C}(\alpha) = \inf_{q \in \mathcal{Q}: \mathcal{A}_I(q) < \infty} \{\log 2 + \mathcal{A}_I(q) - \alpha B_C(q)\}, \quad \alpha > 0.$$

If  $0 < \alpha < r_C$ , the ratio definition implies that  $B_C(q)$  is finite for every finite-cost path, and

$$\log 2 + \mathcal{A}_I(q) - \alpha B_C(q) \geq (r_C - \alpha) B_C(q) \geq (-\log \pi_C)(r_C - \alpha) > 0.$$

The last bound is uniform in  $q$ , so  $s_C(\alpha) > 0$ . If  $\alpha > r_C$ , some trial ratio is smaller than  $\alpha$ , giving  $s_C(\alpha) < 0$ ; if its denominator is infinite, the trial value is  $-\infty$ . Thus

$$s_C(\alpha) > 0 \quad (0 < \alpha < r_C), \quad s_C(\alpha) < 0 \quad (\alpha > r_C).$$

Together with  $\mathcal{P}_{I,\infty,C}(0) = 0$ , these inequalities prove [\(20\)](#), including the possibility  $r_C = 0$  at this stage. The path  $\bar{q}_0$  has  $\mathcal{A}_I(\bar{q}_0) = 0$ . Its scalar recursion has no common Gaussian increment, and hence

$$\Phi_{u_{\beta,C}}(\bar{q}_0) = \log \mathbb{E} e^{-\beta d_C(Z)} \rightarrow \log \pi_C.$$

Therefore  $B_C(\bar{q}_0) = -\log \pi_C$ , giving the upper bound in [\(19\)](#).

Apply [theorem 1.7](#) to any sequence  $M_N/N \rightarrow \alpha > 0$ . For  $\alpha < r_C$ , it gives  $N^{-1} \log \# \mathcal{F}_{M_N,N}^I(C) \rightarrow s_C(\alpha) > 0$  in probability, hence satisfiability with high probability. For  $\alpha > r_C$ , it gives exponential decay of the satisfiability probability. This proves [\(21\)](#).

For strict positivity, choose  $\theta_C$  with  $[\theta_C, \infty) \subseteq C$ . The small-density theorem of Bolthausen–Nakajima–Sun–Xu applies to the half-space indicator  $U(x) = \mathbf{1}_{\{x \geq \theta_C\}}$  and gives satisfiability with high probability at some density  $\alpha_0 > 0$  [[7](#), Theorem 1.1 and Proposition 1.3]. Every such solution also satisfies the constraints in  $C$ . The supercritical conclusion just proved therefore forces  $r_C \geq \alpha_0 > 0$ , completing [\(19\)](#).

For fixed  $N$ , each of the  $2^N$  configurations survives the first  $m$  independent rows with probability  $\pi_C^m$ , and hence

$$\mathbb{P}(\mathfrak{M}_N^I(C) \geq m) \leq 2^N \pi_C^m.$$

Letting  $m \rightarrow \infty$  shows  $\mathfrak{M}_N^I(C) < \infty$  almost surely. For fixed  $0 < a_- < r_C < a_+$ , row-count monotonicity and the fixed-load limits give

$$\mathbb{P}(\mathfrak{M}_N^I(C) \geq \lfloor a_- N \rfloor) \rightarrow 1, \quad \mathbb{P}(\mathfrak{M}_N^I(C) \geq \lceil a_+ N \rceil) \rightarrow 0.$$

Let  $a_- \uparrow r_C$  and  $a_+ \downarrow r_C$ . The integer roundings contribute at most  $N^{-1}$ , proving [\(22\)](#).  $\square$

The same results apply to  $C = [-K, 0] \cup [K, \infty)$ ,  $K > 0$ . Here  $\pi_C = 1/2$ , so  $\alpha_c^I(C) \leq 1$ ; the full-path formula does not by itself establish the optimized value one predicted in [5]. For the square-wave set

$$C = \bigcup_{j \in \mathbb{Z}} [(2j+1)\delta, (2j+2)\delta], \quad \delta > 0,$$

[theorem 1.1](#) gives the soft formula for  $-\beta \text{dist}(\cdot, C)$ , but the hardening theorem does not apply because  $C$  contains no half-line. The present results also do not cover the correlated-label teacher–student model of [4].

### 8.5.2 Spherical half-spaces

In this subsection we abbreviate the half-space quantities by

$$\begin{aligned} u_{\beta, \theta} &= u_{\beta, C_\theta}, & \Phi_{\infty, \theta} &= \Phi_{\infty, C_\theta}, & \mathcal{P}_{\kappa, \infty, \theta} &= \mathcal{P}_{\kappa, \infty, C_\theta}, \\ \alpha_c^I(\theta) &= \alpha_c^I(C_\theta), & \mathcal{F}_{M, N}^\kappa(\theta) &= \mathcal{F}_{M, N}^\kappa(C_\theta), & \mathfrak{M}_N^\kappa(\theta) &= \mathfrak{M}_N^\kappa(C_\theta). \end{aligned}$$

Write  $\varphi_G(x) = (2\pi)^{-1/2} e^{-x^2/2}$  and  $\bar{\Phi}_G(x) = \mathbb{P}(Z \geq x)$ , with  $Z \sim N(0, 1)$ .

$$\mathcal{B}_\theta(q) = -\Phi_{\infty, \theta}(q),$$

$$D_\theta = \mathbb{E}(\theta - Z)_+^2.$$

The hard-volume formula does not itself imply emptiness on the sphere. The next three results supply the threshold geometry, margin transfer, and robust thickening needed to prove [theorem 1.9](#).

**Lemma 8.7** (Constant-path bound and the spherical finiteness threshold). *Fix  $\theta \in \mathbb{R}$ . For  $0 \leq x < 1$ , the constant path satisfies*

$$\mathcal{A}_S(\bar{q}_x) = \frac{1}{2} \left( \frac{x}{1-x} + \log(1-x) \right), \quad (132)$$

$$\Phi_{\infty, \theta}(\bar{q}_x) = \mathbb{E} \log \bar{\Phi}_G \left( \frac{\theta - \sqrt{x}Z}{\sqrt{1-x}} \right). \quad (133)$$

*The set of  $\alpha$  for which  $\mathcal{P}_{S, \infty, \theta}(\alpha) > -\infty$  is downward closed. Hence the spherical hard value is finite for  $\alpha < \alpha_c^S(\theta)$  and equals  $-\infty$  for  $\alpha > \alpha_c^S(\theta)$ . Moreover,*

$$\alpha_c^S(\theta) \leq D_\theta^{-1} < \infty.$$

*Proof.* The spherical identity in (132) follows from (76). For the channel, conditional on the common field  $\sqrt{x}Z$ , the soft normalizer decreases to

$$\bar{\Phi}_G \left( \frac{\theta - \sqrt{x}Z}{\sqrt{1-x}} \right).$$

Its logarithm has integrable negative part by Mills' inequalities, which proves (133). The same inequalities and dominated convergence give, as  $x \uparrow 1$ ,

$$(1-x)\mathcal{A}_S(\bar{q}_x) \longrightarrow \frac{1}{2}, \quad (1-x)\mathcal{B}_\theta(\bar{q}_x) \longrightarrow \frac{1}{2}D_\theta.$$

Thus constant paths make the spherical objective tend to  $-\infty$  when  $\alpha > D_\theta^{-1}$ .

Since  $\Phi_{\infty, \theta}(q) \leq 0$ , the spherical hard value is nonincreasing in  $\alpha$ . Finiteness at  $\alpha_2$  therefore gives a finite lower bound at every  $0 < \alpha_1 < \alpha_2$ , while the trial path  $\bar{q}_0$  gives a finite upper bound. This proves downward closure.  $\square$

**Proposition 8.8** (Supercriticality persists after lowering the margin). *For every  $\theta \in \mathbb{R}$  and every  $\alpha > \alpha_c^S(\theta)$ , there exists  $\delta > 0$  such that*

$$\mathcal{P}_{S,\infty,\theta-\delta}(\alpha) = -\infty. \quad (134)$$

*Proof.* We bound the cost of shifting the margin by the spherical spin cost. First, every path of finite spherical spin cost satisfies

$$\begin{aligned} \mathcal{A}_S(q) &\geq \frac{1}{2} \left( \frac{1}{L(q)} - 1 + \log L(q) \right), \\ \frac{1}{L(q)} &\leq 4(1 + \mathcal{A}_S(q)). \end{aligned} \quad (135)$$

In particular,  $L(q) > 0$ . To prove these, first suppose  $q(1-) < 1$ , put  $\ell = L(q) = \int_0^1 \gamma_q$ , and write  $D(t) = \int_t^1 \gamma_q(s) ds$ . Since  $D(t) \leq \min\{\ell, 1-t\}$ , (75) gives

$$2\mathcal{A}_S(q) = \int_0^1 \left\{ \frac{1}{D(t)} - \frac{1}{1-t} \right\} dt \geq \int_0^{1-\ell} \left\{ \frac{1}{\ell} - \frac{1}{1-t} \right\} dt = \ell^{-1} - 1 + \log \ell.$$

For general  $q$ , cap it at  $1-\eta$  on  $[0, 1)$ , retaining the diagonal value  $q_\eta(1) = 1$ . The conjugate formula (83) gives  $\mathcal{A}_S(q_\eta) \leq \mathcal{A}_S(q)$ , while  $L(q_\eta) \rightarrow L(q)$ . The preceding bound rules out  $L(q) = 0$  and passes to the limit. Applying  $x \leq 4 + 2(x-1 - \log x)$ ,  $x \geq 1$ , proves (135).

The second estimate, for every  $\delta, \varepsilon > 0$ , is

$$\mathcal{B}_{\theta+\delta}(q) \leq (1 + \varepsilon)\mathcal{B}_\theta(q) + \left(1 + \frac{1}{\varepsilon}\right) \frac{\delta^2}{2L(q)}. \quad (136)$$

The scalar recursion and [theorem 8.1](#), as in the finite-temperature passage above, identify  $\Phi_{u_{\beta,\vartheta}}(q) = \mathcal{T}_{\gamma_q} u_{\beta,\vartheta}(0, 0)$  directly for the Lipschitz terminal  $u_{\beta,\vartheta}(x) = -\beta(\vartheta - x)_+$ . Put  $\ell = L(q)$ . For any admissible control  $v$ , set  $X^v = B_1 + \int_0^1 \gamma_q v$  and  $w = v + \delta/\ell$ . The shifted control has finite weighted quadratic cost and is therefore admissible; the optional clipping bound need not be imposed. For  $p > 1$ ,

$$u_{\beta,\theta+\delta}(X^w) = u_{\beta,\theta}(X^v) \geq p u_{\beta,\theta}(X^v), \quad \int_0^1 \gamma_q w^2 \leq p \int_0^1 \gamma_q v^2 + \frac{p\delta^2}{(p-1)\ell}.$$

The first inequality uses nonpositivity of the terminal and the second is Young's inequality. Taking control suprema, then letting  $\beta \rightarrow \infty$  and setting  $p = 1 + \varepsilon$ , proves (136).

Choose  $\alpha_c^S(\theta) < \alpha_0 < \alpha$ , and then  $\delta > 0$  so small that

$$\alpha_0 \left\{ \frac{1+\delta}{\alpha} + 2\delta(1+\delta) \right\} < 1.$$

If the hard value at  $(\alpha, \theta - \delta)$  were finite, some  $C_0 \geq 0$  would give, for every finite-cost path,

$$\mathcal{B}_{\theta-\delta}(q) \leq \frac{\mathcal{A}_S(q) + C_0}{\alpha}.$$

Apply (136) from  $\theta - \delta$  to  $\theta$ , with  $\varepsilon = \delta$ , and use (135). The result is

$$\mathcal{B}_\theta(q) \leq \left\{ \frac{1+\delta}{\alpha} + 2\delta(1+\delta) \right\} \mathcal{A}_S(q) + C_\delta.$$

Thus  $\mathcal{A}_S(q) - \alpha_0 \mathcal{B}_\theta(q)$  is uniformly bounded below, contradicting [theorem 8.7](#). This proves (134).  $\square$

**Proposition 8.9** (Feasibility at margin  $\theta$  gives exponential volume at margin  $\theta - \delta$ ). *Suppose  $M_N/N \rightarrow \alpha \in (0, \infty)$ . For every  $\theta \in \mathbb{R}$  and  $\delta > 0$ , there is a finite  $C = C(\alpha, \theta, \delta)$  such that*

$$\mathbb{P}\left(\mathcal{F}_{M_N, N}^S(\theta) \neq \emptyset, \quad \frac{1}{N} \log \mu_N^S(\mathcal{F}_{M_N, N}^S(\theta - \delta)) < -C\right) \rightarrow 0. \quad (137)$$

*Proof.* Fix  $\bar{\alpha} > \alpha$ . Gaussian norm tails show that, with probability tending to one,  $M_N \leq \bar{\alpha}N$  and  $\max_{k \leq M_N} \|g_k\| \leq 2\sqrt{N}$ . The remaining argument is deterministic in any rows satisfying these bounds. Identify  $S_N$  with the unit sphere by  $s = \sigma/\sqrt{N}$ , so constraints become  $g_k \cdot s \geq \theta$  and  $\mu_N^S$  remains normalized surface measure. For an independent  $W \sim N(0, I_N/N)$ , Šidák's inequality [33, Corollary 1, p. 628] gives

$$\mathbb{P}_W\left(\max_k |g_k \cdot W| \leq L\right) \geq p_{L/2}^{M_N} \geq e^{-\kappa_L N}, \quad p_x = \mathbb{P}(|Z| \leq x), \quad \kappa_L = -\bar{\alpha} \log p_{L/2}.$$

Here each  $g_k \cdot W$  has variance at most four; arbitrary correlations are allowed. Also  $\mathbb{P}_W(\|W\| > 2) \leq e^{-c_0 N}$  for a universal  $c_0 > 0$ . Fix  $L$  large enough that  $\kappa_L < c_0/2$ . For all large  $N$ ,

$$E = \{\max_k |g_k \cdot W| \leq L, \quad \|W\| \leq 2\}, \quad \mathbb{P}_W(E) \geq \frac{1}{2} e^{-\kappa_L N}.$$

Take any feasible unit vector  $s$  and set

$$r = \min\left\{\frac{1}{4}, \frac{\delta}{2(L + 2|\theta| + 1)}\right\}, \quad Y = s + rW.$$

On  $E$ ,  $1 - 2r \leq \|Y\| \leq 1 + 2r$ , and for every row,

$$g_k \cdot \frac{Y}{\|Y\|} - \theta \geq \frac{\theta(1 - \|Y\|) - rL}{\|Y\|} \geq -\frac{r(L + 2|\theta|)}{1 - 2r} \geq -\delta.$$

Let  $A$  be the cone over the feasible set at margin  $\theta - \delta$  and write  $P = N(0, r^2 I_N/N)$ ,  $Q_s = N(s, r^2 I_N/N)$ . Then  $Q_s(A) \geq \frac{1}{2} e^{-\kappa_L N}$ , while isotropy identifies  $P(A)$  with the normalized spherical volume. The Gaussian likelihood ratio satisfies  $\mathbb{E}_P[(dQ_s/dP)^2] = e^{N/r^2}$ , so Cauchy–Schwarz gives

$$\mu_N^S(\mathcal{F}_{M_N, N}^S(\theta - \delta)) = P(A) \geq e^{-N/r^2} Q_s(A)^2 \geq \frac{1}{4} e^{-N(r^{-2} + 2\kappa_L)}.$$

This holds for every feasible  $s$  with the same constants, proving (137).  $\square$

*Proof of theorem 1.9.* Let  $M_N = \lfloor \alpha N \rfloor$ . By theorem 8.7, the spherical variational hard value is finite for  $\alpha < \alpha_c^S(\theta)$ . The hardening theorem then gives convergence of the log-volume to this finite value, so the feasible set, on whose complement the log-volume is  $-\infty$ , is nonempty with probability tending to one. If  $\alpha > \alpha_c^S(\theta)$ , choose  $\delta > 0$  from theorem 8.8. The hard-wall theorem at margin  $\theta - \delta$  and theorem 8.9 imply  $\mathbb{P}(\mathcal{F}_{M_N, N}^S(\theta) \neq \emptyset) \rightarrow 0$ .

For fixed  $N$ , cover the unit sphere by finitely many Euclidean caps  $\{s : \|s - v_j\|_2 < r\}$  with unit centers  $v_j$  and fixed  $0 < r < 1$ . Choose  $K$  so large that  $-K(1 - r^2/2) + 1 < \theta$ . If  $\|g + Kv_j\|_2 < 1$ , then throughout the  $j$ th cap

$$g \cdot s \leq -Kv_j \cdot s + 1 \leq -K(1 - r^2/2) + 1 < \theta.$$

Each such open Gaussian ball has positive probability, so an infinite i.i.d. row sequence almost surely visits all finitely many of them. Once it has done so, every point of the sphere violates at least one of the rows already seen. Thus  $\mathfrak{M}_N^S(\theta) < \infty$  almost surely. Finally, row-count monotonicity sandwiches the lower tail of  $\mathfrak{M}_N^S(\theta)/N$  by failure of satisfiability at  $\alpha_c^S(\theta) - \varepsilon/2$  and its upper tail by satisfiability at  $\alpha_c^S(\theta) + \varepsilon/2$ ; the first event is omitted when the lower load is nonpositive. The fixed-load statements make both probabilities vanish, proving (23).  $\square$

## References

- [1] A. Adhikari, *Spin distributions for generic spherical spin glasses*, Electron. J. Probab. 27 (2022); doi:10.1214/22-EJP755; arXiv:1907.03847.
- [2] M. Aizenman, R. Sims, and S. L. Starr, *Extended variational principle for the Sherrington–Kirkpatrick spin-glass model*, Phys. Rev. B 68 (2003), no. 21, 214403; doi:10.1103/PhysRevB.68.214403.
- [3] B. Aubin, W. Perkins, and L. Zdeborová, *Storage capacity in symmetric binary perceptrons*, J. Phys. A: Math. Theor. 52 (2019), no. 29, 294003; doi:10.1088/1751-8121/ab227a; arXiv:1901.00314.
- [4] M. Benedetti, A. Bogdanov, E. M. Malatesta, M. Mézard, G. Perrupato, A. Rosen, N. I. Schwartzbach, and R. Zecchina, *Overlap gap and computational thresholds in the square wave perceptron*, J. Stat. Mech. Theory Exp. (2025), no. 12, 123303; doi:10.1088/1742-5468/ae23be; arXiv:2506.05197.
- [5] G. J. Bex, R. Serneels, and C. Van den Broeck, *Storage capacity and generalization error for the reversed-wedge Ising perceptron*, Phys. Rev. E 51 (1995), no. 6, 6309–6312; doi:10.1103/PhysRevE.51.6309.
- [6] E. Bolthausen and N. Kistler, *A quenched large deviation principle and a Parisi formula for a perceptron version of the GREM*, in *Probability in Complex Physical Systems*, Springer Proceedings in Mathematics, vol. 11, Springer, Berlin, Heidelberg, 2012, pp. 425–442; doi:10.1007/978-3-642-23811-6\_17; arXiv:1011.5686.
- [7] E. Bolthausen, S. Nakajima, N. Sun, and C. Xu, *Gardner formula for Ising perceptron models at small densities*, Proceedings of Machine Learning Research 178 (2022), 1787–1911; arXiv:2111.02855.
- [8] E. Bolthausen and A.-S. Sznitman, *On Ruelle’s probability cascades and an abstract cavity method*, Commun. Math. Phys. 197 (1998), 247–276; doi:10.1007/s002200050450.
- [9] H.-B. Chen, V. Issa, and J.-C. Mourrat, *Free energy of non-convex multi-species spin glasses with centered Ising spins*, arXiv:2606.16636, 2026.
- [10] W.-K. Chen, *The Aizenman–Sims–Starr scheme and Parisi formula for mixed  $p$ -spin spherical models*, Electron. J. Probab. 18 (2013), article no. 94, 1–14; doi:10.1214/EJP.v18-2580; arXiv:1204.5115.
- [11] T. M. Cover, *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*, IEEE Trans. Electron. Comput. EC-14 (1965), no. 3, 326–334; doi:10.1109/PGEC.1965.264137.
- [12] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, second edition, corrected printing of the 1998 edition, Stochastic Modelling and Applied Probability, vol. 38, Springer, Berlin, Heidelberg, 2010; doi:10.1007/978-3-642-03311-7.
- [13] S. Franz and G. Parisi, *The simplest model of jamming*, J. Phys. A: Math. Theor. 49 (2016), no. 14, 145001; doi:10.1088/1751-8113/49/14/145001; arXiv:1501.03397.
- [14] E. Gardner, *The space of interactions in neural network models*, J. Phys. A: Math. Gen. 21 (1988), no. 1, 257–270; doi:10.1088/0305-4470/21/1/030.
- [15] E. Gardner and B. Derrida, *Optimal storage properties of neural network models*, J. Phys. A: Math. Gen. 21 (1988), no. 1, 271–284; doi:10.1088/0305-4470/21/1/031.
- [16] G. Györgyi and P. Reimann, *Beyond storage capacity in a single model neuron: continuous replica symmetry breaking*, J. Stat. Phys. 101 (2000), nos. 1–2, 679–702; doi:10.1023/A:1026441500710; arXiv:cond-mat/0003360.
- [17] A. Jagannath and I. Tobasco, *A dynamic programming approach to the Parisi functional*, Proc. Amer. Math. Soc. 144 (2016), no. 7, 3135–3150; doi:10.1090/proc/12968; arXiv:1502.04398.

- [18] W. Krauth and M. Mézard, *Storage capacity of memory networks with binary couplings*, J. Phys. (Paris) 50 (1989), no. 20, 3057–3066; doi:10.1051/jphys:0198900500200305700.
- [19] A. Montanari and K. Zhou, *Which exceptional low-dimensional projections of a Gaussian point cloud can be found in polynomial time?*, Ann. Probab., to appear; arXiv:2406.02970v3, 2025.
- [20] A. Montanari and K. Zhou, *Replica symmetry breaking and algorithmic thresholds in empirical risk minimization under multi-index model*, arXiv:2606.28573v1, 2026.
- [21] J.-C. Mourrat and D. Panchenko, *Extending the Parisi formula along a Hamilton–Jacobi equation*, Electron. J. Probab. 25 (2020), Paper No. 23, 17 pp.; doi:10.1214/20-EJP432; arXiv:1910.14482.
- [22] S. Nakajima and N. Sun, *Sharp threshold sequence and universality for Ising perceptron models*, Proceedings of the 2023 Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), 638–674, SIAM, 2023; doi:10.1137/1.9781611977554.ch28; arXiv:2204.03469.
- [23] OpenAI, *The free energy of the Ising random perceptron*, manuscript, September 24, 2026; [OpenAI mathematics repository](#).
- [24] OpenAI, *The free energy of the spherical random perceptron*, manuscript, September 24, 2026; [OpenAI mathematics repository](#).
- [25] OpenAI, *Microscopic jamming in the negative spherical perceptron*, OpenAI Math Release preprint, September 24, 2026; [OpenAI mathematics repository](#).
- [26] D. Panchenko, *The Sherrington–Kirkpatrick Model*, Springer Monographs in Mathematics, Springer, 2013; doi:10.1007/978-1-4614-6289-7.
- [27] D. Panchenko, *The free energy in a multi-species Sherrington–Kirkpatrick model*, Ann. Probab. 43 (2015), no. 6, 3494–3513; doi:10.1214/14-AOP967; arXiv:1310.6679.
- [28] D. Panchenko and M. Talagrand, *On one property of Derrida–Ruelle cascades*, C. R. Math. Acad. Sci. Paris 345 (2007), no. 11, 653–656; doi:10.1016/j.crma.2007.10.035.
- [29] D. Panchenko and M. Talagrand, *Guerra’s interpolation using Derrida–Ruelle cascades*, preprint, arXiv:0708.3641 (2007).
- [30] F. Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain*, Psychol. Rev. 65 (1958), no. 6, 386–408; doi:10.1037/h0042519.
- [31] D. Ruelle, *A mathematical reformulation of Derrida’s REM and GREM*, Commun. Math. Phys. 108 (1987), no. 2, 225–239; doi:10.1007/BF01210613.
- [32] M. Shcherbina and B. Tirozzi, *Rigorous solution of the Gardner problem*, Commun. Math. Phys. 234 (2003), no. 3, 383–422; doi:10.1007/s00220-002-0783-3; arXiv:math-ph/0112003.
- [33] Z. Šidák, *Rectangular confidence regions for the means of multivariate normal distributions*, J. Amer. Statist. Assoc. 62 (1967), no. 318, 626–633; doi:10.1080/01621459.1967.10482935.
- [34] M. Talagrand, *Mean Field Models for Spin Glasses, Volume I: Basic Examples*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 54, Springer, 2011; doi:10.1007/978-3-642-15202-3.
- [35] M. Talagrand, *Mean Field Models for Spin Glasses, Volume II: Advanced Replica-Symmetry and Low Temperature*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 55, Springer, 2011; doi:10.1007/978-3-642-22253-5.
- [36] M. Talagrand, *Free energy of the spherical mean field model*, Probab. Theory Related Fields 134 (2006), 339–382; doi:10.1007/s00440-005-0433-8.