

hexagon:2610.00177v1

An $N^{-7/10}$ Bound for the Finite-Size Bias of the Sherrington–Kirkpatrick Free Energy

P. M. Aronow

Patrick Lopatto

AUTHORS' NOTE

With the exception of this note and small edits to the abstract, the entire document is machine written. We have taken care to supervise this writing to ensure that the proofs meet a minimum standard of readability and that previous literature is appropriately cited. However, the resulting document remains lacking in exposition and overall coherence. We are nonetheless releasing this manuscript on Hexagon to disseminate the result as quickly as possible, because we believe it may be of interest to the mathematics and physics communities. We welcome suggestions concerning mathematical corrections, omitted citations, and attribution of ideas.

P.L. was partially supported by NSF grant DMS-2450004.

ABSTRACT

For every fixed inverse temperature $\beta > 1$, we prove that the expected free energy per spin of the zero-field Gaussian Sherrington–Kirkpatrick model differs from its Parisi limit by at most $C_\beta N^{-7/10}$. This improves the exponent $2/3$ of our earlier work (arXiv:2609.09103). We also compare the bound with the finite-size correction of order $N^{-3/4}$ predicted by Franz, Parisi, and Ricci-Tersenghi, and with the strict improvement over $2/3$ that would follow from the ideas in OpenAI’s recently released manuscripts.

Contents

1	Introduction	1
1.1	The Sherrington–Kirkpatrick model and the Parisi formula	1
1.2	Finite-size corrections.	2
1.3	Main result.	2
1.4	Relation to physical predictions.	4
1.5	Related work	7
1.6	Comparison with the OpenAI preprints.	9
1.7	Organization and conventions	12
2	Overview of the proof	14
2.1	The structure of the argument.	14
2.2	Reduction to the fractional-moment error	15
2.3	The interpolation and the defect	16
2.4	The thermal terms	17
2.5	The self-overlap variance for small ρ	18
2.6	The cavity contraction	20
2.7	Stability of the scalar response	23
2.8	Localization of the minimizing profile.	24
2.9	Closing the implicit inequality	26
2.10	Minimizers, arbitrary profiles, and atoms	28
2.11	A map of the proof	28
PART I	The interpolation	31
3	The interpolation and the defect	32
3.1	Completion and the constrained scalar value	33
3.2	The interpolating functional	34
3.3	The tilted joint law and overlap observables.	36
3.4	Martingale and contact identities	38
3.5	The integrated variance bound and the overlap error	41
3.6	The optimized interpolation and the defect	44
3.7	The scalar trial profile and the trial budget	47
3.8	Reduction of the mean bias	48

4	Gaussian tools for arbitrary profiles	49
4.1	Generalized laws and a posterior variance bound	49
4.2	Differentiating normalized Gaussian path expectations	55
PART II	Overlap estimates at a minimizer	62
5	Self-overlap moments on quantile intervals	63
5.1	Quantile windows and kernels	63
5.2	Two-sided conditional comparison	65
5.3	Variance on a quantile interval	68
5.4	A cubic-moment bound.	71
6	The mean–covariance term	74
6.1	Standing assumptions	75
6.2	Covariance along an earlier measurable direction	75
6.3	Mean increments toward a support point	77
6.4	A dyadic decomposition of the mean increment	78
6.5	An integrated bound for j	80
PART III	The minimizing profile and the scalar response	84
7	Isolation of the minimizing profile	85
7.1	Energy normalization and variational facts	85
7.2	Smoothing and coercivity.	90
7.3	A density bound for the scalar measure	94
7.4	Distance from the scalar trial profile.	96
8	Localization of the lowest quantiles	102
8.1	A local deficit certificate	102
8.2	Root localization	106
9	Scalar response and stability	110
9.1	The scalar law	111
9.2	Continuity and kernel bounds	116
9.3	The response operator	119
9.4	Domination by the scalar multiplier.	126
9.5	The reference multiplier	129
9.6	Stability at a minimizing profile.	131
PART IV	The cavity contraction	133
10	The cavity interpolation and its tree laws	134
10.1	Removing the incident row	135

10.2	Laws of finitely many descendants	137
10.3	Products of three overlap defects	152
11	The derivative at the independent endpoint	156
11.1	Linear response at the independent endpoint	157
11.2	The integrated response errors	168
12	Closing the cavity inequality	174
12.1	The mechanism of the contraction.	175
12.2	Comparison with coefficient one and the trace identity	176
12.3	The third moment of the rest overlap	179
12.4	Closing the variance inequality	180
PART V	Proof of the main theorem	184
13	Proof of the main theorem	185
13.1	The interval near $\rho = 0$	185
13.2	Closure at the critical tilt	188
	References	193

Introduction

1.1 The Sherrington–Kirkpatrick model and the Parisi formula

Let $(g_{ij})_{1 \leq i < j \leq N}$ be independent standard Gaussian variables. The Sherrington–Kirkpatrick (SK) model [76] at zero external field and inverse temperature $\beta > 0$ has Hamiltonian

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{1 \leq i < j \leq N} g_{ij} \sigma_i \sigma_j, \quad \sigma \in \{-1, 1\}^N.$$

Its logarithmic partition function and expected pressure are

$$F_{N,\beta} = \log \sum_{\sigma \in \{-1,1\}^N} e^{\beta H_N(\sigma)}, \quad p_{N,\beta} = \frac{\mathbb{E} F_{N,\beta}}{N}.$$

For a probability measure μ on $[0, 1]$ with distribution function $\alpha(q) = \mu([0, q])$, let u_α be the solution of the Parisi equation on $[0, 1] \times \mathbb{R}$, and define the Parisi functional $\mathcal{P}_\beta$ and the Parisi value P_β by

$$\begin{aligned} \partial_q u_\alpha &= -\frac{\beta^2}{2} (\partial_{xx} u_\alpha + \alpha(q) (\partial_x u_\alpha)^2), & u_\alpha(1, x) &= \log(2 \cosh x), \\ \mathcal{P}_\beta(\alpha) &= u_\alpha(0, 0) - \frac{\beta^2}{2} \int_0^1 q \alpha(q) \, dq, & P_\beta &= \inf_{\mu \in \mathcal{P}([0,1])} \mathcal{P}_\beta(\alpha), \end{aligned} \tag{1.1}$$

where $\mathcal{P}([0, 1])$ is the set of probability measures on $[0, 1]$. For a general μ , the equation is understood in the weak sense of [43, Theorem 4].

Sherrington and Kirkpatrick introduced this model and studied it by a replica-symmetric calculation, which fails at low temperature [76]; a longer treatment, with numerical comparisons, is given in [45]. De Almeida and Thouless showed that the replica-symmetric saddle point becomes unstable [34]. Parisi proposed a hierarchical breaking of the replica symmetry in [62], and he developed a sequence of approximate solutions and the order parameter on $[0, 1]$ that leads to (1.1) in [63, 64]. He later interpreted this order parameter through the distribution of overlaps between states [65]. Mézard, Parisi, Soutoul, and Virasoro then described the organization of the states, with sample-dependent weights and ultrametric overlaps, in [50].

The Parisi formula states that $p_{N,\beta} \rightarrow P_\beta$ as $N \rightarrow \infty$. Guerra proved the upper bound $p_{N,\beta} \leq P_\beta$ for every N [40], and Talagrand proved the matching lower bound in

the limit [78]. The infimum in (1.1) is attained at a unique measure μ_β [10], the Parisi measure, whose distribution function we denote by α_β . For $\beta > 1$, its support is an interval $[0, q_\beta]$ with $0 < q_\beta < 1$ [48, Theorem 1.1]; for $\beta > 1$ close to 1, this was proved earlier by Zhou [86].

1.2 Finite-size corrections

In this paper, we study the rate of convergence in the Parisi formula at low temperature. By Guerra's bound, the *mean bias* $P_\beta - p_{N,\beta}$ is nonnegative for every N . Up to a factor $1/\beta$, it is the finite-size correction to the expected free energy per spin (Section 1.3).

At high temperature, the mean bias is small. For $0 < \beta < 1$, Guerra and Toninelli showed by a quadratic coupling of two replicas that it is of order at most N^{-1} [42, Section 3]. With a nonzero external field, Lopatto proved that the limiting pressure equals the replica-symmetric value up to and including the de Almeida–Thouless line [49, Theorem 1.1], and Kusuoka and Nakajima proved that the pressure is within $O(N^{-1})$ of this value on compact subsets of the strict de Almeida–Thouless region [47, Theorem 1.1].

Below the critical temperature, the mean bias is expected to decay more slowly. In [4, Theorem 1.1, (1.5)], we proved that $0 \leq P_\beta - p_{N,\beta} \leq C\beta N^{-1/2}$ for every $\beta > 0$ and $N \geq 1$, with a numerical constant C . For fixed $\beta > 1$, we used the interval support of μ_β to prove the sharper upper bound $C_\beta N^{-2/3}$ and the lower bound $c_\beta N^{-11/12}$, for all sufficiently large N [4, Theorem 1.4]. In physics, the correction is expected to decay like $N^{-\omega}$ with $\omega < 1$ at low temperature. Two values have been proposed for the full model, $\omega = 2/3$ and $\omega = 3/4$. A calculation retaining only longitudinal fluctuations gives $\omega = 4/5$ [72]. We review these predictions in Section 1.4. We improve the exponent of the upper bound from $2/3$ to $7/10$.

1.3 Main result

Our main result is the following.

Theorem 1.1. *For every $\beta > 1$ there is a constant $C_\beta < \infty$ such that*

$$P_\beta - p_{N,\beta} \leq C_\beta N^{-7/10} \quad (N \geq 2).$$

Since $P_\beta - p_{N,\beta} \geq 0$ for every N [40], Theorem 1.1 is a two-sided estimate. In physical units, the expected free energy density is $\mathbb{E}[-F_{N,\beta}/(\beta N)]$, and its limit is $-P_\beta/\beta$. By Theorem 1.1, we have

$$0 \leq \mathbb{E}\left[-\frac{F_{N,\beta}}{\beta N}\right] - \left(-\frac{P_\beta}{\beta}\right) = \frac{P_\beta - p_{N,\beta}}{\beta} \leq \frac{C_\beta}{\beta} N^{-7/10} \quad (N \geq 2).$$

For the expected logarithm of the partition function itself, the corresponding bound is $0 \leq NP_\beta - \mathbb{E}F_{N,\beta} \leq C_\beta N^{3/10}$. The constant C_β depends on β , and we do not control it as $\beta \downarrow 1$ or as $\beta \rightarrow \infty$. In particular, we do not obtain a corresponding bound for the ground-state energy.

Together with the lower bound in [4, Theorem 1.4], Theorem 1.1 implies that for every fixed $\beta > 1$,

$$c_\beta N^{-11/12} \leq P_\beta - p_{N,\beta} \leq C_\beta N^{-7/10}$$

for all sufficiently large N . The exponent $7/10$ lies between the exponent $2/3$ of [4, Theorem 1.4] and the exponent $3/4$ predicted by Franz, Parisi, and Ricci-Tersenghi [38], which we discuss in Section 1.4. In Section 1.6, we compare Theorem 1.1 with two conditional consequences of results stated in the OpenAI manuscripts [54, 56]. An uncentered comparison stated there would give an unspecified strict improvement over $2/3$, and the variance asymptotic stated there would imply the lower bound $P_\beta - p_{N,\beta} \geq c_\beta N^{-5/6}$.

We now outline the proof; Section 2 gives a detailed overview. We follow the interpolation scheme of [4], which is a quantitative form of the Guerra interpolation [40] that Talagrand used for the positive fractional moments of the partition function [80]. For a tilt $0 < \lambda \leq \lambda_0$, let $e_{N,\beta}(\lambda) \geq 0$ be the error, in energy units, between the constrained Parisi value and the normalized logarithmic moment of order λ/β of the completed partition function. The constraint requires mass at least λ/β at zero (Section 3). Imposing this mass raises the Parisi value by at most $C_\beta \lambda^5$ in energy units [4, Lemma 7.2]. Using this bound and the convexity of Gaussian logarithmic moments [26, Theorem 1], as in the proof of [4, Theorem 7.8, Step 1], we find

$$0 \leq \frac{P_\beta - p_{N,\beta}}{\beta} \leq C_\beta \lambda^5 + 2e_{N,\beta}(\lambda).$$

For sufficiently large N , we take $\lambda = N^{-7/50}$, so it suffices to show that $e_{N,\beta}(N^{-7/50}) \leq C_\beta N^{-7/10}$.

Up to a factor of four, $e_{N,\beta}(\lambda)$ is the integral over $\rho \in (0, 1]$ of a defect $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ along the interpolation, evaluated at a minimizing profile of an N -dimensional variational problem. The terms A and J are controlled directly (Sections 3 and 6). The difficulty is V , the integrated variance of a self-overlap. For ρ above a cutoff ρ_0 , we bound V by a one-site cavity argument (Part IV). Removing one spin leads to an inequality $V \leq (1 - \rho)(1 + \varepsilon_N)V + \text{errors}$, where $1 + \varepsilon_N$ bounds a scalar response operator at the minimizing profile, and this inequality contracts once $\varepsilon_N \leq \rho/4$. To prove this bound, we show that the response operator at the Parisi measure is bounded by the identity, which reflects the marginal stability of μ_β , and that the minimizing profile is close to a scalar trial profile (Part III). For $\rho \leq \rho_0$, we bound V by the quantile-shell estimates of [4, Section 3], refined in Section 5, and by a localization of the lowest quantiles of the

minimizer (Section 8), which refines [4, Lemma 7.6]. The cutoff is $\rho_0 = \min\{1, K_\beta \eta^{1/7}\}$, where the trial budget $\eta = e_{N,\beta}(\lambda) + C_\beta \lambda^5$ contains the unknown error. The result is an implicit inequality of the form

$$e_{N,\beta}(\lambda) \lesssim N^{-2/3} (e_{N,\beta}(\lambda) + \lambda^5)^{1/21} + \text{smaller terms.}$$

If $e_{N,\beta}(\lambda)$ and λ^5 are both of order N^{-a} , this inequality is balanced when $a = 2/3 + a/21$, that is, $a = 7/10$.

1.4 Relation to physical predictions

Expectations and early calculations. Below the critical temperature, physicists expect the finite-size correction to the free energy density to be nonanalytic, of order $N^{-\omega}$ with $\omega < 1$. This expectation is related to marginal stability. In a truncated replica model near criticality, De Dominicis and Kondor found that the stability matrix of the Parisi solution has nonnegative eigenvalues and many zero modes [35, pp. 606–608]. These zero modes make the expansion in fluctuations around this solution singular; see [38, Section 2]. Parisi, Ritort, and Slanina studied finite-size corrections at criticality [71], where the free energy density has a correction $(\log N)/(12N)$ (see [14]), as well as slightly below criticality and along the de Almeida–Thouless line, including corrections to overlap distributions and to ultrametricity [70].

The exponent 2/3. At zero temperature, Bouchaud, Krzakala, and Martin found numerically, for Gaussian couplings, that the mean ground-state energy per spin approaches its limit like $N^{-2/3}$ [19, Section VI B]. In a study of the SK model with binary couplings, Boettcher fits the corrections to the mean of the ground-state energy per spin and to the width of its distribution with exponents close to 2/3 and 3/4, respectively [15, Section 3]. In a later study of diluted versions of the model, the fit for the undiluted model is again close to 2/3 [17, Table I]. At positive temperature, Billoire estimated the correction numerically for binary couplings, by integrating overlap observables along the interpolation of Guerra and Toninelli [41], and found data compatible with $\omega = 2/3$ [14]. Aspelmeier, Billoire, Marinari, and Moore derived $\omega = 2/3$ from a heuristic argument: a system of N spins resolves about $N^{1/6}$ steps of replica symmetry breaking, and truncation after K steps changes the free energy density by order K^{-4} [8, Section 2].

Longitudinal fluctuations and the exponent 4/5. Parisi and Biscari obtained an earlier contribution of order $N^{-3/4} \log N$ from the positive longitudinal eigenvalues near criticality, neglecting interactions between modes [66, Section 1 and Eq. (5.25)]. Their contribution from the longitudinal zero modes diverges when treated separately from the omitted transverse modes [66, discussion following Eq. (5.27)]; see [72, Section 1]. Parisi, Sarra, and Talamanca later used a resummation-based scaling argument for the

longitudinal fluctuations of a truncated free-energy functional for the SK model near criticality and predicted $\omega = 4/5$ [72, Section 3, Eqs. (16)–(19)]. They omit the transverse fluctuations, so this exponent describes their longitudinal approximation. To determine the exponent for the full model, one must also control the omitted modes.

The exponent $3/4$. Franz, Parisi, and Ricci-Tersenghi [38] argue that $\omega = 2/3$ is incorrect and that the correct value is $\omega = 3/4$, possibly up to logarithmic factors [38, Eq. (14)]. Their calculation for the Gaussian SK model matches two regimes: one in which the overlaps of two coupled systems lock and the Guerra–Toninelli integrand decays like $(Nt)^{-1/2}$, and a perturbative regime in which this integrand is of order $t^{-3/2}/N$ [38, Eqs. (9)–(12) and (15)–(17)]. They estimate the large- z locking exponent semianalytically by sampling random ultrametric trees generated by the Bolthausen–Sznitman coalescent and extrapolating in the number of leaves [38, Materials and Methods]. Matching this behavior to the perturbative regime gives the scaling prediction. The fit of the energy uses random regular graphs of degree four [38, Eq. (18) and Fig. 2], and the simulations of the mode-locking integral use rewired random regular graphs of degree four with binary couplings (see the subsection “Direct simulations of the mode-locking integral” of [38, Materials and Methods]). These models differ from the Gaussian SK model in their connectivity, and the latter also in the law of the couplings.

Consequences of Theorem 1.1. Theorem 1.1 implies that $P_\beta - p_{N,\beta} = O(N^{-7/10})$ for every fixed $\beta > 1$; in particular, the correction is $o(N^{-2/3})$. We therefore exclude a nonzero leading correction of order $N^{-2/3}$ for the Gaussian SK model at every fixed temperature below the critical one. This is consistent with the prediction of Franz, Parisi, and Ricci-Tersenghi. Theorem 1.1 does not determine the leading exponent ω . Since our constants are not uniform as $\beta \rightarrow \infty$, it also says nothing about the numerical results at zero temperature described above.

The estimates on which our proof rests have a similar structure with two regimes, although they concern a different interpolation. If we ignore logarithms and the dependence on the tilt λ , the defect $\mathcal{D}_N(\rho)$ is at most of order $(N\rho)^{-2/3}$ for small ρ and of order $(N\rho^2)^{-1}$ on the interval where our cavity argument applies. If the second bound held for all $\rho \geq N^{-1/4}$, where the two bounds are equal, integration would give $N^{-3/4}$. In our proof, the cavity interval begins only at $\rho_0 \geq N^{-1/10}$, and the tilt $\lambda = N^{-7/50}$ enters both bounds; these two features account for the difference between $7/10$ and $3/4$.

Marginal stability also enters our proof, through a scalar statement: at the Parisi measure, the multiplier that dominates our scalar response operator equals one on $\text{supp } \mu_\beta$ (Section 9). Crisanti and Rizzo wrote the corresponding identity in physical variables, on the part of the order parameter that varies [32, Eqs. (22)–(23)]. This scalar statement does not describe the full stability spectrum of the replica calculation.

Mean, width, and large deviations. The exponents ω for the finite-size correction discussed above concern the mean. The sample-to-sample fluctuations of the free energy and its large deviations are governed by different exponents. In a truncated replica model near criticality, Kondor obtained a fifth-order dependence on n along the Parisi branch for small positive n [46, pp. L129–L130, Eq. (11)]. Interpreted as a formal prediction for the SK model, this says that the replica-normalized pressure in the thermodynamic limit, $\lim_{N \rightarrow \infty} (Nn)^{-1} \log \mathbb{E} e^{nF_{N,\beta}}$, differs from P_β first at order n^5 as $n \downarrow 0$ with $n > 0$ (see [67, pp. 1 and 3]). Parisi and Rizzo extended this calculation throughout the low-temperature phase in [67] and [68, Appendix B]. Equivalently, in this formal expansion, the limit of $N^{-1} \log \mathbb{E} e^{n(F_{N,\beta} - NP_\beta)}$ begins at order n^6 , and its Legendre transform gives a rate with power $6/5$ for upward deviations of $F_{N,\beta}/N$ [67]. Crisanti, Paladin, Sommers, and Vulpiani interpreted the replica number as a tilt over the disorder and predicted that the free energy per spin fluctuates on the scale $N^{-5/6}$ [33]. Under an asymmetric scaled-law assumption, the same calculation also predicts a correction to the mean density on this scale [33, p. 1331, Eq. (34) and following text]. This conclusion uses the truncated replica theory near criticality and a matching of the disorder tails to the typical peak. In our convention, in which $F_{N,\beta}$ is the logarithm of the partition function, Parisi and Rizzo’s replica predictions give upward density deviations with speed N and power $6/5$, and downward density deviations with speed N^2 and power $12/5$ [69, Eqs. (12)–(15)]. Smoothly matching these tails to the finite-size peak predicts the typical density scale $N^{-5/6}$. This matching is an additional assumption; the limits $n \downarrow 0$ and $N \rightarrow \infty$ cannot simply be interchanged [68, Section VI]. Aspelmeier, Billoire, Marinari, and Moore also predicted a variance of order $N^{1/3}$ for $F_{N,\beta}$ [8]. Aspelmeier derived exact Gaussian interpolation identities relating the variance of the free energy to two- and four-replica overlaps between systems with correlated disorder [7, Eqs. (14) and (31)]. Combining these identities with replica estimates of the bond-chaos overlap distribution gives a predicted upper bound $1/4$ for the extensive standard-deviation exponent [6, 7, Section 4.3]. The identities are exact; this exponent estimate uses replica approximations and concerns the width. At zero temperature, the width of the distribution of the ground-state energy has been studied numerically in [15, 16, 58]. For Gaussian couplings, Palassini finds data consistent with a mean correction exponent $2/3$ and, after subleading corrections, favors the exponent $5/6$ over $3/4$ for the standard deviation per spin [58, Sections 3–4 and 6]. Boettcher discusses the difficulty of distinguishing the exponents $3/4$ and $5/6$ after corrections to scaling [16, Section IV]. Monthus and Garel studied rare-disorder scenarios and the matching between typical and large-deviation tails for the ground-state energy of the SK model at zero temperature [53, Section III C–F]. Their alternative width–tail relations illustrate the additional assumptions involved in such matching.

1.5 Related work

The Parisi formula. Besides the proofs of Guerra and Talagrand, the rigorous theory of the Parisi formula includes the following results. Guerra and Toninelli proved by interpolation that the thermodynamic limit exists [41], and Aizenman, Sims, and Starr formulated an extended variational principle in terms of random overlap structures [2]. Panchenko proved that the Ghirlanda–Guerra identities [39] imply ultrametricity of overlaps [60, Theorem 1], and he proved the Parisi formula for mixed p -spin models, including mixtures with odd interactions [61]. We refer to the books of Talagrand [81, 82] and Panchenko [59] for accounts of this theory; Panchenko develops it through Ruelle probability cascades.

The Parisi measure. Toninelli proved that the replica-symmetric value is not optimal for $\beta > 1$ at zero field [84]. Auffinger and Chen proved that the minimizer is unique [10], that zero belongs to its support, and that the largest point of its support is an atom for $1 < \beta \leq 3/2$ [9, Theorems 1 and 4]. They also proved that its distribution function is smooth on $[a, b)$ whenever (a, b) is contained in its support, for $0 \leq a < b \leq 1$ [9, Theorem 2]. They expressed the Parisi functional through a stochastic control problem, which they used in their uniqueness proof, and they proved martingale identities for the derivatives of u_α [10, Theorem 3 and Lemma 2]. An earlier Hamilton–Jacobi–Bellman representation appears in [20, Section 6.2], and Jagannath and Tobasco developed a dynamic programming treatment of the functional [43, Lemma 18]. First-variation formulas along segments and optimality conditions on the support are given by Jagannath and Tobasco [44, Lemma 3.2 and Proposition 1.1] and by Chen [25, Theorem 2 and Proposition 1]. Lopatto proved, in addition to the interval support for every $\beta > 1$, that μ_β is the sum of a density that is smooth up to the endpoints of $[0, q_\beta]$ and an atom at q_β [48, Theorem 1.1]. Auffinger, Chen, and Zeng proved that the Parisi minimizer at zero temperature has infinitely many support points [11]. H.-B. Chen proved that $q_\beta \rightarrow 1$ as $\beta \rightarrow \infty$, and that the Stieltjes measure associated with the Parisi order parameter at zero temperature has a smooth density and support $[0, 1)$ [23, Theorems 1.2 and 1.3]. The full support at zero temperature is also stated in [55, Theorem 1.1]. We use the interval support at fixed $\beta > 1$ throughout, and the regularity of μ_β in Part III.

Fluctuations of the free energy. The fluctuations of $F_{N,\beta}$ depend on β and on the external field. For $0 < \beta < 1$, Aizenman, Lebowitz, and Ruelle proved a Gaussian limit for its fluctuations, which are of order one [1], and Comets and Neveu studied this limit by stochastic calculus [30]. With a nonzero external field, Chen, Dey, and Panchenko proved that the variance is of order N and that a central limit theorem holds, under assumptions on the interactions that include the SK model [28, Theorems 1 and 2]. At $\beta = 1$, Chen and Lam proved a variance bound of order $\log^2 N$ [27, Theorem 1.1]. Schertzer proved that the variance is at most $(\log N)/4 + C$ and at least of order $\log \log N$ [75, Theorems 1

and 2], and Du and Huang proved that $\text{Var}(F_{N,1}) = (\log N)/6 + O(1)$ and that the fluctuations have a Gaussian limit [37, Theorem 1.3]. Dey and Kang obtained variance asymptotics and a Gaussian limit as β approaches 1 from below along sequences $\beta = \beta_N$ satisfying $N^{1/3}(1 - \beta_N^2) \rightarrow c \in (0, \infty]$ [36, Theorem 1.1]. For fixed $\beta > 1$, Chatterjee proved the superconcentration bound $\text{Var}(F_{N,\beta}) \leq C_\beta N / \log N$ and related it to disorder chaos [22, Theorem 1.5]. In [4, Theorem 1.4], we proved that $\text{Var}(F_{N,\beta})$ lies between constant multiples of $N^{4/15}$ and $N^{7/15}$ for large N . The OpenAI manuscripts state the order $N^{1/3+o(1)}$, a positive limit for $N^{-1/3} \text{Var}(F_{N,\beta})$, and a limiting law for the normalized fluctuations; see Section 1.6. A variance bound alone does not bound the mean bias from above, but a lower bound on the variance gives a lower bound on the mean bias through [4, Lemma 7.4]. For the spherical SK model, Baik and Lee proved that the free energy per spin has Gaussian fluctuations of order N^{-1} at high temperature and Tracy–Widom fluctuations of order $N^{-2/3}$ at low temperature, governed by the largest eigenvalue of the coupling matrix [12, Theorems 2.10 and 2.11].

Gaussian-process comparisons. Chu and Raginsky use Gaussian-process geometry and information-theoretic tensorization to bound expected soft maxima [29]. Their application to the SK model gives a finite-volume comparison with the Parisi variational value up to universal multiplicative constants [29, Section 6, Theorem 10]. It uses a uniform spin prior and normalizes the Hamiltonian differently. Its conclusion is a multiplicative comparison; the present theorem controls the additive error to the limiting value.

Fractional moments. Talagrand computed the limits of normalized logarithms of positive moments of the partition function [80]. For moments of order $0 < s < 1$, the limit of $(Ns)^{-1} \log \mathbb{E} e^{sF_{N,\beta}}$ is the minimum $P_\beta(s)$ of $\mathcal{P}_\beta$ over measures μ with $\mu(\{0\}) \geq s$. For the SK model at zero field, this formula is contained in the one-level specialization of the multiscale formula of Contucci and Mingione [31, Theorem 2.1]. Chen, Guionnet, Ko, Lacroix-A-Chez-Toine, and Mourrat stated it in this form and used it to study one-sided large deviations of the ground-state energy [24, Theorem 2.1]. In [4, Lemma 7.2 and Proposition 7.3], we proved that for fixed $\beta > 1$, the increase $P_\beta(s) - P_\beta$ is of order s^5 as $s \downarrow 0$. This is a rigorous counterpart of the replica calculations of Kondor and of Parisi and Rizzo described in Section 1.4, and Pei identified the leading coefficient of this increase [73, Theorem 1(a)]. The tilt $\lambda = N^{-7/50}$ in our proof balances this cost against the fractional-moment error.

Origins of the methods. The interpolation of [4, Section 2] is a finite-dimensional, quantitative form of Guerra’s interpolation, with which he proved the upper bound in the Parisi formula [40]. For posterior estimates under its tilted laws, we combine the Brascamp–Lieb inequality [21, Theorem 4.1], in its inverse-Hessian form [4, Proposition 2.2], with a joint-source argument from the proof of [54, Lemma 2.3] (Section 4). Near the singular end of the interpolation, our estimates refine the quantile-window bound of [4, Lemma 3.3]

and the localization of lowest quantiles in [4, Lemma 7.6]. We close the final implicit inequality as in the proof of [4, Proposition 7.7].

The cavity method originates in physics. An early precursor is the single-site description of Thouless, Anderson, and Palmer [83, Section 3, Eqs. (10)–(12)], and Mézard, Parisi, and Virasoro formulated the cavity construction for broken replica symmetry in [51]; see also the exposition in [52]. We use Talagrand’s rigorous version of the one-site argument [81, Section 1.6]. The expansion under hierarchical Gaussian laws relies on normalized replica trees and covariance insertions of the type developed in [54, Section 2]. In the scalar stability argument, the first variation of $\mathcal{P}_\beta$ under transports of levels is due to Talagrand [79, Lemma 3.7]; see [9, Section 3, (19)]. We compute the second variation and compare it with a response multiplier, using convexity of $\mathcal{P}_\beta$ along segments of distribution functions [10]. This multiplier equals one on $\text{supp } \mu_\beta$, by an identity of Talagrand [79], in the form stated in [9, Theorem 5] and [48, Lemma 2.6], and by the interval support [48, Theorem 1.1]. Above the support, we bound it by one using the curvature of the forward density and its heat evolution from [48, Proposition 4.1 and Lemma 2.4(ii)].

Our related work. This paper continues [4], where we proved the quantitative Parisi formulas recalled in Section 1.2, the variance bounds above, and upper-tail estimates for $F_{N,\beta}$ with exponent $6/5$ [4, Theorem 1.5]. The interval-support theorem used in both papers was proved by the second author [48]. In [5], we studied a different approximation of the SK pressure. On d -dimensional tori of side length $L \geq 3$ and on hypercubes, with independent Gaussian couplings of variance $1/(2d)$ and $1/d$, respectively, the expected pressure per site converges to the SK pressure as $d \rightarrow \infty$, for every fixed inverse temperature $\beta \geq 0$ and every uniform field $h \in \mathbb{R}$; on tori, the convergence is uniform in the side length. The present paper concerns the finite-size bias of the SK model itself.

1.6 Comparison with the OpenAI preprints

We obtained the exponent $7/10$ on September 23, 2026, following a suggestion from Silvio Franz on September 15, 2026. The OpenAI mathematics repository [57] contains *The low-temperature Sherrington–Kirkpatrick fluctuation scale* [54] and its companion *The low-temperature Sherrington–Kirkpatrick free-energy limiting law* [56], both dated September 24, 2026. We refer to their initial repository version.

Stated results. For every fixed $\beta > 1$, [54, Theorem 1.1] states that $\text{Var}(F_{N,\beta}) = N^{1/3+o(1)}$. The companion paper states that $N^{-1/3} \text{Var}(F_{N,\beta})$ converges to a positive limit and that $F_{N,\beta}$, standardized by its exact mean and variance, converges in law to a nondegenerate limit [56, Theorems 1.1 and 1.2]. These results concern fluctuations about the finite-size

mean $\mathbb{E}F_{N,\beta}$, and they do not by themselves estimate $P_\beta - p_{N,\beta}$. We do not use these results, or the comparison discussed next, in the proof of Theorem 1.1.

A conditional upper bound. An improvement of the exponent $2/3$ for the mean bias is implicit in an uncentered comparison stated in [56]. We use the notation of that paper. Let $T = \beta^2$ and $x = N^{-1/6}$, and fix an admissible choice of its auxiliary exponents and reference parameters. For its direct, unmasked, and unseeded reference paths, [56, Proposition 6.1, (G1)] states that there exists $c_G > 0$ such that

$$0 \leq F_*(\lambda) - f_N(b_\lambda, h_{*,\lambda}) \leq x^{4+c_G} \quad (1.2)$$

for all sufficiently large N . Here F_* is a scalar trial objective, f_N is a self-corrected hierarchical pressure evaluated at the clocks $(b_\lambda, h_{*,\lambda})$, and λ is a perturbation strength that is unrelated to our tilt. We take terminal mass one and the allowed strength $\lambda = x^{12}$. From the definitions in [56, Section 3.1], one checks that the scalar objective at $\lambda = 0$ equals $P_\beta - \log 2 - T/4$. Using the perturbation estimates for the direct scalar lift [56, Section 3.2, proof of Lemma 3.2] and the construction of the clocks [56, Section 3.4, (P3)–(P4)], we find

$$\|b_\lambda - T\|_\infty + \|h_{*,\lambda}\|_\infty \leq C_\beta \lambda, \quad \left| F_*(\lambda) - \left(P_\beta - \log 2 - \frac{T}{4} \right) \right| \leq C_\beta \lambda.$$

With the uniform spin prior and centered Gaussian completion of [56, Section 2.1], the unperturbed clocks satisfy $f_N(T, 0) = p_{N,\beta} - \log 2 - T/4$. Integrating the pressure variation formula [56, Lemma 2.1, (A1)] along the segment between the two pairs of clocks, and using that its integrands lie in $[0, 1]$ at terminal mass one, we obtain

$$|f_N(b_\lambda, h_{*,\lambda}) - f_N(T, 0)| \leq \frac{1}{4} \|b_\lambda - T\|_1 + \frac{1}{2} \|h_{*,\lambda}\|_1 \leq C_\beta \lambda.$$

We write $P_\beta - p_{N,\beta}$ as the sum of $P_\beta - \log 2 - T/4 - F_*(\lambda)$, $F_*(\lambda) - f_N(b_\lambda, h_{*,\lambda})$, and $f_N(b_\lambda, h_{*,\lambda}) - f_N(T, 0)$. We bound the first and last differences by $C_\beta \lambda$ and the middle one by (1.2). Since $x^{4+c_G} = N^{-2/3-c_G/6}$ and $x^{12} = N^{-2}$, we obtain

$$0 \leq P_\beta - p_{N,\beta} \leq x^{4+c_G} + C_\beta x^{12} \leq C_\beta N^{-2/3-\delta_\beta}, \quad \delta_\beta = c_G/6 > 0, \quad (1.3)$$

after decreasing c_G so that $c_G < 8$ if necessary. The lower bound is Guerra's, and (1.3) holds for the finitely many remaining $N \geq 2$ after we enlarge C_β . This bound follows from (1.2) and the perturbation estimates above; it is not a theorem on the mean bias stated in either paper. The proof of [56, Proposition 6.1] leaves c_G unspecified, and we do not extract from it the saving $c_G \geq 1/5$ that the exponent $7/10$ would require. The budgets of order x^{5+c} in [56, Proposition 4.1] concern increments of a centered pressure [56, (A4)], from which the finite-size mean has been subtracted, so they do not bound the mean bias.

Conditional lower bounds. In the other direction, the variance–mean-deficit inequality of [4, Lemma 7.4 and Remark 7.14, (7.68)] states that for every $N \geq 1$,

$$P_\beta - p_{N,\beta} \geq c_\beta \left(\frac{\text{Var}(F_{N,\beta}) + \beta^2/2}{N} \right)^{5/4}.$$

Inserting $\text{Var}(F_{N,\beta}) \geq N^{1/3-4\epsilon/5}$, we see that the variance exponent stated in [54, Theorem 1.1] would imply, for every fixed $\epsilon > 0$, that $P_\beta - p_{N,\beta} \geq c_{\beta,\epsilon} N^{-5/6-\epsilon}$ for all sufficiently large N . The positive limit of $N^{-1/3} \text{Var}(F_{N,\beta})$ stated in [56, Theorem 1.2] would give the stronger bound

$$P_\beta - p_{N,\beta} \geq c_\beta N^{-5/6} \quad \text{for all sufficiently large } N. \quad (1.4)$$

For (1.4), an eventual lower bound $\text{Var}(F_{N,\beta}) \geq k_\beta N^{1/3}$ with $k_\beta > 0$ would suffice. Combined with Theorem 1.1, the variance exponent of [54] would imply that if $P_\beta - p_{N,\beta}$ is of exact order N^{-a} , then $7/10 \leq a \leq 5/6$.

Verification status. We have checked the statements quoted above and the conditional deductions from them. We have not independently reconstructed complete proofs of the global comparison and fluctuation estimates of [54, 56], and we do not assume their conclusions. The repository also contains Lean sources on the fluctuations of the SK model. Their main theorem on the fluctuation exponents, as we inspected it, takes as an explicit hypothesis a two-sided bound on the centered logarithmic Laplace transform of the completed free energy, `PositiveLaplaceWindow`; in the fluctuation-scale paper, this bound is the conclusion of [54, Proposition 7.5]. The conclusion of that theorem concerns the fluctuation exponents, not a positive variance prefactor or a limiting law.¹ Neither this conditional formal statement nor our incomplete reconstruction shows that the stated results are false. Our inspection did not identify an invalid step in the portions examined.

Proof techniques. The limiting-law paper contains a one-site cavity expansion whose structure is similar to that of Part IV. In [56, Lemma 2.2], one site is chosen by exchangeability and removed, and its Gaussian kernels are interpolated to those of an independent scalar model on one site; the linear cavity kernel is half of a connected four-spin correlation of the removed spin [56, (A8)]. In the proof of [56, Proposition 6.1],

¹We inspected the 151 files of the directory `lean/OAI/Probability/SKFluctuations` at commit `adc7f1241b42e322a6451854ab7e4b4c146bf78a`. The theorem `main_of_positiveLaplaceWindow`, at line 294 of `EventuallyNatPos.lean`, takes a hypothesis `hw` of type `PositiveLaplaceWindow`, which is defined at lines 453–459 of `Spin.lean`; the conclusion of the theorem, `MainFluctuationStatement`, is defined at lines 531–540 of the same file. We found no proof supplying the hypothesis in this directory. The directory was unchanged at commit `fd4aeeb2ee4fc729c18d98444fed42fd0529eeeb`, which we checked on October 9, 2026. This was an inspection of the sources, not a compilation.

the resulting linear equation is inverted under a gap condition on the scalar response [56, Proposition 5.1, (H1)], and [56, Lemma 5.2, (H2)] writes the response on the additive sector as a susceptibility multiplier minus a positive semidefinite Hessian kernel. In our argument, the cavity inequality contracts because the scalar response operator is at most $1 + \varepsilon_N$ with $\varepsilon_N \leq \rho/4$, while the interpolation contributes the factor $1 - \rho$. We derive the bound on the response operator from the second variation of the Parisi functional under transports of levels (Lemma 9.7); this operator acts on a space of transports, not on the additive sector of [56]. The fluctuation-scale paper also contains a one-site closure for the mean overlap [54, Proposition 4.1]. We use two tools from that paper. Part IV relies on normalized replica trees and covariance insertions of the type developed in [54, Section 2] and used in [56, Sections 2.1–2.3]. In Lemma 4.2, we adapt the joint-source Brascamp–Lieb argument of the proof of [54, Lemma 2.3], which is also used in the proof of [56, Lemma 2.1], to squared conditional means on selected coordinates and to anisotropic laws.

1.7 Organization and conventions

Section 2 contains an overview of the proof and a map of its dependencies. The proof itself occupies five parts.

- Part I concerns the interpolation. Section 3 recalls the interpolation of [4], its tilted law, and the decomposition of the defect, and reduces the mean bias to the fractional-moment error. Section 4 proves two Gaussian tools for tilted laws defined by arbitrary profiles, which we use in Parts III and IV.
- Part II contains overlap estimates at a minimizing profile. Section 5 estimates moments of the self-overlap on quantile intervals, and Section 6 bounds the mean–covariance term.
- Part III concerns the minimizing profile and the scalar response. Section 7 shows that the minimizing profile is close in L^1 to the scalar trial profile, and Section 8 localizes its lowest quantiles. Section 9 studies the scalar response operator and proves its stability at the minimizing profile.
- Part IV contains the cavity argument. Section 10 sets up the cavity interpolation and its tree laws, Section 11 computes the derivative at the independent endpoint, and Section 12 closes the cavity inequality and bounds the defect.
- In Part V, Section 13 integrates these estimates and proves Theorem 1.1.

We rely on [4] for the interpolation identities and the scalar floor estimates, and on [48] and [9] for the structure and regularity of μ_β . We restate these results where they are used. Locators in [4] refer to its arXiv version 1, locators in [48] to its arXiv version 4, and locators in the OpenAI manuscripts to their initial repository version.

Throughout, $\beta > 1$ is fixed, and C_β denotes a constant depending only on β whose value may change from line to line; a numerical constant does not depend on any parameter. All profiles defining a tilted law are deterministic: minimizers may depend on N, ρ , and λ but not on the couplings. For symmetric matrices or self-adjoint operators, $B \preceq C$ means that $C - B$ is positive semidefinite. We write $\log_+ x = \log(\max\{1, x\})$ for $x \geq 0$. Distribution functions, such as α and the profiles a of Section 3, are right-continuous. Atoms of measures are retained in Stieltjes integrals, and an interval $(q, 1]$ excludes an atom at q .

Overview of the proof

In this section, we explain the proof of Theorem 1.1, which occupies Sections 3 to 13. For each step, we state its main estimate in the form in which it is proved, give the precise reference, and explain briefly why the estimate holds and where it is used. Section 2.11 contains a map of the proof, and Section 2.10 discusses three conventions that recur in every part of it. Throughout, $\beta > 1$ is fixed. We write $X \lesssim Y$ if $X \leq C_\beta Y$ for a constant C_β that depends only on β . Our estimates are uniform in the system size N , the interpolation parameter ρ , the tilt λ , the minimizing profile, and the positions and masses of its atoms. Much of the technical work goes into this uniformity.

2.1 The structure of the argument

We follow the interpolation scheme of [4]; what changes is how we estimate the self-overlap variance along the interpolation. The argument has four steps.

First, we reduce Theorem 1.1 to a bound on a fractional-moment error $e_{N,\beta}(\lambda)$, which compares a normalized logarithmic moment of order λ/β of the partition function with a Parisi problem constrained to have mass at least λ/β at zero. By Proposition 3.9, the mean bias divided by β is at most $2e_{N,\beta}(\lambda)$ plus a term of order λ^5 , and it suffices to prove that $e_{N,\beta}(N^{-7/50}) \lesssim N^{-7/10}$ (Section 2.2).

Second, we write $4e_{N,\beta}(\lambda)$ as the integral over $\rho \in (0, 1]$ of a defect $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ along an interpolation, evaluated at a minimizing profile of a variational problem in dimension N (Section 2.3). The terms A and J are at most of order $\log^2(2 + N)/(N\rho\lambda)$ (Section 2.4), and the difficulty is the self-overlap variance V .

Third, we bound V in two regimes. Near $\rho = 0$, we combine estimates on quantile intervals of the minimizing profile with a localization of its lowest quantiles (Section 2.5). For ρ above a cutoff ρ_0 , we bound V by a one-site cavity argument (Section 2.6), which contracts if a scalar response operator at the minimizing profile is at most $1 + \rho/4$. We prove this bound by showing that the operator is at most the identity at the Parisi measure (Section 2.7) and that the minimizing profile is close to a scalar trial profile (Section 2.8). The two regimes give the following bounds. For almost every $\rho \in [\rho_0, 1]$, we have

$$\mathcal{D}_N(\rho) \lesssim \frac{\log^2(2 + N)}{N\lambda\rho^2} \quad (2.1)$$

by Corollary 12.5, and for almost every $\rho \in (0, 1]$, we have

$$\mathcal{D}_N(\rho) \lesssim \frac{\log^2(2+N)}{N\rho\lambda} + (N\rho)^{-2/3} + (N\rho)^{-2/3}\lambda^{-1/3}\omega(\rho, \eta)^{2/3} \quad (2.2)$$

by Proposition 13.1. Here $\eta = e_{N,\beta}(\lambda) + L_s\lambda^5$ is the trial budget of Lemma 3.8, with $L_s \geq 1$ depending only on β , and $\omega(\rho, \eta) = \min\{1, (\eta/\rho)^{1/6}\}$.

Fourth, the cutoff is $\rho_0 = \min\{1, K\eta^{1/7}\}$, with K depending only on β . Note that ρ_0 is defined through the trial budget, which contains the unknown error. Integrating (2.1) over $[\rho_0, 1]$ and (2.2) over $(0, \rho_0]$, we obtain an implicit inequality whose leading term is $N^{-2/3}\eta^{1/21}$. At $\lambda = N^{-7/50}$, this inequality implies that $e_{N,\beta}(\lambda) \lesssim N^{-7/10}$ (Section 2.9).

We take the first two steps from [4], where the same scheme gives the rate $N^{-2/3}$ [4, Theorem 1.4], and our estimates near $\rho = 0$ refine those of [4, Sections 3 and 7]. We use the cavity argument, and the stability and isolation estimates that it requires, only above the cutoff.

2.2 Reduction to the fractional-moment error

As in [4, Section 1.1], we complete the Hamiltonian by an independent diagonal Gaussian term D_N , so that the completed Hamiltonian has covariance exactly $NR^2/2$, where R is the overlap (Section 3.1). For $0 < \lambda < \beta$, the fractional-moment error is

$$e_{N,\beta}(\lambda) = E_\beta(\lambda) - \psi(\lambda), \quad \psi(\lambda) = \frac{1}{N\lambda} \log \mathbb{E}_{\text{ref}} e^{\lambda(F_{N,\beta}/\beta + D_N)},$$

where $E_\beta(\lambda) = \beta^{-1}P_\beta(\lambda/\beta)$, and $P_\beta(s)$ is the minimum of the Parisi functional over probability measures μ on $[0, 1]$ with $\mu(\{0\}) \geq s$ (3.1). Talagrand showed that the normalized moments converge to this constrained value [80]; we need a rate.

The reduction is the midpoint argument of [4, Theorem 7.8, Step 1]. Since $F_{N,\beta}/\beta + D_N$ is a convex Lipschitz function of the Gaussian variables (g_{ij}, z_i) , Chen's convexity theorem [26, Theorem 1] implies that ψ extends to a convex function on $[0, \infty)$ with $\psi(0) = p_{N,\beta}/\beta$ [4, Lemma 5.2]. Using $\psi(0) \geq 2\psi(\lambda) - \psi(2\lambda)$ and $e_{N,\beta}(2\lambda) \geq 0$, we find that

$$\frac{P_\beta - p_{N,\beta}}{\beta} \leq E_\beta(0) - 2E_\beta(\lambda) + E_\beta(2\lambda) + 2e_{N,\beta}(\lambda).$$

The second difference of E_β is controlled by the cost of the prescribed mass at zero. By [4, Lemma 7.2], the Parisi profile $\gamma_\beta = \beta\alpha_\beta$ grows linearly near 0. As a consequence, the scalar trial profile $\gamma_\lambda^{\text{tr}} = \max\{\lambda, \gamma_\beta\}$ differs from it only on an interval of length of order λ , on which the first variation vanishes at γ_β . The cost of the trial profile is at most $L_s\lambda^5$ (3.17). Combining these facts, we arrive at the bound of Proposition 3.9,

$$0 \leq \frac{P_\beta - p_{N,\beta}}{\beta} \leq 32L_s\lambda^5 + 2e_{N,\beta}(\lambda) \quad (0 < \lambda \leq \lambda_0),$$

with $\lambda_0 > 0$ depending only on β . We take $\lambda = N^{-7/50}$ for N so large that $\lambda \leq \lambda_0$, and then $\lambda^5 = N^{-7/10}$. The tilt must be positive, because our bounds on the defect degenerate as $\lambda \downarrow 0$; we fix its exponent in Section 2.9 by matching λ^5 with the final bound on the error.

2.3 The interpolation and the defect

The interpolation of [4, Section 2] is a finite-dimensional, quantitative version of Guerra's interpolation [40], in the form that Talagrand used for positive fractional moments [80]. For $0 < \rho \leq 1$, let $t = 1 - \rho$, $c = \beta^2 \rho$, and $M = Nc$. An admissible profile is a deterministic probability measure $\mu = s\delta_0 + \nu$ on $[0, 1]$, where $s = \lambda/\beta$ and $\nu \geq 0$, with distribution function $a(q) = \mu([0, q])$. We solve the Parisi equation in $\mathbb{R}^N$,

$$\partial_q U = -\frac{c}{2}(\Delta U + a(q)|\nabla U|^2), \quad U(1, x; W) = \log \sum_{\sigma \in \{-1, 1\}^N} e^{\beta\sqrt{t}\sigma^T W \sigma / 2 + x \cdot \sigma},$$

and we minimize the functional $\mathcal{F}_{t,s}(\mu)$ of (3.3) over admissible profiles. At $t = 0$, the equation separates into N scalar Parisi equations, and the minimum is $P_\beta(s)$; at $t = 1$, the minimum is the completed fractional moment (3.4). The fractional-moment error is β^{-1} times the decrease of the minimum along the interpolation.

The tilted law. Each profile defines a joint law of the couplings, an auxiliary field $X_q \in \mathbb{R}^N$, $q \in [0, 1]$, and a terminal spin σ . In this law, W is tilted by $e^{sU(0,0;W)}$, the field is a diffusion with drift $ca(q)\nabla U(q, X_q; W)$ and noise coefficient $\sqrt{c}$, and σ has the Gibbs law in the external field X_1 (Section 3.3). We write $\mathcal{G}_q$ for the information revealed up to level q , m_q and C_q for the conditional mean and covariance of σ given $\mathcal{G}_q$, $H_q = \nabla^2 U(q, X_q; W)$, and $S_q = |m_q|^2/N$. By Proposition 3.1, m_q is a martingale with $dm_q = \sqrt{c}H_q dB_q$, and $a(q)C_q \preceq H_q \preceq C_q$. If μ is a minimizer, every point q of $\text{supp } \nu$, called a contact level, satisfies $\mathbb{E}S_q = q$ and $\mathbb{E} \text{Tr } H_q^2 \leq N/c$, which are the counterparts of the optimality conditions for the Parisi measure.

The defect. Let R_q be the overlap of two replicas that are conditionally independent given $\mathcal{G}_q$. By Proposition 3.6, the derivative of $\mathcal{F}_{t,s}(\mu)$ in t is $-(\beta^2/4) \int \mathbb{E}(R_q - q)^2 \nu(dq)$ for every profile. It is a square because the completed Hamiltonian has covariance $NR^2/2$, and only ν appears because the outer moment of order s cancels the prescribed mass $s\delta_0$. At a contact level, $R_q - S_q$ splits into three conditionally orthogonal pieces, which gives $\mathbb{E}(R_q - q)^2 = a_q^{\text{th}} + 2j_q + \text{Var}(S_q)$ (Lemma 3.5). Integrating against ν , we obtain Corollary 3.7, which states that at almost every ρ and at every minimizer,

$$\mathcal{D}_N(\rho) = -\frac{4}{\beta} f'_{N,s}(1 - \rho) = \beta(A + 2J + V), \quad 4e_{N,\beta}(\lambda) = \int_0^1 \mathcal{D}_N(\rho) d\rho,$$

where $f_{N,s}(t)$ is the minimum of $\mathcal{F}_{t,s}$ and

$$A = \int \frac{\mathbb{E} \operatorname{Tr} C_q^2}{N^2} \nu(\mathrm{d}q), \quad J = \int \frac{\mathbb{E}[m_q^T C_q m_q]}{N^2} \nu(\mathrm{d}q), \quad V = \int \operatorname{Var}(S_q) \nu(\mathrm{d}q).$$

The terms A and J measure conditional fluctuations of the terminal spin given $\mathcal{G}_q$, while V measures fluctuations of the information that the field has revealed by level q . Each term may depend on the minimizer, but their combination does not. We bound all three at an arbitrary minimizer.

The posterior variance bound. Let $L_q = (\sigma \cdot m_q - |m_q|^2/2)/N$, whose conditional mean given $\mathcal{G}_q$ is $S_q/2$. By Proposition 3.3, we have, at a minimizer and for every deterministic $k \in L^2(\nu)$,

$$\operatorname{Var}\left(\int k(q) L_q \nu(\mathrm{d}q)\right) \leq \frac{1}{M} \int \frac{k(q)^2}{a(q)} \nu(\mathrm{d}q), \quad A \leq \frac{1}{Ms}.$$

With the terminal spin fixed, the Gaussian coordinates that generate W and the field have a uniformly log-concave law, and the bound follows from the Brascamp–Lieb inequality [21], in the form of [4, Proposition 2.2], and the contact identities. Note that the factor $a(q)^{-1} \leq s^{-1}$ is where the prescribed mass enters. We integrate such factors through the quantile function Q of ν (3.6), which pushes Lebesgue measure on $(0, 1-s)$ forward to ν and satisfies $a(Q(v)) \geq s+v$.

The trial budget. In the energy normalization, we set $\gamma = \beta a$ and $F_\rho(\gamma) = \beta^{-1} \mathcal{F}_{1-\rho,s}(\mu)$. By Lemma 3.8, we have, for every $\rho \in (0, 1]$,

$$0 \leq F_\rho(\gamma_\lambda^{\text{tr}}) - \min_\gamma F_\rho(\gamma) \leq e_{N,\beta}(\lambda) + L_s \lambda^5 = \eta.$$

Indeed, $\mathcal{F}_{t,s}(\mu)$ and its minimum are nonincreasing in t , so that $F_\rho(\gamma_\lambda^{\text{tr}})$ is at most the scalar value of the trial profile, which is at most $E_\beta(0) + L_s \lambda^5$, while $\min F_\rho \geq \psi(\lambda) = E_\beta(\lambda) - e_{N,\beta}(\lambda)$ and $E_\beta(\lambda) \geq E_\beta(0)$. In words, the scalar trial profile has value at most η above the minimum at every ρ . This budget controls how far a minimizer can be from the trial profile (Sections 7 and 8), and it fixes the cutoff ρ_0 (Section 13). Since it contains the unknown error, the final inequality is implicit.

2.4 The thermal terms

At a contact level, by the covariance order, we have $\operatorname{Tr} H_q^2 \geq a(q) \operatorname{Tr}(H_q C_q) \geq a(q)^2 \operatorname{Tr} C_q^2$, and the contact bound on $\mathbb{E} \operatorname{Tr} H_q^2$ implies that $\mathbb{E} \operatorname{Tr} C_q^2 / N^2 \leq (M a(q)^2)^{-1}$. In the quantile variable, we have

$$A \leq \frac{1}{M} \int_0^{1-s} \frac{\mathrm{d}v}{a(Q(v))^2} \leq \frac{1}{M} \int_0^{1-s} \frac{\mathrm{d}v}{(s+v)^2} \leq \frac{1}{Ms}.$$

The mean–covariance term needs more work, because $j_q = N^{-2}\mathbb{E}[m_q^T C_q m_q]$ involves the conditional covariance in the direction m_q , which is revealed at level q itself. For a direction b that is measurable with respect to $\mathcal{G}_{q'}$ for some $q' < q$, we show by a scalar Riccati inequality that $\mathbb{E}[b^T C_q b]$ is at most $\mathbb{E}|b|^2$ divided by the accumulated clock $c \int_{q'}^q a(u)^2 du$ (Lemma 6.1). However, this bound does not apply to $b = m_q$. To overcome this obstacle, we write $m_{Q(v)}$ as a weighted average of the earlier means $m_{Q(w)}$, $w < v$, plus increments. First-order optimality implies that $\mathbb{E}S_r$ is 2-Lipschitz toward every point q of $\text{supp } \nu$, which bounds $\mathbb{E}|m_q - m_r|^2$ by $2N|q - r|$ (Lemma 6.2). We control the increments by combining this bound with Lemma 6.1 in a dyadic decomposition (Lemma 6.3). The conditional covariance in the direction of the average is controlled by the posterior variance bound. With a quadratic weight $(s + w)^2$, both contributions are at most of order $(s + v)^{-2}$. We obtain

$$j_{Q(v)} \leq \frac{341(1 + \log_+ M)^2}{M(s + v)^2}, \quad J \leq \frac{341(1 + \log_+ M)^2}{Ms}$$

(Proposition 6.4). Since $Ms = \beta N \rho \lambda$ and $M \leq \beta^2 N$, we conclude that

$$\beta(A + 2J) \leq \frac{1 + 682(1 + \log_+ M)^2}{N \rho \lambda} \lesssim \frac{\log^2(2 + N)}{N \rho \lambda}.$$

Integrated over $\rho \in [1/N, 1]$, this is at most of order $\log^3(2 + N)/(N\lambda)$, which equals $N^{-43/50} \log^3(2 + N)$ at $\lambda = N^{-7/50}$ and is smaller than $N^{-7/10}$ by a power of N .

2.5 The self-overlap variance for small ρ

The posterior variance bound controls averages of L_q over levels. However, V involves S_q at single levels, which may carry little mass of ν . Following [4, Lemma 3.3], we compare S_q with averages of L_p over windows of quantile levels next to q (Section 5).

Two comparisons. For $p \geq q$, the conditional Jensen inequality for the martingale m implies that $S_q \leq 2\mathbb{E}[L_p \mid \mathcal{G}_q]$, and for $p \leq q$, we have $\mathbb{E}[L_p \mid \mathcal{G}_q] = S_q/2 - |m_q - m_p|^2/(2N)$. Averaging over a window of quantile levels after or before q , with mean level $\bar{p}$, we show in Lemma 5.1 that, at a minimizer, $S_q - q$ is at most (for a window after q) or at least (for a window before q) the sum of $\bar{p} - q$ and twice the conditional expectation of a centered average of L_p . If the window has quantile length r and $a \geq u$ on it, this conditional expectation has second moment at most $(Mur)^{-1}$. Note that conditioning on $\mathcal{G}_q$ removes the centered part $N^{-1}(\sigma - m_q) \cdot m_q$ of L_q , which in [4, Lemma 3.3] produces additional residual terms; we bound its variance j_q separately (Section 2.4).

Variance on a quantile interval. Let $I \subset (0, 1 - s)$ be an interval of length h such that $a(Q(v)) \geq u$ for $v \in I$, let ℓ be the diameter of $Q(I)$, and set $\epsilon = (Mu)^{-1}$. For a threshold

τ , we take windows of radius of order ϵ/τ , whose random parts have second moment less than τ . A point whose two windows lie in I and for which $\text{Var}(S_{Q(v)}) > \tau$ must then see an increment of Q at least of order $\tau^{1/2}$ across its window, and since Q is nondecreasing with total increase ℓ on I , such points form a set of measure at most of order $\epsilon\ell\tau^{-3/2}$. The remaining points lie within distance of order ϵ/τ of the ends of I . Integrating $\min\{h, \epsilon\ell\tau^{-3/2}\}$ over τ , we find the main term in the bound of Proposition 5.2,

$$\int_I \text{Var}(S_{Q(v)}) dv \leq 128 \left[\min\{h, \epsilon(1 + \log_+(h/\epsilon))\} + (\epsilon\ell)^{2/3} h^{1/3} \right]. \quad (2.3)$$

The first term is at most ϵ up to a logarithm, and the interval may contain part of the quantile interval of an atom.

Dyadic shells and root localization. In the proof of Proposition 13.1, we cover $(0, 1 - s)$ by dyadic shells $I_u = \{v : u \leq s + v < 2u\}$, $u = 2^j s < 1$, on which $a(Q(v)) \geq u$, $h \leq u$, and $\epsilon = (Mu)^{-1}$. If we only use $\ell \leq 1$, the second term in (2.3) is at most $M^{-2/3}u^{-1/3}$, and summing over shells, with $\beta M^{-2/3}s^{-1/3} = (N\rho)^{-2/3}\lambda^{-1/3}$, we would find only

$$\beta V \lesssim (N\rho)^{-2/3}\lambda^{-1/3} + \frac{\log^2(2 + N)}{N\rho\lambda}.$$

The bound $\beta V \leq C(N\rho)^{-2/3}\lambda^{-1/3}$ that follows from [4, (4.2)] has the same form. At $\lambda = N^{-7/50}$, the factor $\lambda^{-1/3} = N^{7/150}$ would leave a positive power of N after the normalization of Section 2.9. To remove it, we use the root localization of Corollary 8.4, which states that at every minimizer and for every trial budget η ,

$$Q(v) \leq C_Q(s + v + \omega(\rho, \eta)), \quad \omega(\rho, \eta) = \min\{1, (\eta/\rho)^{1/6}\},$$

with C_Q depending only on β . Then $\ell(I_u) \leq 2C_Q(u + \omega)$, the second term in (2.3) is at most of order $M^{-2/3}(u^{1/3} + u^{-1/3}\omega^{2/3})$, and its sum over shells is at most of order $M^{-2/3}(1 + s^{-1/3}\omega^{2/3})$. Together with the thermal terms, this proves (2.2), that is, (13.2). Its leading term $(N\rho)^{-2/3}$ does not depend on λ , and the correction is integrable at $\rho = 0$, because $\omega^{2/3} \leq (\eta/\rho)^{1/9}$. Below $\rho = 1/N$, we rely instead on the bound $\mathcal{D}_N(\rho) \leq C(N\rho)^{-1/2}$ of [4, Lemma 4.1], recalled in (3.16). Integrating, we arrive at (13.4), which states that for $0 < \rho_0 \leq 1$,

$$\int_0^{\rho_0} \mathcal{D}_N(\rho) d\rho \lesssim N^{-1} + \frac{\log^3(2 + N)}{N\lambda} + N^{-2/3}\rho_0^{1/3} + N^{-2/3}\lambda^{-1/3}\eta^{1/9}\rho_0^{2/9}. \quad (2.4)$$

The barrier at $2/3$. For $\rho_0 = 1$, the third term in (2.4) is of order $N^{-2/3}$. This is the mechanism of [4, Proposition 7.7], in which quantile estimates [4, Lemma 3.3] and a localization of low quantiles [4, Lemma 7.6] lead to $e_{N,\beta}(\lambda) \lesssim N^{-2/3} + N^{-3/4}\lambda^{-1/2} + \log(2 + N)/(N\lambda)$ [4, (7.43)], and the tilt $\lambda = N^{-2/15}$ gives the rate $N^{-2/3}$ of [4, Theorem 1.4]. When ρ is of order one, however, $(N\rho)^{-2/3}$ is much larger than the cavity bound (2.1), as we discuss heuristically in Section 1.4. We therefore use (2.4) only below the cutoff, where it contributes $N^{-2/3}\rho_0^{1/3}$.

2.6 The cavity contraction

Let μ be a minimizer at (ρ, s) with distribution function a , let $d = \beta^2(1 - t/N)$, and let $J_{a,d}^{\text{sc}}$ be the scalar response operator of Section 2.7, which acts on a subspace $\mathcal{H}_\nu$ of $L^2(\nu)$. In Proposition 12.4, we show that if

$$J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I \quad \text{on } \mathcal{H}_\nu, \quad 0 \leq \varepsilon_N \leq \frac{\rho}{4}, \quad (2.5)$$

then

$$\frac{3\rho}{4}V \leq C_{\text{cav}}[N^{-1} + A + J + H_3], \quad H_3 = \int \mathbb{E}|S_q - q|^3 \nu(\mathrm{d}q),$$

with C_{cav} depending only on β . The argument occupies Part IV, and Section 12.1 explains how the inequality closes before the proofs begin.

Exchangeability. Let $n = N - 1$, let $m_N(p)$ be the conditional magnetization of the last spin given $\mathcal{G}_p$, and let A_p^- be the self-overlap of the conditional mean of the other n spins, which we call the rest self-overlap. Then $NS_p = m_N(p)^2 + nA_p^-$. Since the tilted law is invariant under permutations of the sites, we have $\text{Cov}(S_p, m_N(p)^2) = \text{Var}(S_p)$ (12.1), and it follows that

$$\text{Var}(S_p) = \frac{1}{N} \text{Var}(m_N(p)^2) + \frac{n}{N} \text{Cov}(m_N(p)^2, A_p^-).$$

Up to an error of order N^{-1} , the variance V is the integrated covariance between the squared magnetization of one spin and the self-overlap of the others. Exchangeability is used in the same way in the cavity method of [81, Section 1.6].

The cavity interpolation. For $u \in [0, 1]$, we give each coupling between the last spin and the others variance $\alpha u/n$, and the last coordinate of the field variance coefficient $b(u) = d - \alpha u$, where $\alpha = \beta^2 t n / N$ and $d = c + \alpha$ (10.1). At $u = 1$, we recover the tilted law of the full system. At $u = 0$, the last spin is independent of the others: the law is a product of the tilted law of an n -spin system, with the same profile, and the scalar law of Section 9 with coefficient d . In particular, $\mathcal{C}_{pp}(u) = \text{Cov}_u(m_N(p)^2, A_p^-)$ vanishes at $u = 0$. By Taylor's formula, we have

$$\mathcal{C}_{pp}(1) = \mathcal{C}'_{pp}(0) + \int_0^1 (1 - u) \mathcal{C}''_{pp}(u) \mathrm{d}u.$$

Note that the n -spin system at $u = 0$ carries the profile of the full system, which need not minimize its own functional.

Tree laws and the remainder. Derivatives in u produce products of conditional expectations at several levels. We represent these products by laws of finitely many descendants

on marked trees, in which deleting a leaf does not change the law of the others (Section 10.2). Normalized replica trees of this kind were developed in [54, Section 2]. By Lemma 10.5, the derivative in u of a tree moment is $\alpha/2$ times a combination of moments with one more overlap defect $D_{ab} = \epsilon_a \epsilon_b (R_{ab}^- - q_{ab})$, plus a term from the normalization of the tilt. The row contributes the rest overlap R_{ab}^- of two leaves, and the decrease of $b(u)$ subtracts the level q_{ab} at which they split. It follows that non-negative tree moments at different values of u agree up to a factor $K_{\beta,k}$ that depends only on β and the number k of leaves (Corollary 10.6), and that every term of $\mathcal{C}_{pp}''$ contains three defects. By Lemmas 10.7, 10.8, and 12.3, the Taylor remainder is at most $C_\beta \sup_u \mathcal{M}_3(u)$, where $\mathcal{M}_3(u) = \int \mathbb{E}_{2,q,u} |R_{12}^- - q|^3 \nu(dq)$, and this supremum is at most $K_{\beta,2}[64(A+J) + 16H_3 + 32N^{-3}]$.

The linear term. Along the interpolation, the fields acting on the last spins of two leaves have covariance $dq_{ab} + \alpha u(R_{ab}^- - q_{ab})$ (Section 12.1). In the scalar law, if we replace the covariance $d(q \wedge r)$ of the field by $dT(q \wedge r)$ with $T(q) = q + \theta\varphi(q)$, the mean square magnetization at level p changes at the rate $(J_{a,d}\varphi)(p)$ (Lemma 9.6). Heuristically, the row acts on the last spin as a random transport of levels with displacement $(\alpha u/d)(A_q^- - q)$, and we expect the main term of $\mathcal{C}_{pp}'(0)$ to be $(\alpha/d)(J_{a,d}^{\text{sc}}\Sigma^-)(p,p)$, where $\Sigma^-(p,q) = \text{Cov}^-(A_p^-, A_q^-)$ is the covariance of the rest self-overlap at $u = 0$. In Lemma 11.2, we prove the exact identity

$$\int \mathcal{C}_{pp}'(0) \nu(dp) = \alpha \left[\frac{1}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-) - \frac{1}{2}\mathcal{P} + \int T_p \nu(dp) \right], \quad \mathcal{P} \geq 0.$$

The term $-\mathcal{P}/2$ comes from the difference between conditional and unconditional covariances of the rest self-overlap; we discard it. The errors T_p raise an obstacle: they contain covariances of centered variables of the rest system with A_p^- , and if we applied the Cauchy–Schwarz inequality at a single level, the bound would involve $\text{Var}^-(A_p^-)$, whose integral is what we want to control. To overcome this obstacle, we first integrate in p against the deterministic kernels of the scalar law, so that the covariance order and the bound (10.4) on the variance of ν -averages of A^- apply. After a comparison with the full system through Corollary 10.6, we find in Lemma 11.3 that

$$\alpha \left| \int T_p \nu(dp) \right| \leq 192\beta^2 K_{\beta,3} [A + J + N^{-2}].$$

The factor $1 - \rho$. Combining these identities, we obtain the trace identity of Lemma 12.2,

$$V = \frac{V^{\text{site}}}{N} + \frac{n}{N} \left[\frac{\alpha}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-) - \frac{\alpha}{2}\mathcal{P} + \alpha \int T_p \nu(dp) + \mathcal{R}_{\text{cav}} \right],$$

where $0 \leq V^{\text{site}} \leq 1/4$ and $\mathcal{R}_{\text{cav}}$ is the Taylor remainder. The operator Σ^- is positive, with range in $\mathcal{H}_\nu$ and trace $V^- = \int \text{Var}^-(A_p^-) \nu(dp)$. Under (2.5), the trace $\text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-)$

is therefore at most $(1 + \varepsilon_N)V^-$. The contraction comes from the ratio

$$\frac{\alpha}{d} = t \frac{N-1}{N-t} \leq t = 1 - \rho$$

(12.3). Indeed, the field acting on the last spin has variance coefficient d , and at $u = 1$, the part $c = \beta^2 \rho$ of it comes from the auxiliary field, whose covariance is deterministic; only the part α that comes from the couplings feels the fluctuations of the rest overlap. We conclude that $V \leq (1 - \rho)(1 + \varepsilon_N)V^- + \text{errors}$.

Comparison with coefficient one. It remains to compare V^- with V , and the comparison must have coefficient one. If we compared the endpoints through Corollary 10.6, we would obtain only $V^- \leq KV + \text{errors}$ with some $K > 1$, and the inequality $V \leq (1 - \rho)K(1 + \varepsilon_N)V + \text{errors}$ bounds V only if $(1 - \rho)K(1 + \varepsilon_N) < 1$, which fails for $\rho \leq 1 - 1/K$. In Lemma 12.1, we prove instead that

$$V^- \leq V + C_\beta \sup_{0 \leq u \leq 1} \mathcal{M}_3(u) + \frac{1}{4n^2}.$$

For each p , the second moment $\mathbb{E}_u(A_p^- - p)^2$ is the mean of a product of two overlap defects on four leaves that split at p . Its derivative in u contains three defects, and after integration against ν , the values at $u = 0$ and $u = 1$ differ by at most $C_\beta \sup_u \mathcal{M}_3(u)$. At $u = 0$, it bounds $\text{Var}^-(A_p^-)$, even though we do not know that $\mathbb{E}^- A_p^- = p$. At $u = 1$, exchangeability and the contact identity $\mathbb{E} S_p = p$ show that it is at most $\text{Var}(S_p) + 1/(4n^2)$.

Closing. As $(1 - \rho)(1 + \rho/4) \leq 1 - 3\rho/4$, these bounds imply that $V \leq (1 - 3\rho/4)V + C_\beta[N^{-1} + A + J + H_3]$, which is Proposition 12.4. By Corollary 5.3, if $M \geq 1$, then

$$H_3 \leq zV + \frac{425}{Ms} \log(2M) \log \frac{2}{z} \quad (0 < z \leq 1/4).$$

We take z proportional to ρ and absorb zV in the margin $3\rho/4$. Inserting the bounds on A and J and using $Ms = \beta N \rho \lambda$, we arrive at Corollary 12.5, which states that at every minimizer that satisfies (2.5), we have

$$\beta(A + 2J + V) \leq C_\beta \frac{\log^2(2 + N)}{N \lambda \rho^2} \quad (1/N \leq \rho \leq 1).$$

One factor ρ^{-1} comes from the margin, and the other from Ms .

The framework of differentiation, comparison, and contraction on one site is that of [81, Section 1.6]. A one-site expansion with an independent scalar endpoint also appears in [56, Lemma 2.2], where the resulting linear equation is inverted under a gap condition on the scalar response [56, Proposition 5.1]; we compare the two arguments in Section 1.6. We use no gap condition. The multiplier that dominates our scalar response equals one on the support of the Parisi measure, and the contraction comes from the ratio α/d .

2.7 Stability of the scalar response

At the independent endpoint, the last spin follows the scalar law with the profile a_N of the N -spin minimizer and the coefficient $d_N = \beta^2(1 - t/N)$, not with the Parisi measure. In Section 9, we verify (2.5) with an explicit ε_N .

The response operator. For a profile a and $d \in [\beta^2/2, \beta^2]$, let $u = u_{a,d}$ solve the scalar Parisi equation with coefficient d (9.1), and let Y be its diffusion. With $m_q = \partial_x u(q, Y_q)$ and $h_q = \partial_{xx} u(q, Y_q)$, we define

$$\Gamma_{a,d}(q) = \mathbb{E}^0 m_q^2, \quad k_{a,d}(q) = d \mathbb{E}^0 h_q^2,$$

so that $\Gamma'_{a,d} = k_{a,d}$ (9.6); we call $k_{a,d}$ the scalar multiplier. By Lemma 9.6, the operator $J_{a,d}$ of (9.19) is the derivative of $\Gamma_{a,d}$ under transports of covariance levels: if we replace the covariance $d(q \wedge r)$ of the field by $dT_\theta(q \wedge r)$, where $T_\theta(q) = q + \theta\varphi(q)$ and $\varphi(0) = \varphi(1) = 0$, the mean square magnetization at level p has derivative $(J_{a,d}\varphi)(p)$ at $\theta = 0$. The scalar response operator $J_{a,d}^{\text{sc}}$ is the compression of $J_{a,d}$ to the space $\mathcal{H}_\nu$ of functions in $L^2(\nu)$ that vanish at 0 and 1 when these are atoms of ν . Its kernel has no sign, and a direct estimate bounds its norm only by C_β .

Domination by the multiplier. The following observation reduces the operator bound to a pointwise bound. In Lemma 9.7, we show that, for every profile, $J_{a,d}^{\text{sc}}$ is self-adjoint and

$$J_{a,d}^{\text{sc}} \preceq M_{k_{a,d}}, \quad (M_{k_{a,d}}v)(q) = k_{a,d}(q)v(q).$$

To prove it, we follow the scalar Parisi functional $\mathcal{P}_d$ along a transport of levels, $\Pi(\theta) = \mathcal{P}_d(a \circ T_\theta^{-1})$. By Lemma 9.6, we have $\Pi''(0) = (d/2)\langle \varphi, (I - J_{a,d})\varphi \rangle$ (9.33); the first derivative is the first variation under transports of levels due to Talagrand [79] (see [9, Section 3, (19)]). Since the functional is convex along segments of distribution functions [10], its increment along the transport is at least its first variation at a , applied to $a \circ T_\theta^{-1} - a$. We expand this lower bound to second order in θ , which produces the integral of $(1 - k_{a,d})\varphi^2$ against ν . Comparing the two expansions, we find that $\langle \varphi, J_{a,d}\varphi \rangle \leq \langle \varphi, k_{a,d}\varphi \rangle$. Note that Π need not be convex, because $\theta \mapsto a \circ T_\theta^{-1}$ is not affine; the term $1 - k_{a,d}$ accounts for this. It therefore suffices to show that $\sup_q k_{a_N, d_N}(q) \leq 1 + \varepsilon_N$.

The multiplier at the Parisi measure. In Lemma 9.8, we prove that

$$k_{\alpha_\beta, \beta^2}(q) = 1 \quad (0 \leq q \leq q_\beta), \quad k_{\alpha_\beta, \beta^2}(q) \leq e^{-\beta^2(q - q_\beta)/2} \quad (q_\beta \leq q \leq 1).$$

On the support, we rely on an identity of Talagrand [79], in the form of [9, Theorem 5] and [48, Lemma 2.6], which states that $\Gamma(q) = q$ on $\text{supp } \mu_\beta$. As the support is the interval $[0, q_\beta]$ [48, Theorem 1.1], we may differentiate this identity along it, which gives $k = \Gamma' = 1$. Above the support, we have $\alpha_\beta = 1$, and the solution u is explicit. A

transformed density z_q of Y_q evolves there by the heat equation [48, Lemma 2.4(ii)], and z_{q_β} is even and log-concave [48, Proposition 4.1]. In particular, z_{q_β} is decreasing on $(0, \infty)$. The multiplier is an explicit multiple of the integral of z_q against sech^3 (9.37), and Gaussian smoothing of an even decreasing density does not increase its integral against an even decreasing function; this gives the exponential bound. We need it on all of $[0, 1]$, because a finite-dimensional minimizer may place mass above q_β .

Marginality. At the Parisi measure, the multiplier equals one on the whole support, and the domination gives no contraction there. The cavity inequality contracts because of the ratio $\alpha/d \leq 1 - \rho$, which leaves a margin of order ρ only. This is why we need $\varepsilon_N \leq \rho/4$ in (2.5), and why stability holds only above a cutoff. The equality $k = 1$ on the support expresses, in scalar form, the marginal stability that De Dominicis and Kondor found in a truncated replica model near the critical temperature [35, pp. 606–608]; see Section 1.4.

Transfer to the minimizer. By Lemma 9.3, the multiplier depends Lipschitz continuously on the profile and coefficient, in the sense that

$$\sup_{0 \leq q \leq 1} |k_{a,d}(q) - k_{\tilde{a},\tilde{d}}(q)| \leq C_\beta (\|a - \tilde{a}\|_{L^1(0,1)} + |d - \tilde{d}|).$$

We have $|d_N - \beta^2| \leq \beta^2/N$ and $\|\gamma_\lambda^{\text{tr}} - \gamma_\beta\|_{L^1(0,1)} \lesssim \lambda^2$ (9.39). Combining these facts with the isolation estimate of Section 2.8, we obtain Proposition 9.9, which states that on the space $\mathcal{H}_{\nu_N}$ of the minimizer itself,

$$J_{a_N, d_N}^{\text{sc}} \preceq (1 + \varepsilon_N)I, \quad \varepsilon_N \leq C_{\text{st}} [\rho^{-3/4} \eta^{1/4} + \lambda^2 + N^{-1}],$$

with C_{st} depending only on β . We compare only the multipliers of the two profiles, and the operator inequality comes from Lemma 9.7 at a_N .

2.8 Localization of the minimizing profile

We need two bounds on the distance between a minimizer γ_* of F_ρ and the scalar trial profile γ^{tr} , in terms of a trial budget η : a bound in $L^1(0, 1)$ for the stability of the scalar response (Section 7), and a localization of the lowest quantiles for small ρ (Section 8).

Convexity and response. From [4, Section 7.4 and Appendix A], we import the following facts (Proposition 7.1). The functional F_ρ is convex along segments $\gamma_\theta = (1 - \theta)\gamma_0 + \theta\gamma_1$. Let $f(\theta) = F_\rho(\gamma_\theta)$ and $\zeta_\theta(q) = \beta^2 N^{-1} \mathbb{E}_{\gamma_\theta} \text{Tr } H_q^2$, so that $\rho\zeta_\theta$ is the derivative of the expected self-overlap. The second variation controls how the self-overlap responds to $\gamma_1 - \gamma_0$:

$$\rho \left| \int_0^1 (\gamma_1 - \gamma_0) \zeta_\theta \varphi \, dq \right| \leq \left(\frac{f''(\theta)}{\rho} \right)^{1/2} [Y^{1/2} + 2(Y + Z)^{1/2}]$$

(7.3), for Lipschitz cutoffs φ , where Y is a weighted Dirichlet energy of φ and Z is the integral of $\rho\varphi^2\zeta_\theta$ against $d\gamma_\theta$. Further, $\zeta_\theta \geq k_\beta^2$ for a constant $k_\beta > 0$ (7.4). The midpoint inequality (7.9) bounds the integral of $\min\{\theta, 1 - \theta\}f''(\theta)$ over $(0, 1)$ by $F_\rho(\gamma_0) + F_\rho(\gamma_1) - 2\min F_\rho$. If both endpoints are nearly minimal, the segment has little curvature, while a difference of one sign and of definite size on some region forces curvature (Lemma 7.4).

Isolation in L^1 . A minimizer may have atoms, and Z contains $d\gamma_\theta$ on the support of the cutoff. To overcome this difficulty, we first spread the free mass of γ_* on a scale h by a reflected regularization (Lemma 7.3). It creates a density at most β/h at a cost at most $\rho\beta h^2/12$; the cost is quadratic because the first-variation potential is semiconcave, as in [4, Lemma 7.5]. Along the segment from the regularized minimizer to γ^{tr} , we use cutoffs of one sign built from one-sided averages of superlevel sets of the difference, away from the atom of γ^{tr} at q_β ; below q_β , the trial profile is Lipschitz (Lemma 7.5). Using the layer-cake formula, we find that $f'' \gtrsim \rho^2 \min\{\rho D^4, hD^2\}$, where $D = \|\gamma_* - \gamma^{\text{tr}}\|_{L^1(0,1)}$ (7.20). For h proportional to ρD^2 , we have $f'' \gtrsim \rho^3 D^4$, and the smoothing cost is a small fraction of $\rho^3 D^4$. The midpoint inequality implies that $\rho^3 D^4 \lesssim \eta$, which is the bound of Proposition 7.6,

$$\|\gamma_* - \gamma^{\text{tr}}\|_{L^1(0,1)} \leq C \min\{1, \rho^{-3/4} \eta^{1/4}\},$$

for every minimizer and every trial budget, with C depending only on β .

Root localization. In the isolation argument, the lower bound $\rho \int \zeta_\theta \varphi \geq \rho k_\beta^2 \int \varphi$ on the response carries a factor ρ , which is too costly near $\rho = 0$. To avoid it, we use the local deficit certificate of Lemma 8.1. Suppose that two nearly minimal profiles have bounded densities on an interval I of length ℓ , that the density of the second is also bounded below, and that the second exceeds the first by a definite amount on I . Then the certificate bounds η/ρ from below by an explicit expression in ℓ , the gap, and the density bounds. In its proof, we move mass of the profile toward the middle of I along a bump. Since both endpoints are nearly minimal and F_ρ is convex, the response integrated against the bump must be at least of order ℓ , without the factor ρ . To localize a quantile $Q_{\gamma_*}(v)$, we bound its truncation $q = \min\{Q_{\gamma_*}(v), T_1\}$ at the level T_1 of Lemma 8.2, which depends only on β ; this suffices, because $Q_{\gamma_*}(v) \leq q/T_1$. If q is large compared with $\lambda + v$, we regularize the free mass of the minimizer below q only. The result is small on $(q/4, 3q/4)$, where the trial profile has density at least $k_0/2$ (Lemma 8.2) and exceeds it by an amount of order q . Using the certificate, we find that $\eta/\rho \gtrsim \min\{q^6, q^4/(m/h + C)\}$, where m/h is the density created by the regularization. With a width adapted to the moved mass, we obtain Proposition 8.3, which states that

$$Q_{\gamma_*}(v) \leq C[\lambda + v + \min\{1, (\eta/\rho)^{1/6}\}] \quad (0 < v < \beta - \lambda),$$

together with its form (8.11) in the probability normalization. We follow the strategy of [4, Lemma 7.6]. Compared with [4, (7.30)], the bound depends on η and ρ only through η/ρ , and the certificate is what gives this dependence.

2.9 Closing the implicit inequality

By Proposition 9.9, the stability assumption (2.5) holds as soon as $C_{\text{st}}[\rho^{-3/4}\eta^{1/4} + \lambda^2 + N^{-1}] \leq \rho/4$. Recall that $\eta \geq L_s \lambda^5 \geq \lambda^5$, so that $\eta^{1/7} \geq \lambda^{5/7} = N^{-1/10}$, which is at least λ^2 and N^{-1} . The first term satisfies $\rho^{-3/4}\eta^{1/4} \lesssim \rho$ exactly when $\rho \gtrsim \eta^{1/7}$. With $K = 12C_{\text{st}}$, we set

$$\rho_0 = \min\{1, K\eta^{1/7}\}$$

(13.10). If $\rho_0 < 1$, then $\varepsilon_N \leq \rho/4$ for every $\rho \in [\rho_0, 1]$ and every minimizer, and (2.1) applies; if $\rho_0 = 1$, the stable interval is empty. Using $\rho_0 \geq \lambda^{5/7}$, we find that

$$\int_{\rho_0}^1 \mathcal{D}_N(\rho) \, d\rho \lesssim \frac{\log^2(2+N)}{N\lambda\rho_0} \lesssim \frac{\log^2(2+N)}{N\lambda^{12/7}} = N^{-19/25} \log^2(2+N).$$

On $(0, \rho_0]$, we insert $\rho_0 \leq K\eta^{1/7}$ into (2.4), which gives $\rho_0^{1/3} \lesssim \eta^{1/21}$ and $\eta^{1/9}\rho_0^{2/9} \lesssim \eta^{1/7}$. Adding the two contributions and recalling that $4e_{N,\beta}(\lambda) = \int_0^1 \mathcal{D}_N(\rho) \, d\rho$, we obtain the bound (13.12),

$$e_{N,\beta}(\lambda) \lesssim N^{-19/25} \log^3(2+N) + N^{-2/3}\eta^{1/21} + N^{-2/3}\lambda^{-1/3}\eta^{1/7}. \quad (2.6)$$

But there is an apparent circularity here, since the cutoff ρ_0 is defined through $\eta = e_{N,\beta}(\lambda) + L_s \lambda^5$, which contains the quantity that we want to bound. The way out is that the error enters (2.6) only through powers of η smaller than one. Let $r = N^{-7/10} = \lambda^5$ and $x = \eta/r$, which is finite and at least $L_s \geq 1$. Dividing (2.6) by r and adding L_s to both sides, we find (13.13),

$$x \leq L_s + C_\beta [N^{-3/50} \log^3(2+N) + x^{1/21} + N^{-1/50} x^{1/7}] \leq C_\beta x^{1/7},$$

which implies that $x \leq C_\beta^{7/6}$ and $e_{N,\beta}(\lambda) \leq \eta = rx \lesssim N^{-7/10}$. We need no preliminary bound on the error. The proof of [4, Proposition 7.7] closes an inequality in which the unknown error enters the trial budget in the same way. By Proposition 3.9, we conclude that $P_\beta - p_{N,\beta} \lesssim N^{-7/10}$ for all large N , and the finitely many remaining values of N only change the constant.

The exponent. If η is of order N^{-a} , the leading term $N^{-2/3}\eta^{1/21}$ in (2.6) is of the same order as η when

$$a = \frac{2}{3} + \frac{a}{21}, \quad \text{that is,} \quad a = \frac{7}{10},$$

Term	Origin	After division by $N^{-7/10}$
N^{-1}	$\rho \leq 1/N$, by (3.16)	$N^{-3/10}$
$N^{-43/50} \log^3(2+N)$	thermal and logarithmic terms on $[1/N, \rho_0]$, (13.2)	$N^{-4/25} \log^3(2+N)$
$N^{-19/25} \log^2(2+N)$	stable interval $[\rho_0, 1]$, Corollary 12.5	$N^{-3/50} \log^2(2+N)$
$N^{-2/3} \eta^{1/21}$	quantile shells on $(0, \rho_0]$, Proposition 5.2	$x^{1/21}$
$N^{-2/3} \lambda^{-1/3} \eta^{1/7}$	root localization correction, Corollary 8.4	$N^{-1/50} x^{1/7}$

Table 2.1. The terms of the bound on $e_{N,\beta}(\lambda)$ at $\lambda = N^{-7/50}$, with $x = \eta/N^{-7/10}$. The first three terms are combined into one in (2.6); see Section 13.2.

Quantity	Where it comes from	Statement
λ^5	cost of the prescribed mass at zero, at the scalar end	(3.17), [4, Lemma 7.2]
$(N\rho)^{-2/3}$	variance on dyadic quantile shells, after root localization	Proposition 5.2, (13.2)
$(\eta/\rho)^{1/6}$	localization of the lowest quantiles	Proposition 8.3, Corollary 8.4
$\rho^{-3/4} \eta^{1/4}$	L^1 isolation of the minimizing profile	Proposition 7.6
$\rho/4$	admissible error in the scalar response; it leaves the margin $3\rho/4$	(12.10)
$\eta^{1/7}$	cutoff: $\rho^{-3/4} \eta^{1/4} \lesssim \rho$ when $\rho \gtrsim \eta^{1/7}$	(13.10)
$(N\lambda\rho^2)^{-1}$	cavity bound: margin $3\rho/4$ and $Ms = \beta N\rho\lambda$	Corollary 12.5
$\rho_0^{1/3}$	integral of $(N\rho)^{-2/3}$ over $(0, \rho_0]$	(13.4)
$\eta^{1/21}$	$\rho_0^{1/3}$ with $\rho_0 \leq K\eta^{1/7}$	(13.12)
$a = 7/10$	balance $a = 2/3 + a/21$	Remark 13.2
$\lambda = N^{-7/50}$	$\lambda^5 = N^{-7/10}$	(13.8)

Table 2.2. The exponents of the proof.

and the cost λ^5 is of the same order when $\lambda = N^{-7/50}$ (Remark 13.2). Table 2.1 lists the terms of (2.6) after division by $N^{-7/10}$. Only the leading term has no polynomial slack; the others carry the factors $N^{-1/50}$, $N^{-3/50}$, $N^{-4/25}$, and $N^{-3/10}$, up to logarithms. After the closure, we have $L_s N^{-7/10} \leq \eta \lesssim N^{-7/10}$, and the cutoff is then comparable to $N^{-1/10}$ for large N . Table 2.2 collects the exponents of the proof and the statements in which they arise.

What would improve the exponent. The exponent $1/21$ is the product of the exponent $1/3$ of ρ_0 in the integrated shell estimate and the exponent $1/7$ of η in the cutoff, which comes from the isolation estimate. A larger exponent in either place would improve the

balance. A smaller cutoff, however, increases the contribution $\log^2(2 + N)/(N\lambda\rho_0)$ of the stable interval, and such an improvement requires rechecking the other terms of Table 2.1. If the cavity bound (2.1) held down to ρ of order $N^{-1/4}$, and if we ignored logarithms and the tilt, the two regimes would give $N^{-3/4}$; we discuss this heuristic in Section 1.4.

2.10 Minimizers, arbitrary profiles, and atoms

Three conventions recur throughout the proof.

Which statements use minimality. Estimates that use the contact identities (3.9) or the trial budget hold at minimizers: the main results of Sections 5 to 8, Step 2 of the proof of Lemma 12.1, and everything built on them. The Gaussian tools of Section 4, the lemmas of Section 9 on the scalar law of an arbitrary profile, and the cavity calculus of Sections 10 and 11 hold for every deterministic profile (with $\nu(\{1\}) = 0$ in Section 11). The reference estimate Lemma 9.8 concerns the Parisi measure μ_β itself, and Proposition 9.9 concerns minimizing profiles. We need this generality, since at the independent endpoint, the n -spin system carries the profile of the full system.

Every minimizer. A minimizer need not be unique, and it may depend on N , ρ , and λ , but not on the couplings. Since (3.15) holds at every minimizer, and since our bounds on A , J , and V hold at every minimizer with the same constants, we can integrate in ρ without a measurable choice of minimizers.

Atoms. Minimizing profiles may have atoms in $[0, 1)$, and ν may charge 0, but $\nu(\{1\}) = 0$ by Proposition 3.1. Distribution functions are right-continuous, so that $a(q)$ includes an atom at q while $(q, 1]$ excludes it. We compute integrals against ν through quantiles, and a quantile window may contain part of an atom (Section 5.1). We handle scalar and tree laws through exact path likelihoods (Lemmas 9.2 and 10.3) and through the heat identity of Lemma 4.3, in which profile measures are held fixed with their atoms. In Sections 7 and 8, we spread the free mass before using cutoffs.

2.11 A map of the proof

Table 2.3 lists, for each section, its main results, its main inputs, and the places where its results are used. The proof of Theorem 1.1 in Section 13.2 begins with a table of its inputs and of the parameters at which they are used. In terms of these results, the proof has the following structure.

- Proposition 3.9, Corollary 3.7, and Lemma 3.8 reduce the theorem to a bound on the integral of $\mathcal{D}_N$, evaluated at any minimizer, and give the trial budget η that defines the cutoff ρ_0 .

- On $(0, \rho_0]$, Proposition 13.1 bounds the defect, using Proposition 5.2, Corollary 8.4, (3.10), Proposition 6.4, and (3.16).
- On $[\rho_0, 1]$, Corollary 12.5 bounds the defect, using Proposition 12.4, Corollary 5.3, and Proposition 6.4. The contraction of Proposition 12.4 combines Lemmas 10.8, 11.2, 11.3, and 12.1 to 12.3, and its stability hypothesis (12.10) is verified by Proposition 9.9, which combines Lemmas 9.3, 9.7, and 9.8 and Proposition 7.6.
- The closure in Section 13.2 combines the two bounds.

Sec.	Main results	Main inputs	Used in
3	defect identity (3.15); contact and variance bounds, Propositions 3.1 and 3.3; trial budget, Lemma 3.8; reduction, Proposition 3.9	[4, Section 2, Lemma 7.2, Theorem 7.8], [26]	all later sections
4	posterior bound for arbitrary profiles, Lemma 4.2; heat identity, Lemma 4.3	[4, Section 2], [54, Lemma 2.3]	Sections 9 to 11
5	variance on quantile intervals, Proposition 5.2; cubic moment, Corollary 5.3	Propositions 3.1 and 3.3	Proposition 13.1 and Corollary 12.5
6	mean–covariance term, Proposition 6.4	Propositions 3.1 and 3.3	Proposition 13.1 and Corollary 12.5
7	isolation, Proposition 7.6	[4, Section 7.4, Appendix A], [48, Theorem 1.1], Lemma 3.8	Section 8, Proposition 9.9
8	root localization, Proposition 8.3 and Corollary 8.4	Proposition 7.1 and Lemmas 7.3 and 7.4	Proposition 13.1
9	response identity, Lemma 9.6; domination, Lemma 9.7; reference multiplier, Lemma 9.8; stability, Proposition 9.9	Section 4, Proposition 7.6, [10, 48]	Lemma 11.2 and Proposition 12.4; proof of Theorem 1.1
10	tree calculus, Lemma 10.5 and Corollary 10.6; remainder, Lemmas 10.7 and 10.8	Section 4, scalar law of Section 9	Sections 11 and 12
11	linear response, Lemma 11.2; errors, Lemma 11.3	Section 10, (9.19), Lemma 4.3	Lemma 12.2 and Proposition 12.4
12	coefficient one, Lemma 12.1; trace identity, Lemma 12.2; contraction, Proposition 12.4; defect, Corollary 12.5	Sections 10 and 11, Lemma 9.7, Corollary 5.3 and Proposition 6.4	proof of Theorem 1.1
13	small ρ , Proposition 13.1; Theorem 1.1	Sections 3, 5, 6, 8, 9, and 12	—

Table 2.3. The sections of the proof, their main results and inputs, and where the results are used.

PART I

The interpolation

The interpolation and the defect

In this section, we recall the interpolation of [4, Section 2], which compares the fractional moments of the partition function with the constrained scalar Parisi problem. We also fix the normalizations, the tilted law, and the observables that we use in the rest of the paper. The later sections build on two results. The first is the defect identity (3.15) of Corollary 3.7. It states that four times the fractional-moment error $e_{N,\beta}(\lambda)$ is the integral over the interpolation of a defect $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$, whose three terms are overlap observables at a minimizing profile. The second is Proposition 3.9, where we show that $(P_\beta - p_{N,\beta})/\beta$ is at most $2e_{N,\beta}(\lambda)$ plus a term of order λ^5 . Along the way, we record the martingale and contact identities in Proposition 3.1, the integrated variance bound in Proposition 3.3, the decomposition of the overlap error in Lemma 3.5, the interpolation identity in Proposition 3.6, and the trial budget in Lemma 3.8.

Most of these facts are proved in [4]. We state each of them in full, with its location in [4]. Where the argument is short, we also sketch the proof, so that this section can be read without [4] at hand. Besides [4], this section uses only the scalar Parisi functional (1.1) and the notation of Section 1.1, together with the support theorem [48, Theorem 1.1]. Its results are used throughout the paper. Propositions 3.1 and 3.3 and Lemma 3.5 are used in Sections 5, 6, 10, and 12; the energy normalization and the trial budget in Sections 7 to 9; and (3.15), (3.16), and Proposition 3.9 in Sections 12 and 13. We treat the laws with anisotropic fields in Section 4, where we also prove the two Gaussian tools that we need for profiles that do not minimize the functional.

Throughout this section, we fix $\beta > 1$ and $N \geq 1$, and we assume no relation between the tilt λ and N . The completion, the interpolating functional, and the identities of Sections 3.3 to 3.6 hold for every $\beta > 0$. We use the hypothesis $\beta > 1$ for the support of the Parisi measure in Section 3.1 and for the cost of the scalar floor in Section 3.7.

The interpolation of [4, Section 2] is a quantitative form of the Guerra interpolation that Talagrand used for the positive fractional moments of the partition function, whose limits he determined in [80]; see [4, Section 1.2]. For fixed $0 < s < 1$, the limit of the normalized logarithmic moment of order s is the minimum of the Parisi functional over probability measures with mass at least s at zero. For the SK model at zero field, this formula is also contained in the one-level case of [31, Theorem 2.1], and it is stated in this form in [24, Theorem 2.1], in a normalization of the spins that changes the functional by the constant $\log 2$. The finite- N functional of Section 3.2 carries the same constraint, in

the sense that its profiles have the prescribed mass s at 0.

3.1 Completion and the constrained scalar value

Let W be the symmetric $N \times N$ matrix with entries $W_{ij} = g_{ij}/\sqrt{N}$ for $i < j$ and $W_{ii} = \sqrt{2/N} z_i$, where $z_1, \dots, z_N$ are standard Gaussian variables, independent of each other and of the couplings (g_{ij}) . We write $\mathbb{E}_{\text{ref}}$ for expectation over (g_{ij}, z_i) under this Gaussian law; the subscript distinguishes it from the tilted law of Section 3.3. Following [4, Section 1.1], we set

$$D_N = \frac{1}{2} \text{Tr } W, \quad F_{N,\beta}^c = \log \sum_{\sigma \in \{-1,1\}^N} e^{\beta \sigma^T W \sigma / 2},$$

and we recall that $F_{N,\beta} = \log \sum_{\sigma} e^{\beta H_N(\sigma)}$ is the logarithmic partition function of Section 1.1. Note that $\sigma^T W \sigma / 2 = H_N(\sigma) + D_N$, because $\sigma_i^2 = 1$, so that $F_{N,\beta}^c = F_{N,\beta} + \beta D_N$. The variable $D_N = (2N)^{-1/2} \sum_i z_i$ is centered Gaussian with variance $1/2$ and is independent of $F_{N,\beta}$.

The diagonal term completes the covariance of the Hamiltonian to an exact function of the overlap $R(\sigma, \sigma') = N^{-1} \sigma \cdot \sigma'$. Indeed, writing $\tau_i = \sigma_i \sigma'_i$, we have

$$\mathbb{E}_{\text{ref}}[H_N(\sigma) H_N(\sigma')] = \frac{1}{N} \sum_{i < j} \tau_i \tau_j = \frac{1}{2N} \left[\left(\sum_{i=1}^N \tau_i \right)^2 - N \right] = \frac{N}{2} R(\sigma, \sigma')^2 - \frac{1}{2},$$

and adding $\text{Var}(D_N) = 1/2$, we find that the completed Hamiltonian $\sigma^T W \sigma / 2$ has covariance $NR(\sigma, \sigma')^2/2$. This is why the derivative of the interpolation in Proposition 3.6 is exactly an integral of squared overlap deviations. Because D_N is independent of $F_{N,\beta}$, the completion changes the logarithmic moments only by an explicit term. For $s > 0$, the normalized logarithmic moment of the completed partition function is [4, (1.3)]

$$\phi_{N,\beta}^c(s) = \frac{1}{Ns} \log \mathbb{E}_{\text{ref}} e^{s F_{N,\beta}^c} = \frac{1}{Ns} \log \mathbb{E}_{\text{ref}} e^{s F_{N,\beta}} + \frac{\beta^2 s}{4N},$$

where we used $\log \mathbb{E}_{\text{ref}} e^{s \beta D_N} = \beta^2 s^2 / 4$. As $s \downarrow 0$, we have $\phi_{N,\beta}^c(s) \rightarrow p_{N,\beta}$, since $F_{N,\beta}$ has exponential moments of every order.

Recall the scalar Parisi functional $\mathcal{P}_\beta$ of (1.1), which we evaluate at the distribution function of a probability measure on $[0, 1]$. For $0 \leq s \leq 1$ and $0 \leq \lambda \leq \beta$, we define

$$P_\beta(s) = \min_{\mu \in \mathcal{P}([0,1]): \mu(\{0\}) \geq s} \mathcal{P}_\beta(\mu([0, \cdot])), \quad E_\beta(\lambda) = \beta^{-1} P_\beta(\lambda/\beta).$$

The minimum is attained [4, Section 1.1], and $P_\beta(0) = P_\beta$. The function E_β is the constrained value in the energy normalization of [4, Section 7.1], in which the order s of

the moment is replaced by the tilt $\lambda = \beta s$. Note that E_β is nondecreasing, because the admissible class shrinks as s grows; in particular, $E_\beta(\lambda) \geq E_\beta(0)$.

The unconstrained minimizer μ_β is unique [10]. We write $\alpha_\beta(q) = \mu_\beta([0, q])$ and $\gamma_\beta = \beta\alpha_\beta$. Since $\beta > 1$, we have $\text{supp } \mu_\beta = [0, q_\beta]$ with $0 < q_\beta < 1$, by [48, Theorem 1.1]. We recall the further structure of μ_β in Section 7.

For $0 < \lambda < \beta$ and $s = \lambda/\beta$, we define the *fractional-moment error* by

$$e_{N,\beta}(\lambda) = \frac{P_\beta(s) - \phi_{N,\beta}^c(s)}{\beta} = E_\beta(\lambda) - \frac{1}{N\lambda} \log \mathbb{E}_{\text{ref}} e^{\lambda(F_{N,\beta}/\beta + D_N)}. \quad (3.1)$$

In the second expression, we used $sF_{N,\beta}^c = \lambda(F_{N,\beta}/\beta + D_N)$; the completion contributes $\lambda/(4N)$ to the logarithmic moment on the right. As β is fixed, we also write $e_N(\lambda) = e_{N,\beta}(\lambda)$. From now on, we assume that $s = \lambda/\beta \leq 1/2$, which is the range of the results of [4, Section 2]. We will see in Corollary 3.7 that the error $e_{N,\beta}(\lambda)$ is nonnegative.

3.2 The interpolating functional

For $0 \leq \rho \leq 1$, we set

$$t = 1 - \rho, \quad c = \beta^2 \rho, \quad M = Nc.$$

The interpolation reaches the fractional moment at $\rho = 0$, and $\rho = 1$ is the scalar end. We parametrize the interpolation by ρ because the estimates of the later sections degenerate as $\rho \downarrow 0$. We work with the tilted law of Section 3.3 for $0 < \rho \leq 1$, where $c > 0$.

An *admissible profile* is a deterministic probability measure $\mu = s\delta_0 + \nu$ on $[0, 1]$, where $\nu \geq 0$ has mass $1 - s$ and may charge 0. Equivalently, μ is a deterministic probability measure such that $\mu(\{0\}) \geq s$. We call $s\delta_0$ the prescribed root mass and ν the free measure of μ . We write $a(q) = \mu([0, q])$ for the cumulative function. It is nondecreasing and right-continuous, with $a \geq s$ and $a(1) = 1$. Conditional on W , let $U = U_{\mu,t}$ solve

$$\partial_q U = -\frac{c}{2}(\Delta U + a(q)|\nabla U|^2), \quad U(1, x; W) = \log \sum_{\sigma \in \{-1, 1\}^N} e^{\beta\sqrt{t}\sigma^T W\sigma/2 + x \cdot \sigma}. \quad (3.2)$$

This is the Parisi equation in $\mathbb{R}^N$; its terminal value is the logarithm of the completed partition function at inverse temperature $\beta\sqrt{t}$ in the external field x . We define the interpolating functional and its minimum over admissible profiles by

$$\mathcal{F}_{t,s}(\mu) = \frac{1}{Ns} \log \mathbb{E}_{\text{ref}} e^{sU(0,0;W)} - \frac{c}{2} \int_0^1 qa(q) dq, \quad f_{N,s}(t) = \min_{\mu} \mathcal{F}_{t,s}(\mu). \quad (3.3)$$

By [4, Section 2.6 and Proposition 2.3], the equation (3.2) has a solution for every admissible profile, and $\mathcal{F}_{t,s}$ is continuous under weak convergence of profiles. For a step profile, the solution is given by the Hopf–Cole recursion that we recall in Section 3.3. For

a general profile, it is the uniform limit of the solutions for step profiles whose cumulative functions converge in $L^1([0, 1])$. Indeed, if U_a and $U_{\tilde{a}}$ are the solutions for the cumulative functions a and $\tilde{a}$, then [4, Section 2.6, (2.19)]

$$\sup_{q, x, W} |U_a(q, x; W) - U_{\tilde{a}}(q, x; W)| \leq \frac{cN}{2} \int_0^1 |a(u) - \tilde{a}(u)| du.$$

The admissible profiles form a weakly compact set, since they are the image of the compact set of measures of mass $1 - s$ under $\nu \mapsto s\delta_0 + \nu$, and weak convergence of profiles implies L^1 convergence of their cumulative functions. We conclude that the minimum in (3.3) is attained.

We note that minimizers are deterministic. The Gaussian expectation is taken inside the functional, which means that $\mathcal{F}_{t,s}(\mu)$ is a number for each profile. A minimizer may depend on N , ρ , and λ but not on the couplings. This matters below, because the tilted law of Section 3.3 is constructed from a deterministic profile, and the estimates at a minimizer combine that law with the variational inequality that the same profile satisfies. A minimizer need not be unique; at each use, we say which minimizer is meant.

The prescribed root mass $s\delta_0$ matches the order s of the moment. In the derivative of the interpolation, which we compute in Proposition 3.6, the outer logarithmic moment contributes a term that cancels the root mass exactly, so that only the free measure ν remains. In the posterior densities of Section 4, the outer tilt cancels it in the same way.

The two endpoints of the interpolation are explicit. At $t = 0$, the terminal value in (3.2) is $\sum_i \log(2 \cosh x_i)$ and $c = \beta^2$. Then the equation separates into N copies of the scalar Parisi equation of (1.1), and we have $U(q, x; W) = \sum_i u_a(q, x_i)$, where u_a is the solution of that equation with $\alpha = a$. This shows that $U(0, 0; W) = Nu_a(0, 0)$ is deterministic and that $\mathcal{F}_{0,s}(\mu) = \mathcal{P}_\beta(a)$. At $t = 1$, we have $c = 0$, which implies that $U(0, 0; W) = U(1, 0; W) = F_{N,\beta}^c$ and that $\mathcal{F}_{1,s}(\mu) = \phi_{N,\beta}^c(s)$ for every profile. Minimizing over admissible profiles, we obtain [4, Proposition 2.3]

$$f_{N,s}(0) = P_\beta(s), \quad f_{N,s}(1) = \phi_{N,\beta}^c(s). \quad (3.4)$$

For a fixed profile, the map $t \mapsto \mathcal{F}_{t,s}(\mu)$ is nonincreasing (Proposition 3.6). By (3.1), the fractional-moment error is $(f_{N,s}(0) - f_{N,s}(1))/\beta$, which is the decrease of $f_{N,s}/\beta$ along the interpolation. The defect of Section 3.6 is four times the rate of this decrease.

The energy normalization. We vary profiles in the energy normalization of [4, Section 7.1]. For an admissible profile μ with cumulative function a , we set

$$\gamma = \beta a, \quad \nu_\gamma = \beta \nu, \quad F_\rho(\gamma) = \beta^{-1} \mathcal{F}_{1-\rho, \lambda/\beta}(\mu).$$

Then $d\gamma$ is a measure of mass β with $d\gamma \geq \lambda\delta_0$; we also call γ admissible. We have $\min_\gamma F_\rho(\gamma) = \beta^{-1} f_{N,s}(1 - \rho)$, where the minimum is over admissible γ . The name

refers to the following rescaling of the spatial variable. The function $\tilde{U}(q, y; W) = \beta^{-1}U(q, \beta y; W)$ solves

$$\partial_q \tilde{U} = -\frac{\rho}{2}(\Delta \tilde{U} + \gamma(q)|\nabla \tilde{U}|^2),$$

with terminal value β^{-1} times the logarithm of the completed partition function in the field βy , and we have

$$F_\rho(\gamma) = \frac{1}{N\lambda} \log \mathbb{E}_{\text{ref}} e^{\lambda \tilde{U}(0,0;W)} - \frac{\rho}{2} \int_0^1 q \gamma(q) dq.$$

This is the functional of [4, (7.2)] with $r = \rho$, in which the diagonal of W has been integrated out explicitly [4, Section 7.1]. Note that the rescaling leaves the gradient unchanged, $\nabla \tilde{U}(q, y; W) = \nabla U(q, \beta y; W)$, and that it multiplies the Hessian by β .

We define the *scalar trial profile* by $\gamma_\lambda^{\text{tr}}(q) = \max\{\lambda, \gamma_\beta(q)\}$. It is nondecreasing and right-continuous, and it satisfies $\lambda \leq \gamma_\lambda^{\text{tr}} \leq \beta$ and $\gamma_\lambda^{\text{tr}}(1) = \beta$. In other words, $\gamma_\lambda^{\text{tr}}/\beta$ is the cumulative function of an admissible profile. When λ is fixed, we write $\gamma^{\text{tr}} = \gamma_\lambda^{\text{tr}}$. We discuss its cost and its use as a comparison profile in Section 3.7.

3.3 The tilted joint law and overlap observables

Fix $0 < \rho \leq 1$ and an admissible profile μ . We recall the joint law of [4, Section 2], under which all later expectations are taken. It is built in three stages.

- (i) The Gaussian law of W is tilted by the density $e^{sU(0,0;W)}/\mathbb{E}_{\text{ref}} e^{sU(0,0;W)}$.
- (ii) Given W , the auxiliary field solves

$$dX_q = ca(q)\nabla U(q, X_q; W) dq + \sqrt{c} dB_q, \quad X_0 = 0, \quad (3.5)$$

where B is a standard Brownian motion in $\mathbb{R}^N$ under the conditional law given W .

- (iii) Given W and the path X , a terminal spin $\sigma \in \{-1, 1\}^N$ is drawn with probabilities proportional to $\exp\{\beta\sqrt{t}\sigma^T W\sigma/2 + X_1 \cdot \sigma\}$.

We write $\mathbb{P}$ and $\mathbb{E}$ for this joint law, and $\mathcal{G}_q = \sigma(W, X_r : 0 \leq r \leq q)$ for the information revealed up to level q ; note that it does not contain the terminal spin. At fixed N , the drift in (3.5) is bounded and Lipschitz in x [4, Section 2.6.1], and the equation has a unique strong solution. The tilt in (i) changes the law of W , which is no longer Gaussian under $\mathbb{P}$, but it does not change the conditional law of B given W .

For a step profile, the law has a discrete description, which explains the construction. The Gaussian Hopf–Cole recursion in the Parisi setting is described in [40, Section 3, (10)–(13)], and the finite-dimensional tilted law used here is the formulation of [4, Section 2.1]. Suppose that the profile has grid points $0 = q_0 < q_1 < \dots < q_K = 1$ and that $a = a_\ell$ on $[q_\ell, q_{\ell+1})$. Then (3.2) is solved by the Hopf–Cole recursion

$$U(q_\ell, x; W) = \frac{1}{a_\ell} \log \mathbb{E}_g \exp\{a_\ell U(q_{\ell+1}, x + \sqrt{c(q_{\ell+1} - q_\ell)} g; W)\},$$

where g is a standard Gaussian vector in $\mathbb{R}^N$, and we read the right side as $\mathbb{E}_g U(q_{\ell+1}, x + \sqrt{c(q_{\ell+1} - q_\ell)} g; W)$ when $a_\ell = 0$. Given W and X_{q_ℓ} , the increment $X_{q_{\ell+1}} - X_{q_\ell}$ has the law of $\sqrt{c(q_{\ell+1} - q_\ell)} g$, tilted by the density $\exp\{a_\ell(U(q_{\ell+1}, X_{q_{\ell+1}}; W) - U(q_\ell, X_{q_\ell}; W))\}$, which integrates to one by the recursion. The outer tilt in (i) has the same form, with mass s at level 0. In Section 4, we multiply these densities and use the resulting formula for the joint density.

Observables. We denote the conditional mean, the conditional covariance, and the spatial Hessian by

$$m_q = \mathbb{E}[\sigma \mid \mathcal{G}_q], \quad C_q = \text{Cov}(\sigma \mid \mathcal{G}_q), \quad H_q = \nabla_x^2 U(q, X_q; W).$$

We will see in Proposition 3.1 that $m_q = \nabla U(q, X_q; W)$ and that $a(q)C_q \preceq H_q \preceq C_q$. With the same terminal spin at every level, we define

$$S_q = \frac{|m_q|^2}{N}, \quad L_q = \frac{\sigma \cdot m_q - |m_q|^2/2}{N}, \quad j_q = \frac{\mathbb{E}[m_q^T C_q m_q]}{N^2}, \quad a_q^{\text{th}} = \frac{\mathbb{E} \text{Tr } C_q^2}{N^2};$$

we use the superscript in a_q^{th} to distinguish the covariance trace from the cumulative function a . For each q , let σ^1 and σ^2 be conditionally independent given $\mathcal{G}_q$, each with the conditional law of σ given $\mathcal{G}_q$, and let $R_q = R(\sigma^1, \sigma^2)$. Note that the self-overlap S_q of the conditional mean is the conditional expectation of R_q given $\mathcal{G}_q$, and that L_q has conditional expectation $S_q/2$ given $\mathcal{G}_q$. In Proposition 3.3, we bound the variance of averages of L_q over levels. We also have $0 \leq S_q \leq 1$ and $|L_q| \leq 3/2$.

The processes m , C , and H have continuous paths. Indeed, we have $m_q = \nabla U(q, X_q; W)$, $H_q = \nabla^2 U(q, X_q; W)$, and $\mathbb{E}[\sigma \sigma^T \mid \mathcal{G}_q] = \Sigma(q, X_q; W)$, where the conditional second moment Σ solves a linear backward equation with drift $ca \nabla U$, and ∇U , $\nabla^2 U$, and Σ are jointly continuous in (q, x) [4, Section 2.6.1]; the path X is continuous. In particular, S_q and L_q are jointly measurable in the level and the outcome, and Fubini's theorem applies to their integrals against ν . By bounded convergence, the functions $q \mapsto j_q$, a_q^{th} , $\mathbb{E} S_q$, $\text{Var}(S_q)$, and $\mathbb{E}|S_q - q|^3$ are continuous on $[0, 1]$, and so is $q \mapsto \mathbb{E}(R_q - q)^2$, because $\mathbb{E}[R_q \mid \mathcal{G}_q] = S_q$ and $\mathbb{E}[R_q^2 \mid \mathcal{G}_q] = N^{-2} \sum_{i,j} \Sigma_{ij}(q, X_q; W)^2$.

We also define the integrated quantities

$$A = \int a_q^{\text{th}} \nu(\mathrm{d}q), \quad J = \int j_q \nu(\mathrm{d}q), \\ V = \int \text{Var}(S_q) \nu(\mathrm{d}q), \quad H_3 = \int \mathbb{E}|S_q - q|^3 \nu(\mathrm{d}q).$$

We note that they depend on N , ρ , λ , and the profile. The trace term A and the mean-covariance term J are controlled by the conditional covariance. The integrated variance V of the self-overlap is the main term of the defect, and H_3 is the corresponding cubic moment, which enters in Section 12.

Quantiles and contact levels. The quantile function

$$Q(v) = \inf\{q \in [0, 1] : \nu([0, q]) \geq v\}, \quad 0 < v < 1 - s, \quad (3.6)$$

pushes Lebesgue measure on $(0, 1 - s)$ forward to ν , and $a(Q(v)) = s + \nu([0, Q(v)]) \geq s + v$, by the right continuity of $q \mapsto \nu([0, q])$. We use quantiles to integrate against ν in a variable in which the lower bound on a is explicit, whether or not ν has a density. The image under Q of Lebesgue measure on a Borel subset of $(0, 1 - s)$ is a sub-measure of ν , whose density with respect to ν takes values in $[0, 1]$. For an interval of values of v , this sub-measure may contain only part of an atom of ν , and by integration over such an interval, we mean integration against this sub-measure.

If μ minimizes $\mathcal{F}_{t,s}$, we call a point of $\text{supp } \nu$ a *contact level* of μ . The contact levels are where the free measure sits, and Proposition 3.1 shows that the expected self-overlap $\mathbb{E}S_q$ equals q at each of them.

3.4 Martingale and contact identities

In the next proposition, we collect the finite-dimensional martingale, first-variation, and contact identities of [4, Proposition 2.1]. For the scalar Parisi functional, the stochastic-control representation and the martingale identities are in [10, Theorem 3 and Lemma 2], with a dynamic programming treatment in [43]. The first variation and the optimality conditions on the support are in [44, Lemma 3.2 and Proposition 1.1] and [25, Theorem 2 and Proposition 1].

Proposition 3.1 (Martingale, covariance order, and contact). *Let $0 < \rho \leq 1$, and let $\mu = s\delta_0 + \nu$ be an admissible profile. For every $q \in [0, 1]$,*

$$m_q = \nabla U(q, X_q; W), \quad m_0 = 0, \quad dm_q = \sqrt{c} H_q dB_q, \quad a(q)C_q \preceq H_q \preceq C_q, \quad (3.7)$$

and $q \mapsto \mathbb{E}S_q$ is continuously differentiable on $[0, 1]$ with derivative $c \mathbb{E} \text{Tr } H_q^2 / N$. If $\nu' \geq 0$ has mass $1 - s$, then the right derivative of $\theta \mapsto \mathcal{F}_{t,s}(s\delta_0 + (1 - \theta)\nu + \theta\nu')$ at $\theta = 0$ is

$$D\mathcal{F}_{t,s}(\mu)[\nu' - \nu] = \frac{c}{2} \int_0^1 G(x) (\nu' - \nu)(dx), \quad G(x) = \int_x^1 (\mathbb{E}S_y - y) dy. \quad (3.8)$$

If μ minimizes $\mathcal{F}_{t,s}$, then $\text{supp } \nu$ is contained in the set of minimum points of G on $[0, 1]$, and

$$\text{supp } \nu \subset [0, 1), \quad \mathbb{E}S_q = q, \quad \mathbb{E} \text{Tr } H_q^2 \leq \frac{N}{c} \quad \text{for every contact level } q. \quad (3.9)$$

Remark 3.2. A contact level is a point of the support of the free measure ν . At $q = 0$, we have $\mathbb{E}S_0 = 0$ for every profile, since $m_0 = 0$. The bound on $\mathbb{E} \text{Tr } H_0^2$ in (3.9) holds when $0 \in \text{supp } \nu$; the prescribed mass $s\delta_0$ alone does not give it, because the minimization does not act on that mass.

Proof of Proposition 3.1. We recall the argument of [4, Proposition 2.1 and Section 2.6]. The proof has four steps. We first prove the martingale identities and the covariance order for step profiles. We then derive the first variation and deduce the support property and (3.9) at a minimizer. Finally, we extend these statements to general profiles.

Step 1: martingale identities for step profiles. Let μ be a step profile. Then U is smooth in x , every spatial derivative of fixed order is bounded uniformly in (q, x, W) at fixed N , and U is smooth in q between consecutive grid points [4, Section 2.6.1]. Differentiating (3.2) in x and applying Itô's formula along (3.5) on each interval between grid points, we obtain [4, (2.5)]

$$d \nabla U(q, X_q; W) = \sqrt{c} H_q dB_q, \quad dH_q = -ca(q)H_q^2 dq + \sqrt{c} \nabla H_q dB_q,$$

where $\nabla H_q dB_q$ denotes the contraction of the third spatial derivative of U at (q, X_q) with dB_q . The stochastic integrals are martingales, since their integrands are bounded. At $q = 1$, the gradient $\nabla U(1, X_1; W)$ is the Gibbs mean of the terminal spin given (W, X_1) , that is, $\mathbb{E}[\sigma \mid \mathcal{G}_1]$. By the tower property, we have $\nabla U(q, X_q; W) = \mathbb{E}[\sigma \mid \mathcal{G}_q] = m_q$ for all q . Replacing σ by $-\sigma$ in the sum that defines $U(1, x; W)$, we see that $U(1, \cdot; W)$ is even. The equation preserves this symmetry, which implies that $U(q, \cdot; W)$ is even and that $m_0 = \nabla U(0, 0; W) = 0$.

For the covariance order, we first note that $C_1 = H_1$, because the Hessian of the logarithm of a partition function is the covariance of its Gibbs measure. Applying the martingale isometry to m , and taking conditional expectations in the equation for H , we find

$$C_q = \mathbb{E}[H_1 \mid \mathcal{G}_q] + c \int_q^1 \mathbb{E}[H_u^2 \mid \mathcal{G}_q] du, \quad H_q = \mathbb{E}[H_1 \mid \mathcal{G}_q] + c \int_q^1 a(u) \mathbb{E}[H_u^2 \mid \mathcal{G}_q] du.$$

Observe that every matrix in these expressions is positive semidefinite, and that $a(q) \leq a(u) \leq 1$ for $u \geq q$. Subtracting, we obtain

$$C_q - H_q = c \int_q^1 (1 - a(u)) \mathbb{E}[H_u^2 \mid \mathcal{G}_q] du \succeq 0$$

and

$$H_q - a(q)C_q = (1 - a(q)) \mathbb{E}[H_1 \mid \mathcal{G}_q] + c \int_q^1 (a(u) - a(q)) \mathbb{E}[H_u^2 \mid \mathcal{G}_q] du \succeq 0,$$

which proves $a(q)C_q \preceq H_q \preceq C_q$ [4, (2.6)]. Finally, Itô's formula gives $d|m_q|^2 = 2m_q \cdot dm_q + c \operatorname{Tr} H_q^2 dq$. Taking expectations, we find that the derivative of $q \mapsto \mathbb{E} S_q$ is $c \mathbb{E} \operatorname{Tr} H_q^2 / N$ [4, (2.7)], which is continuous in q .

Step 2: the first variation. Let $\nu' \geq 0$ have mass $1 - s$, and let $\delta a(q) = (\nu' - \nu)([0, q])$, so that the profile $s\delta_0 + (1 - \theta)\nu + \theta\nu'$ has cumulative function $a + \theta\delta a$. The derivative $\dot{U}$ of the solution in θ at $\theta = 0$ solves the linearized equation

$$\partial_q \dot{U} + \frac{c}{2} \Delta \dot{U} + ca(q) \nabla U \cdot \nabla \dot{U} = -\frac{c}{2} \delta a(q) |\nabla U|^2, \quad \dot{U}(1, \cdot; W) = 0.$$

Its generator is the generator of (3.5). By the Feynman–Kac formula, we have

$$\dot{U}(0, 0; W) = \frac{c}{2} \mathbb{E} \left[\int_0^1 \delta a(q) |m_q|^2 dq \mid W \right].$$

When we differentiate the outer logarithmic moment in (3.3), this identity is averaged under the tilted law of W , and the penalty contributes $-(c/2) \int_0^1 q \delta a(q) dq$. We obtain

$$D\mathcal{F}_{t,s}(\mu)[\nu' - \nu] = \frac{c}{2} \int_0^1 (\mathbb{E} S_q - q) \delta a(q) dq.$$

We write $\delta a(q) = \int \mathbf{1}\{x \leq q\} (\nu' - \nu)(dx)$ and exchange the order of integration, which is allowed because the integrand is bounded and $\nu' - \nu$ has finite total variation. This gives (3.8), which is [4, (2.8)]. We note that its derivation does not use minimality. In the energy normalization, the identity is stated for every admissible profile in [4, Section 7.4, (7.22)].

Step 3: support and contact at a minimizer. Let μ minimize $\mathcal{F}_{t,s}$, and let $\nu' \geq 0$ have mass $1 - s$. The profiles $s\delta_0 + (1 - \theta)\nu + \theta\nu'$ with $0 \leq \theta \leq 1$ are admissible, and $\theta = 0$ is a minimum point of $\mathcal{F}_{t,s}$ along them. This implies that the right derivative in (3.8) is nonnegative. Taking $\nu' = (1 - s)\delta_x$, we find

$$(1 - s)G(x) \geq \int G d\nu \quad \text{for every } x \in [0, 1].$$

The function G is continuous on $[0, 1]$. Choosing for x a minimum point of G , we see that $\int (G - \min_{[0,1]} G) d\nu \leq 0$, so that $G = \min_{[0,1]} G$ holds ν -almost everywhere. The set where $G > \min_{[0,1]} G$ is relatively open in $[0, 1]$ and ν -null, and we conclude that $\text{supp } \nu$ is contained in the set of minimum points of G .

By Step 1, G is twice continuously differentiable, with

$$G'(q) = q - \mathbb{E} S_q, \quad G''(q) = 1 - \frac{c}{N} \mathbb{E} \text{Tr } H_q^2.$$

If $q \in \text{supp } \nu \cap (0, 1)$, then q is an interior minimum point of G , and therefore $G'(q) = 0$ and $G''(q) \geq 0$. These are the last two relations in (3.9). At $q = 0$, we have $G'(0) = -\mathbb{E} S_0 = 0$, because $m_0 = 0$. If $0 \in \text{supp } \nu$, then 0 is a minimum point of G , and $G''(0) \geq 0$, since otherwise $G'' < 0$ near 0 and $G(h) < G(0)$ for small $h > 0$. At $q = 1$, the terminal Gibbs measure gives positive probability to every spin configuration. Then $|m_{1,i}| < 1$ for all i , and $G'(1) = 1 - \mathbb{E} S_1 > 0$, which implies that $G(1 - h) < G(1)$ for small $h > 0$ and that 1 is not a minimum point of G . This proves $\text{supp } \nu \subset [0, 1)$ and (3.9), which is [4, (2.9)].

Step 4: general profiles. Let μ be an arbitrary admissible profile. Following [4, Section 2.6], we approximate ν by its pushforwards ν_j onto finite meshes that contain 0, 1, and any prescribed finite set of levels, with mesh size tending to zero. These pushforwards are taken under the map that sends each point to the smallest mesh point greater than or equal to it. Let $\mu_j = s\delta_0 + \nu_j$. Then $\mu_j \Rightarrow \mu$, the cumulative functions converge in $L^1([0, 1])$, and they agree with a at the mesh points. The spatial derivatives of fixed order are bounded uniformly over step profiles at fixed N . We apply the L^1 estimate recalled in Section 3.2 to the solutions, and the same comparison argument to their differentiated equations. Using also the uniform bound on the derivatives and Grönwall's inequality, we find that the solutions and their first and second spatial derivatives converge uniformly. We couple the laws for μ_j and μ through the same matrix W under its Gaussian reference law and the same Brownian motion B . We also use the same uniform variable to sample the terminal spin. Since the drifts are bounded and Lipschitz, Grönwall's inequality implies that X , m , and H for μ_j converge to those for μ uniformly in q , in probability. The conditional second moments of the terminal spin, which solve linear backward equations with drift $ca\nabla U$, converge in the same way, and so does C_q . The ratios of the tilt densities for μ_j to the density for μ converge uniformly to one, since these ratios lie between e^{-2sd_j} and e^{2sd_j} , where $d_j = (cN/2)\|a_j - a\|_{L^1} \rightarrow 0$. Then the densities converge in L^1 of the Gaussian reference law. Since the conditional moments and spatial derivatives used here are bounded at fixed N , their convergence in probability, uniformly in the level, implies that they also converge in every finite L^p norm needed below. Using these convergences, we pass (3.7), the formula for the derivative of $\mathbb{E}S_q$, its continuity in q , and (3.8) to the limit [4, Section 2.6.2]. If μ is a minimizer, we apply the limiting first variation directly to μ ; the approximating step profiles need not be minimizers. Step 3 then applies to μ without change. $\square$

3.5 The integrated variance bound and the overlap error

At a minimizer, we use the contact identities to turn the posterior variance bound of [4, Proposition 2.2] into bounds that depend only on the cumulative function a .

Proposition 3.3 (Integrated variance and trace bounds). *Let $0 < \rho \leq 1$, and let $\mu = s\delta_0 + \nu$ minimize $\mathcal{F}_{t,s}$. For every deterministic $k \in L^2(\nu)$,*

$$\text{Var} \left(\int k(q) L_q \nu(\mathrm{d}q) \right) \leq \frac{1}{M} \int \frac{k(q)^2}{a(q)} \nu(\mathrm{d}q), \quad A \leq \frac{1}{Ms}. \quad (3.10)$$

The first bound is [4, Proposition 2.2, (2.16)], and the second is [4, Section 2.5, (2.18)]. We briefly recall the argument. We fix the terminal spin and apply the Brascamp–Lieb inequality to the law of the independent Gaussian coordinates that generate W and the

field. For a fixed spin, this law is uniformly log-concave, although the marginal law of W under $\mathbb{P}$ is not Gaussian.

Proof of Proposition 3.3. We first show that, for every admissible profile and every deterministic $k \in L^2(\nu)$,

$$\text{Var} \left(\int k(q) L_q \nu(\mathrm{d}q) \right) \leq \frac{1}{N^2} \int k(q)^2 \mathbb{E} \text{Tr}(H_q C_q) \nu(\mathrm{d}q), \quad (3.11)$$

which is [4, (2.15)]. Let μ be a step profile. For $\varepsilon > 0$, we start the field at $X_0 = \sqrt{\varepsilon} z_0$, where z_0 is a standard Gaussian vector included in the outer tilt. We collect the independent standard Gaussian coordinates that generate W , z_0 , and the field increments in one vector Z . With the terminal spin fixed, Z has a density proportional to $e^{-\Phi_\sigma(Z)}$, and the Hessian of Φ_σ is $\mathcal{Q} = I + \int \mathcal{A}_q \nu(\mathrm{d}q) \succeq I$, where $\mathcal{A}_q = \nabla_Z^2 U(q, X_q; W) \succeq 0$ [4, (2.1)–(2.3)]. We derive this density for more general laws in the proof of Lemma 4.2. We work in these Gaussian reference coordinates, in which, at the grid points q_ℓ of the profile,

$$X_{q_\ell} = \sqrt{\varepsilon} z_0 + \sum_{h=1}^{\ell} \sqrt{c(q_h - q_{h-1})} z_h;$$

the tilted law is obtained by reweighting these coordinates. In this representation, the field depends on z_0 only through the term $\sqrt{\varepsilon} z_0$, which gives $\nabla_{z_0} U(q, X_q; W) = \sqrt{\varepsilon} m_q$ and shows that the block of $\mathcal{A}_q$ for the coordinates of z_0 is εH_q . Let v_q be the vector whose z_0 block is $(\sigma - m_q) / \sqrt{\varepsilon}$ and whose other blocks vanish. Differentiating L_q with the spin fixed, we find

$$\nabla_Z L_q = \frac{1}{N} \mathcal{A}_q v_q, \quad v_q^T \mathcal{A}_q v_q = (\sigma - m_q)^T H_q (\sigma - m_q).$$

For $Y_k = \int k L_q \mathrm{d}\nu$ and $b = \int k(q) \mathcal{A}_q v_q \nu(\mathrm{d}q)$, we have $\nabla_Z Y_k = b/N$. For every vector w , using the Cauchy–Schwarz inequality for the forms $\mathcal{A}_q$ and for the integral against ν , together with $\int \mathcal{A}_q \mathrm{d}\nu = \mathcal{Q} - I \preceq \mathcal{Q}$, we have

$$|w^T b|^2 \leq \left(\int w^T \mathcal{A}_q w \nu(\mathrm{d}q) \right) \int k(q)^2 v_q^T \mathcal{A}_q v_q \nu(\mathrm{d}q) \leq (w^T \mathcal{Q} w) \int k(q)^2 v_q^T \mathcal{A}_q v_q \nu(\mathrm{d}q).$$

The choice $w = \mathcal{Q}^{-1} b$ yields $b^T \mathcal{Q}^{-1} b \leq \int k^2 v_q^T \mathcal{A}_q v_q \mathrm{d}\nu$. Then the Brascamp–Lieb inequality [21], in the form [4, Proposition 2.2, (2.14)], implies that the conditional variance of Y_k given σ is at most $N^{-2} \mathbb{E}[b^T \mathcal{Q}^{-1} b \mid \sigma]$. By the spin-gauge symmetry of [4, Section 2.1], the conditional law of Y_k given σ does not depend on σ . Using the law of total variance, we obtain the same bound for $\text{Var}(Y_k)$. Conditioning on $\mathcal{G}_q$, we have $\mathbb{E}[(\sigma - m_q)^T H_q (\sigma - m_q) \mid \mathcal{G}_q] = \text{Tr}(H_q C_q)$. We remove the regularization by dominated convergence as $\varepsilon \downarrow 0$, using $|L_q| \leq 3/2$, $\text{Tr}(H_q C_q) \leq N^2$, and the Gaussian domination

of the outer tilt. The approximation of Step 4 in the proof of Proposition 3.1 then extends (3.11) to all profiles, first for continuous k and then by density for $k \in L^2(\nu)$ [4, Section 2.6.2].

Now suppose that μ is a minimizer. Using the fact that the trace of the product of two positive semidefinite matrices is nonnegative, together with $H_q - a(q)C_q \succeq 0$, we have

$$\mathrm{Tr} H_q^2 = \mathrm{Tr}(H_q(H_q - a(q)C_q)) + a(q) \mathrm{Tr}(H_q C_q) \geq a(q) \mathrm{Tr}(H_q C_q),$$

and similarly

$$\mathrm{Tr}(H_q C_q) = \mathrm{Tr}((H_q - a(q)C_q)C_q) + a(q) \mathrm{Tr} C_q^2 \geq a(q) \mathrm{Tr} C_q^2.$$

By (3.9), $\mathbb{E} \mathrm{Tr} H_q^2 \leq N/c$ at every contact level, and ν is carried by its support. Dividing by $a(q) \geq s > 0$, we find that $\mathbb{E} \mathrm{Tr}(H_q C_q) \leq N/(ca(q))$ and $\mathbb{E} \mathrm{Tr} C_q^2 \leq N/(ca(q)^2)$ for ν -almost every q . Inserting the first bound into (3.11) and using $M = Nc$, we obtain the first bound in (3.10). For the second, the quantile function (3.6) and $a(Q(v)) \geq s + v$ give

$$A \leq \frac{1}{M} \int \frac{\nu(dq)}{a(q)^2} = \frac{1}{M} \int_0^{1-s} \frac{dv}{a(Q(v))^2} \leq \frac{1}{M} \int_0^{1-s} \frac{dv}{(s+v)^2} = \frac{1}{M} \left(\frac{1}{s} - 1 \right) \leq \frac{1}{Ms}. \quad \square$$

Remark 3.4. Since k in (3.10) is arbitrary, the bound applies to sub-measures of ν , such as the image of Lebesgue measure on an interval of quantiles that contains only part of an atom. If h is a Borel function with $0 \leq h \leq 1$, then applying (3.10) with kh in place of k and using $h^2 \leq h$, we obtain

$$\mathrm{Var}\left(\int k(q)h(q)L_q \nu(dq)\right) \leq \frac{1}{M} \int \frac{k(q)^2 h(q)}{a(q)} \nu(dq).$$

We now show that the overlap error splits into the three terms of the defect.

Lemma 3.5 (Decomposition of the overlap error). *Let $0 < \rho \leq 1$. For every admissible profile and every $q \in [0, 1]$, we have $\mathrm{Var}(L_q) = j_q + \mathrm{Var}(S_q)/4$. If q is a contact level of a minimizer, then*

$$\mathbb{E}(R_q - q)^2 = a_q^{\mathrm{th}} + 2j_q + \mathrm{Var}(S_q). \quad (3.12)$$

Proof. Both identities are displayed in [4, Section 2.5] before (2.17). Let $\tilde{\xi}^1 = \sigma^1 - m_q$ and $\tilde{\xi}^2 = \sigma^2 - m_q$. Conditionally on $\mathcal{G}_q$, these vectors are independent and centered, with covariance C_q , and

$$R_q = S_q + \frac{1}{N} (m_q \cdot \tilde{\xi}^1 + m_q \cdot \tilde{\xi}^2 + \tilde{\xi}^1 \cdot \tilde{\xi}^2).$$

The three terms in parentheses have conditional mean zero, they are pairwise orthogonal given $\mathcal{G}_q$, and their conditional second moments are $m_q^T C_q m_q$, $m_q^T C_q m_q$, and $\mathrm{Tr} C_q^2$. Taking expectations, we find that for every admissible profile and every q ,

$$\mathbb{E}(R_q - S_q)^2 = a_q^{\mathrm{th}} + 2j_q. \quad (3.13)$$

Since $\mathbb{E}[R_q - S_q \mid \mathcal{G}_q] = 0$, this implies that

$$\mathbb{E}(R_q - q)^2 = a_q^{\text{th}} + 2j_q + \mathbb{E}(S_q - q)^2 = a_q^{\text{th}} + 2j_q + \text{Var}(S_q) + (\mathbb{E}S_q - q)^2.$$

At a contact level of a minimizer, $\mathbb{E}S_q = q$ by (3.9), which proves (3.12). Similarly, $L_q = S_q/2 + N^{-1}(\sigma - m_q) \cdot m_q$, where the second term has conditional mean zero and conditional variance $N^{-2}m_q^T C_q m_q$ given $\mathcal{G}_q$. By the law of total variance, we have $\text{Var}(L_q) = j_q + \text{Var}(S_q)/4$. $\square$

3.6 The optimized interpolation and the defect

We next recall from [4, Proposition 2.3] that the derivative of the interpolating functional in t is an integral of squared overlap deviations.

Proposition 3.6 (Interpolation identity). *Let $N \geq 1$ and $0 < s \leq 1/2$.*

- (a) *For every admissible profile μ , the map $t \mapsto \mathcal{F}_{t,s}(\mu)$ is Lipschitz on $[0, 1]$ and continuously differentiable on $(0, 1)$, with*

$$\partial_t \mathcal{F}_{t,s}(\mu) = -\frac{\beta^2}{4} \int \mathbb{E}(R_q - q)^2 \nu(\mathrm{d}q), \quad (3.14)$$

where the expectation is under the tilted law of μ at t . The integral in (3.14) lies in $[0, 4(1-s)]$.

- (b) *The function $f_{N,s}$ is Lipschitz on $[0, 1]$. At every $t \in (0, 1)$ at which $f_{N,s}$ is differentiable, every minimizer μ of $\mathcal{F}_{t,s}$ satisfies $f'_{N,s}(t) = \partial_t \mathcal{F}_{t,s}(\mu)$.*

Proof. Part (a) is [4, Proposition 2.3 and (2.12)]. The regularity in t and the extension from step profiles to general profiles are proved in [4, Section 2.6.3]. We sketch the computation for a step profile, which shows where the square comes from. The parameter t enters $\mathcal{F}_{t,s}$ in three places: the factor $\beta\sqrt{t}$ of the matrix interaction, the coefficient $c = \beta^2(1-t)$ of the field in (3.2), and the penalty. We write $B(q) = \mathbb{E}R_q^2$ and $\Gamma(q) = \mathbb{E}S_q$, and note that $\mathbb{E}R_q = \Gamma(q)$.

Matrix interaction. We write $\beta\sigma^T W \sigma / 2 = \sum_e g_e v_e(\sigma)$, where e runs over the pairs $i < j$, with $g_e = g_{ij}$ and $v_e(\sigma) = \beta\sigma_i\sigma_j / \sqrt{N}$, and over the diagonal pairs $e = (i, i)$, with $g_e = z_i$ and $v_e = \beta / \sqrt{2N}$. By the computation of Section 3.1, we have $\sum_e v_e(\sigma)v_e(\sigma') = \beta^2 N R(\sigma, \sigma')^2 / 2$. We regard $y_e = \sqrt{t}g_e$ as Gaussian coordinates of variance t . By the Gaussian heat identity, at fixed c , the derivative in t of the outer logarithmic moment is

$$\frac{1}{2N} \mathbb{E} \sum_e (\partial_{y_e}^2 U(0, 0; W) + s(\partial_{y_e} U(0, 0; W))^2).$$

For a function $b(\sigma)$ of the terminal spin, we add $y b(\sigma)$ to the terminal exponent. Differentiating the Hopf–Cole recursion twice in y at $y = 0$, we find [4, (2.10)]

$$\partial_y U(0, 0) = \mathbb{E}[b(\sigma) \mid \mathcal{G}_0], \quad \partial_y^2 U(0, 0) = \mathbb{E}[b(\sigma)^2 \mid \mathcal{G}_0] - \int \mathbb{E}[\mathbb{E}[b(\sigma) \mid \mathcal{G}_q]^2 \mid \mathcal{G}_0] \mu(\mathrm{d}q).$$

Applying this with $b = v_e$, summing over e , and using $\mathbb{E}[v_e(\sigma) \mid \mathcal{G}_q]^2 = \mathbb{E}[v_e(\sigma^1)v_e(\sigma^2) \mid \mathcal{G}_q]$ and $\mu = s\delta_0 + \nu$, we find that the matrix interaction contributes

$$\frac{\beta^2}{4} \left(1 - \int B(q) \mu(\mathrm{d}q) + sB(0) \right) = \frac{\beta^2}{4} \left(1 - \int B(q) \nu(\mathrm{d}q) \right)$$

[4, (2.11)]. The term $s(\partial_{y_e} U)^2$, which comes from the outer moment of order s , cancels the prescribed root mass.

Field. As in Step 2 of the proof of Proposition 3.1, the derivative of U in c solves the linearized equation with source $-(\Delta U + a|\nabla U|^2)/2$, and by the Feynman–Kac formula, we have

$$\partial_c \left(\frac{1}{N_s} \log \mathbb{E}_{\text{ref}} e^{sU(0,0;W)} \right) = \frac{1}{2} \int_0^1 (h(q) + a(q)\Gamma(q)) \mathrm{d}q, \quad h(q) = \frac{1}{N} \mathbb{E} \operatorname{Tr} H_q.$$

By the equation for H_q in Step 1 of the proof of Proposition 3.1 and the derivative of Γ , we have $h' = -a\Gamma'$. Further, $h(1) = 1 - \Gamma(1)$, because $H_1 = C_1$ and $\operatorname{Tr} C_1 = N - |m_1|^2$. Integrating by parts, and using that h and Γ are continuous and that a is right-continuous with $a(1) = 1$, we find

$$h(q) + a(q)\Gamma(q) = 1 - \int_{(q,1]} \Gamma(r) \mu(\mathrm{d}r) \quad (0 \leq q \leq 1).$$

The interval is open at q because $a(q)$ already contains an atom of μ at q . Integrating in q and using $\mathrm{d}c/\mathrm{d}t = -\beta^2$, we see that the field contributes

$$-\frac{\beta^2}{2} \left(1 - \int q\Gamma(q) \nu(\mathrm{d}q) \right);$$

here we may replace μ by ν because the integrand vanishes at $q = 0$.

Penalty. By Fubini's theorem, we have $2 \int_0^1 qa(q) \mathrm{d}q = 1 - \int q^2 \mu(\mathrm{d}q)$. It follows that the penalty contributes

$$\frac{\beta^2}{4} \left(1 - \int q^2 \nu(\mathrm{d}q) \right).$$

Adding the three contributions, in which the constants cancel, we obtain

$$\partial_t \mathcal{F}_{t,s}(\mu) = -\frac{\beta^2}{4} \int (B(q) - 2q\Gamma(q) + q^2) \nu(\mathrm{d}q),$$

which is (3.14). The integral lies in $[0, 4(1-s)]$ because $|R_q - q| \leq 2$ and ν has mass $1-s$.

For (b), the bound in (a) implies that the functions $t \mapsto \mathcal{F}_{t,s}(\mu)$ are Lipschitz with the common constant $\beta^2(1-s)$. Their minimum $f_{N,s}$ is then Lipschitz as well. Let $t \in (0,1)$ be a point at which $f_{N,s}$ is differentiable, and let μ be a minimizer at t . For every h such that $t+h \in [0,1]$, we have $f_{N,s}(t+h) \leq \mathcal{F}_{t+h,s}(\mu)$ and $f_{N,s}(t) = \mathcal{F}_{t,s}(\mu)$. Dividing the difference by $h > 0$ and letting $h \downarrow 0$, we find $f'_{N,s}(t) \leq \partial_t \mathcal{F}_{t,s}(\mu)$; dividing by $h < 0$ gives the reverse inequality. This is [4, (2.20)]. $\square$

In particular, $t \mapsto \mathcal{F}_{t,s}(\mu)$ and $f_{N,s}$ are nonincreasing. We define the *defect*

$$\mathcal{D}_N(\rho) = -\frac{4}{\beta} f'_{N,s}(1 - \rho)$$

at every $\rho \in (0, 1)$ such that $f_{N,s}$ is differentiable at $1 - \rho$. Since $f_{N,s}$ is Lipschitz, Rademacher's theorem implies that this holds for almost every ρ , and $\mathcal{D}_N$ is a measurable function defined almost everywhere. All pointwise statements about $\mathcal{D}_N$ are understood at these values of ρ . The defect depends on N , β , ρ , and λ , and not on a choice of minimizer.

Corollary 3.7 (The defect identity). *Let $N \geq 1$ and $0 < \lambda \leq \beta/2$, and let $s = \lambda/\beta$. At every $\rho \in (0, 1)$ at which $\mathcal{D}_N(\rho)$ is defined, and for every minimizer μ of $\mathcal{F}_{1-\rho,s}$,*

$$\mathcal{D}_N(\rho) = \beta(A + 2J + V), \quad 4e_{N,\beta}(\lambda) = \int_0^1 \mathcal{D}_N(\rho) \, d\rho, \quad (3.15)$$

where A , J , and V are computed under the tilted law of μ . In particular, $\mathcal{D}_N \geq 0$ and $e_{N,\beta}(\lambda) \geq 0$.

Proof. Let ρ be as in the statement, and let μ be a minimizer at $t = 1 - \rho$. By Proposition 3.6(b) and (3.14), we have $\mathcal{D}_N(\rho) = \beta \int \mathbb{E}(R_q - q)^2 \nu(\,dq)$. Recall that the measure ν is carried by the contact levels of μ , at which (3.12) holds. Integrating (3.12) against ν , we obtain the first identity. For the second, since $f_{N,s}$ is Lipschitz, we have

$$\int_0^1 \mathcal{D}_N(\rho) \, d\rho = -\frac{4}{\beta} \int_0^1 f'_{N,s}(t) \, dt = \frac{4}{\beta} (f_{N,s}(0) - f_{N,s}(1)) = \frac{4}{\beta} (P_\beta(s) - \phi_{N,\beta}^c(s)),$$

where we used (3.4) in the last step. By (3.1), the right side is $4e_{N,\beta}(\lambda)$; see also [4, Proposition 2.3, (2.21)]. $\square$

The individual terms A , J , and V may depend on the choice of minimizer, but by (3.15), their combination $A + 2J + V$ does not. In the later sections, we bound A , J , and V at an arbitrary minimizer; such bounds then hold for $\mathcal{D}_N(\rho)$ at almost every ρ .

Near $\rho = 0$, we use a cruder bound. For every $N \geq 1$, $0 < s \leq 1/2$, $0 < \rho \leq 1$, and every minimizer μ of $\mathcal{F}_{1-\rho,s}$, we have

$$\int \mathbb{E}(R_q - q)^2 \nu(\,dq) \leq C \min\{1, M^{-1/2}, M^{-2/3} s^{-1/3}\}$$

with a numerical constant C . The first two terms are [4, Lemma 4.1, (4.1)], and the third comes from the proof of [4, Theorem 1.1]; see [4, (4.2)]. Keeping the first two terms and using $\beta M^{-1/2} = (N\rho)^{-1/2}$, we obtain

$$\begin{aligned} \int \mathbb{E}(R_q - q)^2 \nu(\,dq) &\leq C \min\{1, M^{-1/2}\}, \\ 0 \leq \mathcal{D}_N(\rho) &\leq C \min\{\beta, (N\rho)^{-1/2}\} \quad (0 < \rho \leq 1), \end{aligned} \quad (3.16)$$

where the second bound holds at every ρ at which $\mathcal{D}_N(\rho)$ is defined.

3.7 The scalar trial profile and the trial budget

The scalar trial profile $\gamma_\lambda^{\text{tr}} = \max\{\lambda, \gamma_\beta\}$ raises the Parisi profile γ_β to the floor λ . By [4, Lemma 7.2], γ_β grows linearly near 0, and the two profiles differ only on an interval of length of order λ , on which the first variation vanishes at γ_β . The cost of the floor is of order λ^5 . A short lower plateau and its fifth-order variational cost also occur in the replica calculation of [68, Appendix B, (B1)–(B3)]. Here $\lambda = \beta s$ specifies a mass-floor constraint; we do not identify it with the replica parameter in that calculation. The uniform bound used here is the rigorous estimate of [4, Lemma 7.2]. Precisely, by [4, Lemma 7.2, (7.3)], applied with $\beta_0 = \beta$, which uses $\beta > 1$, there exist $\lambda_* > 0$ and $C_* < \infty$, depending only on β , such that

$$\beta^{-1} \mathcal{P}_\beta(\gamma_\lambda^{\text{tr}}/\beta) \leq E_\beta(0) + C_* \lambda^5 \quad (0 < \lambda \leq \lambda_*).$$

We set $L_s = \max\{1, C_*\}$ and $\lambda_0 = \min\{1, \lambda_*/2, \beta/4\}$. Recall that E_β is nondecreasing, and note that the profile with cumulative function $\gamma_\lambda^{\text{tr}}/\beta$ has mass at least λ/β at zero. We obtain

$$0 \leq E_\beta(\lambda) - E_\beta(0) \leq \beta^{-1} \mathcal{P}_\beta(\gamma_\lambda^{\text{tr}}/\beta) - E_\beta(0) \leq L_s \lambda^5 \quad (0 < \lambda \leq 2\lambda_0). \quad (3.17)$$

The bound $\lambda_0 \leq \lambda_*/2$ allows us to use (3.17) at 2λ , and $\lambda_0 \leq \beta/4$ ensures that $2\lambda/\beta \leq 1/2$ for $\lambda \leq \lambda_0$, so that the results above apply with $s = 2\lambda/\beta$. The bound $\lambda_0 \leq 1$ is used in Sections 7 and 13.

Lemma 3.8 (Trial budget). *For every $N \geq 1$, $0 < \rho \leq 1$, and $0 < \lambda \leq \lambda_0$,*

$$0 \leq F_\rho(\gamma_\lambda^{\text{tr}}) - \min_\gamma F_\rho(\gamma) \leq e_{N,\beta}(\lambda) + L_s \lambda^5. \quad (3.18)$$

Proof. The lower bound holds because $\gamma_\lambda^{\text{tr}}$ is admissible. Let μ^{tr} be the admissible profile with cumulative function $\gamma_\lambda^{\text{tr}}/\beta$. By Proposition 3.6(a), the map $t \mapsto \mathcal{F}_{t,s}(\mu^{\text{tr}})$ is non-increasing, and $\mathcal{F}_{0,s}(\mu^{\text{tr}}) = \mathcal{P}_\beta(\gamma_\lambda^{\text{tr}}/\beta)$ by Section 3.2. Hence $F_\rho(\gamma_\lambda^{\text{tr}}) \leq \beta^{-1} \mathcal{P}_\beta(\gamma_\lambda^{\text{tr}}/\beta)$. Recall that the optimized value $f_{N,s}$ is also nonincreasing. Using (3.4) and (3.1), we find

$$\min_\gamma F_\rho(\gamma) = \beta^{-1} f_{N,s}(1 - \rho) \geq \beta^{-1} f_{N,s}(1) = \beta^{-1} \phi_{N,\beta}^e(s) = E_\beta(\lambda) - e_{N,\beta}(\lambda).$$

Subtracting the two bounds, we obtain

$$F_\rho(\gamma_\lambda^{\text{tr}}) - \min_\gamma F_\rho(\gamma) \leq e_{N,\beta}(\lambda) + \beta^{-1} \mathcal{P}_\beta(\gamma_\lambda^{\text{tr}}/\beta) - E_\beta(\lambda) \leq e_{N,\beta}(\lambda) + L_s \lambda^5,$$

where in the last step we used (3.17) and $E_\beta(\lambda) \geq E_\beta(0)$. This is [4, Proposition 7.7, (7.44)]. $\square$

We call a number $\eta \geq 0$ that bounds the left side of (3.18) a *trial budget*. By Lemma 3.8, $\eta = e_{N,\beta}(\lambda) + L_s \lambda^5$ is a trial budget simultaneously for every $\rho \in (0, 1]$ and every minimizer. In Sections 7 and 8, we use the trial budget to control the distance between a minimizer and the scalar trial profile, and in Section 13, it determines the cutoff between the two regimes of the proof.

3.8 Reduction of the mean bias

We now reduce the mean bias to the fractional-moment error by the midpoint argument of Step 1 in the proof of [4, Theorem 7.8]. We define

$$\psi(\lambda) = \frac{1}{N\lambda} \log \mathbb{E}_{\text{ref}} e^{\lambda(F_{N,\beta}/\beta + D_N)} \quad (\lambda > 0), \quad \psi(0) = \frac{p_{N,\beta}}{\beta},$$

so that $\psi(\lambda) = E_\beta(\lambda) - e_{N,\beta}(\lambda)$ for $0 < \lambda < \beta$, by (3.1). As a function of the $N(N+1)/2$ independent standard Gaussian coordinates (g_{ij}, z_i) , the completed free energy $F_{N,\beta}^c/\beta$ is convex, because it is β^{-1} times the logarithm of a sum of exponentials of linear functions. Its partial derivatives are $N^{-1/2}\langle\sigma_i\sigma_j\rangle$ in g_{ij} , where $\langle\cdot\rangle$ is the Gibbs average of $F_{N,\beta}^c$, and $(2N)^{-1/2}$ in z_i . We find that the squared norm of its gradient is at most

$$\frac{1}{N} \cdot \frac{N(N-1)}{2} + N \cdot \frac{1}{2N} = \frac{N}{2},$$

and the function is $\sqrt{N/2}$ -Lipschitz. By [4, Lemma 5.2], which is a special case of Chen's convexity theorem [26, Theorem 1], the function $\lambda \mapsto \lambda^{-1} \log \mathbb{E}_{\text{ref}} e^{\lambda F_{N,\beta}^c/\beta} = N\psi(\lambda)$ extends to a convex function on $[0, \infty)$ whose value at zero is $\mathbb{E}_{\text{ref}} F_{N,\beta}^c/\beta = N\psi(0)$, where we used $\mathbb{E}_{\text{ref}} D_N = 0$. In particular, ψ is convex on $[0, \infty)$.

Proposition 3.9 (Mean bias). *For every $N \geq 1$ and $0 < \lambda \leq \lambda_0$,*

$$0 \leq \frac{P_\beta - p_{N,\beta}}{\beta} \leq E_\beta(0) - 2E_\beta(\lambda) + E_\beta(2\lambda) + 2e_{N,\beta}(\lambda) \leq 32L_s\lambda^5 + 2e_{N,\beta}(\lambda). \quad (3.19)$$

Proof. The lower bound is [4, Theorem 1.1, (1.5)]; it is also Guerra's bound [40]. Since ψ is convex on $[0, \infty)$, we have $\psi(\lambda) \leq (\psi(0) + \psi(2\lambda))/2$, that is, $\psi(0) \geq 2\psi(\lambda) - \psi(2\lambda)$. Recall that $\psi(\lambda) = E_\beta(\lambda) - e_{N,\beta}(\lambda)$. Since $\lambda \leq \lambda_0 \leq \beta/4$, we have $2\lambda/\beta \leq 1/2$. Then Corollary 3.7 applies at the tilt 2λ and implies that $e_{N,\beta}(2\lambda) \geq 0$, that is, $\psi(2\lambda) \leq E_\beta(2\lambda)$. Using $E_\beta(0) = P_\beta/\beta$, we obtain

$$\frac{P_\beta - p_{N,\beta}}{\beta} = E_\beta(0) - \psi(0) \leq E_\beta(0) - 2E_\beta(\lambda) + E_\beta(2\lambda) + 2e_{N,\beta}(\lambda),$$

which is the middle inequality in (3.19). For the last inequality, we use $E_\beta(\lambda) \geq E_\beta(0)$ and (3.17) at $2\lambda \leq 2\lambda_0$:

$$E_\beta(0) - 2E_\beta(\lambda) + E_\beta(2\lambda) \leq E_\beta(2\lambda) - E_\beta(0) \leq L_s(2\lambda)^5 = 32L_s\lambda^5. \quad \square$$

In Section 13, we take $\lambda = N^{-7/50}$, so that $\lambda^5 = N^{-7/10}$. Proposition 3.9 then reduces Theorem 1.1 to the bound $e_{N,\beta}(N^{-7/50}) \leq C_\beta N^{-7/10}$ for all sufficiently large N .

Gaussian tools for arbitrary profiles

The results of Section 3 concern the law of Section 3.3, and the sharper ones hold only at a minimizing profile. In Sections 9 to 11, we also need laws whose profile minimizes no functional, and laws whose auxiliary field has different variances in different coordinates. In this section, we introduce a class of generalized laws that contains these cases (Definition 4.1), and we prove two Gaussian tools for them.

The first tool is Lemma 4.2, which shows that the martingale identities and the covariance order hold for every generalized law. It also bounds the variance of averages over levels of the squared conditional mean on a set of coordinates. We note that this bound holds for every profile and that its proof does not use contact identities. We use it in Section 9 for the scalar law, which is the generalized law with $N = 1$, field variance d , no couplings, and root mass $s = 0$. We also use it in Sections 10 and 11 for the cavity laws, which have diffusion matrix $\text{diag}(cI_n, b(u))$ and root mass $s = \lambda/\beta$.

The second tool is Lemma 4.3, a version of the Gaussian heat identity for normalized expectations. Here we take expectations of functionals of finitely many Gaussian variables and continuous Gaussian paths, and we hold the profile measure fixed. We use the lemma in Section 9 to differentiate the scalar law under transports of its levels, and in Sections 10 and 11 to differentiate the cavity laws in the cavity parameter. In this section, we use only the definitions of Section 3, the arguments of [4, Section 2], and the proof of Proposition 3.1, which we adapt.

Note that some letters have a local meaning in this section. The letter ω denotes Gaussian couplings, while in Section 10, ω denotes the weights of a tree. In Section 4.2, the derivative measures M and Q are unrelated to $M = Nc$, to the quantile function (3.6), and to the vector M_q of Section 10.

4.1 Generalized laws and a posterior variance bound

In Section 10, we decouple one spin from the others. The remaining spins then evolve under a law whose profile need not minimize any functional and whose auxiliary field has unequal variances. The following class of laws contains both situations.

Definition 4.1 (Generalized law). Let $\Lambda = \text{diag}(\vartheta_1, \dots, \vartheta_N)$ with $\vartheta_i > 0$, and let $(\kappa_{ij})_{1 \leq i \leq j \leq N}$ be real numbers; Λ and κ are deterministic. Let $\omega = (\omega_{ij})_{1 \leq i \leq j \leq N}$ be

independent standard Gaussian variables, and write $\mathbb{E}_{\text{ref}}$ for expectation under their law. Let $\mu = s\delta_0 + \nu$ be either an admissible profile in the sense of Section 3.2, so that $0 < s \leq 1/2$, or, with $s = 0$, an arbitrary deterministic probability measure $\mu = \nu$ on $[0, 1]$. Let $a(q) = \mu([0, q])$, and let U solve

$$\partial_q U = -\frac{1}{2}(\text{Tr}(\Lambda \nabla^2 U) + a(q) \nabla U^T \Lambda \nabla U), \quad U(1, x; \omega) = \log \sum_{\sigma \in \{-1, 1\}^N} e^{\mathcal{H}_\sigma(\omega, x)}, \quad (4.1)$$

where $\mathcal{H}_\sigma(\omega, x) = \sum_{i \leq j} \kappa_{ij} \omega_{ij} \sigma_i \sigma_j + x \cdot \sigma$. The *generalized law* of (Λ, κ, μ) is the joint law of (ω, X, σ) built in three stages.

- (i) The Gaussian law of ω is tilted by the density $e^{sU(0,0;\omega)} / \mathbb{E}_{\text{ref}} e^{sU(0,0;\omega)}$.
- (ii) Given ω , the field solves $dX_q = a(q) \Lambda \nabla U(q, X_q; \omega) dq + \Lambda^{1/2} dB_q$ with $X_0 = 0$, where B is a standard Brownian motion in $\mathbb{R}^N$.
- (iii) Given ω and X , the spin σ is drawn with probabilities proportional to $e^{\mathcal{H}_\sigma(\omega, X_1)}$.

As in Section 3.3, we write $\mathbb{P}$ and $\mathbb{E}$ for this law, and we define $\mathcal{G}_q = \sigma(\omega, X_r : 0 \leq r \leq q)$, $m_q = \mathbb{E}[\sigma \mid \mathcal{G}_q]$, $C_q = \text{Cov}(\sigma \mid \mathcal{G}_q)$, and $H_q = \nabla_x^2 U(q, X_q; \omega)$.

When $s = 0$, the outer density in (i) is identically one. Recall that a Hopf–Cole step of mass zero is an ordinary Gaussian expectation. The fractional-moment functionals of Section 3 still have $s > 0$. The equation (4.1) has a solution for every such profile, which is obtained from step profiles by the Hopf–Cole recursion and uniform approximation, as we show in Step 6 of the proof of Lemma 4.2. With $\Lambda = cI$, with $\omega_{ij} = g_{ij}$ and $\kappa_{ij} = \beta \sqrt{t/N}$ for $i < j$, and with $\omega_{ii} = z_i$ and $\kappa_{ii} = \beta \sqrt{t/(2N)}$, we have $\mathcal{H}_\sigma(\omega, x) = \beta \sqrt{t} \sigma^T W \sigma / 2 + x \cdot \sigma$. In this case, the generalized law is the law of Section 3.3. The scalar law of Section 9 is the generalized law with $N = 1$, $\Lambda = d$, $\kappa = 0$, and $s = 0$. The cavity laws of Section 10 are generalized laws with $\Lambda = \text{diag}(cI_n, b(u))$ and independent Gaussian couplings with deterministic variances.

In the next lemma, we adapt the argument of the proof of [54, Lemma 2.3], which is also used in the proof of [56, Lemma 2.1]. That argument combines joint convexity in the Gaussian drivers and a spin source, the inverse-Hessian Brascamp–Lieb inequality, and a gauge symmetry that removes the conditioning on the terminal spin. We use it for averages of squared conditional means on selected coordinates, and for diagonal anisotropic laws.

Lemma 4.2 (Averaged self-overlap under an arbitrary profile). *Let $\mu = s\delta_0 + \nu$ be an admissible profile, and consider either its joint law from Section 3.3, for some $0 < \rho \leq 1$, or a generalized law; a generalized law also allows $s = 0$ and an arbitrary probability measure $\mu = \nu$. Then $m_q = \nabla U(q, X_q) = \mathbb{E}[\sigma \mid \mathcal{G}_q]$, $m_0 = 0$, and $a(q)C_q \preceq H_q \preceq C_q$ for every $q \in [0, 1]$. Fix a*

set of $n \geq 1$ coordinates, let m_q^- be the restriction of m_q to these coordinates, and let H_q^{--} and C_q^{--} be the corresponding principal submatrices of H_q and C_q . For every deterministic $k \in L^2(\nu)$,

$$\begin{aligned} \text{Var} \left(\int k(q) \frac{|m_q^-|^2}{n} \nu(dq) \right) &\leq \frac{4}{n^2} \int k(q)^2 \mathbb{E}[(m_q^-)^T H_q^{--} m_q^-] \nu(dq) \\ &\leq \frac{4}{n^2} \int k(q)^2 \mathbb{E}[(m_q^-)^T C_q^{--} m_q^-] \nu(dq). \end{aligned} \quad (4.2)$$

Recall that the law of Section 3.3 is a generalized law, so that it suffices to prove the lemma for generalized laws. The proof also gives the martingale differential

$$dm_q = H_q \Lambda^{1/2} dB_q \quad (0 \leq q \leq 1) \quad (4.3)$$

for every generalized law, as we will see in Steps 1 and 6 below.

Proof of Lemma 4.2. The proof has six steps. We first prove the martingale identities and the covariance order for step profiles. Next, still for step profiles, we write the law in terms of independent Gaussian coordinates. We compare the curvature of the posterior density with the gradient of the squared conditional mean, and we apply the Brascamp–Lieb inequality with the terminal spin fixed. We then remove the conditioning on the spin by a gauge symmetry. Finally, we pass to general profiles and general k .

Step 1: martingale and covariance order. Let μ be a step profile. We adapt Step 1 of the proof of Proposition 3.1, replacing cI by Λ . The generator of the field is $\text{Tr}(\Lambda \nabla^2)/2 + a(q)(\Lambda \nabla U) \cdot \nabla$. Differentiating (4.1) in x and applying Itô's formula along the field on each interval between grid points, we find that

$$d \nabla U(q, X_q; \omega) = H_q \Lambda^{1/2} dB_q, \quad dH_q = -a(q)H_q \Lambda H_q dq + \nabla H_q \Lambda^{1/2} dB_q,$$

where the last term is the contraction of the third spatial derivative of U with $\Lambda^{1/2} dB_q$. The stochastic integrals are martingales, since the spatial derivatives are bounded at fixed N . The gradient $\nabla U(1, X_1; \omega)$ is the Gibbs mean of the terminal spin. By the tower property, we have $\nabla U(q, X_q; \omega) = \mathbb{E}[\sigma \mid \mathcal{G}_q] = m_q$, which proves (4.3) for step profiles. Using $C_1 = H_1$, the martingale isometry, and the equation for H , we find

$$\begin{aligned} C_q &= \mathbb{E}[H_1 \mid \mathcal{G}_q] + \int_q^1 \mathbb{E}[H_u \Lambda H_u \mid \mathcal{G}_q] du, \\ H_q &= \mathbb{E}[H_1 \mid \mathcal{G}_q] + \int_q^1 a(u) \mathbb{E}[H_u \Lambda H_u \mid \mathcal{G}_q] du. \end{aligned}$$

Note that the matrices $H_u \Lambda H_u$ are positive semidefinite and that $a(q) \leq a(u) \leq 1$ for $u \geq q$. Subtracting the two identities as in Step 1 of the proof of Proposition 3.1, we obtain $a(q)C_q \preceq H_q \preceq C_q$. Note that $\mathcal{H}_{-\sigma}(\omega, x) = \mathcal{H}_\sigma(\omega, -x)$. Replacing σ by $-\sigma$ in the terminal sum, we see that $U(1, \cdot; \omega)$ is even. The equation preserves this symmetry, which implies that $U(q, \cdot; \omega)$ is even and that $m_0 = 0$.

Step 2: Gaussian coordinates and the posterior density. Let $0 = q_0 < q_1 < \dots < q_K = 1$ and $a = a_\ell$ on $[q_\ell, q_{\ell+1})$, with $a_\ell \geq s \geq 0$. Then the cumulative function is constant between grid points, and the free measure ν is carried by the grid points, with $\nu(\{q_0\}) = a_0 - s$, $\nu(\{q_\ell\}) = a_\ell - a_{\ell-1}$ for $1 \leq \ell \leq K-1$, and $\nu(\{q_K\}) = 1 - a_{K-1}$. Let $z_1, \dots, z_K$ be independent standard Gaussian vectors in $\mathbb{R}^N$, let

$$X_{q_\ell} = \sum_{h=1}^{\ell} (q_h - q_{h-1})^{1/2} \Lambda^{1/2} z_h \quad (0 \leq \ell \leq K),$$

and let Z collect the coordinates of ω and of $z_1, \dots, z_K$. We write $U_q(Z) = U(q, X_q; \omega)$ at the grid points. We recall that (4.1) is solved by the Hopf–Cole recursion $U_{q_\ell} = a_\ell^{-1} \log \mathbb{E}_{z_{\ell+1}} e^{a_\ell U_{q_{\ell+1}}}$ when $a_\ell > 0$, and by $U_{q_\ell} = \mathbb{E}_{z_{\ell+1}} U_{q_{\ell+1}}$ when $a_\ell = 0$. We follow the construction of [4, (2.1)–(2.3)]. We claim that, restricted to ω , to the field at the grid points, and to the spin, the generalized law is the law of $(\omega, (X_{q_\ell})_\ell, \sigma)$ when (Z, σ) has the density

$$\frac{\exp\{\mathcal{H}_\sigma(\omega, X_1) - \int U_q(Z) \nu(dq)\}}{\mathbb{E}_{\text{ref}} e^{sU(0,0;\omega)}}$$

with respect to the standard Gaussian law of Z and counting measure on $\{-1, 1\}^N$. Indeed, the outer tilt contributes sU_{q_0} to the exponent, the transition from q_ℓ to $q_{\ell+1}$ contributes $a_\ell(U_{q_{\ell+1}} - U_{q_\ell})$, and the terminal Gibbs probability contributes $\mathcal{H}_\sigma(\omega, X_1) - U_{q_K}$. In the sum, we see that the coefficients of $U_{q_0}, U_{q_1}, \dots, U_{q_K}$ are

$$s - a_0, \quad a_0 - a_1, \quad \dots, \quad a_{K-2} - a_{K-1}, \quad a_{K-1} - 1,$$

which are the negatives of the masses of ν at the grid points. A transition of mass zero has likelihood one, and the calculation also holds at $s = 0$, where the denominator equals one. For a fixed spin σ , we find that the negative logarithm of the conditional density of Z is, up to a constant, the posterior potential

$$\Phi_\sigma(Z) = \frac{1}{2}|Z|^2 - \mathcal{H}_\sigma(\omega, X_1) + \int U_q(Z) \nu(dq).$$

Note that $\mathcal{H}_\sigma(\omega, X_1)$ is linear in Z and that each U_q is convex in Z . To see the convexity, observe that the terminal value in (4.1) is a log-sum-exp of linear functions of (ω, x) , that each Hopf–Cole step of positive mass preserves convexity by Hölder’s inequality, and that Gaussian expectation at mass zero preserves it as well. Since X_q is linear in Z , we obtain

$$\nabla_Z^2 \Phi_\sigma = \mathcal{Q} = I + \int \mathcal{A}_q \nu(dq) \succeq I, \quad \mathcal{A}_q = \nabla_Z^2 U_q \succeq 0.$$

The matrix $\mathcal{Q}$ does not depend on σ .

Step 3: joint source curvature. Let $\iota: \mathbb{R}^n \rightarrow \mathbb{R}^N$ embed a vector in the selected coordinates, and let $\mathcal{U}_q(Z, y) = U(q, X_q(Z) + \iota y; \omega)$ for $y \in \mathbb{R}^n$. This function is jointly convex in (Z, y) , since $U(q, x; \omega)$ is convex in (ω, x) by Step 2 and the map $(Z, y) \mapsto (\omega, X_q(Z) + \iota y)$ is linear. We write $\mathcal{B}_q = \nabla_Z \nabla_y \mathcal{U}_q(Z, 0)$, a matrix with n columns. Since $\nabla_y \mathcal{U}_q(Z, 0) = m_q^-$ and $\nabla_y^2 \mathcal{U}_q(Z, 0) = H_q^{--}$, joint convexity and the chain rule give

$$\begin{pmatrix} \mathcal{A}_q & \mathcal{B}_q \\ \mathcal{B}_q^T & H_q^{--} \end{pmatrix} \succeq 0, \quad \nabla_Z |m_q^-|^2 = 2\mathcal{B}_q m_q^-. \quad (4.4)$$

For a vector w and $r \in \mathbb{R}$, we evaluate this positive semidefinite form at (rw, m_q^-) . The resulting quadratic polynomial in r is nonnegative. This implies that its discriminant is nonpositive:

$$|w^T \mathcal{B}_q m_q^-|^2 \leq (w^T \mathcal{A}_q w) (m_q^-)^T H_q^{--} m_q^-.$$

The same statements hold at $q = 0$, where $m_0^- = 0$.

Step 4: the Brascamp–Lieb inequality. Let $Y = n^{-1} \int k(q) |m_q^-|^2 \nu(\mathrm{d}q)$, a finite sum over grid points, and let $b = \int k(q) \mathcal{B}_q m_q^- \nu(\mathrm{d}q)$, so that $\nabla_Z Y = 2b/n$ by (4.4). For every vector w , using the last display, the Cauchy–Schwarz inequality for the integral against ν , and $\int \mathcal{A}_q \nu(\mathrm{d}q) = \mathcal{Q} - I \preceq \mathcal{Q}$, we obtain

$$\begin{aligned} |w^T b|^2 &\leq \left(\int w^T \mathcal{A}_q w \nu(\mathrm{d}q) \right) \int k(q)^2 (m_q^-)^T H_q^{--} m_q^- \nu(\mathrm{d}q) \\ &\leq (w^T \mathcal{Q} w) \int k(q)^2 (m_q^-)^T H_q^{--} m_q^- \nu(\mathrm{d}q). \end{aligned}$$

Taking $w = \mathcal{Q}^{-1}b$ and dividing by $b^T \mathcal{Q}^{-1}b$ when it is positive, we find that

$$b^T \mathcal{Q}^{-1}b \leq \int k(q)^2 (m_q^-)^T H_q^{--} m_q^- \nu(\mathrm{d}q);$$

if $b^T \mathcal{Q}^{-1}b = 0$, this inequality holds trivially. Fix σ . The potential Φ_σ is smooth, its Hessian $\mathcal{Q}$ is at least I , and it has Gaussian tails. By the Brascamp–Lieb inequality [21], in the form [4, Proposition 2.2, (2.14)], we have $\mathrm{Var}(Y \mid \sigma) \leq \mathbb{E}[\nabla_Z Y^T \mathcal{Q}^{-1} \nabla_Z Y \mid \sigma]$, and we conclude that

$$\mathrm{Var}(Y \mid \sigma) \leq \frac{4}{n^2} \mathbb{E} \left[\int k(q)^2 (m_q^-)^T H_q^{--} m_q^- \nu(\mathrm{d}q) \mid \sigma \right].$$

Step 5: removing the conditioning. For $\tau \in \{-1, 1\}^N$, we consider the map

$$\sigma_i \mapsto \tau_i \sigma_i, \quad \omega_{ij} \mapsto \tau_i \tau_j \omega_{ij}, \quad z_{h,i} \mapsto \tau_i z_{h,i}.$$

This map preserves the Gaussian reference law. Because Λ is diagonal, it multiplies each X_q coordinatewise by τ . Further, $\mathcal{H}_{\tau\sigma}$ evaluated at the transformed (ω, x) equals

$\mathcal{H}_\sigma(\omega, x)$, and the map leaves every U_q invariant. It therefore preserves the joint density of Step 2. It maps $m_{q,i}$ to $\tau_i m_{q,i}$ and $(H_q)_{ij}$ to $\tau_i \tau_j (H_q)_{ij}$, so that Y and $(m_q^-)^T H_q^- m_q^-$ are invariant. The map acts transitively on spin configurations, which implies that the conditional law of these quantities given σ does not depend on σ . In particular, $\mathbb{E}[Y \mid \sigma]$ is constant, and the law of total variance implies that $\text{Var}(Y) = \mathbb{E} \text{Var}(Y \mid \sigma)$. Combining this with Step 4, we obtain the first inequality in (4.2) for step profiles. The second inequality follows from $H_q^- \preceq C_q^-$, since principal submatrices preserve the order $H_q \preceq C_q$.

Step 6: general profiles and general k . We approximate μ by step profiles with the same root mass, as in Step 4 of the proof of Proposition 3.1 and [4, Section 2.6]. That is, we push ν forward to finite meshes that contain 0, 1, and any prescribed finite set of levels, with mesh size tending to zero. The argument of [4, Section 2.6] uses only that the spins are bounded, that spatial derivatives of fixed order are bounded uniformly over step profiles at fixed N , that the solution is Lipschitz in the profile in L^1 , that the outer tilt has Gaussian domination, and that the disorder, the Brownian motion, and the terminal spin can be coupled. These facts carry over to a bounded diagonal Λ and to couplings $\kappa_{ij} \omega_{ij}$. For the Lipschitz bound, let U and $\tilde{U}$ be the solutions for the cumulative functions a and $\tilde{a}$, and let $\varphi = U - \tilde{U}$. Subtracting the two equations, we find

$$\partial_q \varphi + \frac{1}{2} \text{Tr}(\Lambda \nabla^2 \varphi) + \frac{a}{2} (\nabla U + \nabla \tilde{U})^T \Lambda \nabla \varphi = -\frac{a - \tilde{a}}{2} \nabla \tilde{U}^T \Lambda \nabla \tilde{U}, \quad \varphi(1, \cdot; \omega) = 0.$$

Since each coordinate of $\nabla \tilde{U}$ is a conditional spin mean, we have $\nabla \tilde{U}^T \Lambda \nabla \tilde{U} \leq \text{Tr} \Lambda$. By the backward representation of φ , we have

$$\sup_{x, \omega} |\varphi(q, x; \omega)| \leq \frac{\text{Tr} \Lambda}{2} \int_q^1 |a(u) - \tilde{a}(u)| \, du \leq \frac{N}{2} \max_i \vartheta_i \|a - \tilde{a}\|_{L^1([0,1])}.$$

This replaces the constant $cN/2$ of [4, (2.19)]. It follows that the ratio of the outer densities for μ and for an approximating step profile converges to one uniformly in ω . At $s = 0$, the outer density is identically one, and the approximation requires no division by a root mass.

Under the common coupling of [4, Section 2.6.1], the processes X , m , H , and C for the step profiles converge to those for μ , uniformly in q , in probability. Then the identities of Step 1, including (4.3), pass to the limit, where for (4.3) we use the Itô isometry and the bounded convergence of H . For continuous k , the integrands $q \mapsto k(q) |m_q^-|^2$ and $q \mapsto k(q)^2 (m_q^-)^T H_q^- m_q^-$ are bounded and continuous along each path. From any subsequence we may extract a further subsequence on which the process convergence is almost surely uniform in q . On that subsequence the integrands converge uniformly, and their integrals then converge by weak convergence of the deterministic free measures. The integrands are bounded, the ratio of the outer densities converges to one uniformly

in ω , and, since $U(0,0;\omega)$ grows at most linearly in ω , the outer densities are dominated, uniformly along the approximation, by a function that is integrable under the Gaussian reference law. This allows passage to the expectations and variances. Thus (4.2) holds for continuous k . Next, we show that both sides of (4.2) are continuous in $k \in L^2(\nu)$. Since $|m_q^-|^2 \leq n$, we have

$$\left\| \frac{1}{n} \int (k_1(q) - k_2(q)) |m_q^-|^2 \nu(dq) \right\|_{L^2(\mathbb{P})} \leq \nu([0,1])^{1/2} \|k_1 - k_2\|_{L^2(\nu)},$$

which shows that the left side of (4.2) is continuous in k . Further, $(m_q^-)^T H_q^- m_q^- \leq (m_q^-)^T C_q^- m_q^- \leq n \|C_q\|_{\text{op}} \leq nN$, which implies that the right sides change by at most $4n^{-1}N \|k_1 - k_2\|_{L^2(\nu)} \|k_1 + k_2\|_{L^2(\nu)}$. Since continuous functions are dense in $L^2(\nu)$, this proves the lemma. $\square$

4.2 Differentiating normalized Gaussian path expectations

In Sections 9 to 11, we differentiate normalized expectations under Gaussian field laws while keeping their profile measures fixed. The parameter changes the covariance of the Gaussian variables. In Section 9, it changes the clock of the field path, and in Sections 10 and 11, it changes the variance of a row of couplings and the clock of one field coordinate. The densities that we encounter are exponentials of integrals of smooth functions of the field against the profile measure, which may have atoms. We use the following version of the Gaussian heat identity, for finitely many Gaussian variables and continuous field paths. Its finite-dimensional input is the classical formula for the derivative of a Gaussian expectation with respect to the covariance. A general version, under hypotheses different from ours, is proved in [85, Theorem 1]. In the proof, we reduce to finite Gaussian vectors by Bernstein approximation of the paths, while we keep the profile measure, including its atoms, exact.

Setting. Let $I \subset \mathbb{R}$ be a nondegenerate compact interval with interior I° , and let $z \in I$ be the parameter. Under a reference probability measure, with expectation $\mathbb{E}_{\text{ref}}$, we consider the following random variables.

- (i) *Finite coordinates* $G_z^{\text{fin}} = (v_k(z)^{1/2} \xi_k)_{1 \leq k \leq n_0}$, where $\xi_1, \dots, \xi_{n_0}$ are independent standard Gaussian variables and $v_k: I \rightarrow [0, \infty)$.
- (ii) *Field paths* on a fixed finite rooted tree. Each edge e carries the level interval $[0, 1]$, after an affine change of variable, and d components. For each edge e and component $j \leq d$, let $B_{e,j}$ be a standard Brownian motion, and let $T_{e,j,z}: [0, 1] \rightarrow [0, \infty)$ be a deterministic nondecreasing clock such that $T_{e,j,z}(0) = 0$. The Brownian motions for different edges and components are mutually independent. The path on the edge e is

$$G_z(e, j, v) = G_z(e', j, 1) + B_{e,j}(T_{e,j,z}(v)), \quad 0 \leq v \leq 1,$$

where e' is the parent edge of e , and the first term is 0 if e starts at the root.

- (iii) A finite family ω of independent standard Gaussian variables that do not depend on z and are independent of the variables in (i) and (ii), and a variable σ that ranges over a nonempty finite set and is integrated against counting measure.

The variables ξ_k and the Brownian motions $B_{e,j}$ are mutually independent. Thus the finite coordinates and the field paths are independent under the reference law. This setting includes all the scalar and cavity applications below. We write $\mathcal{I}$ for the index space formed by the finite coordinates $1 \leq k \leq n_0$ and the field points and components (e, j, v) . It is a compact metric space, the disjoint union of a finite set and finitely many copies of $[0, 1]$. The Gaussian coordinates $G_z = (G_z^{\text{fin}}, G_z^{\text{path}})$ form a random element of the space $\mathcal{C}$ of functions on $\mathcal{I}$ that are continuous on each copy of $[0, 1]$, with the norm $|G^{\text{fin}}| + \|G^{\text{path}}\|_\infty$. We denote their covariance by $K_z(i, i') = \mathbb{E}_{\text{ref}}[G_z(i)G_z(i')]$ for $i, i' \in \mathcal{I}$. For example, for one edge and one component, $K_z(v, v') = T_z(\min\{v, v'\})$.

Let $\ell_z(G) = \ell_z(G, \omega, \sigma)$ and $\Phi_z(G) = \Phi_z(G, \omega, \sigma)$ be real functions of $z \in I$, $G \in \mathcal{C}$, ω , and σ . We write $\dot{\ell}_z$ and $\dot{\Phi}_z$ for their derivatives in z , with (G, ω, σ) fixed, and $\dot{K}_z$ for the derivative of K_z in z . For a continuous kernel A on $\mathcal{I}^2$ and a finite signed measure B on $\mathcal{I}^2$, we write $\langle A, B \rangle = \iint A(i, i')B(di, di')$, and $S \otimes P$ denotes the product of two signed measures on $\mathcal{I}$.

Lemma 4.3 (Normalized Gaussian path heat identity). *In the setting above, assume the following.*

- (A1) *Variances and clocks. Each v_k and each $T_{e,j,z}(v)$ is continuous in $z \in I$, jointly with v for the clocks, and continuously differentiable in $z \in I^\circ$. For $z \in I^\circ$, each nonconstant v_k is positive and each clock is strictly increasing in v . Coordinates with $v_k \equiv 0$ are omitted.*
- (A2) *Covariance. The covariance $K_z(i, i')$ is differentiable in $z \in I^\circ$, and the maps $(z, i, i') \mapsto K_z(i, i')$ and $(z, i, i') \mapsto \dot{K}_z(i, i')$ are jointly continuous on $I \times \mathcal{I}^2$, where $\dot{K}_z$ at the endpoints of I denotes the continuous extension.*
- (A3) *Functionals. For each σ , the functions $\ell_z(G, \omega, \sigma)$, $\Phi_z(G, \omega, \sigma)$, $\dot{\ell}_z(G, \omega, \sigma)$, and $\dot{\Phi}_z(G, \omega, \sigma)$ are jointly continuous in $(z, G, \omega) \in I \times \mathcal{C} \times \mathbb{R}^{n_1}$, where n_1 is the number of auxiliary Gaussian coordinates, with $\dot{\ell}_z$ and $\dot{\Phi}_z$ extended continuously to the endpoints of I . The functions Φ_z , $\dot{\Phi}_z$, and ℓ_z are bounded uniformly in all variables, and there exists $C < \infty$ such that*

$$|\ell_z(G)| \leq C(1 + |\omega| + |G^{\text{fin}}| + \|G^{\text{path}}\|_\infty).$$

- (A4) *Derivative measures. For fixed (z, ω, σ) , the maps $G \mapsto \ell_z(G)$ and $G \mapsto \Phi_z(G)$ are twice continuously differentiable on $\mathcal{C}$, and their derivatives are represented by finite signed measures*

$$S = D\ell_z, \quad M = D^2\ell_z, \quad P = D\Phi_z, \quad Q = D^2\Phi_z$$

on $\mathcal{I}$ and $\mathcal{I}^2$, in the sense that, for example,

$$D\ell_z(G)[h] = \int h \, dS, \quad D^2\ell_z(G)[h, h'] = \iint h(i)h'(i') M(di, di') \quad (h, h' \in \mathcal{C}).$$

For fixed (ω, σ) , the four measures depend continuously on (z, G) in the total-variation norm, and their total variations are bounded uniformly in all variables.

Let $Z_z = \mathbb{E}_{\text{ref}} e^{\ell_z(G_z)}$, where $\mathbb{E}_{\text{ref}}$ includes the sum over σ , and let $\mathbb{E}_z \Phi = Z_z^{-1} \mathbb{E}_{\text{ref}} [e^{\ell_z(G_z)} \Phi]$. Then, for every $z \in I$,

$$\begin{aligned} \frac{d}{dz} \mathbb{E}_z \Phi_z &= \mathbb{E}_z \dot{\Phi}_z + \text{Cov}_z(\Phi_z, \dot{\ell}_z) \\ &\quad + \frac{1}{2} \mathbb{E}_z \left\langle \dot{K}_z, Q + S \otimes P + P \otimes S + (\Phi_z - \mathbb{E}_z \Phi_z)(M + S \otimes S) \right\rangle, \end{aligned} \quad (4.5)$$

where the functionals and measures are evaluated at G_z . The derivative is one-sided at the endpoints of I , including endpoints at which a finite-coordinate variance vanishes. The same differentiation gives

$$\frac{d}{dz} \log Z_z = \mathbb{E}_z \dot{\ell}_z + \frac{1}{2} \mathbb{E}_z \langle \dot{K}_z, M + S \otimes S \rangle. \quad (4.6)$$

Remark 4.4. We show in the proof that the right sides of (4.5) and (4.6) are continuous in $z \in I$, so that $z \mapsto \mathbb{E}_z \Phi_z$ and $z \mapsto \log Z_z$ are continuously differentiable on I . The assumptions hold, in particular, for the smooth local functionals that we use in Sections 9 to 11, integrated against a fixed finite measure. For instance, let ω be a finite measure on the path part of $\mathcal{I}$, possibly with atoms, and let $\ell_z(G) = - \int V_z(i, G(i)) \, \omega(di)$, where $V_z(i, x)$ is twice continuously differentiable in x , and $\partial_x V_z$ and $\partial_x^2 V_z$ are bounded and jointly continuous in (z, i, x) . Then

$$S = -\partial_x V_z(\cdot, G(\cdot)) \, \omega, \quad M = -\text{diag}_\#(\partial_x^2 V_z(\cdot, G(\cdot)) \, \omega),$$

where $\text{diag}_\#$ is the pushforward under $i \mapsto (i, i)$. The total variation of the difference of the scores at G and G' is at most $\omega(\mathcal{I}) \sup_i |\partial_x V_z(i, G(i)) - \partial_x V_z(i, G'(i))|$, and the same holds for the Hessians with $\partial_x^2 V_z$. We note that these functions are uniformly continuous on compact sets and that G is bounded, which gives the continuity in (A4) for this term. An atom of ω produces an atom of M on the diagonal, and its contribution to $\langle \dot{K}_z, M \rangle$ involves $\dot{K}_z$ on the diagonal.

Proof of Lemma 4.3. The proof has seven steps. We first prove the identities for finite Gaussian vectors. We then approximate the paths by their Bernstein polynomials and couple the approximations on one probability space. Next, we show that the derivative measures and the covariance kernels converge, and we establish a common integrable envelope and a lower bound for the normalizers. Finally, we pass to the limit in the interior of I and treat the endpoints.

Step 1: the finite-dimensional identity. We begin with a finite Gaussian vector. Let J be an open interval, and let Y_z , $z \in J$, be a centered Gaussian vector in $\mathbb{R}^p$ whose covariance Σ_z is continuously differentiable in z . We allow the matrix Σ_z to be singular. Let $F_z: \mathbb{R}^p \rightarrow \mathbb{R}$ be continuously differentiable in z and twice continuously differentiable in $y \in \mathbb{R}^p$, and assume that F_z , $\partial_z F_z$, ∇F_z , and $\nabla^2 F_z$ are bounded by $Ce^{C|y|}$, locally uniformly in z . We claim that

$$\frac{d}{dz} \mathbb{E} F_z(Y_z) = \mathbb{E} \partial_z F_z(Y_z) + \frac{1}{2} \mathbb{E} \langle \dot{\Sigma}_z, \nabla^2 F_z(Y_z) \rangle. \quad (4.7)$$

Let $\varepsilon > 0$, let ζ be a standard Gaussian vector in $\mathbb{R}^p$, and let $A_z = (\Sigma_z + \varepsilon I)^{1/2}$. The matrix square root is smooth on positive definite matrices, so A_z is continuously differentiable in z . By dominated convergence, we may differentiate $\mathbb{E} F_z(A_z \zeta)$ under the expectation. The chain rule gives, besides $\mathbb{E} \partial_z F_z(A_z \zeta)$, the term $\mathbb{E} [\nabla F_z(A_z \zeta) \cdot \dot{A}_z \zeta]$. Using Gaussian integration by parts, $\mathbb{E} [\zeta_l g(\zeta)] = \mathbb{E} [\partial_l g(\zeta)]$, we find that this term equals $\mathbb{E} \text{Tr}(\dot{A}_z \nabla^2 F_z(A_z \zeta) A_z)$. Since $\nabla^2 F_z$ is symmetric and $\dot{A}_z A_z + A_z \dot{A}_z = \dot{\Sigma}_z$, this is $\mathbb{E} \langle \dot{\Sigma}_z, \nabla^2 F_z(A_z \zeta) \rangle / 2$. We integrate this identity over $[z_1, z_2] \subset J$ and let $\varepsilon \downarrow 0$. The vector $A_z \zeta$ has the law of $Y_z + \sqrt{\varepsilon} \zeta'$, where ζ' is a standard Gaussian vector independent of Y_z . Since the growth bounds give domination, we obtain the integrated form of (4.7). Realizing $Y_z = \Sigma_z^{1/2} \zeta$, which is continuous in z , we see by the same domination that the right side of (4.7) is continuous in z . The fundamental theorem of calculus then gives (4.7). This is the classical formula. Note that we never invert Σ_z .

We will apply (4.7) to $e^{\ell z}$ and to $e^{\ell z} \Phi_z$, composed with a linear map from $\mathbb{R}^p$ to $\mathcal{C}$, as in Step 2. By the product rule, we have $De^\ell = e^\ell S$, $D^2 e^\ell = e^\ell (M + S \otimes S)$, and

$$D^2(e^\ell \Phi) = e^\ell [Q + S \otimes P + P \otimes S + \Phi(M + S \otimes S)].$$

Inserting these derivatives into (4.7) for the numerator $\mathbb{E}_{\text{ref}}[e^{\ell z} \Phi_z]$ and the denominator $\mathbb{E}_{\text{ref}} e^{\ell z}$, and using the quotient rule, we obtain (4.5) for a finite Gaussian vector. Observe that the centering $\Phi_z - \mathbb{E}_z \Phi_z$ comes from the derivative of the denominator. Dividing the derivative of the denominator by the denominator gives (4.6).

Step 2: Bernstein approximation and kernel convergence. We use the Bernstein approximation introduced in [13]. Let $m \geq 1$ and $v_l = l/m$ for $0 \leq l \leq m$. For a continuous function f on $[0, 1]$, the Bernstein polynomial is

$$\Pi_m f(v) = \sum_{l=0}^m b_{m,l}(v) f(v_l), \quad b_{m,l}(v) = \binom{m}{l} v^l (1-v)^{m-l}.$$

The operator Π_m is linear, and $\Pi_m f(0) = f(0)$ and $\Pi_m f(1) = f(1)$. The weights are nonnegative and sum to one, which gives $\|\Pi_m f\|_\infty \leq \|f\|_\infty$. Further, for every $\delta > 0$, we have

$$\|\Pi_m f - f\|_\infty \leq \omega_f(\delta) + \frac{\|f\|_\infty}{2m\delta^2},$$

where ω_f is the modulus of continuity of f . This follows from Chebyshev's inequality for a binomial variable with m trials, whose variance is at most $m/4$. In particular, $\Pi_m f \rightarrow f$ uniformly over any bounded equicontinuous family of functions f . The same holds for the tensor product $\Pi_m \otimes \Pi_m$ acting on continuous functions on $[0, 1]^2$, with the bound $\omega_f(\delta) + \|f\|_\infty / (m\delta^2)$.

On each edge and component, we let Π_m act on the path, and we leave the finite coordinates unchanged. We denote the resulting operator on $\mathcal{C}$ also by Π_m . Then $\Pi_m G_z$ is a linear function of the finite Gaussian vector Y_z formed by the finite coordinates and the sampled values $G_z(e, j, v_l)$. A value at a vertex of the tree may appear in several coordinates of Y_z . Step 1 allows such repeated coordinates, since it does not require an invertible covariance. The covariance of Y_z consists of the values of K_z at the sampled points. By (A2), it is continuously differentiable on I° , with derivative given by the values of $\dot{K}_z$. The growth conditions of Step 1 hold by (A3) and (A4), since $\|\Pi_m y\|_{\mathcal{C}} \leq 2|y|$.

We set

$$\ell_z^{(m)} = \ell_z(\Pi_m G_z), \quad \Phi_z^{(m)} = \Phi_z(\Pi_m G_z), \quad Z_z^{(m)} = \mathbb{E}_{\text{ref}} e^{\ell_z^{(m)}},$$

and we normalize the approximate moment $\mathbb{E}_z^{(m)} \Phi_z^{(m)} = (Z_z^{(m)})^{-1} \mathbb{E}_{\text{ref}}[e^{\ell_z^{(m)}} \Phi_z^{(m)}]$ by its own normalizer $Z_z^{(m)}$. By the chain rule, the derivative of $\ell_z(\Pi_m y)$ in the sampled coordinate (e, j, v_l) is $\int b_{m,l} dS$, where $b_{m,l}$ is placed on the edge e and the component j . We obtain the second derivatives in the same way from M . We use Step 1 for fixed (ω, σ) , and we then integrate over ω and sum over σ . The domination of Step 5 justifies this. We conclude that (4.5) and (4.6) hold for the approximations at every $z \in I^\circ$, with $\mathbb{E}_z$ replaced by $\mathbb{E}_z^{(m)}$ and with the functionals and measures evaluated at $\Pi_m G_z$. In these identities, $\dot{K}_z$ is replaced by the kernel

$$\dot{K}_z^{(m)}(v, w) = \sum_{l, l'} b_{m,l}(v) b_{m,l'}(w) \dot{K}_z(v_l, v_{l'})$$

on each pair of edges and components. Blocks between finite coordinates are unchanged, and blocks between a finite coordinate k and a path use the same sampling in their path variable:

$$\dot{K}_z^{(m)}(k, (e, j, v)) = \sum_l b_{m,l}(v) \dot{K}_z(k, (e, j, v_l)).$$

By (A2), the kernels K_z and $\dot{K}_z$ are uniformly continuous on the compact set $I \times \mathcal{I}^2$. Hence the families $\{K_z\}_{z \in I}$ and $\{\dot{K}_z\}_{z \in I}$ are bounded and equicontinuous. Using the Bernstein bound, we obtain

$$\sup_{z \in I} \|\dot{K}_z^{(m)} - \dot{K}_z\|_\infty \longrightarrow 0, \quad \sup_{z \in I} \|K_z^{(m)} - K_z\|_\infty \longrightarrow 0, \quad (4.8)$$

where $K_z^{(m)}$, the covariance of $\Pi_m G_z$, is defined in the same way.

Step 3: coupling. In Step 1, we used only the law of Y_z at each fixed z . To pass to the limit, we use the realization of the setting, with one Brownian motion $B_{e,j}$ for each edge and component, evaluated at the varying clocks. The finite coordinates are $v_k(z)^{1/2}\xi_k$ for fixed ξ_k , independent of these Brownian motions as in the setting. By (A1), the map $(z, v) \mapsto G_z(e, j, v)$ is almost surely continuous on the compact set $I \times [0, 1]$. It is therefore uniformly continuous, and the family $\{G_z(e, j, \cdot)\}_{z \in I}$ is equicontinuous. By Step 2, we have $\sup_{z \in I} \|\Pi_m G_z - G_z\|_{\mathcal{C}} \rightarrow 0$ almost surely. Let $\bar{T}$ be the maximum of the clocks over $I \times [0, 1]$, with $\bar{T} = 0$ if there are no field paths. Let $\bar{v} = \max(\{0\} \cup \{v_k(z) : 1 \leq k \leq n_0, z \in I\})$, so that $\bar{v} = 0$ if there are no finite coordinates. Since the value of a path at a point of the edge e is a sum of increments of the Brownian motions on e and its ancestor edges, we find that

$$\sup_{z \in I} \|\Pi_m G_z\|_{\mathcal{C}} \leq \sup_{z \in I} \|G_z\|_{\mathcal{C}} \leq \bar{v}^{1/2}|\xi| + \sum_{e,j} \sup_{0 \leq \tau \leq \bar{T}} |B_{e,j}(\tau)|.$$

Step 4: convergence of the derivative measures. By (A4) and Step 3, we see that the measures S , M , P , and Q evaluated at $\Pi_m G_z$ converge in total variation to the same measures evaluated at G_z , almost surely, for every $z \in I$. Combining this with (4.8) and the uniform bounds on the total variations, we find that every contraction in the identity of Step 2 converges. For example, writing $M^{(m)}$ for M evaluated at $\Pi_m G_z$, we have

$$|\langle \dot{K}_z^{(m)}, M^{(m)} \rangle - \langle \dot{K}_z, M \rangle| \leq \|\dot{K}_z^{(m)} - \dot{K}_z\|_{\infty} \|M^{(m)}\|_{\text{TV}} + \|\dot{K}_z\|_{\infty} \|M^{(m)} - M\|_{\text{TV}},$$

and for product measures we use $\|S' \otimes P' - S \otimes P\|_{\text{TV}} \leq \|S' - S\|_{\text{TV}} \|P'\|_{\text{TV}} + \|S\|_{\text{TV}} \|P' - P\|_{\text{TV}}$. This includes the diagonal contributions of atoms, as in Remark 4.4, since only the field values are approximated and $\dot{K}_z^{(m)}(v, v) \rightarrow \dot{K}_z(v, v)$ uniformly.

Step 5: domination and the normalizers. By (A3) and Step 3, we have $e^{\ell_z^{(m)}} \leq \mathcal{W}$ for all m and z , where

$$\mathcal{W} = \exp \left\{ C \left(1 + |\omega| + \bar{v}^{1/2}|\xi| + \sum_{e,j} \sup_{0 \leq \tau \leq \bar{T}} |B_{e,j}(\tau)| \right) \right\}$$

and C is the constant of (A3). All differentiated terms in the identity of Step 2 are products of a weight with quantities bounded by (A3), (A4), and (4.8). Hence a constant multiple of $\mathcal{W}$ is a common envelope. It is integrable to every fixed positive power, because the suprema of Brownian motion over a bounded interval have Gaussian tails, and the sum over σ is finite. We now show that the approximate normalizers have a common positive lower bound. We retain one value σ_0 of the counting variable. Using Jensen's inequality and the bound on $|\ell_z|$ in (A3), we obtain

$$Z_z^{(m)} \geq \mathbb{E}_{\text{ref}} e^{\ell_z(\Pi_m G_z, \omega, \sigma_0)} \geq \exp(\mathbb{E}_{\text{ref}} \ell_z(\Pi_m G_z, \omega, \sigma_0)) \geq e^{-C'},$$

where the expectations are over the Gaussian variables only, and C' does not depend on m or z .

Step 6: passage to the limit in the interior. Let $\mathcal{E}_m(z) = \mathbb{E}_z^{(m)} \Phi_z^{(m)}$, and let $\Psi_m(z)$ be the right side of the identity of Step 2. Let $[z_1, z_2] \subset I^\circ$. By Step 2 and the fundamental theorem of calculus, we have

$$\mathcal{E}_m(z_2) - \mathcal{E}_m(z_1) = \int_{z_1}^{z_2} \Psi_m(z) \, dz.$$

By Steps 3–5 and dominated convergence, we see that $\mathcal{E}_m(z) \rightarrow \mathbb{E}_z \Phi_z$ and $\Psi_m(z) \rightarrow \Psi(z)$ for all $z \in I$, where $\Psi(z)$ is the right side of (4.5). Here we used the lower bound on the normalizers and the continuity of ℓ_z , Φ_z , $\dot{\ell}_z$, and $\dot{\Phi}_z$ in G . The functions Ψ_m are bounded uniformly in m and z . Using dominated convergence in z , we obtain

$$\mathbb{E}_{z_2} \Phi_{z_2} - \mathbb{E}_{z_1} \Phi_{z_1} = \int_{z_1}^{z_2} \Psi(z) \, dz.$$

The same argument with $z_k \rightarrow z$ in place of $m \rightarrow \infty$ shows that Ψ and $z \mapsto \mathbb{E}_z \Phi_z$ are continuous on I . Indeed, by the coupling, $G_{z_k} \rightarrow G_z$ uniformly, the measures converge in total variation by (A4), $\dot{K}_{z_k} \rightarrow \dot{K}_z$ uniformly by (A2), and the envelope of Step 5 dominates. By the fundamental theorem of calculus, we conclude that (4.5) holds on I° . The same argument with $\Phi \equiv 1$, applied to $\log Z_z^{(m)}$ instead of a ratio, gives (4.6) on I° .

Step 7: endpoints. Let z_- be an endpoint of I , say the left one. The integrated identity of Step 6 holds for $z_- < z_1 < z_2$ in I° . As $z_1 \downarrow z_-$, the left side converges because $z \mapsto \mathbb{E}_z \Phi_z$ is continuous on I , and the integral converges because Ψ is continuous on I . This also holds if a finite variance vanishes at z_- . In that case, the corresponding coordinate converges to zero, while the derivative measures and $\dot{K}_z$ have continuous limits, and we never invert the covariance matrix. We obtain $\mathbb{E}_{z_2} \Phi_{z_2} - \mathbb{E}_{z_-} \Phi_{z_-} = \int_{z_-}^{z_2} \Psi(z) \, dz$, and the continuity of Ψ gives the one-sided derivative at z_- . We treat the right endpoint and (4.6) in the same way. $\square$

PART II

Overlap estimates at a minimizer

Self-overlap moments on quantile intervals

In this section, we bound two moments of the self-overlap at a minimizing profile. The first is the integrated variance $V = \int \text{Var}(S_q) \nu(\mathrm{d}q)$, one of the three terms of the defect in (3.15). The second is the integrated cubic moment $H_3 = \int \mathbb{E}|S_q - q|^3 \nu(\mathrm{d}q)$, which appears in the cavity argument of Section 10. We control both by comparing S_q with averages of L_p over windows of quantile levels on either side of q ; the variances of these averages are bounded by (3.10).

The section contains three results. In Lemma 5.1, we prove two comparisons: conditionally on $\mathcal{G}_q$, the self-overlap S_q is at most twice an average of L_p over later levels $p \geq q$, and it equals twice an average of L_p over earlier levels $p \leq q$ plus a nonnegative term. Proposition 5.2 bounds the integral of $\text{Var}(S_{Q(v)})$ over a quantile interval in terms of its length, a lower bound for the cumulative mass on it, and the diameter of its image under Q . Corollary 5.3 shows that H_3 is at most an arbitrarily small multiple zV of V plus an error of order $(Ms)^{-1} \log(2M) \log(2/z)$.

We use the martingale and contact identities (3.7) and (3.9) of Proposition 3.1, the integrated variance bound (3.10) of Proposition 3.3, and the quantile function (3.6). We use Proposition 5.2 in the proof of Proposition 13.1, which bounds the defect for small ρ , and Corollary 5.3 in the proof of Corollary 12.5, where the cavity inequality is closed. Lemma 5.1 is used only in this section.

We refine the quantile-window argument of [4, Lemma 3.3], which estimates integrals of $\text{Var}(L_{Q(v)})$ over quantile intervals. By Lemma 3.5, we have $\text{Var}(L_q) = j_q + \text{Var}(S_q)/4$, where j_q is the variance of the conditionally centered part $N^{-1}(\sigma - m_q) \cdot m_q$ of L_q . An estimate for $\text{Var}(L_q)$ must therefore also control this centered part, and in [4, Lemma 3.3] it produces additional residual terms. Here we compare conditional expectations given $\mathcal{G}_q$, in which the centered part vanishes. We estimate j_q separately in Section 6.

5.1 Quantile windows and kernels

Standing assumptions. Throughout this section, we fix $N \geq 1$, $0 < \rho \leq 1$, and $0 < s \leq 1/2$, so that $t = 1 - \rho \in [0, 1)$ and $M = Nc > 0$. We let $\mu = s\delta_0 + \nu$ be a deterministic admissible profile, and expectations refer to its joint law from Section 3.3. Each statement says whether μ must minimize $\mathcal{F}_{t,s}$; a statement about a minimizer holds for every minimizer, with the same constants. All constants in this section are absolute: they do

not depend on N, β, ρ, s , or the profile.

Recall that $\sigma \in \{-1, 1\}^N$ and $m_q = \mathbb{E}[\sigma \mid \mathcal{G}_q]$. Then $|m_q|^2 \leq N$ and $|\sigma \cdot m_q| \leq N$, and it follows that

$$0 \leq S_q \leq 1, \quad |L_q| \leq \frac{3}{2} \quad (0 \leq q \leq 1).$$

The observables S_q and L_q are jointly measurable in the level and the outcome, because the process $q \mapsto m_q$ is continuous. They are also bounded, and Fubini's theorem applies to their integrals against deterministic functions in $L^1(\nu)$.

The variance bound. The estimates of this section rest on the first bound in (3.10), and its factor $a(q)^{-1}$ is the reason for working with quantiles. In the proof of Proposition 3.3, this factor comes from two facts. The first is the bound $N^{-2} \int k(q)^2 \mathbb{E} \operatorname{Tr}(H_q C_q) \nu(dq)$ for the variance, which holds for every admissible profile [4, Proposition 2.2, (2.15)] and rests on the Brascamp–Lieb inequality for the posterior law of the Gaussian coordinates. The second is the inequality $\operatorname{Tr}(H_q C_q) \leq a(q)^{-1} \operatorname{Tr} H_q^2$ at contact levels, which follows from $H_q - a(q)C_q \succeq 0$, together with $\mathbb{E} \operatorname{Tr} H_q^2 \leq N/c$.

Quantiles. Recall that the quantile function Q of (3.6) pushes Lebesgue measure on $(0, 1-s)$ forward to ν . We write $F(p) = \nu([0, p])$ and $F_-(p) = \nu([0, p))$, so that $a(p) = s + F(p)$. The infimum in (3.6) is attained because F is right-continuous, that is, $F(Q(v)) \geq v$; in particular, $a(Q(v)) \geq s + v$. We also use

$$a(q) \geq s \quad (0 \leq q \leq 1), \quad Q(v) \in \operatorname{supp} \nu \quad (0 < v < 1-s). \quad (5.1)$$

The first bound holds because $\mu \geq s\delta_0$. For the second, let $q_* = Q(v)$, so that $\nu([0, q_*]) \geq v > 0$. If $q_* = 0$, then 0 is an atom of ν and lies in $\operatorname{supp} \nu$. If $q_* > 0$ and the interval $(q_* - r, q_* + r)$ were ν -null for some $0 < r < q_*$, then $\nu([0, q_* - r/2]) = \nu([0, q_*]) \geq v$, which contradicts the definition of $Q(v)$ as an infimum.

If μ is a minimizer, then every point of $\operatorname{supp} \nu$ is a contact level. By (5.1) and (3.9), we then have $\mathbb{E} S_{Q(v)} = Q(v)$ for every $v \in (0, 1-s)$. In that case, we define

$$\xi_v = S_{Q(v)} - Q(v) \quad (0 < v < 1-s).$$

Note that $|\xi_v| \leq 1$, $\mathbb{E} \xi_v = 0$, and $\mathbb{E} \xi_v^2 = \operatorname{Var}(S_{Q(v)})$. Using that Q pushes Lebesgue measure forward to ν , we can write

$$V = \int_0^{1-s} \mathbb{E} \xi_v^2 dv, \quad H_3 = \int_0^{1-s} \mathbb{E} |\xi_v|^3 dv.$$

The integrands are Borel functions of v , by the joint measurability noted above, Fubini's theorem, and the monotonicity of Q .

Windows. Let $I \subset (0, 1 - s)$ be an interval of length $|I| > 0$. The image of Lebesgue measure on I under Q is dominated by the image of Lebesgue measure on $(0, 1 - s)$, which is ν . As in [4, Section 3], it has a density f_I with respect to ν , with

$$0 \leq f_I \leq 1, \quad \int f_I d\nu = |I|.$$

The density is explicit at an atom. If p is an atom of ν , with mass $r_p = \nu(\{p\})$, then $\{v : Q(v) = p\}$ is an interval with endpoints $F_-(p)$ and $F(p)$, and

$$f_I(p) = \frac{|I \cap (F_-(p), F(p))|}{r_p}.$$

In particular, a window may contain only part of the quantile interval of an atom; the atom then receives the corresponding fraction of its mass. We define the normalized density $\kappa_I = f_I/|I|$. It satisfies

$$0 \leq \kappa_I \leq |I|^{-1}, \quad \int \kappa_I d\nu = 1, \quad \int \varphi \kappa_I d\nu = \frac{1}{|I|} \int_I \varphi(Q(v)) d\nu$$

for every bounded Borel function φ . We say that I has *threshold* $u > 0$ if $a(Q(v)) \geq u$ for almost every $v \in I$. Every such interval has threshold s , by (5.1).

Kernels. More generally, we call a deterministic Borel function $\kappa \geq 0$ on $[0, 1]$ with $\int \kappa d\nu = 1$ and $\int \kappa^2 d\nu < \infty$ a *kernel*. For a kernel κ , we define its mean level and the average of $L_p - p/2$, which is centered when μ is a minimizer,

$$\bar{p}_\kappa = \int p \kappa(p) \nu(dp), \quad \Lambda_\kappa = \int \kappa(p) \left(L_p - \frac{p}{2} \right) \nu(dp).$$

We write $\bar{p}_I$ and Λ_I when $\kappa = \kappa_I$. Then

$$\bar{p}_I = \frac{1}{|I|} \int_I Q(v) d\nu, \quad \Lambda_I = \frac{1}{|I|} \int_I \left(L_{Q(v)} - \frac{Q(v)}{2} \right) d\nu.$$

Fixed kernels. Each assertion of Lemma 5.1 below concerns one level q and one kernel κ , both deterministic. In Proposition 5.2 and Corollary 5.3, the level and the kernels depend on a quantile parameter v . We then apply Lemma 5.1 separately for each value of v , and we integrate over v only the resulting deterministic expectations. In this way, we never need versions of the conditional expectations that are chosen simultaneously for all v .

5.2 Two-sided conditional comparison

For a later level $p \geq q$, we show by the conditional Jensen inequality that S_q is at most $2\mathbb{E}[L_p \mid \mathcal{G}_q]$. For an earlier level $p \leq q$, the conditional expectation of L_p given $\mathcal{G}_q$

is $S_q/2$ minus the nonnegative quantity $|m_q - m_p|^2/(2N)$. We average these relations over p against a kernel and obtain two one-sided comparisons between $S_q - q$ and the conditional expectation of Λ_κ given $\mathcal{G}_q$. At a minimizer, Λ_κ is centered, and its variance is controlled by (3.10).

Lemma 5.1 (Two-sided conditional comparison). *Let μ be an admissible profile, let $q \in [0, 1]$, and let κ be a kernel.*

(a) *If $\kappa = 0$ ν -almost everywhere on $[0, q)$, then almost surely*

$$S_q - q \leq (\bar{p}_\kappa - q) + 2\mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q].$$

(b) *If $\kappa = 0$ ν -almost everywhere on $(q, 1]$, then almost surely*

$$\begin{aligned} S_q - q &= (\bar{p}_\kappa - q) + 2\mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q] + \frac{1}{N} \int \kappa(p) |m_q - m_p|^2 \nu(\mathrm{d}p) \\ &\geq (\bar{p}_\kappa - q) + 2\mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q]. \end{aligned}$$

If μ minimizes $\mathcal{F}_{t,s}$, then in addition the following hold.

(c) *We have $\mathbb{E}\Lambda_\kappa = 0$ and*

$$\mathbb{E}[(\mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q])^2] \leq \mathrm{Var}(\Lambda_\kappa) \leq \frac{1}{M} \int \frac{\kappa(p)^2}{a(p)} \nu(\mathrm{d}p).$$

(d) *Let $\vartheta > 0$. If κ satisfies the hypothesis of (a) and $\bar{p}_\kappa - q \leq \vartheta/2$, then*

$$\mathbb{E}[(S_q - q)^2 \mathbf{1}_{\{S_q - q > \vartheta\}}] \leq 16 \mathrm{Var}(\Lambda_\kappa).$$

If κ satisfies the hypothesis of (b) and $q - \bar{p}_\kappa \leq \vartheta/2$, then

$$\mathbb{E}[(S_q - q)^2 \mathbf{1}_{\{S_q - q < -\vartheta\}}] \leq 16 \mathrm{Var}(\Lambda_\kappa).$$

(e) *If $I \subset (0, 1 - s)$ is an interval of length $x > 0$ with threshold u , then $\mathrm{Var}(\Lambda_I) \leq 1/(Mux)$.*

For every admissible profile and every interval $I \subset (0, 1 - s)$ of positive length, κ_I satisfies the hypothesis of (a) if $Q \geq q$ on I , and that of (b) if $Q \leq q$ on I .

Proof. Parts (a) and (b) are statements about conditional expectations. We prove each first for a single level p and then integrate against $\kappa(p) \nu(\mathrm{d}p)$. Parts (c)–(e) add the contact identity and (3.10). Let

$$\Lambda' = \int \kappa(p) L_p \nu(\mathrm{d}p) = \Lambda_\kappa + \frac{\bar{p}_\kappa}{2}.$$

Taking the conditional expectation given $\mathcal{G}_p$ in the definition of L_p , and recalling that $m_p = \mathbb{E}[\sigma \mid \mathcal{G}_p]$ is $\mathcal{G}_p$ -measurable, we find that

$$\mathbb{E}[L_p \mid \mathcal{G}_p] = \frac{1}{2} S_p \quad (0 \leq p \leq 1). \quad (5.2)$$

Part (a). Let $p \geq q$. The process m is a martingale by (3.7), and $m_q = \mathbb{E}[m_p \mid \mathcal{G}_q]$. Using the conditional Jensen inequality for $x \mapsto |x|^2$, the tower property, and (5.2), we have

$$S_q \leq \mathbb{E}[S_p \mid \mathcal{G}_q] = 2\mathbb{E}[L_p \mid \mathcal{G}_q].$$

Let $\zeta \geq 0$ be a bounded $\mathcal{G}_q$ -measurable random variable. Then $\mathbb{E}[\zeta S_q] \leq 2\mathbb{E}[\zeta L_p]$ for every $p \geq q$. We integrate this inequality against the probability measure $\kappa(p) \nu(dp)$, which does not charge $[0, q]$. By Fubini's theorem, which applies because $|\zeta L_p|$ is bounded, this yields $\mathbb{E}[\zeta S_q] \leq 2\mathbb{E}[\zeta \Lambda']$. As ζ is arbitrary and S_q is $\mathcal{G}_q$ -measurable, we conclude that $S_q \leq 2\mathbb{E}[\Lambda' \mid \mathcal{G}_q]$ almost surely. Subtracting q from both sides and recalling that $\Lambda' = \Lambda_\kappa + \bar{p}_\kappa/2$, we obtain (a).

Part (b). Let $p \leq q$. Both m_p and m_q are $\mathcal{G}_q$ -measurable, and $\mathbb{E}[\sigma \mid \mathcal{G}_q] = m_q$. Expanding $|m_q - m_p|^2$, we find

$$\mathbb{E}[L_p \mid \mathcal{G}_q] = \frac{m_q \cdot m_p - |m_p|^2/2}{N} = \frac{S_q}{2} - \frac{|m_q - m_p|^2}{2N}.$$

The right side is bounded and jointly measurable in $p \in [0, q]$ and the outcome. Its integral against $\kappa(p) \nu(dp)$, which is concentrated on $[0, q]$, is $\mathcal{G}_q$ -measurable, because m is continuous and adapted. Testing against bounded $\mathcal{G}_q$ -measurable variables as in (a), we find that

$$\mathbb{E}[\Lambda' \mid \mathcal{G}_q] = \frac{S_q}{2} - \frac{1}{2N} \int \kappa(p) |m_q - m_p|^2 \nu(dp).$$

Multiplying by two, subtracting q , and using $\Lambda' = \Lambda_\kappa + \bar{p}_\kappa/2$, we obtain the identity in (b). The inequality holds because the last term is nonnegative.

Part (c). Let μ be a minimizer. By (5.2) and (3.9), we have $\mathbb{E}L_p = \mathbb{E}S_p/2 = p/2$ for every $p \in \text{supp } \nu$, and in particular for ν -almost every p . Then $\mathbb{E}\Lambda_\kappa = 0$ by Fubini's theorem. By the conditional Jensen inequality, we have

$$\mathbb{E}[(\mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q])^2] \leq \mathbb{E}\Lambda_\kappa^2 = \text{Var}(\Lambda_\kappa).$$

As Λ_κ and Λ' differ by a constant, the last bound in (c) is the first bound in (3.10) with $k = \kappa$.

Part (d). Let $\xi = S_q - q$ and $Y = \mathbb{E}[\Lambda_\kappa \mid \mathcal{G}_q]$. Suppose first that κ satisfies the hypothesis of (a) and $\bar{p}_\kappa - q \leq \vartheta/2$. On the event $\{\xi > \vartheta\}$, part (a) implies that

$$2Y \geq \xi - (\bar{p}_\kappa - q) \geq \xi - \frac{\vartheta}{2} > \frac{\xi}{2} > 0,$$

so that $\xi^2 < 16Y^2$. Next, suppose that κ satisfies the hypothesis of (b) and $q - \bar{p}_\kappa \leq \vartheta/2$. On the event $\{\xi < -\vartheta\}$, part (b) implies that

$$2Y \leq \xi + (q - \bar{p}_\kappa) \leq \xi + \frac{\vartheta}{2} < \frac{\xi}{2} < 0,$$

and again $\zeta^2 < 16Y^2$. In both cases, the expectation in (d) is at most $16\mathbb{E}Y^2$, which is at most $16\text{Var}(\Lambda_\kappa)$ by (c).

Part (e). The bound $\kappa_I \leq 1/x$ implies that $\kappa_I^2 \leq \kappa_I/x$. The function $1/a$ is Borel and bounded by $1/s$, by (5.1). Using the identity for $\int \varphi \kappa_I d\nu$ with $\varphi = 1/a$, we have

$$\int \frac{\kappa_I(p)^2}{a(p)} \nu(dp) \leq \frac{1}{x} \int \frac{\kappa_I(p)}{a(p)} \nu(dp) = \frac{1}{x^2} \int_I \frac{dv}{a(Q(v))} \leq \frac{1}{ux},$$

where we used the threshold in the last step. Inserting this bound into (c), we find that $\text{Var}(\Lambda_I) \leq 1/(Mux)$. This is the bound [4, Section 3, (3.6)], because Λ_I and $x^{-1} \int_I L_{Q(v)} dv$ differ by the constant $\bar{p}_I/2$.

The hypotheses for κ_I . Suppose that $Q \geq q$ on I . Then

$$\int \mathbf{1}_{[0,q)} \kappa_I d\nu = \frac{1}{|I|} |\{v \in I : Q(v) < q\}| = 0,$$

so κ_I satisfies the hypothesis of (a). The case $Q \leq q$ is the same, with $(q, 1]$ in place of $[0, q)$. $\square$

5.3 Variance on a quantile interval

We now use the comparisons of Lemma 5.1 to bound the integral of $\text{Var}(S_{Q(v)})$ over a quantile interval. For each variance threshold, we compare every interior point v of the interval with the two windows of one common radius r on either side of v . The random part of each comparison is twice a conditional expectation whose second moment is at most $(Mur)^{-1}$, which decreases as r grows. The deterministic part is at most the increment of Q across the window $(v-r, v+r)$. After we extend Q by constants beyond the ends of the interval, the integral of this increment over $\mathbb{R}$ equals $2r$ times the diameter of the image of the interval, because Q is nondecreasing; its integral over the interior points is at most this quantity. This bounds the measure of the set of points where the variance exceeds the threshold, and integrating over the threshold gives the result. For an interval $I \subset (0, 1-s)$, we let

$$\ell(I) = \text{ess sup}_{v,w \in I} |Q(v) - Q(w)|$$

denote the diameter of its image.

Proposition 5.2 (Variance on a quantile interval). *Let μ be a minimizer of $\mathcal{F}_{t,s}$, and let $I \subset (0, 1-s)$ be an interval of length $h > 0$ with threshold $u > 0$ and image diameter $\ell = \ell(I)$. Let $\epsilon = (Mu)^{-1}$. Then*

$$\int_I \text{Var}(S_{Q(v)}) dv \leq 128 \left[\min\{h, \epsilon(1 + \log_+(h/\epsilon))\} + (\epsilon\ell)^{2/3} h^{1/3} \right]. \quad (5.3)$$

The constant is absolute, and I may contain part of the quantile interval of an atom of Q .

Proof. The proof has four steps. First, we record an identity for the increments of Q . Second, we bound $\text{Var}(S_{Q(v)})$ at an interior point v by comparing it with the two windows of radius r next to v . Third, for each threshold $\tau \in (0, 1)$, we choose r depending on τ and bound the measure of the set where the variance exceeds τ . Fourth, we integrate over τ .

Step 1: increments of the quantile function. Discarding endpoints, which changes no integral below, we write $I = (\underline{v}, \bar{v})$, with $h = \bar{v} - \underline{v}$. We set

$$\chi(v) = \mathbb{E}\zeta_v^2 = \text{Var}(S_{Q(v)}) \in [0, 1] \quad (v \in I).$$

As Q is nondecreasing, its essential infimum and supremum on I are $q_- = \lim_{v \downarrow \underline{v}} Q(v)$ and $q_+ = \lim_{v \uparrow \bar{v}} Q(v)$, and $q_+ - q_- = \ell$. We extend the restriction of Q to I to a nondecreasing function $\tilde{Q}$ on $\mathbb{R}$ by setting $\tilde{Q} = q_-$ on $(-\infty, \underline{v}]$ and $\tilde{Q} = q_+$ on $[\bar{v}, \infty)$. Let $d\tilde{Q}$ be the distributional derivative of $\tilde{Q}$, a positive Lebesgue–Stieltjes measure of total mass $q_+ - q_- = \ell$. We claim that

$$\int_{\mathbb{R}} (\tilde{Q}(v+r) - \tilde{Q}(v-r)) dv = 2r\ell \quad (r > 0). \quad (5.4)$$

Indeed, $\tilde{Q}$ has at most countably many discontinuities. For fixed $r > 0$ and v outside the two translates of this countable set by r and by $-r$, both $v-r$ and $v+r$ are continuity points, and $\tilde{Q}(v+r) - \tilde{Q}(v-r)$ is the $d\tilde{Q}$ -measure of $[v-r, v+r]$. By Tonelli's theorem,

$$\int_{\mathbb{R}} d\tilde{Q}([v-r, v+r]) dv = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{|x-v| \leq r\}} dv d\tilde{Q}(x) = 2r\ell.$$

In words, each unit of mass of $d\tilde{Q}$ is counted for the centers v within distance r of it, and these form a set of length $2r$. The identity uses only that $\tilde{Q}$ is nondecreasing; in particular, it holds when Q has jumps or intervals of constancy.

Step 2: the two windows at a fixed radius. Let $0 < r < h/2$ and $v \in (\underline{v} + r, \bar{v} - r)$. The windows $I_- = (v-r, v)$ and $I_+ = (v, v+r)$ lie in I , have length r , and have threshold u . We set

$$\Delta_r(v) = Q(v+r) - Q(v-r) \geq 0, \quad Y_{\pm} = \mathbb{E}[\Lambda_{I_{\pm}} \mid \mathcal{G}_{Q(v)}].$$

Note that $Q \geq Q(v)$ on I_+ and $Q \leq Q(v)$ on I_- , because Q is nondecreasing. It follows from the last assertion of Lemma 5.1 that κ_{I_+} and κ_{I_-} satisfy the hypotheses of Lemma 5.1(a) and (b), respectively, with $q = Q(v)$. Further, by monotonicity,

$$\begin{aligned} 0 &\leq \bar{p}_{I_+} - Q(v) \leq Q(v+r) - Q(v) \leq \Delta_r(v), \\ 0 &\leq Q(v) - \bar{p}_{I_-} \leq Q(v) - Q(v-r) \leq \Delta_r(v). \end{aligned}$$

Then Lemma 5.1(a) and (b) imply that

$$\xi_v \leq \Delta_r(v) + 2Y_+, \quad -\xi_v \leq \Delta_r(v) - 2Y_-.$$

On the event $\{\xi_v > 0\}$, the first inequality implies that $\xi_v^2 \leq (\Delta_r(v) + 2|Y_+|)^2$, and this is at most $2\Delta_r(v)^2 + 8Y_+^2$ because $(x + y)^2 \leq 2x^2 + 2y^2$. On the event $\{\xi_v < 0\}$, we have $\xi_v^2 \leq 2\Delta_r(v)^2 + 8Y_-^2$ in the same way, by the second inequality. Since these two events are disjoint, we have

$$\xi_v^2 \leq 2\Delta_r(v)^2 + 8Y_+^2 + 8Y_-^2.$$

Lemma 5.1(c) and (e) imply that $\mathbb{E}Y_\pm^2 \leq \text{Var}(\Lambda_{I_\pm}) \leq (Mur)^{-1} = \epsilon/r$. Taking expectations, we conclude that

$$\chi(v) \leq 2\Delta_r(v)^2 + \frac{16\epsilon}{r} \quad \text{for every } v \in (\underline{v} + r, \bar{v} - r). \quad (5.5)$$

Note that v and r are fixed here. The windows and the level $Q(v)$ are then deterministic, and Lemma 5.1 applies as stated.

Step 3: the measure of the set where the variance is large. Fix $\tau \in (0, 1)$ and let $r = 32\epsilon/\tau$, so that $16\epsilon/r = \tau/2$. Suppose first that $r < h/2$. If $v \in (\underline{v} + r, \bar{v} - r)$ and $\chi(v) > \tau$, then (5.5) implies that $2\Delta_r(v)^2 > \tau/2$, that is, $\Delta_r(v) > \sqrt{\tau}/2$. On this interior interval, we have $\Delta_r(v) = \tilde{Q}(v + r) - \tilde{Q}(v - r)$. By Markov's inequality and (5.4), the measure of the interior points with $\chi(v) > \tau$ is at most

$$\frac{2}{\sqrt{\tau}} \int_{\mathbb{R}} (\tilde{Q}(v + r) - \tilde{Q}(v - r)) dv = \frac{4r\ell}{\sqrt{\tau}} = \frac{128\epsilon\ell}{\tau^{3/2}}.$$

The two boundary strips $(\underline{v}, \underline{v} + r]$ and $[\bar{v} - r, \bar{v})$ have total length $2r = 64\epsilon/\tau$. If instead $r \geq h/2$, then $h \leq 2r = 64\epsilon/\tau$, and the same bound holds without any interior windows, because the whole interval has length at most $2r$. In both cases, using also the trivial bound h , we have

$$|\{v \in I : \chi(v) > \tau\}| \leq \min\left\{h, \frac{64\epsilon}{\tau} + \frac{128\epsilon\ell}{\tau^{3/2}}\right\} \leq \min\left\{h, \frac{64\epsilon}{\tau}\right\} + \min\left\{h, \frac{128\epsilon\ell}{\tau^{3/2}}\right\}.$$

Step 4: integration over the threshold. By the layer-cake formula, and since $0 \leq \chi \leq 1$, we have

$$\int_I \chi(v) dv = \int_0^1 |\{v \in I : \chi(v) > \tau\}| d\tau.$$

We integrate the two terms of the last bound of Step 3 separately. For the first, we use $\min\{h, 64y\} \leq 64 \min\{h, y\}$ for $y \geq 0$. If $\epsilon \geq h$, then $\epsilon/\tau \geq h$ for $\tau \leq 1$, and the integral of $\min\{h, \epsilon/\tau\}$ over $(0, 1)$ equals h . If $\epsilon < h$, we split the integral at $\tau = \epsilon/h$ and obtain $\epsilon + \epsilon \log(h/\epsilon)$, which is at most h because $1 + \log y \leq y$ for $y \geq 1$. In both cases, we have

$$\int_0^1 \min\left\{h, \frac{64\epsilon}{\tau}\right\} d\tau \leq 64 \min\{h, \epsilon(1 + \log_+(h/\epsilon))\}.$$

For the second term, let $B = 128\epsilon\ell$. If $\ell = 0$, the term vanishes. Otherwise, we extend the integral to $(0, \infty)$ and split it at $\tau_0 = (B/h)^{2/3}$, where $B\tau_0^{-3/2} = h$. This gives

$$\begin{aligned} \int_0^\infty \min\{h, B\tau^{-3/2}\} d\tau &= h\tau_0 + 2B\tau_0^{-1/2} = 3B^{2/3}h^{1/3} \\ &= 3 \cdot 128^{2/3}(\epsilon\ell)^{2/3}h^{1/3} \leq 128(\epsilon\ell)^{2/3}h^{1/3}. \end{aligned}$$

In the last inequality, we used that $3 \cdot 128^{2/3} \leq 128$, which is equivalent to $27 \leq 128$. Adding the two estimates, we obtain (5.3). $\square$

5.4 A cubic-moment bound

The cavity argument produces third moments of $S_q - q$. Fluctuations of ξ_v of size at most z contribute at most zV to H_3 . For fluctuations above a level ϑ , we divide the range of levels into bins of width $\vartheta/4$. Within a bin, the mean level of the window after a point v is at least $Q(v)$ and less than $Q(v) + \vartheta/4$, and the mean level of the window before v is at most $Q(v)$ and greater than $Q(v) - \vartheta/4$. We use Lemma 5.1(d) to control both tails. Integrating the reciprocal window lengths over v produces one logarithm, and integrating over ϑ produces the second.

Corollary 5.3 (Cubic self-overlap moment). *Let μ be a minimizer of $\mathcal{F}_{t,s}$, and assume that $M \geq 1$. Then for every $0 < z \leq 1/4$,*

$$H_3 \leq zV + \frac{425}{Ms} \log(2M) \log \frac{2}{z}. \quad (5.6)$$

The constant is absolute; atoms of v , including an atom at zero, are allowed.

Proof. The proof has three steps. We first reduce H_3 to integrals of the tails of ξ_v , then bound these tail integrals using level bins, and finally integrate the tail bound.

Step 1: reduction to tails. For $|y| \leq 1$ and $0 < z \leq 1/4$, we have

$$|y| \leq z + \int_z^1 \mathbf{1}_{\{|y| > \vartheta\}} d\vartheta.$$

Indeed, the right side is at least z , and if $|y| > z$, the integral equals $|y| - z$. Multiplying this inequality by y^2 gives

$$|y|^3 \leq zy^2 + \int_z^1 y^2 \mathbf{1}_{\{|y| > \vartheta\}} d\vartheta.$$

We use this inequality with $y = \xi_v$, take expectations, and integrate over $v \in (0, 1-s)$. We define

$$T(\vartheta) = \int_0^{1-s} \mathbb{E}[\xi_v^2 \mathbf{1}_{\{|\xi_v| > \vartheta\}}] dv \quad (0 < \vartheta \leq 1).$$

Using Tonelli's theorem and the identities $V = \int_0^{1-s} \mathbb{E} \zeta_v^2 dv$ and $H_3 = \int_0^{1-s} \mathbb{E} |\zeta_v|^3 dv$, we obtain

$$H_3 \leq zV + \int_z^1 T(\vartheta) d\vartheta. \quad (5.7)$$

Step 2: the tail bound. Fix $\vartheta \in (0, 1]$. For $0 \leq k \leq \lfloor 4/\vartheta \rfloor$, let $\mathcal{B}_k = [k\vartheta/4, (k+1)\vartheta/4]$. These level bins cover $[0, 1]$, and there are at most $4/\vartheta + 1 \leq 5/\vartheta$ of them. The preimages

$$I_k = \{v \in (0, 1-s) : Q(v) \in \mathcal{B}_k\}$$

are intervals, because Q is nondecreasing. Some of them may be empty, and we discard those. For the others, we write $\underline{v}_k \leq \bar{v}_k$ for the endpoints of I_k and $h_k = \bar{v}_k - \underline{v}_k \leq 1$ for its length. The intervals I_k partition $(0, 1-s)$, and those with $h_k = 0$ do not contribute to $T(\vartheta)$.

Fix k with $h_k > 0$ and $v \in (\underline{v}_k, \bar{v}_k)$. The later window $(v, \bar{v}_k)$ lies in I_k . For w in this window, we have $Q(v) \leq Q(w) < Q(v) + \vartheta/4$, since $Q(w)$ and $Q(v)$ lie in the same bin of width $\vartheta/4$. In particular, the mean level of the window is at least $Q(v)$ and at most $Q(v) + \vartheta/4$. Likewise, the earlier window $(\underline{v}_k, v)$ satisfies $Q \leq Q(v)$, and its mean level is at least $Q(v) - \vartheta/4$. Both windows have threshold s by (5.1). By the last assertion of Lemma 5.1, their normalized densities satisfy the hypotheses of (a) and (b), respectively, with $q = Q(v)$. Since $\vartheta/4 \leq \vartheta/2$, Lemma 5.1(d) and (e) now imply that

$$\mathbb{E}[\zeta_v^2 \mathbf{1}_{\{\zeta_v > \vartheta\}}] \leq \frac{16}{Ms(\bar{v}_k - v)}, \quad \mathbb{E}[\zeta_v^2 \mathbf{1}_{\{\zeta_v < -\vartheta\}}] \leq \frac{16}{Ms(v - \underline{v}_k)}.$$

Suppose that $h_k > 1/M$. We integrate the first bound over $(\underline{v}_k, \bar{v}_k - 1/M)$ and use $|\zeta_v| \leq 1$ on the remaining strip of length $1/M$. This gives

$$\begin{aligned} \int_{I_k} \mathbb{E}[\zeta_v^2 \mathbf{1}_{\{\zeta_v > \vartheta\}}] dv &\leq \frac{16}{Ms} \int_{\underline{v}_k}^{\bar{v}_k - 1/M} \frac{dv}{\bar{v}_k - v} + \frac{1}{M} = \frac{16}{Ms} \log(Mh_k) + \frac{1}{M} \\ &\leq \frac{16 \log(2M)}{Ms} + \frac{1}{M}, \end{aligned}$$

where we used $h_k \leq 1$ in the last line. The lower tail satisfies the same bound, with the strip of length $1/M$ at the left end of I_k . If $h_k \leq 1/M$, then both tails together contribute at most $h_k \leq 1/M$, since the two events are disjoint and $|\zeta_v| \leq 1$. Since $M \geq 1$ and $s \leq 1/2$, we have

$$\frac{2}{M} \leq \frac{1}{Ms} \leq \frac{2 \log(2M)}{Ms},$$

where we used $\log 2 \geq 1/2$ in the second inequality. Each bin therefore contributes at most $32 \log(2M)/(Ms) + 2/M \leq 34 \log(2M)/(Ms)$ to $T(\vartheta)$. Summing over the at most $5/\vartheta$ bins, we conclude that

$$T(\vartheta) \leq \frac{5}{\vartheta} \cdot \frac{34 \log(2M)}{Ms} = \frac{170 \log(2M)}{Ms \vartheta} \quad (0 < \vartheta \leq 1).$$

Step 3: integration. The tail bound holds on the whole range $(z, 1]$ of the integral in (5.7). Inserting it, we obtain

$$H_3 \leq zV + \frac{170 \log(2M)}{Ms} \int_z^1 \frac{d\vartheta}{\vartheta} = zV + \frac{170 \log(2M)}{Ms} \log \frac{1}{z} \leq zV + \frac{425 \log(2M)}{Ms} \log \frac{2}{z}.$$

This proves (5.6). In fact, we have proved the stronger bound with $170 \log(1/z)$ in place of $425 \log(2/z)$. $\square$

The mean–covariance term

In this section, we bound the mean–covariance term $J = \int j_q \nu(dq)$ at a minimizing profile. Here $j_q = N^{-2} \mathbb{E}[m_q^T C_q m_q]$ is the expected conditional variance of the spin in the direction of its conditional mean. Our main result, Proposition 6.4, states that $J \leq 341(1 + \log_+ M)^2 / (Ms)$ at every minimizer. Together with the trace bound $A \leq 1 / (Ms)$ of (3.10), this controls the thermal part $A + 2J$ of the defect $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$. As $Ms = \beta^2 N \rho \cdot \lambda / \beta = \beta N \rho \lambda$, the two bounds imply that

$$\beta(A + 2J) \leq \frac{1 + 682(1 + \log_+ M)^2}{N \rho \lambda}.$$

We use Proposition 6.4 in this form in the proof of Corollary 12.5, for those ρ where the cavity inequality contracts, and in (13.1) in the proof of Proposition 13.1, for small ρ .

The proof has four steps.

1. In Lemma 6.1, we show by a scalar Riccati inequality that conditional covariances decay in every direction that is measurable at an earlier level: if b is $\mathcal{G}_{q'}$ -measurable, then $\mathbb{E}[b^T C_q b]$ is at most $\mathbb{E}|b|^2$ divided by the accumulated clock $c \int_{q'}^q a(u)^2 du$. This holds for every admissible profile.
2. In Lemma 6.2, we use first-order optimality to show that $S(r) = \mathbb{E}S_r$ is 2-Lipschitz toward every point q of $\text{supp } \nu$, also when r lies in a gap of the support. By (6.2) below, this bounds $\mathbb{E}|m_q - m_r|^2$ by $2N|q - r|$.
3. In Lemma 6.3, we combine these two estimates through a dyadic decomposition of $m_q - m_p$, for $q \in \text{supp } \nu$ and $p \leq q$. We bound $\mathcal{R}(p, q) = N^{-2} \mathbb{E}[(m_q - m_p)^T C_q (m_q - m_p)]$ by a multiple of $(1 + \log_+ M)^2 / (Ma((p + q)/2)^2)$.
4. In Proposition 6.4, we average earlier means $m_{Q(w)}$, $0 < w < v$, over the whole quantile prefix with a quadratic weight, together with a deterministic weight on $m_0 = 0$. We control the conditional covariance in the direction of this average by (3.10), and its distance from $m_{Q(v)}$ by Lemma 6.3. The weight produces a factor $(s + v)^{-2}$, whose integral over $v \in (0, 1 - s)$ is at most $1/s$.

We use Proposition 3.1, namely the martingale equation and covariance order (3.7), the first variation (3.8), and the contact identities (3.9); Proposition 3.3; and facts about the quantile function from Section 5.1, including (5.1). Lemmas 6.1 to 6.3 are used only in this section.

6.1 Standing assumptions

Throughout this section, we fix $N \geq 1$, $0 < \rho \leq 1$, and $0 < s \leq 1/2$, so that $t = 1 - \rho \in [0, 1)$, $c > 0$, and $M = Nc > 0$. Expectations refer to the joint law of the deterministic admissible profile $\mu = s\delta_0 + \nu$ under consideration. Lemma 6.1 holds for every admissible profile. The later results assume that μ minimizes $\mathcal{F}_{t,s}$; they hold for every minimizer, with absolute constants.

We write $S(q) = \mathbb{E}S_q$ for the expected self-overlap. Recall from (3.7) that $a(q)C_q \preceq H_q \preceq C_q$, and in particular $0 \preceq H_q \preceq C_q$. The diagonal entries of C_q are $1 - m_{q,i}^2 \leq 1$. These two facts give

$$\|H_q\|_{\text{op}} \leq \|C_q\|_{\text{op}} \leq \text{Tr } C_q \leq N, \quad 0 \leq S(q) \leq 1, \quad (6.1)$$

where $\|\cdot\|_{\text{op}}$ is the operator norm. The process m is a bounded martingale, and $\mathbb{E}[m_q \mid \mathcal{G}_p] = m_p$ for $p \leq q$. Then $\mathbb{E}|m_q|^2 = \mathbb{E}|m_p|^2 + \mathbb{E}|m_q - m_p|^2$. In particular, S is nondecreasing, and

$$\mathbb{E}|m_q - m_p|^2 = N(S(q) - S(p)) \quad (0 \leq p \leq q \leq 1). \quad (6.2)$$

For $0 \leq p \leq q \leq 1$, we define

$$\mathcal{R}(p, q) = N^{-2} \mathbb{E}[(m_q - m_p)^T C_q (m_q - m_p)],$$

which is the analog of j_q for the increment $m_q - m_p$ in place of m_q . Note that $\mathcal{R}(0, q) = j_q$, since $m_0 = 0$ by (3.7).

6.2 Covariance along an earlier measurable direction

Fix a direction that is measurable at some level. As more of the auxiliary field is revealed, the expected conditional variance of the spin in this direction can only decrease. The martingale equation determines the rate of decrease, and the covariance order bounds this rate from below by a multiple of the square of the variance itself. We integrate the resulting scalar Riccati inequality and obtain decay that is inversely proportional to the accumulated clock.

Lemma 6.1 (Directional covariance decay). *Let $\mu = s\delta_0 + \nu$ be any deterministic admissible profile, let $0 \leq q' < q \leq 1$, and let*

$$\mathcal{I} = c \int_{q'}^q a(u)^2 du.$$

Then $\mathcal{I} \geq cs^2(q - q') > 0$, and for every $\mathcal{G}_{q'}$ -measurable random vector $b \in \mathbb{R}^N$ with $\mathbb{E}|b|^2 < \infty$,

$$\mathbb{E}[b^T C_q b] \leq \frac{\mathbb{E}|b|^2}{\mathcal{I}}.$$

Proof. The lower bound for $\mathcal{I}$ holds because $a \geq s$. We first prove the main bound for bounded b , by deriving a differential inequality for the function $r \mapsto \mathbb{E}[b^T C_r b]$ on $[q', q]$, and we then remove the boundedness by monotone convergence.

Step 1: the rate of decrease. Suppose that b is bounded, and define $\varphi(r) = \mathbb{E}[b^T C_r b]$ for $r \in [q', q]$. For such r , the vector b is $\mathcal{G}_r$ -measurable, so that $b^T C_r b = \mathbb{E}[(b \cdot (\sigma - m_r))^2 \mid \mathcal{G}_r]$ and $\varphi(r) = \mathbb{E}(b \cdot (\sigma - m_r))^2$. We write

$$b \cdot (\sigma - m_{q'}) = b \cdot (\sigma - m_r) + b \cdot (m_r - m_{q'}).$$

The second term is $\mathcal{G}_r$ -measurable, and the first has conditional mean zero given $\mathcal{G}_r$. The cross term has expectation zero, and

$$\varphi(q') = \varphi(r) + \mathbb{E}(b \cdot (m_r - m_{q'}))^2.$$

By (3.7), and since b is $\mathcal{G}_{q'}$ -measurable and bounded, we have $b \cdot (m_r - m_{q'}) = \sqrt{c} \int_{q'}^r (H_u b) \cdot dB_u$. The integrand is progressive and bounded by $N|b|$, by (6.1). Itô's isometry gives

$$\varphi(r) = \varphi(q') - c \int_{q'}^r \mathbb{E}|H_u b|^2 du \quad (q' \leq r \leq q).$$

Then φ is absolutely continuous and nonincreasing, with $\varphi'(r) = -c\mathbb{E}|H_r b|^2$ for almost every $r \in [q', q]$.

Step 2: the Riccati inequality. On the event $\{b \neq 0\}$, the Cauchy–Schwarz inequality and the covariance order $0 \leq a(r)b^T C_r b \leq b^T H_r b$ imply that

$$|H_r b|^2 \geq \frac{(b^T H_r b)^2}{|b|^2} \geq \frac{a(r)^2 (b^T C_r b)^2}{|b|^2}.$$

Let $Z_r = |b|^{-1} b^T C_r b$ on $\{b \neq 0\}$ and $Z_r = 0$ otherwise. Then $\mathbb{E}|H_r b|^2 \geq a(r)^2 \mathbb{E}Z_r^2$. Further, $\varphi(r) = \mathbb{E}[Z_r |b|]$, and by the Cauchy–Schwarz inequality, $\varphi(r)^2 \leq \mathbb{E}Z_r^2 \mathbb{E}|b|^2$. If $\mathbb{E}|b|^2 > 0$, we conclude that

$$\varphi'(r) \leq -\frac{c a(r)^2 \varphi(r)^2}{\mathbb{E}|b|^2} \quad \text{for almost every } r \in [q', q].$$

Step 3: integration. If $\mathbb{E}|b|^2 = 0$ or $\varphi(q) = 0$, there is nothing to prove. Otherwise, $\varphi \geq \varphi(q) > 0$ on $[q', q]$, because φ is nonincreasing. The function $1/\varphi$ is then absolutely continuous, with $(1/\varphi)' = -\varphi'/\varphi^2 \geq c a(r)^2 / \mathbb{E}|b|^2$ almost everywhere. Integrating over $[q', q]$, we find

$$\frac{1}{\varphi(q)} \geq \frac{1}{\varphi(q')} + \frac{\mathcal{I}}{\mathbb{E}|b|^2} \geq \frac{\mathcal{I}}{\mathbb{E}|b|^2},$$

which is the claim for bounded b . The profile a may have jumps, but only Lebesgue-almost every level enters this integration.

Step 4: general directions. For general b with $\mathbb{E}|b|^2 < \infty$, let $b_n = b\mathbf{1}_{\{|b| \leq n\}}$, which is bounded and $\mathcal{G}_{q'}$ -measurable. The quantity $b_n^T C_q b_n = (b^T C_q b)\mathbf{1}_{\{|b| \leq n\}}$ increases to $b^T C_q b$ as $n \rightarrow \infty$, because $C_q \succeq 0$, and $|b_n|^2$ increases to $|b|^2$. The bound holds for each b_n , and monotone convergence as $n \rightarrow \infty$ yields the claim for b . $\square$

6.3 Mean increments toward a support point

In the dyadic argument below, we need $\mathbb{E}|m_q - m_r|^2$ to be at most of order $N|q - r|$ when q is a point of $\text{supp } \nu$, including when r lies in a gap of the support. In view of (6.2), this is a Lipschitz bound for S toward q , and it follows from the global first-order optimality of a minimizer. Let

$$G(y) = \int_y^1 (S(u) - u) \, du \quad (0 \leq y \leq 1)$$

be the function in the first variation (3.8). Proposition 3.1 states that if μ is a minimizer, then $\text{supp } \nu$ is contained in the set of minimum points of G on $[0, 1]$, that is,

$$G(y) \geq G(q) \quad (q \in \text{supp } \nu, 0 \leq y \leq 1). \quad (6.3)$$

We recall the argument from the proof of [4, Proposition 2.1], because we will use (6.3) at points y outside $\text{supp } \nu$. For $y \in [0, 1]$, the profiles $s\delta_0 + (1 - \theta)\nu + \theta(1 - s)\delta_y$, $0 \leq \theta \leq 1$, are admissible. At a minimizer, the right derivative of $\mathcal{F}_{t,s}$ along them at $\theta = 0$, which is (3.8) with $\nu' = (1 - s)\delta_y$, is nonnegative. Dividing by $c/2$, we find that

$$(1 - s)G(y) \geq \int G(p) \nu(\mathrm{d}p) \quad (0 \leq y \leq 1).$$

The function G is continuous. Taking y to be a minimum point of G and recalling that ν has mass $1 - s$, we find that $\int (G - \min G) \, \mathrm{d}\nu \leq 0$. Then the open set $\{G > \min G\}$ is ν -null and therefore does not meet $\text{supp } \nu$.

Lemma 6.2 (Lipschitz bound toward the support). *Let μ be a minimizer of $\mathcal{F}_{t,s}$. For every $q \in \text{supp } \nu$ and every $r \in [0, 1]$,*

$$|S(r) - S(q)| \leq 2|r - q|; \quad \text{in particular,} \quad 0 \leq S(q) - S(r) \leq 2(q - r) \quad (0 \leq r \leq q).$$

Proof. Let $q \in \text{supp } \nu$, so that $S(q) = q$ by (3.9). Recall that S is nondecreasing with values in $[0, 1]$, and note that $G(y_1) - G(y_2) = \int_{y_1}^{y_2} (S(u) - u) \, \mathrm{d}u$ for $y_1 \leq y_2$. Let $r \in [0, 1]$ and $\xi = S(r) \in [0, 1]$. We compare the values of G at ξ and at q . The point ξ need not lie in $\text{supp } \nu$, and this is why we need (6.3) for every $y \in [0, 1]$.

The case $r \leq q$. By monotonicity, $S(q) - S(r) \geq 0$. If $\xi \geq r$, then $S(q) - S(r) = q - \xi \leq q - r$. If $\xi < r$, then $S \leq S(r) = \xi$ on $[\xi, r]$ and $S \leq S(q) = q$ on $[r, q]$, by monotonicity. By (6.3) with $y = \xi$, we have

$$\begin{aligned} 0 \leq G(\xi) - G(q) &= \int_{\xi}^r (S(u) - u) du + \int_r^q (S(u) - u) du \\ &\leq \int_{\xi}^r (\xi - u) du + \int_r^q (q - u) du = \frac{(q - r)^2 - (r - \xi)^2}{2}. \end{aligned}$$

Then $r - \xi \leq q - r$, and $S(q) - S(r) = (q - r) + (r - \xi) \leq 2(q - r)$.

The case $r > q$. By monotonicity, $S(r) - S(q) \geq 0$. If $\xi \leq r$, then $S(r) - S(q) = \xi - q \leq r - q$. If $\xi > r$, then $S \geq q$ on $[q, r]$ and $S \geq \xi$ on $[r, \xi]$, and in the same way

$$\begin{aligned} 0 \leq G(\xi) - G(q) &= - \int_q^r (S(u) - u) du - \int_r^{\xi} (S(u) - u) du \\ &\leq \int_q^r (u - q) du + \int_r^{\xi} (u - \xi) du = \frac{(r - q)^2 - (\xi - r)^2}{2}. \end{aligned}$$

We obtain $\xi - r \leq r - q$, and $S(r) - S(q) = (r - q) + (\xi - r) \leq 2(r - q)$. This completes the proof. $\square$

Together, (6.2) and Lemma 6.2 imply that

$$\mathbb{E}|m_q - m_r|^2 \leq 2N(q - r) \quad (q \in \text{supp } v, 0 \leq r \leq q).$$

6.4 A dyadic decomposition of the mean increment

Lemma 6.1 does not apply directly to the increment $m_q - m_p$, because this increment is not measurable at any level before q . Instead, we split it into increments whose right endpoints approach q geometrically. Each of them except the last is measurable strictly before q , and Lemma 6.1 applies to it. By Lemma 6.2, its second moment is at most proportional to the remaining distance to q , which cancels the length of the clock in Lemma 6.1. The last, short increment is controlled by $C_q \preceq NI$.

Lemma 6.3 (Covariance of a mean increment). *Let μ be a minimizer of $\mathcal{F}_{t,s}$, let $q \in \text{supp } v$ and $0 \leq p \leq q$, and let $a_* = a((p + q)/2)$. For every integer $K \geq 1$,*

$$\mathcal{R}(p, q) \leq (K + 1) \left[\frac{4K}{Ma_*^2} + 2(q - p)2^{-K} \right]. \quad (6.4)$$

Moreover,

$$\mathcal{R}(p, q) \leq \frac{56(1 + \log_+(M(q - p)))^2}{Ma_*^2}. \quad (6.5)$$

Proof. We may assume that $p < q$, since for $p = q$ the left side vanishes and both bounds hold. The proof has three steps. We split $m_q - m_p$ into $K + 1$ dyadic pieces; we bound the first K pieces by Lemmas 6.1 and 6.2 and the last piece directly; and we then choose K .

For $0 \leq k \leq K$, let $q_k = q - (q - p)2^{-k}$, so that $q_0 = p$, $q_1 = (p + q)/2$, and $q - q_k = (q - p)2^{-k}$. We set $b_k = m_{q_{k+1}} - m_{q_k}$ for $0 \leq k < K$ and $b_K = m_q - m_{q_K}$. Then $m_q - m_p = \sum_{k=0}^K b_k$. Applying the Cauchy-Schwarz inequality to the quadratic form of C_q , which is positive semidefinite, we find

$$(m_q - m_p)^T C_q (m_q - m_p) \leq (K + 1) \sum_{k=0}^K b_k^T C_q b_k.$$

The points q_k need not be contact levels; we use Lemma 6.2 only with the endpoint $q \in \text{supp } \nu$.

Step 1: the first K pieces. Let $0 \leq k < K$. The vector b_k is $\mathcal{G}_{q_{k+1}}$ -measurable, and $q_{k+1} < q$. Since a is nondecreasing and $q_{k+1} \geq q_1$, we have $a(u) \geq a(q_1) = a_*$ for $u \geq q_{k+1}$. The clock of Lemma 6.1 with $q' = q_{k+1}$ satisfies

$$\mathcal{I} = c \int_{q_{k+1}}^q a(u)^2 du \geq ca_*^2(q - q_{k+1}),$$

and Lemma 6.1 implies that $\mathbb{E}[b_k^T C_q b_k] \leq \mathbb{E}|b_k|^2 / (ca_*^2(q - q_{k+1}))$. By (6.2), the monotonicity of S , and Lemma 6.2 with $r = q_k$, we have

$$\mathbb{E}|b_k|^2 = N(S(q_{k+1}) - S(q_k)) \leq N(S(q) - S(q_k)) \leq 2N(q - q_k) = 4N(q - q_{k+1}).$$

The last two bounds imply that

$$N^{-2} \mathbb{E}[b_k^T C_q b_k] \leq \frac{4}{Nca_*^2} = \frac{4}{Ma_*^2} \quad (0 \leq k < K).$$

Step 2: the last piece. Using $\|C_q\|_{\text{op}} \leq N$ from (6.1), then (6.2) and Lemma 6.2 with $r = q_K$, we have

$$N^{-2} \mathbb{E}[b_K^T C_q b_K] \leq N^{-1} \mathbb{E}|b_K|^2 = S(q) - S(q_K) \leq 2(q - q_K) = 2(q - p)2^{-K}.$$

We sum the bounds of Steps 1 and 2 over the $K + 1$ pieces and insert them into the Cauchy-Schwarz bound above. This proves (6.4).

Step 3: the choice of K . Let $z = M(q - p)$, $\theta = \log_+ z$, and $K = \lceil \log_2(2 + z) \rceil$, so that $K \geq 1$. Then $2^K \geq 2 + z > z$, which gives

$$2(q - p)2^{-K} = \frac{2z}{M2^K} \leq \frac{2}{M} \leq \frac{2}{Ma_*^2},$$

where we used $a_* \leq 1$. With this choice of K , (6.4) implies that

$$\mathcal{R}(p, q) \leq \frac{(K+1)(4K+2)}{Ma_*^2}.$$

Since $2+z \leq 3 \max\{1, z\}$, we have $\log_2(2+z) \leq \log_2 3 + \theta/\log 2 < 2+2\theta$, where we used $\log_2 3 < 2$ and $1/\log 2 < 2$. It follows that $K < 1 + \log_2(2+z) < 3+2\theta$, and

$$(K+1)(4K+2) \leq (4+2\theta)(14+8\theta) = 56 + 60\theta + 16\theta^2 \leq 56(1+\theta)^2.$$

This proves (6.5). $\square$

The two powers of the logarithm in (6.5) have different sources: there are K increments of comparable cost, and the Cauchy–Schwarz inequality for their sum contributes the factor $K+1$.

6.5 An integrated bound for j

With $p = 0$, Lemma 6.3 bounds $j_q = \mathcal{R}(0, q)$ by a multiple of $(1 + \log_+ M)^2 / (Ma(q/2)^2)$. To integrate against ν , we need a bound in terms of the cumulative mass $s+v$ at $q = Q(v)$, rather than $a(q/2)$, which may be as small as s . We therefore average the earlier means $m_{Q(w)}$, $0 < w < v$, with density proportional to $(s+w)^2$, and assign the remaining weight to $m_0 = 0$. The conditional covariance in the direction of this average is controlled by the integrated variance bound (3.10), and its distance from $m_{Q(v)}$ by Lemma 6.3. With the quadratic weight, both estimates are of order $(s+v)^{-2}$, also for $v < s$, and no separate argument is needed near the root.

Proposition 6.4 (Covariance in the mean direction). *Let $\mu = s\delta_0 + \nu$ be a minimizer of $\mathcal{F}_{t,s}$. For every $v \in (0, 1-s)$,*

$$j_{Q(v)} \leq \frac{341(1 + \log_+ M)^2}{M(s+v)^2}, \quad (6.6)$$

and

$$J = \int j_q \nu(dq) \leq \frac{341(1 + \log_+ M)^2}{Ms}. \quad (6.7)$$

The constant is absolute. The bounds hold for all $N \geq 1$, $c > 0$, $0 < s \leq 1/2$, and every minimizing profile, including profiles with atoms.

Proof. Fix $v \in (0, 1-s)$, and let $q = Q(v)$ and $\mathcal{L} = 1 + \log_+ M$. By (5.1), we have $q \in \text{supp } \nu$, and $Q(w) \leq q$ for $w < v$ because Q is nondecreasing. If $q = 0$, then $m_q = m_0 = 0$ and $j_q = 0$, so we may assume that $q > 0$. The proof has four steps. First, we write m_q as a weighted average $\bar{m}$ of earlier means plus a remainder $\bar{\Delta}$, and we prove the bound (6.8) below. Second, we bound the conditional covariance in the direction of $\bar{m}$ by (3.10), using the conditional Jensen inequality on quantile fibers. Third, we bound the remainder by Lemma 6.3. Fourth, we combine the bounds and integrate over v .

Step 1: the decomposition. We use the deterministic weights

$$r_0 = \frac{s^3}{(s+v)^3}, \quad k_v(w) = \frac{3(s+w)^2}{(s+v)^3} \quad (0 < w < v),$$

which satisfy

$$r_0 + \int_0^v k_v(w) \, dw = \frac{s^3 + ((s+v)^3 - s^3)}{(s+v)^3} = 1.$$

We use the probability measure $r_0\delta_0 + k_v(w) \, dw$ on $[0, v)$ to average the vector $m_0 = 0$ and the vectors $m_{Q(w)}$, $0 < w < v$. We define

$$\begin{aligned} \bar{m} &= \int_0^v k_v(w) m_{Q(w)} \, dw, & \bar{L} &= \int_0^v k_v(w) L_{Q(w)} \, dw, \\ \bar{\Delta} &= r_0 m_q + \int_0^v k_v(w) (m_q - m_{Q(w)}) \, dw. \end{aligned}$$

By the normalization of the weights, $m_q = \bar{m} + \bar{\Delta}$. The vector $\bar{m}$ and the number $D = 2^{-1} \int_0^v k_v(w) |m_{Q(w)}|^2 \, dw$ are $\mathcal{G}_q$ -measurable, because $Q(w) \leq q$ for $w < v$, and $N\bar{L} = \sigma \cdot \bar{m} - D$. It follows that $\text{Var}(\bar{L} \mid \mathcal{G}_q) = N^{-2} \bar{m}^T C_q \bar{m}$, and by the law of total variance,

$$N^{-2} \mathbb{E}[\bar{m}^T C_q \bar{m}] = \mathbb{E} \text{Var}(\bar{L} \mid \mathcal{G}_q) \leq \text{Var}(\bar{L}).$$

Next, we write $\bar{\Delta} = r_0(m_q - m_0) + \int_0^v k_v(w)(m_q - m_{Q(w)}) \, dw$, which is the average of the vectors $m_q - m_0$ and $m_q - m_{Q(w)}$ with respect to the probability measure above. By Jensen's inequality for the convex quadratic form of C_q , we have

$$\bar{\Delta}^T C_q \bar{\Delta} \leq r_0(m_q - m_0)^T C_q (m_q - m_0) + \int_0^v k_v(w)(m_q - m_{Q(w)})^T C_q (m_q - m_{Q(w)}) \, dw.$$

The integrand is bounded by $4N^2$, by (6.1). We take expectations, use Fubini's theorem, and divide by N^2 . This gives

$$N^{-2} \mathbb{E}[\bar{\Delta}^T C_q \bar{\Delta}] \leq r_0 \mathcal{R}(0, q) + \int_0^v k_v(w) \mathcal{R}(Q(w), q) \, dw.$$

Finally, since $C_q \succeq 0$, we have $m_q^T C_q m_q \leq 2\bar{m}^T C_q \bar{m} + 2\bar{\Delta}^T C_q \bar{\Delta}$. Combining the last three bounds, we obtain

$$j_q \leq 2 \text{Var}(\bar{L}) + 2r_0 \mathcal{R}(0, q) + 2 \int_0^v k_v(w) \mathcal{R}(Q(w), q) \, dw. \quad (6.8)$$

The quantity $\mathcal{R}(0, q) = j_q$ appears on the right side. This is not circular: in Step 3, we bound it by Lemma 6.3, independently of (6.8).

Step 2: the variance of the weighted prefix. We extend k_v by zero to $(0, 1-s)$, and note that $k_v \leq 3/(s+v)$. The image of the measure $k_v(w) dw$ under Q is then at most $3/(s+v)$ times ν , and it has a density κ_v with respect to ν , with $0 \leq \kappa_v \leq 3/(s+v)$; in particular, $\kappa_v \in L^2(\nu)$. By the definition of κ_v and the pushforward property of Q , we have, for every bounded Borel function φ ,

$$\int_0^{1-s} k_v(w) \varphi(Q(w)) dw = \int \kappa_v \varphi d\nu = \int_0^{1-s} \kappa_v(Q(w)) \varphi(Q(w)) dw.$$

Applied for each outcome to the bounded function $p \mapsto L_p$, the first identity shows that $\bar{L} = \int \kappa_v(p) L_p \nu(dp)$. The identity as a whole says that, under the normalized Lebesgue measure on $(0, 1-s)$, the function $\kappa_v \circ Q$ is the conditional expectation of k_v given the σ -algebra generated by Q . Explicitly, if p is an atom of ν with mass r_p , then

$$\kappa_v(p) = \frac{1}{r_p} \int_{(F_-(p), F(p))} k_v(w) dw,$$

with the notation of Section 5.1. A prefix $(0, v)$ that ends inside the quantile interval of an atom gives the atom the average of k_v over that interval. Since $1/a(Q)$ is a bounded nonnegative function of Q , we may apply the conditional Jensen inequality and find

$$\int \frac{\kappa_v(p)^2}{a(p)} \nu(dp) = \int_0^{1-s} \frac{\kappa_v(Q(w))^2}{a(Q(w))} dw \leq \int_0^{1-s} \frac{k_v(w)^2}{a(Q(w))} dw = \int_0^v \frac{k_v(w)^2}{a(Q(w))} dw.$$

As μ is a minimizer, Proposition 3.3 applies with $k = \kappa_v$. Together with $a(Q(w)) \geq s+w$, this yields

$$\begin{aligned} \text{Var}(\bar{L}) &\leq \frac{1}{M} \int_0^v \frac{k_v(w)^2}{s+w} dw = \frac{9}{M(s+v)^6} \int_0^v (s+w)^3 dw \\ &= \frac{9}{4M(s+v)^2} \left[1 - \frac{s^4}{(s+v)^4} \right] \leq \frac{9}{4M(s+v)^2}. \end{aligned}$$

Step 3: the remainder. Let $0 < w < v$ and $p = Q(w)$. Then $p \leq q$, and since a is nondecreasing, $a((p+q)/2) \geq a(p) \geq s+w$. For the root term, we have $a(q/2) \geq s$ by (5.1). All level differences are at most one, and hence $\log_+(M(q-p)) \leq \log_+ M$. Applying Lemma 6.3 with the endpoint $q \in \text{supp } \nu$ and the lower endpoints 0 and $Q(w)$, we have

$$\mathcal{R}(0, q) \leq \frac{56\mathcal{L}^2}{Ms^2}, \quad \mathcal{R}(Q(w), q) \leq \frac{56\mathcal{L}^2}{M(s+w)^2} \quad (0 < w < v).$$

Using $r_0 s^{-2} = s/(s+v)^3$ and $k_v(w)(s+w)^{-2} = 3/(s+v)^3$, we find that

$$r_0 \mathcal{R}(0, q) + \int_0^v k_v(w) \mathcal{R}(Q(w), q) dw \leq \frac{56\mathcal{L}^2}{M} \cdot \frac{s+3v}{(s+v)^3} \leq \frac{168\mathcal{L}^2}{M(s+v)^2},$$

where we used $s+3v \leq 3(s+v)$.

Step 4: conclusion. Inserting the bounds of Steps 2 and 3 into (6.8) and using $\mathcal{L} \geq 1$, we obtain

$$j_{Q(v)} \leq \frac{9}{2M(s+v)^2} + \frac{336\mathcal{L}^2}{M(s+v)^2} \leq \frac{(9/2 + 336)\mathcal{L}^2}{M(s+v)^2} \leq \frac{341\mathcal{L}^2}{M(s+v)^2},$$

which is (6.6). Finally, $J = \int_0^{1-s} j_{Q(v)} \, dv$, because Q pushes Lebesgue measure on $(0, 1-s)$ forward to ν , and (6.7) follows from

$$\int_0^{1-s} \frac{dv}{(s+v)^2} = \frac{1}{s} - 1 \leq \frac{1}{s}.$$

This completes the proof. □

PART III

The minimizing profile and the scalar response

Isolation of the minimizing profile

In this section and the next, we study an arbitrary minimizer γ_* of the finite-dimensional functional F_ρ in the energy normalization of Section 3. We compare γ_* with the scalar trial profile $\gamma_\lambda^{\text{tr}} = \max\{\lambda, \gamma_\beta\}$. We state all estimates in terms of a trial budget η , that is, a bound on the excess $F_\rho(\gamma_\lambda^{\text{tr}}) - F_\rho(\gamma_*)$. The main result of this section is Proposition 7.6, which bounds the L^1 distance between γ_* and $\gamma_\lambda^{\text{tr}}$ by $C \min\{1, \rho^{-3/4} \eta^{1/4}\}$ for every minimizer and every trial budget, with a constant C depending only on β . In Proposition 9.9, we use this bound to prove the stability of the scalar response operator at the minimizing profile. Section 8 uses the tools of this section to localize the lowest quantiles of the free mass of γ_* .

We use the energy normalization, the tilted joint law, and the martingale identities of Section 3, in particular Proposition 3.1, and the trial budget of Lemma 3.8. From [4], we import the convexity of F_ρ along segments of profiles, its first variation, the response inequality for the second variation, and an averaged curvature bound (Proposition 7.1); the regularization of [4, Lemma 7.5] (Lemma 7.3(a)); and the linear growth of the scalar Parisi profile near the origin (7.5). From [48], we use the structure of the scalar Parisi measure (7.6). To prove Proposition 7.6, we first smooth the minimizer (Lemma 7.3(b)) and then join it to the trial profile by a segment. With cutoffs of one sign, built from one-sided averages of the superlevel sets of the difference of the two profiles, we turn the response inequality into lower bounds on the second derivative of F_ρ along the segment (Lemma 7.4). The midpoint inequality (7.9) then compares these lower bounds with the trial budget.

Proposition 7.1, Lemma 7.3(a), Lemma 7.4, the midpoint inequality (7.9), and the facts (7.5), (7.6), and (7.7) are used again in Section 8. Section 9 uses (7.5), (7.6), (7.7), and Proposition 7.6.

7.1 Energy normalization and variational facts

Throughout this section and Section 8, we fix $1 < \beta < \infty$, $N \geq 1$, $0 < \rho \leq 1$, $t = 1 - \rho$, and $0 < \lambda \leq \lambda_0$, where $\lambda_0 \leq \min\{1, \beta/4\}$ is the constant fixed with (3.17), and $s = \lambda/\beta$. Unless stated otherwise, constants depend only on β . In particular, they do not depend on N , ρ , λ , the profile, or the positions and masses of its atoms.

Admissible profiles and segments. We use the energy normalization of Section 3. An *admissible profile* is a function $\gamma = \beta a$, where $a(q) = \mu([0, q])$ is the cumulative function of an admissible profile $\mu = s\delta_0 + \nu$ of Section 3. The function γ is right-continuous and nondecreasing with $\gamma(1) = \beta$. With the convention $\gamma(0-) = 0$, its derivative measure $d\gamma$ has total mass β and satisfies $d\gamma \geq \lambda\delta_0$. We call

$$\nu_\gamma = d\gamma - \lambda\delta_0 = \beta\nu$$

the *free measure* of γ . It has mass $\beta - \lambda$ and may charge 0. The admissible profiles form a convex set. We write $\min F_\rho$ for the minimum of F_ρ over this set, which is attained (Section 3). An admissible profile γ_* is a *minimizer* if $F_\rho(\gamma_*) = \min F_\rho$. A *segment* is a family $\gamma_\theta = (1 - \theta)\gamma_0 + \theta\gamma_1$, $0 \leq \theta \leq 1$, with admissible endpoints γ_0, γ_1 . Along a segment, the free measure $\nu_{\gamma_\theta} = (1 - \theta)\nu_{\gamma_0} + \theta\nu_{\gamma_1}$ is affine in θ .

Self-overlap, trace, and curvature. For an admissible profile γ , we write $\mathbb{E}_\gamma$ for expectation under the joint law of Section 3.3 with profile $a = \gamma/\beta$. We define

$$S_\gamma(q) = \mathbb{E}_\gamma S_q, \quad b_\gamma(q) = \frac{\beta}{N} \mathbb{E}_\gamma \operatorname{Tr} H_q, \quad \zeta_\gamma(q) = \frac{\beta^2}{N} \mathbb{E}_\gamma \operatorname{Tr} H_q^2.$$

By Proposition 3.1, we have $m_0 = 0$, $dm_q = \sqrt{c} H_q dB_q$, and $a(q)C_q \preceq H_q \preceq C_q$, and the function $q \mapsto \mathbb{E}_\gamma S_q$ is continuously differentiable with derivative $(c/N)\mathbb{E}_\gamma \operatorname{Tr} H_q^2$. Recalling that $c = \beta^2\rho$, we see that this derivative is $\rho\zeta_\gamma(q)$. The coordinates of m_q lie in $[-1, 1]$, which gives $0 \leq S_q \leq 1$ and $\operatorname{Tr} C_q = \sum_i (1 - m_{q,i}^2) \leq N$. Further, $0 \preceq H_q \preceq C_q$ implies that $0 \leq \operatorname{Tr} H_q \leq \operatorname{Tr} C_q$. We obtain

$$S_\gamma(0) = 0, \quad 0 \leq S_\gamma \leq 1, \quad S'_\gamma = \rho\zeta_\gamma \geq 0, \quad 0 \leq b_\gamma \leq \beta. \quad (7.1)$$

Note that $S_\gamma(q) = \rho \int_0^q \zeta_\gamma(r) dr$, and that S_γ is nondecreasing. We also note that $0 \preceq H_q \preceq C_q$ implies that $\operatorname{Tr} H_q^2 \leq \operatorname{Tr} C_q^2 \leq (\operatorname{Tr} C_q)^2 \leq N^2$, and hence that $0 \leq \zeta_\gamma \leq \beta^2 N$. At fixed N , the functions S_γ , b_γ , and ζ_γ are bounded.

The rescaled functional of [4]. The variational facts we need are proved in [4, Sections 7.1 and 7.4 and Appendix A] for the rescaled form of F_ρ recalled in Section 3, and we now check this identification in detail. Let γ be admissible, let $U = U_{\mu,t}$ be the solution of (3.2) with $a = \gamma/\beta$, and let $\tilde{U}(q, y; W) = \beta^{-1}U(q, \beta y; W)$ be the rescaled solution of Section 3; when we need to indicate the profile, we write $\tilde{U}_\gamma$. Recall from Section 3 that $\tilde{U}$ solves $\partial_q \tilde{U} = -(\rho/2)(\Delta_y \tilde{U} + \gamma(q)|\nabla_y \tilde{U}|^2)$, that $\nabla_y \tilde{U}(q, y; W) = \nabla_x U(q, \beta y; W)$ and $\nabla_y^2 \tilde{U}(q, y; W) = \beta \nabla_x^2 U(q, \beta y; W)$, and that $F_\rho(\gamma)$ equals $(N\lambda)^{-1} \log \mathbb{E}_{\text{ref}} e^{\lambda \tilde{U}(0,0;W)} - (\rho/2) \int_0^1 q\gamma(q) dq$. By (3.2), the terminal value of $\tilde{U}$ is

$$\tilde{U}(1, y; W) = \frac{1}{\beta} \log \sum_{\sigma \in \{-1,1\}^N} \exp \left\{ \beta \left(\sqrt{t} \frac{\sigma^T W \sigma}{2} + y \cdot \sigma \right) \right\}.$$

Let W^{off} be the matrix W with its diagonal set to zero. For $\sigma \in \{-1, 1\}^N$, we have $\sigma^T W \sigma / 2 = \sigma^T W^{\text{off}} \sigma / 2 + D_N$, where $D_N = \text{Tr } W / 2$. Since the equation for $\tilde{U}$ is invariant under the addition of constants, we find that $\tilde{U}(q, y; W) = \sqrt{t} D_N + \tilde{U}^{\text{off}}(q, y)$, where $\tilde{U}^{\text{off}}$ is the solution of the same equation with W^{off} in place of W in the terminal value. This function is the solution u of [4, (7.2)] with $r = \rho$ and coupling matrix $\sqrt{t} W^{\text{off}}$. The variable D_N is centered Gaussian with variance $1/2$ and independent of W^{off} , so that $(N\lambda)^{-1} \log \mathbb{E}_{\text{ref}} e^{\lambda \sqrt{t} D_N} = \lambda t / (4N)$. Inserting this into the representation of F_ρ recalled above, we find

$$F_\rho(\gamma) = \frac{1}{N\lambda} \log \mathbb{E}_{\text{ref}} e^{\lambda \tilde{U}^{\text{off}}(0,0)} - \frac{\rho}{2} \int_0^1 q \gamma(q) dq + \frac{\lambda t}{4N}.$$

This is exactly the functional $F_t^c(\gamma)$ of [4, (7.2)], which is written F_r in [4, Section 7.4]; the term $\lambda t / (4N)$ is part of that functional and does not depend on γ .

The joint laws also agree. The tilt $e^{sU(0,0;W)} = e^{\lambda \sqrt{t} D_N} e^{\lambda \tilde{U}^{\text{off}}(0,0)}$ of Section 3.3 factorizes, and D_N enters neither the field equation (3.5), whose drift involves $\nabla_x U$, nor the conditional law of the terminal spin given W and the field. Under the tilted law, the couplings $\sqrt{t} W^{\text{off}}$ have the law Q_γ of [4, Appendix A], and the rescaled field X_q / β solves

$$d(X_q / \beta) = \rho \gamma(q) \nabla_y \tilde{U}(q, X_q / \beta; W) dq + \sqrt{\rho} dB_q, \quad X_0 / \beta = 0,$$

which is the field equation of [4, Section 7.1]. Under this law, the magnetization $m_q = \nabla_x U(q, X_q; W) = \nabla_y \tilde{U}(q, X_q / \beta; W)$ is unchanged, while the Hessian in the rescaled variable is $\nabla_y^2 \tilde{U}(q, X_q / \beta; W) = \beta H_q$. We conclude that S_γ , b_γ , and ζ_γ are the functions S , b , and ζ of [4, Sections 7.1 and 7.4], and that b_γ and ζ_γ are the functions of the same names in [4, Appendix A]. We use the results proved there in the following form.

Proposition 7.1 (Variational and response facts). *Let γ be admissible and define $G_\gamma(x) = \int_x^1 (S_\gamma(q) - q) dq$ for $0 \leq x \leq 1$.*

- (a) *Let $\nu' \geq 0$ be a measure on $[0, 1]$ of mass $\beta - \lambda$. Then F_ρ is convex along the segment from γ to the profile with free measure ν' , and its one-sided derivative at γ is*

$$D_+ F_\rho(\gamma)[\nu' - \nu_\gamma] = \frac{\rho}{2} \int G_\gamma d(\nu' - \nu_\gamma). \quad (7.2)$$

The function G_γ is continuously differentiable with $G'_\gamma(x) = x - S_\gamma(x)$. In particular, $G'_\gamma(0) = 0$, and $x \mapsto G'_\gamma(x) - x$ is nonincreasing.

- (b) *Let (γ_θ) be a segment, $\psi = \gamma_1 - \gamma_0$, and $f(\theta) = F_\rho(\gamma_\theta)$. Then f is convex and continuous on $[0, 1]$ and twice differentiable on $(0, 1)$. Fix $\theta \in (0, 1)$ and write $b_\theta = b_{\gamma_\theta}$ and $\zeta_\theta = \zeta_{\gamma_\theta}$. For every compactly supported Lipschitz function ϕ on $(0, 1)$, let*

$$Y = \int_0^1 b_\theta |\phi'|^2 dq, \quad Z = \rho \int_{(0,1)} \phi^2 \zeta_\theta d\gamma_\theta.$$

Then

$$\rho \left| \int_0^1 \psi \zeta_\theta \varphi \, dq \right| \leq \left(\frac{f''(\theta)}{\rho} \right)^{1/2} [Y^{1/2} + 2(Y + Z)^{1/2}]. \quad (7.3)$$

(c) There exists $k_\beta > 0$, depending only on β , such that

$$\zeta_\gamma(q) \geq k_\beta^2 \quad (0 \leq q < 1). \quad (7.4)$$

The integral Z is a Stieltjes integral over the open interval $(0, 1)$. It includes the atoms of $d\gamma_\theta$ in $(0, 1)$ and excludes the prescribed atom at 0.

Proof. We prove the three parts in order. By the discussion preceding the proposition, F_ρ is exactly the functional F_r of [4, Section 7.4] with $r = \rho$, and $S_\gamma, b_\gamma, \zeta_\gamma$ are the response functions used there.

Part (a). The segment from γ to the profile with free measure ν' replaces ν_γ by $(1 - \theta)\nu_\gamma + \theta\nu'$. The convexity of F_r along every such segment and the first variation (7.2) are stated in [4, Section 7.4, (7.22)], with G_γ the function defined there. The identity (7.2) is also (3.8) divided by β : the free measures satisfy $\nu_\gamma = \beta\nu$, which gives $\beta^{-1}(c/2) \int G \, d(\nu'/\beta - \nu) = (\rho/2) \int G \, d(\nu' - \nu_\gamma)$, and the function G of (3.8) is G_γ . By (7.1), S_γ is continuous, nondecreasing, and vanishes at 0. By the definition of G_γ , we then have $G'_\gamma(x) = x - S_\gamma(x)$, which is continuous. We also have $G'_\gamma(0) = 0$, and $G'_\gamma(x) - x = -S_\gamma(x)$ is nonincreasing. This proves (a); see also [4, (7.23)].

Part (b). Applying (a) with $\gamma = \gamma_0$ and $\nu' = \nu_{\gamma_1}$, we see that f is convex. The profile-continuity estimate of [4, Section 7.4] states that

$$\sup_{q,y} |\tilde{U}_\gamma^{\text{off}}(q, y) - \tilde{U}_{\gamma'}^{\text{off}}(q, y)| \leq \frac{3\rho N}{2} \|\gamma - \gamma'\|_{L^1(0,1)}$$

for admissible γ, γ' and every value of the couplings. Inserting this bound into the representation of F_ρ above, we find

$$|F_\rho(\gamma) - F_\rho(\gamma')| \leq 2\rho \|\gamma - \gamma'\|_{L^1(0,1)},$$

where the outer logarithmic moment contributes $3\rho/2$ and the penalty contributes $\rho/2$. Then f is Lipschitz on $[0, 1]$, because $\|\gamma_\theta - \gamma_{\theta'}\|_{L^1} = |\theta - \theta'| \|\gamma_1 - \gamma_0\|_{L^1}$.

We now apply [4, Proposition A.1 and Lemma A.2] with $r = \rho$, the endpoints γ_0, γ_1 , and $v = \psi$. Their hypotheses hold: N is finite, $0 \leq t < 1$, the endpoints are profiles of [4, Appendix A] ($\lambda \leq \gamma_i \leq \beta$ on $[0, 1)$, $\gamma_i(0-) = 0$, and $\gamma_i(1) = \beta$) with the same outer tilt λ , and [4, Proposition A.1] allows their additional masses at 0 to differ. By [4, Proposition A.1, (A.1)], f is twice differentiable, and $f''(\theta)$ is the sum of $\rho \int_0^1 e_\theta(q) \, dq$ and a nonnegative variance term, where $e_\theta \geq 0$ is the function defined there. In particular,

$$\int_0^1 e_\theta(q) \, dq \leq \frac{f''(\theta)}{\rho}.$$

The pairing estimate [4, (A.2)], applied to φ and to $-\varphi$, and the Fisher estimate [4, (A.3)] give a nonnegative function $\mathcal{J}_\theta$, also defined in [4, Proposition A.1], such that

$$\rho \left| \int_0^1 \psi \zeta_\theta \varphi \, dq \right| \leq \left(\int_0^1 e_\theta \right)^{1/2} \left[Y^{1/2} + 2 \left(\int_0^1 \mathcal{J}_\theta \varphi^2 \right)^{1/2} \right], \quad \int_0^1 \mathcal{J}_\theta \varphi^2 \leq Y + Z.$$

Combining the last three displays, we obtain (7.3).

Part (c). We use [4, Lemma A.3, (A.5)] with the level in (A.5) equal to β ; the hypothesis $\beta \geq 1$ of that lemma holds. As $\max\{1, \beta\} = \beta$, the lemma states that $\gamma(q) \leq \beta$ and $q < 1$ imply $\mathbb{E}_\gamma[\beta H_q] \geq k_\beta I$, where

$$k_\beta = \frac{e^{-16c'\beta - 2\beta^2}}{2(1 + c')}, \quad c' = \sqrt{\pi/2}.$$

Every admissible profile satisfies $\gamma \leq \beta$. By Jensen's inequality, applied to the convex function $H \mapsto \text{Tr } H^2$, we have $\zeta_\gamma(q) = N^{-1} \mathbb{E}_\gamma \text{Tr}(\beta H_q)^2 \geq N^{-1} \text{Tr}(\mathbb{E}_\gamma \beta H_q)^2 \geq k_\beta^2$, which is (7.4).

In [4], these results are first proved for step profiles, and then extended to all admissible profiles by approximation at fixed N and β [4, Section 7.4 and Appendix A.3]. This includes profiles with atoms, and the Stieltjes integral in Z includes the atoms of $d\gamma_\theta$ in $(0, 1)$. $\square$

Remark 7.2 (Atoms and the Stieltjes term). The term Z is where atoms of the profile enter (7.3). In the proof of [4, Lemma A.2], the Fisher estimate (A.3) follows from the identity of measures $d\mathcal{P} = 2\mathcal{J}_\theta \, dq - \rho \zeta_\theta \, d\gamma_\theta$ on $(0, 1)$ [4, (A.19)], where $\mathcal{P}(q) = N^{-1} \mathbb{E}_\theta \text{Tr } \dot{h}$ and $\dot{h}$ is the derivative in q , at fixed magnetization, of the Hessian of $\tilde{U}_{\gamma_\theta}$ written as a function of the magnetization. At an atom of $d\gamma_\theta$ at q , the function $\mathcal{P}$ jumps by $-\rho \zeta_\theta(q) \, d\gamma_\theta(\{q\})$. In Proposition 7.6 and Lemma 8.1, we choose the cutoffs so that $d\gamma_\theta$ has a bounded density on the set where they are positive.

The scalar reference profile

Let μ_β be the scalar Parisi measure, $\alpha_\beta = \mu_\beta([0, \cdot])$, and $\gamma_\beta = \beta \alpha_\beta$, as in Section 3, and let $\gamma^{\text{tr}} = \max\{\lambda, \gamma_\beta\}$ be the scalar trial profile; we omit the subscript λ of Section 3. The function γ_β is the unconstrained minimizer of the rescaled scalar Parisi functional $\gamma \mapsto \beta^{-1} \mathcal{P}_\beta(\gamma/\beta)$ of [4, Section 7.1]. By [4, Lemma 7.2], applied with $\beta_0 = \beta$, there exist constants $T_*, k_0 > 0$ and $C_0 < \infty$, depending only on β , such that $\gamma_\beta(0) = 0$, γ_β is absolutely continuous on $[0, T_*]$, and

$$k_0 q \leq \gamma_\beta(q) \leq C_0 q, \quad 0 \leq \gamma'_\beta(q) \leq C_0 \quad (0 \leq q \leq T_*),$$

the second bound for almost every q . We set $T_0 = \min\{T_*, 1/4, 2/C_0\}$. On $[0, T_0]$, we have $\gamma_\beta \leq C_0 T_0 \leq 2$, and since $\lambda \leq 1$, also $\gamma^{\text{tr}} \leq 2$. The maximum of the constant λ and

the nondecreasing absolutely continuous function γ_β is absolutely continuous on $[0, T_0]$, with derivative $\gamma'_\beta \mathbf{1}_{\{\gamma_\beta > \lambda\}} \in [0, C_0]$ almost everywhere. We obtain

$$\gamma_\beta(0) = 0, \quad k_0 q \leq \gamma_\beta(q), \quad \gamma^{\text{tr}}(q) \leq 2 \quad (0 \leq q \leq T_0), \quad (7.5)$$

and the restriction of $d\gamma^{\text{tr}}$ to $(0, T_0]$ has density at most C_0 . At 0, the measure $d\gamma^{\text{tr}}$ has the prescribed atom of mass $\gamma^{\text{tr}}(0) = \lambda$.

By [48, Theorem 1.1, (1.4) and (1.5)], there exist $q_\beta \in (0, 1)$, $c_\beta \in (0, 1)$, and a nonnegative function $\varrho_\beta \in C^\infty([0, q_\beta])$ such that

$$\mu_\beta(dq) = \varrho_\beta(q) \mathbf{1}_{[0, q_\beta)}(q) dq + c_\beta \delta_{q_\beta}(dq), \quad \text{supp } \mu_\beta = [0, q_\beta]. \quad (7.6)$$

Here, as in [48], $\varrho_\beta \in C^\infty([0, q_\beta])$ means that ϱ_β is smooth on $(0, q_\beta)$ and that each of its derivatives extends continuously to $[0, q_\beta]$. The only atom of μ_β is at q_β , and $\alpha_\beta = 1$ and $\gamma_\beta = \beta$ on $[q_\beta, 1]$. (The Parisi functional [48, (1.3)] is written with the terminal value $\log \cosh x$ and the additional constant $\log 2$; it equals the functional in (1.1), and μ_β is its minimizer.)

Trial budgets

For a minimizer γ_* of F_ρ , a *trial budget* is a number $\eta \geq 0$ such that

$$F_\rho(\gamma^{\text{tr}}) - F_\rho(\gamma_*) \leq \eta. \quad (7.7)$$

Since $F_\rho(\gamma_*) = \min F_\rho$, this is the notion of trial budget of Section 3, measured in energy units. By (3.18), $\eta = e_{N, \beta}(\lambda) + L_s \lambda^5$ is a trial budget for every minimizer at every $\rho \in (0, 1]$. The estimates of this section and the next hold for every trial budget.

7.2 Smoothing and coercivity

A minimizer may have atoms, and the response inequality (7.3) involves the derivative measure $d\gamma_\theta$ on the set where the cutoff is positive. We therefore first spread the free mass, at a cost that is quadratic in the smoothing width (Lemma 7.3). Part (a) of the next lemma is [4, Lemma 7.5], and part (b) extends the reflection argument of that lemma to the upper endpoint 1. We then show that a difference of one sign between the endpoints of a segment forces curvature along the segment (Lemma 7.4).

Lemma 7.3 (Reflected regularization). *Let γ be admissible and $h > 0$.*

- (a) *Let $0 < T < 1$, $h \leq 1 - T$, and $m = \nu_\gamma([0, T])$. Let $\bar{\gamma}$ be the profile with $d\bar{\gamma} = \lambda \delta_0 + \bar{\nu}$, where $\bar{\nu}$ is obtained from ν_γ by replacing each unit of mass at a point $q < T$ by the image of the uniform distribution on $[q - h, q + h]$ under $x \mapsto |x|$. Then $\bar{\gamma}$ is admissible, $\bar{\gamma}(0) = \lambda$, the restriction of $d\bar{\gamma}$ to $(0, T)$ has density at most m/h , and $F_\rho(\bar{\gamma}) - F_\rho(\gamma) \leq \rho m h^2 / 12$.*

(b) Let $h \leq 1/2$, and let $\pi(x) = |x|$ for $x \leq 1$ and $\pi(x) = 2 - x$ for $x > 1$. Let γ^h be the profile with $d\gamma^h = \lambda\delta_0 + v^h$, where v^h is obtained from v_γ by replacing each unit of mass at a point $q \in [0, 1]$ by the image of the uniform distribution on $[q - h, q + h]$ under π . Then γ^h is admissible, $\gamma^h(0) = \lambda$, its free measure has density at most β/h on $[0, 1]$, and

$$\|\gamma^h - \gamma\|_{L^1(0,1)} \leq \beta h, \quad F_\rho(\gamma^h) - F_\rho(\gamma) \leq \frac{\rho\beta h^2}{12}.$$

Proof. Part (a) is [4, Lemma 7.5], including the cost bound [4, (7.24)]; by Section 7.1, the functional F_r of that lemma is exactly F_ρ with $r = \rho$. Its proof is the argument for (b) below, with the following changes. Only the free mass on $[0, T)$ is moved, and the reflection is $x \mapsto |x|$. The condition $h \leq 1 - T$ ensures that the moved mass stays in $[0, T + h) \subset [0, 1]$. The semiconcavity bound is then needed only on $[-h, T + h] \subset [-1, 1]$, where the reflection at 1 plays no role. The cost bound becomes $(\rho/2) \cdot m \cdot (h^2/6) = \rho m h^2/12$, and the moved mass has density at most m/h on $(0, T)$, while the mass on $[T, 1]$ does not charge $(0, T)$.

For (b), we first verify the properties of γ^h , then the displacement bound, and finally the cost bound. For $q \in [0, 1]$, let κ_q be the image of the uniform distribution on $[q - h, q + h]$ under π . The free measure of γ^h is then $v^h = \int \kappa_q v_\gamma(dq)$. The map π sends $[-h, 1 + h]$ into $[0, 1]$ and is piecewise affine with slopes ± 1 . For $x \in (0, 1)$, the points of $[q - h, q + h]$ that π sends to x are among x , $-x$, and $2 - x$. We find that κ_q has the density

$$\frac{1}{2h} (\mathbf{1}_{\{|x-q|\leq h\}} + \mathbf{1}_{\{x+q\leq h\}} + \mathbf{1}_{\{x+q\geq 2-h\}}), \quad x \in (0, 1).$$

Since $h \leq 1/2 < 2 - h$, the last two indicators cannot both equal one, and this density is at most $1/h$. Thus κ_q is a probability measure on $[0, 1]$ without atoms. Integrating over v_γ , we see that v^h is a measure on $[0, 1]$ of mass $\beta - \lambda$, without atoms, with density at most $(\beta - \lambda)/h \leq \beta/h$. The measure $d\gamma^h = \lambda\delta_0 + v^h$ has mass β , we have $\gamma^h(0) = \lambda$, and γ^h is admissible.

For the displacement bound, let Π be the measure on $[0, 1]^2$ defined by

$$\int F d\Pi = \int_{[0,1]} \frac{1}{2h} \int_{-h}^h F(q, \pi(q+z)) dz v_\gamma(dq)$$

for bounded Borel functions F . Its marginals are v_γ and v^h . The map π is 1-Lipschitz and $\pi(q) = q$ for $q \in [0, 1]$. It follows that $|\pi(q+z) - q| \leq |z| \leq h$, and Π is supported on pairs (q, r) with $|r - q| \leq h$. The prescribed atoms of $d\gamma$ and $d\gamma^h$ coincide, and for $x \in [0, 1]$ we have

$$\gamma^h(x) - \gamma(x) = \int (\mathbf{1}_{\{r \leq x\}} - \mathbf{1}_{\{q \leq x\}}) \Pi(dq, dr).$$

For $q, r \in [0, 1]$, we have $\int_0^1 |\mathbf{1}_{\{r \leq x\}} - \mathbf{1}_{\{q \leq x\}}| dx = |r - q|$. By Tonelli's theorem, we conclude that

$$\|\gamma^h - \gamma\|_{L^1(0,1)} \leq \int |r - q| \Pi(dq, dr) \leq (\beta - \lambda)h \leq \beta h.$$

To bound the cost, let $G = G_{\gamma^h}$ and $S = S_{\gamma^h}$, and define $H(z) = G(\pi(z)) - z^2/2$ for $z \in [-h, 1 + h]$. By Proposition 7.1(a), G is continuously differentiable on $[0, 1]$ with $G'(x) = x - S(x)$. Using $\pi(z) = -z$ for $z < 0$ and $\pi(z) = 2 - z$ for $z > 1$, we find that H is continuous and continuously differentiable away from 0 and 1, with

$$H'(z) = \begin{cases} S(-z), & -h < z < 0, \\ -S(z), & 0 < z < 1, \\ S(2 - z) - 2, & 1 < z < 1 + h. \end{cases}$$

Each piece is nonincreasing, because S is nondecreasing by (7.1). The one-sided derivatives at 0 agree, because $S(0) = 0$. At 1, the derivative jumps from $-S(1)$ to $S(1) - 2 \leq -S(1)$, where we used that $S(1) \leq 1$. We see that H' is nonincreasing, and H is concave on $[-h, 1 + h]$. For $q \in [0, 1]$, the number $G'(q) - q$ is a supergradient of H at q ; at $q = 1$, we take the left derivative $G'(1) - 1 = -S(1)$. By the concavity of H , we have $H(q + z) \leq H(q) + (G'(q) - q)z$, that is,

$$G(\pi(q + z)) \leq G(q) + G'(q)z + \frac{z^2}{2} \quad (q \in [0, 1], |z| \leq h).$$

We average over z in $[-h, h]$. Using that the average of z is 0 and that the average of z^2 is $h^2/3$, we obtain

$$\int G d\kappa_q - G(q) \leq \frac{h^2}{6} \quad (q \in [0, 1]). \quad (7.8)$$

By Proposition 7.1(a), F_ρ is convex along the segment from γ^h to γ , which implies that $F_\rho(\gamma) \geq F_\rho(\gamma^h) + D_+ F_\rho(\gamma^h)[\nu_\gamma - \nu^h]$. Using (7.2) at γ^h and (7.8), we conclude that

$$F_\rho(\gamma^h) - F_\rho(\gamma) \leq \frac{\rho}{2} \int G d(\nu^h - \nu_\gamma) = \frac{\rho}{2} \int \left[\int G d\kappa_q - G(q) \right] \nu_\gamma(dq) \leq \frac{\rho(\beta - \lambda)h^2}{12}. \quad \square$$

The midpoint inequality. We compare the second derivative along a segment with the values at its endpoints. Let f be convex and continuous on $[0, 1]$ and twice differentiable on $(0, 1)$. Then f' is nondecreasing on $(0, 1)$, and f is absolutely continuous on $[0, 1]$: on each interval $[\epsilon, 1 - \epsilon]$ it is Lipschitz, and since f' is monotone and f is continuous at 0 and 1, the monotone convergence theorem shows that $f(y) - f(x) = \int_x^y f'$ for all $0 \leq x < y \leq 1$. Further, a nondecreasing function g on $(0, 1)$ satisfies $g(y) - g(x) \geq \int_x^y g'$ for $0 < x < y < 1$, by Lebesgue's theorem on the differentiation of monotone functions; we apply this with $g = f'$, whose derivative $f'' \geq 0$ exists everywhere on $(0, 1)$. We have

$$f(0) + f(1) - 2f(\tfrac{1}{2}) = \int_0^{1/2} (f'(x + \tfrac{1}{2}) - f'(x)) dx \geq \int_0^{1/2} \int_x^{x+1/2} f''(\theta) d\theta dx.$$

For fixed $\theta \in (0, 1)$, the set of $x \in (0, 1/2)$ with $x < \theta < x + 1/2$ has length $\min\{1/2, \theta\} - \max\{0, \theta - 1/2\} = \min\{\theta, 1 - \theta\}$. By Tonelli's theorem, the right side equals $\int_0^1 \min\{\theta, 1 - \theta\} f''(\theta) d\theta$, and since $f'' \geq 0$, we may restrict this integral to any Borel subset of $(0, 1)$.

Now let γ_* be a minimizer, let (γ_θ) be a segment, and let $f(\theta) = F_\rho(\gamma_\theta)$. By Proposition 7.1(b), f has the properties just used. The midpoint $\gamma_{1/2}$ is admissible, which implies that $f(1/2) \geq F_\rho(\gamma_*)$. Inserting this into the previous inequality, we obtain

$$\int_{\Theta} \min\{\theta, 1 - \theta\} f''(\theta) d\theta \leq F_\rho(\gamma_0) + F_\rho(\gamma_1) - 2F_\rho(\gamma_*) \quad (\Theta \subset (0, 1) \text{ Borel}). \quad (7.9)$$

The weight $\min\{\theta, 1 - \theta\}$ has integral $1/4$ over $(0, 1)$ and integral $2 \int_{1/4}^{1/2} \theta d\theta = 3/16$ over $[1/4, 3/4]$.

Lemma 7.4 (Coercivity from a difference of one sign). *Let (γ_θ) be a segment, $\psi = \gamma_1 - \gamma_0$, and $f(\theta) = F_\rho(\gamma_\theta)$, and fix $\theta \in (0, 1)$. Let $\varphi : [0, 1] \rightarrow [0, 1]$ be Lipschitz, with $\varphi(0) = \varphi(1) = 0$, and not identically zero, and let*

$$\Delta = \rho \int_0^1 \zeta_\theta \varphi dq, \quad Y_0 = \int_0^1 |\varphi'|^2 dq.$$

Suppose that there exist $A > 0$, $d \geq 0$, and $\sigma \in \{-1, 1\}$ such that $\sigma\psi \geq A/2$ wherever $\varphi > 0$, and

$$\rho \int_{(0,1)} \varphi^2 \zeta_\theta d\gamma_\theta \leq d\Delta. \quad (7.10)$$

The last condition holds if the restriction of $d\gamma_\theta$ to the set $\{\varphi > 0\}$ has density at most d . Then

$$f''(\theta) \geq \rho \frac{A^2}{72} \min \left\{ \frac{\Delta^2}{\beta Y_0}, \frac{\Delta}{d} \right\}, \quad (7.11)$$

where $\Delta/0 = +\infty$ if $\Delta > 0$, and the right side is 0 if $\Delta = 0$.

For fixed Y_0 and d , both entries of the minimum in (7.11) are nondecreasing in Δ , so we may insert any lower bound for Δ into the right side.

Proof. We first check the sufficient condition for (7.10). If $d\gamma_\theta \leq d dq$ on the open set $\{\varphi > 0\}$, then, since $0 \leq \varphi \leq 1$ and $\zeta_\theta \geq 0$, we have $\rho \int_{(0,1)} \varphi^2 \zeta_\theta d\gamma_\theta \leq d\rho \int_0^1 \varphi^2 \zeta_\theta dq \leq d\Delta$.

Next, we extend (7.3) to φ , which need not have compact support in $(0, 1)$. For $0 < \epsilon < 1/4$, let χ_ϵ be the piecewise linear function that vanishes within distance ϵ of $\{0, 1\}$, equals one at distance at least 2ϵ , and is linear in between; note that $|\chi'_\epsilon| \leq 1/\epsilon$. Let ℓ_φ be the Lipschitz constant of φ and $\varphi_\epsilon = \chi_\epsilon \varphi$. Then φ_ϵ is Lipschitz with compact support in $(0, 1)$. Since $\varphi(0) = \varphi(1) = 0$, we have $|\varphi(q)| \leq \ell_\varphi \min\{q, 1 - q\}$, which is

at most $2\ell_\varphi\epsilon$ on the two strips where $\chi_\epsilon \neq 1$; these strips have total length 4ϵ . Then $\|\varphi_\epsilon - \varphi\|_\infty \leq 2\ell_\varphi\epsilon$, and since $\varphi'_\epsilon - \varphi' = \chi'_\epsilon\varphi + (\chi_\epsilon - 1)\varphi'$ vanishes outside the strips and is bounded by $3\ell_\varphi$ on them, we also have $\|\varphi'_\epsilon - \varphi'\|_{L^2}^2 \leq 36\ell_\varphi^2\epsilon$. At fixed N and θ , the functions ψ , b_θ , and ζ_θ are bounded by β , β , and β^2N , respectively: the first because both profiles take values in $[\lambda, \beta]$, the second by (7.1), and the third because $0 \preceq H_q \preceq C_q$ and $\text{Tr } C_q \leq N$. Further, $d\gamma_\theta$ is a finite measure. Using the uniform convergence of φ_ϵ and the L^2 convergence of φ'_ϵ , we find that the left side of (7.3) and Y converge, as $\epsilon \rightarrow 0$, to their values for φ . The term Z converges as well, because $\varphi_\epsilon^2 \rightarrow \varphi^2$ uniformly and the Stieltjes integral is taken against a finite measure; this includes the atoms of $d\gamma_\theta$. Passing to the limit, we see that (7.3) holds for φ .

If $\Delta = 0$, the claim follows from the convexity of f . Otherwise, we have $Y_0 > 0$, because φ vanishes at 0 and is not identically zero. In the notation of Proposition 7.1(b), we have $Y \leq \beta Y_0$ by (7.1) and $Z \leq d\Delta$ by (7.10). As ζ_θ and φ are nonnegative, the sign condition implies that $\sigma\rho \int_0^1 \psi\zeta_\theta\varphi dq \geq A\Delta/2$. Inserting this bound into (7.3), we find

$$\frac{A}{2}\Delta \leq \left(\frac{f''(\theta)}{\rho}\right)^{1/2} [Y^{1/2} + 2(Y+Z)^{1/2}] \leq 3\left(\frac{f''(\theta)}{\rho}\right)^{1/2} (Y+Z)^{1/2}.$$

The left side is positive, which implies that $Y+Z > 0$ and $f''(\theta) \geq \rho A^2 \Delta^2 / (36(Y+Z))$. Finally, we have $Y+Z \leq \beta Y_0 + d\Delta \leq 2\max\{\beta Y_0, d\Delta\}$, which gives (7.11). $\square$

7.3 A density bound for the scalar measure

By (7.6), the density ϱ_β of μ_β on $[0, q_\beta)$ extends continuously to $[0, q_\beta]$, and in particular it is bounded. The isolation argument uses this bound in the form of Lemma 7.5. We include a direct derivation of the lemma from the identity that expresses α_β on the support in terms of spatial derivatives of the Parisi solution. The same identity underlies the proof of [9, Theorem 2(ii)]. For the SK model, this identity is [48, Lemma 6.2, (6.1)], and the smoothness of the resulting quotient up to q_β is [48, Lemma 6.3]. In the physics literature, Crisanti and Rizzo gave the corresponding identities in their variables [32, Eqs. (22)–(23)].

Lemma 7.5 (Density bound). *There exists $L_\beta < \infty$ such that*

$$\alpha_\beta(y) - \alpha_\beta(x) \leq L_\beta(y - x) \quad (0 \leq x < y < q_\beta). \quad (7.12)$$

With $L = \max\{1, \beta L_\beta\}$, the trial profile satisfies $\gamma^{\text{tr}}(0) = \lambda$, $\gamma^{\text{tr}} = \beta$ on $[q_\beta, 1]$, and $|\gamma^{\text{tr}}(x) - \gamma^{\text{tr}}(y)| \leq L|x - y|$ for $x, y \in [0, q_\beta)$. In particular, the restriction of $d\gamma^{\text{tr}}$ to $(0, 1] \setminus \{q_\beta\}$ is absolutely continuous, with density at most L on $(0, q_\beta)$ and 0 on $(q_\beta, 1]$.

The bound (7.12) is stated for $y < q_\beta$ because of the atom of μ_β at q_β . The continuous function $\bar{\alpha}$ in the proof below equals α_β on $[0, q_\beta)$. At q_β , it equals the left limit $\mu_\beta([0, q_\beta)) = 1 - c_\beta$, while $\alpha_\beta(q_\beta) = 1$.

Proof. We write α_β as a quotient of two moments along the scalar diffusion, show that the two moments are Lipschitz and that the denominator is bounded below, and then deduce the consequences for γ^{tr} .

Let u be the solution of the scalar Parisi equation in (1.1) with $\alpha = \alpha_\beta$. The solution in [48] has the terminal value $\log \cosh x$ instead of $\log(2 \cosh x)$; since the equation is invariant under the addition of constants, the two solutions differ by $\log 2$, and their spatial derivatives coincide. Let w be a standard one-dimensional Brownian motion, and let ξ solve

$$d\xi_r = \beta^2 \alpha_\beta(r) u_x(r, \xi_r) dr + \beta dw_r, \quad \xi_0 = 0,$$

which is the optimal diffusion [48, (2.4)]. We write $\hat{C}_r = u_{xx}(r, \xi_r)$, $\hat{D}_r = u_{xxx}(r, \xi_r)$, and $\hat{A}_r = u_{xxxx}(r, \xi_r)$, and define

$$\Gamma(r) = \mathbb{E} u_x(r, \xi_r)^2, \quad \Lambda(r) = \mathbb{E} \hat{D}_r^2, \quad \Xi(r) = \mathbb{E} \hat{C}_r^3 \quad (0 \leq r \leq 1).$$

We use three facts from [48, Section 2].

- (i) [48, Proposition 2.1], which combines [9, Propositions 1, 2, and 5]: every spatial derivative of u of positive order is bounded on $[0, 1] \times \mathbb{R}$ by a constant depending only on β and the order, $|u_x| \leq 1$, and $0 < u_{xx} \leq 1$.
- (ii) [48, Lemma 2.6, (2.19)]: $\Gamma(r) = r$ for every $r \in \text{supp } \mu_\beta = [0, q_\beta]$.
- (iii) [48, Lemma 2.5, (2.13) and (2.14)]: the function Γ is continuously differentiable with $\Gamma' = \beta^2 \mathbb{E} \hat{C}^2$, and

$$\Gamma'(y) - \Gamma'(x) = \beta^4 \int_x^y (\Lambda(r) - 2\alpha_\beta(r) \Xi(r)) dr \quad (0 \leq x < y \leq 1).$$

Further, by the same lemma, every polynomial moment of the spatial derivatives along ξ is continuous on $[0, 1]$ and satisfies the integrated identity [48, (2.16)], with the polynomials of [48, (2.18)].

Let $K = \max_{j=3,4} \sup_{[0,1] \times \mathbb{R}} |\partial_x^j u|$, which is finite and depends only on β by (i).

We begin with the moments Λ and Ξ . For finitely supported measures, the equations [48, (2.17)] with $j = 2, 3$ read $d\hat{C}_r = -\beta^2 \alpha_\beta \hat{C}_r^2 dr + \beta \hat{D}_r dw_r$ and $d\hat{D}_r = -3\beta^2 \alpha_\beta \hat{C}_r \hat{D}_r dr + \beta \hat{A}_r dw_r$, and Itô's formula yields

$$\begin{aligned} d(\hat{D}_r^2) &= \beta^2 (\hat{A}_r^2 - 6\alpha_\beta \hat{C}_r \hat{D}_r^2) dr + 2\beta \hat{D}_r \hat{A}_r dw_r, \\ d(\hat{C}_r^3) &= 3\beta^2 (\hat{C}_r \hat{D}_r^2 - \alpha_\beta \hat{C}_r^4) dr + 3\beta \hat{C}_r^2 \hat{D}_r dw_r. \end{aligned}$$

The stochastic integrals are martingales, since their integrands are bounded by (i). For the measure μ_β , which has an atom, the identities we need are the cases $f = u_{xxx}^2$ and $f = u_{xx}^3$ of [48, (2.16)], whose polynomials f_0, f_1 from [48, (2.18)] are the drifts just computed:

$$\begin{aligned} \Lambda(y) - \Lambda(x) &= \beta^2 \int_x^y \mathbb{E} [\hat{A}_r^2 - 6\alpha_\beta(r) \hat{C}_r \hat{D}_r^2] dr, \\ \Xi(y) - \Xi(x) &= 3\beta^2 \int_x^y \mathbb{E} [\hat{C}_r \hat{D}_r^2 - \alpha_\beta(r) \hat{C}_r^4] dr, \end{aligned} \quad (0 \leq x < y \leq 1).$$

Since $0 \leq \alpha_\beta \leq 1$, $0 < \hat{C} \leq 1$, and $|\hat{D}|, |\hat{A}| \leq K$, the integrands are bounded by $7K^2$ and $K^2 + 1$, respectively. We find that Λ and Ξ are Lipschitz on $[0, 1]$, with constants $7\beta^2 K^2$ and $3\beta^2(K^2 + 1)$, and $0 \leq \Lambda \leq K^2$. The function Ξ is continuous by (iii) and positive, since $\hat{C}_r > 0$ by (i). We define

$$\Xi_0 = \min_{0 \leq r \leq 1} \Xi(r) > 0,$$

which depends only on β , because μ_β does. (As in the proof of [48, Lemma 6.2], positivity of Ξ comes from $u_{xx} > 0$.)

The quotient $\bar{\alpha} = \Lambda/(2\Xi)$ is Lipschitz on $[0, 1]$. Indeed, let

$$L_\beta = \frac{7\beta^2 K^2}{2\Xi_0} + \frac{3\beta^2 K^2(K^2 + 1)}{2\Xi_0^2}.$$

For $x, y \in [0, 1]$, using the Lipschitz bounds for Λ and Ξ together with $\Lambda \leq K^2$ and $\Xi \geq \Xi_0$, we obtain

$$|\bar{\alpha}(y) - \bar{\alpha}(x)| \leq \frac{|\Lambda(y) - \Lambda(x)|}{2\Xi(y)} + \Lambda(x) \frac{|\Xi(x) - \Xi(y)|}{2\Xi(x)\Xi(y)} \leq L_\beta |y - x|.$$

Next, we identify α_β with $\bar{\alpha}$ on $[0, q_\beta)$. By (ii), we have $\Gamma(r) = r$ on $[0, q_\beta]$, and since Γ is continuously differentiable, $\Gamma' = 1$ on $(0, q_\beta)$. Then (iii) implies that $\int_x^y (\Lambda - 2\alpha_\beta \Xi) dr = 0$ for $0 < x < y < q_\beta$. The integrand is bounded, and by Lebesgue's differentiation theorem, we have $\Lambda - 2\alpha_\beta \Xi = 0$, that is, $\alpha_\beta = \bar{\alpha}$, almost everywhere on $(0, q_\beta)$. Let $r \in [0, q_\beta)$, and choose points $r_n \downarrow r$ in (r, q_β) at which $\alpha_\beta(r_n) = \bar{\alpha}(r_n)$. As $n \rightarrow \infty$, the left side tends to $\alpha_\beta(r)$ by the right continuity of α_β , and the right side tends to $\bar{\alpha}(r)$ by continuity. We conclude that $\alpha_\beta = \bar{\alpha}$ on $[0, q_\beta)$, and (7.12) follows from the Lipschitz bound for $\bar{\alpha}$.

For the consequences, we have $\gamma^{\text{tr}}(0) = \max\{\lambda, 0\} = \lambda$ by (7.5), and $\gamma^{\text{tr}} = \max\{\lambda, \beta\} = \beta$ on $[q_\beta, 1]$ by (7.6). On $[0, q_\beta)$, the function $\gamma_\beta = \beta\alpha_\beta$ is βL_β -Lipschitz, and so is its maximum with the constant λ ; since $L \geq \beta L_\beta$, this gives the Lipschitz bound for γ^{tr} . A nondecreasing Lipschitz function has an absolutely continuous derivative measure, with density bounded by its Lipschitz constant. This yields the density bound on $(0, q_\beta)$, and on $(q_\beta, 1]$ the profile γ^{tr} is constant. $\square$

7.4 Distance from the scalar trial profile

We now compare a minimizer with γ^{tr} in $L^1(0, 1)$. After the regularization of Lemma 7.3(b), the derivative measures along the segment to γ^{tr} have bounded density in $(0, 1)$ away from the atom of γ^{tr} at q_β . We construct cutoffs of one sign from one-sided averages of superlevel sets, and their masses and Dirichlet energies lead to a single estimate for the size of every superlevel set of the difference of the two profiles.

Proposition 7.6 (Isolation against the trial). *There exists $C = C(\beta) < \infty$ such that, for all $N \geq 1$, $0 < \rho \leq 1$, $0 < \lambda \leq \lambda_0$, every minimizer γ_* of F_ρ , and every trial budget η for γ_* ,*

$$\|\gamma_* - \gamma^{\text{tr}}\|_{L^1(0,1)} \leq C \min\{1, \rho^{-3/4} \eta^{1/4}\}. \quad (7.13)$$

The constant depends on β only through β , k_β , and the constant L of Lemma 7.5.

Proof. The proof has seven steps. We first fix a smoothing width h and replace γ_* by its regularization γ_*^h of Lemma 7.3(b), whose free measure has no atoms and which costs at most $\rho\beta h^2/12$. Along the segment from γ_*^h to γ^{tr} , the difference $\psi = \gamma^{\text{tr}} - \gamma_*^h$ increases at rate at most L below q_β , and it is nonnegative and nonincreasing above q_β . For each level $A > 0$, we construct cutoffs of one sign from one-sided averages of the superlevel sets of ψ below q_β , and we use triangular cutoffs on the two intervals next to q_β ; none of them is positive at the atom q_β of γ^{tr} . Lemma 7.4 bounds the size of each superlevel set in terms of $f''(\theta)$, and the layer-cake formula turns these bounds into a lower bound for $f''(\theta)$ in terms of the L^1 distance. Finally, we choose h proportional to ρ times the square of this distance and compare with the midpoint inequality and the trial budget.

Step 1: Setup and constraints on h . Let $D = \|\gamma_* - \gamma^{\text{tr}}\|_{L^1(0,1)}$. If $D = 0$, there is nothing to prove, so we assume from now on that $D > 0$. Both profiles take values in $[\lambda, \beta]$, which implies that $D \leq \beta$. With $L \geq 1$ as in Lemma 7.5, we fix a width h such that

$$0 < h \leq \frac{1}{2}, \quad h \leq \frac{D}{2\beta}, \quad h \leq \frac{\beta}{L}. \quad (7.14)$$

The final choice, in Step 7, will be a small multiple of ρD^2 .

Step 2: The regularized segment. Let γ_*^h be the regularization of γ_* given by Lemma 7.3(b), and define

$$\gamma_\theta = (1 - \theta)\gamma_*^h + \theta\gamma^{\text{tr}}, \quad \psi = \gamma^{\text{tr}} - \gamma_*^h, \quad f(\theta) = F_\rho(\gamma_\theta) \quad (0 \leq \theta \leq 1).$$

The regularization is defined because $h \leq 1/2$. Recall that the trial profile is admissible (Section 3), so that (γ_θ) is a segment. By Lemma 7.3(b) and the constraint $h \leq D/(2\beta)$, we have

$$\bar{D} = \|\psi\|_{L^1(0,1)} \geq D - \|\gamma_*^h - \gamma_*\|_{L^1(0,1)} \geq D - \beta h \geq \frac{D}{2}.$$

For $\theta \in (0, 1)$, the measure $d\gamma_\theta$ is $(1 - \theta) d\gamma_*^h + \theta d\gamma^{\text{tr}}$. On $(0, 1) \setminus \{q_\beta\}$, the measure $d\gamma_*^h$ has density at most β/h by Lemma 7.3(b), and $d\gamma^{\text{tr}}$ has density at most L by Lemma 7.5. By (7.14), we have $L \leq \beta/h$, and the restriction of $d\gamma_\theta$ to $(0, 1) \setminus \{q_\beta\}$ has density at most $\beta/h + L \leq 2\beta/h$. The only atom of $d\gamma_\theta$ in $(0, 1)$ is the atom of $d\gamma^{\text{tr}}$ at q_β . Further, $\zeta_\theta \geq k_\beta^2$ by (7.4).

Step 3: One-sided bounds on ψ . By Lemma 7.3(b) and Lemma 7.5, we have $\gamma_*^h(0) = \gamma^{\text{tr}}(0) = \lambda$. For $0 \leq z < x < q_\beta$, we have $\gamma^{\text{tr}}(x) - \gamma^{\text{tr}}(z) \leq L(x - z)$ by Lemma 7.5, and $\gamma_*^h(x) - \gamma_*^h(z) \geq 0$, because γ_*^h is nondecreasing. We obtain

$$\psi(0) = 0, \quad \psi(x) - \psi(z) \leq L(x - z), \quad \psi(x) \leq Lx \quad (0 \leq z < x < q_\beta). \quad (7.15)$$

On $[q_\beta, 1]$, we have $\psi = \beta - \gamma_*^h$, which is continuous (the free measure of γ_*^h has no atoms), nonnegative, and nonincreasing. Everywhere on $[0, 1]$, both profiles take values in $[\lambda, \beta]$, so $|\psi| \leq \beta$.

Step 4: Cutoffs for the superlevel sets. We fix $\theta \in (0, 1)$. For $A > 0$, we set $r = A/(4L)$ and define the sets

$$\begin{aligned} U_A &= \{x \in (0, q_\beta) : \psi(x) > A\}, \\ V_A^i &= \{x \in (0, q_\beta - r) : \psi(x) < -A\}, \\ V_A^b &= \{x \in [\max\{0, q_\beta - r\}, q_\beta) : \psi(x) < -A\}, \\ W_A &= \{x \in [q_\beta, 1] : \psi(x) > A\}, \end{aligned}$$

where an interval of nonpositive length is empty. These four sets are disjoint, and since $\psi(0) = 0$ and $\psi \geq 0$ on $[q_\beta, 1]$, their union is $\{x \in [0, 1] : |\psi(x)| > A\}$. In this step, we treat U_A and V_A^i ; Step 5 treats V_A^b and W_A .

For a Borel set $E \subset (0, 1)$, we extend its indicator by zero to $\mathbb{R}$ and define

$$\mathcal{A}_+ \mathbf{1}_E(x) = \frac{1}{r} \int_0^r \mathbf{1}_E(x + u) du, \quad \mathcal{A}_- \mathbf{1}_E(x) = \frac{1}{r} \int_0^r \mathbf{1}_E(x - u) du \quad (x \in \mathbb{R}).$$

With $\Phi_E(x) = |E \cap (-\infty, x]|$, we have $\mathcal{A}_+ \mathbf{1}_E(x) = (\Phi_E(x + r) - \Phi_E(x))/r$ and $\mathcal{A}_- \mathbf{1}_E(x) = (\Phi_E(x) - \Phi_E(x - r))/r$. The function Φ_E is 1-Lipschitz, and by Lebesgue's differentiation theorem, its derivative is $\mathbf{1}_E$ almost everywhere. It follows that $\mathcal{A}_\pm \mathbf{1}_E$ is Lipschitz, with derivative $(\mathbf{1}_E(x + r) - \mathbf{1}_E(x))/r$, respectively $(\mathbf{1}_E(x) - \mathbf{1}_E(x - r))/r$, for almost every x . Using Fubini's theorem for the integral and $|(E - r) \triangle E| \leq 2|E|$ for the energy, we obtain

$$0 \leq \mathcal{A}_\pm \mathbf{1}_E \leq 1, \quad \int_{\mathbb{R}} \mathcal{A}_\pm \mathbf{1}_E = |E|, \quad \int_{\mathbb{R}} |\partial_x \mathcal{A}_\pm \mathbf{1}_E|^2 \leq \frac{2|E|}{r^2}. \quad (7.16)$$

For $E = U_A$, we take $\varphi = \mathcal{A}_+ \mathbf{1}_E$. If $\varphi(x) > 0$, then there exists $y \in U_A$ with $x \leq y \leq x + r$. By (7.15), we have $A < \psi(y) \leq Ly$, so that $y > A/L = 4r$ and $x \geq y - r > 3r > 0$; also $x \leq y < q_\beta$. Then (7.15) implies that $\psi(x) \geq \psi(y) - L(y - x) > A - Lr = 3A/4$. We have shown that $\{\varphi > 0\} \subset (0, q_\beta)$ and that $\psi > 3A/4$ on this set.

For $E = V_A^i$, we take $\varphi = \mathcal{A}_- \mathbf{1}_E$. If $\varphi(x) > 0$, then there exists $y \in V_A^i$ with $y \leq x \leq y + r$, and $y + r < q_\beta$ by the definition of V_A^i . Then $0 < x < q_\beta$, and (7.15) implies that $\psi(x) \leq \psi(y) + L(x - y) < -A + Lr = -3A/4$. Again, we have $\{\varphi > 0\} \subset (0, q_\beta)$, and $\psi < -3A/4$ on this set.

In both cases, φ is a Lipschitz function on $[0, 1]$ with values in $[0, 1]$, and it vanishes at 0 and on $[q_\beta, 1]$. By (7.16), its integral over $[0, 1]$ is $|E|$ and its energy over $[0, 1]$ is at most $2|E|/r^2$. The set $\{\varphi > 0\}$ does not contain q_β , and by Step 2, the restriction of $d\gamma_\theta$ to this set has density at most $d = 2\beta/h$. If $|E| > 0$, then φ is not identically zero, and Lemma 7.4 applies with $\sigma = 1$ for U_A and $\sigma = -1$ for V_A^i . Its quantities satisfy

$$\Delta = \rho \int_0^1 \zeta_\theta \varphi \, dq \geq \rho k_\beta^2 |E|, \quad Y_0 \leq \frac{2|E|}{r^2}, \quad d = \frac{2\beta}{h},$$

by (7.4) and (7.16). Inserting these bounds into (7.11) and using $r = A/(4L)$, we obtain

$$f''(\theta) \geq \frac{\rho^2 A^2 |E|}{72} \min \left\{ \frac{\rho k_\beta^4 A^2}{32\beta L^2}, \frac{k_\beta^2 h}{2\beta} \right\} \quad (E = U_A, V_A^i). \quad (7.17)$$

When $|E| = 0$, the right side is zero, and the bound holds by convexity.

Step 5: The boundary sets. The set V_A^b lies within distance r of q_β , where an average of $\mathbf{1}_{V_A^b}$ over a window of length r could be positive at q_β . We use a single triangular cutoff instead. For an interval $(x_0, x_0 + z)$ with $z > 0$, let

$$T_{x_0, z}(x) = \left(1 - \frac{4|x - x_0 - z/2|}{z}\right)_+.$$

This function takes values in $[0, 1]$, is positive exactly on the middle half $(x_0 + z/4, x_0 + 3z/4)$, and satisfies

$$\int_{\mathbb{R}} T_{x_0, z} = \frac{z}{4}, \quad \int_{\mathbb{R}} |T'_{x_0, z}|^2 = \frac{8}{z}.$$

Suppose that V_A^b is nonempty, and let $x_A = \inf V_A^b$ and $l = q_\beta - x_A$. Note that $0 < l \leq r$ and $|V_A^b| \leq l$. For every $z \in (x_A, q_\beta)$, there exists $x \in V_A^b$ with $x < z$, and (7.15) implies that $\psi(z) \leq \psi(x) + L(z - x) < -A + Lr = -3A/4$. We take $\varphi = T_{x_A, l}$. Its positivity set is the middle half of (x_A, q_β) , whose closure lies in $(0, q_\beta)$; on this set, $\psi < -3A/4$ and $d\gamma_\theta$ has density at most $2\beta/h$. Lemma 7.4, applied with $\sigma = -1$, $\Delta \geq \rho k_\beta^2 l/4$, $Y_0 = 8/l$, and $d = 2\beta/h$, yields

$$f''(\theta) \geq \frac{\rho^2 A^2 l}{72} \min \left\{ \frac{\rho k_\beta^4 l^2}{128\beta}, \frac{k_\beta^2 h}{8\beta} \right\}. \quad (7.18)$$

Indeed, $\Delta^2/(\beta Y_0) \geq \rho^2 k_\beta^4 l^3/(128\beta)$ and $\Delta/d \geq \rho k_\beta^2 l h/(8\beta)$.

By Step 3, the function ψ is continuous, nonnegative, and nonincreasing on $[q_\beta, 1]$. Hence W_A is an interval $[q_\beta, q_\beta + w)$ of length $w \leq 1$. If $w > 0$, we take $\varphi = T_{q_\beta, w}$, whose positivity set has closure in $(q_\beta, 1)$. On this set, $\psi > A$, the measure $d\gamma^{\text{tr}}$ vanishes, and $d\gamma_\theta$ has density at most $2\beta/h$. The same computation, with $\sigma = 1$, gives (7.18) with w in place of l . If $w = 0$, there is nothing to prove.

Step 6: Layer-cake integration. Let

$$c_1 = \frac{1}{72} \min \left\{ \frac{k_\beta^4}{32\beta L^2}, \frac{k_\beta^4}{128\beta}, \frac{k_\beta^2}{8\beta} \right\}, \quad K = \frac{f''(\theta)}{c_1 \rho^2}.$$

Since $\min\{ax, by\} \geq \min\{a, b\} \min\{x, y\}$ for nonnegative numbers, (7.17) and (7.18) imply that

$$A^2|E| \min\{\rho A^2, h\} \leq K \quad (E = U_A, V_A^i), \quad A^2 z \min\{\rho z^2, h\} \leq K \quad (z = l, w).$$

For $E = U_A$ or $E = V_A^i$, one of $|E| \leq K/(\rho A^4)$ and $|E| \leq K/(hA^2)$ holds, and $|E| \leq 1$. In either case, we have

$$|E| \leq \min\{1, K/(\rho A^4)\} + \min\{1, K/(hA^2)\}.$$

For the boundary interval, one of $l \leq (K/(\rho A^2))^{1/3}$ and $l \leq K/(hA^2)$ holds. Combining these alternatives with $|V_A^b| \leq l \leq r = A/(4L)$, we obtain

$$|V_A^b| \leq \left(\frac{K}{\rho A^2}\right)^{1/3} + \min\left\{\frac{A}{4L}, \frac{K}{hA^2}\right\}.$$

The same reasoning with $w \leq 1$ yields

$$|W_A| \leq \left(\frac{K}{\rho A^2}\right)^{1/3} + \min\left\{1, \frac{K}{hA^2}\right\}.$$

All four sets are empty for $A \geq \beta$, since $|\psi| \leq \beta$. We use the elementary integrals

$$\begin{aligned} \int_0^\infty \min\left\{1, \frac{x}{A^4}\right\} dA &= \frac{4}{3} x^{1/4}, & \int_0^\infty \min\left\{1, \frac{x}{A^2}\right\} dA &= 2x^{1/2}, \\ \int_0^\infty \min\left\{\frac{A}{4L}, \frac{x}{A^2}\right\} dA &= C_L x^{2/3}, & \int_0^\beta \left(\frac{x}{A^2}\right)^{1/3} dA &= 3\beta^{1/3} x^{1/3}, \end{aligned}$$

where $C_L = 3/(2(4L)^{1/3})$ and $x \geq 0$. The case $x = 0$ is immediate. For $x > 0$, the first three follow by splitting at the point where the two entries of the minimum agree; for the third, this point is $A_0 = (4Lx)^{1/3}$, and the integral is $A_0^2/(8L) + x/A_0 = 3x/(2A_0)$. By the layer-cake formula, we have $\bar{D} = \int_0^\beta |\{|\psi| > A\}| dA$. Integrating the bounds on the four sets over $0 < A < \beta$, we find

$$\bar{D} \leq \frac{8}{3} \left(\frac{K}{\rho}\right)^{1/4} + 6 \left(\frac{K}{h}\right)^{1/2} + 6\beta^{1/3} \left(\frac{K}{\rho}\right)^{1/3} + C_L \left(\frac{K}{h}\right)^{2/3}. \quad (7.19)$$

Let $C_w = \max\{6\beta^{1/3}, C_L\}$. Each coefficient in (7.19) is at most C_w , because $\beta > 1$.

We now show that $K \geq c_2 \min\{\rho D^4, hD^2\}$, where $c_2 = \min\{(8C_w)^{-4}, \beta^{-4}\}$. Suppose first that $K \leq \rho$ and $K \leq h$. Then $(K/\rho)^{1/3} \leq (K/\rho)^{1/4}$ and $(K/h)^{2/3} \leq (K/h)^{1/2}$, and

(7.19) gives $\bar{D} \leq 2C_w[(K/\rho)^{1/4} + (K/h)^{1/2}]$. Using $\bar{D} \geq D/2$, we see that one of the two terms in the bracket is at least $D/(8C_w)$. This implies that $K \geq \rho D^4/(8C_w)^4$ or $K \geq hD^2/(8C_w)^2$, and since $8C_w \geq 1$, in both cases $K \geq (8C_w)^{-4} \min\{\rho D^4, hD^2\}$. If $K > \rho$, then $D \leq \beta$ gives $\min\{\rho D^4, hD^2\} \leq \rho \beta^4 < \beta^4 K$. If $K > h$, then $\min\{\rho D^4, hD^2\} \leq h\beta^2 < \beta^4 K$, since $\beta > 1$. This proves the claim in all cases. Since $\theta \in (0, 1)$ was arbitrary, we have proved that

$$f''(\theta) \geq c_3 \rho^2 \min\{\rho D^4, hD^2\} \quad (0 < \theta < 1), \quad c_3 = c_1 c_2. \quad (7.20)$$

Step 7: Choice of h and conclusion. We now take $h = c_4 \rho D^2$, where

$$c_4 = \min\left\{1, \frac{1}{2\beta^2}, \frac{1}{\beta L}, \frac{3c_3}{2\beta}\right\}.$$

Since $D \leq \beta$ and $\rho \leq 1$, we have $h \leq c_4 \beta^2 \leq 1/2$, $h \leq c_4 \beta D \leq D/(2\beta)$, and $h \leq c_4 \beta^2 \leq \beta/L$. These are the constraints (7.14). Because $c_4 \leq 1$, the minimum in (7.20) equals $c_4 \rho D^4$, and (7.20) gives $f'' \geq c_3 c_4 \rho^3 D^4$ on $(0, 1)$. We use the midpoint inequality (7.9) with $\Theta = (0, 1)$, where the weight has integral $1/4$. Together with the regularization cost of Lemma 7.3(b) and the trial budget (7.7), it yields

$$\frac{c_3 c_4}{4} \rho^3 D^4 \leq F_\rho(\gamma_*^h) + F_\rho(\gamma^{\text{tr}}) - 2F_\rho(\gamma_*) \leq \frac{\rho \beta h^2}{12} + \eta \leq \frac{c_3 c_4}{8} \rho^3 D^4 + \eta.$$

In the last inequality, we used that $\rho \beta h^2/12 = (\beta c_4^2/12) \rho^3 D^4$ and $c_4 \leq 3c_3/(2\beta)$. We conclude that $\eta \geq (c_3 c_4/8) \rho^3 D^4$, that is, $D \leq (8/(c_3 c_4))^{1/4} \rho^{-3/4} \eta^{1/4}$. Together with $D \leq \beta$, this implies (7.13) with $C = \max\{\beta, (8/(c_3 c_4))^{1/4}\}$. The constants c_1, c_2, c_3, c_4 depend only on β, k_β , and L . $\square$

Localization of the lowest quantiles

In this section, we show that the lowest quantiles of the free mass of every minimizer of F_ρ lie close to the origin. Let $\mu = s\delta_0 + \nu$ be a minimizing profile and η a trial budget. By Corollary 8.4, the quantile $Q(v)$ of ν at level v is at most $C_Q(s + v + \omega(\rho, \eta))$, where $\omega(\rho, \eta) = \min\{1, (\eta/\rho)^{1/6}\}$ and C_Q depends only on β . Proposition 8.3 is the same estimate in the energy normalization. In Proposition 13.1, we use Corollary 8.4 to bound the self-overlap variance on dyadic quantile shells for the interpolation parameters near $\rho = 0$.

We keep the standing assumptions and the notation of Section 7.1. From Section 7, we use the convexity and the first variation of F_ρ (Proposition 7.1(a)), the identities (7.1), the coercivity bound of Lemma 7.4, and the midpoint inequality (7.9). We also use the partial regularization of Lemma 7.3(a), the scalar root bounds (7.5), the structure (7.6) of μ_β , and the trial budget (7.7). The isolation estimate Proposition 7.6 and the density bound Lemma 7.5 are not used here.

The main tool is a local deficit certificate (Lemma 8.1). It combines a variation that moves mass of a profile toward the middle of an interval with the response inequality (7.3). If the variation violates its room or sign condition, this failure gives a lower bound on η/ρ as well, and the certificate includes this bound as an alternative. In Section 8.2, we follow the strategy of the proof of [4, Lemma 7.6]. We regularize only the free mass below the selected quantile. Then we compare the result with the scalar trial profile on an interval where the latter has a positive density (Lemma 8.2). The certificate gives the dependence on η/ρ in (8.7).

8.1 A local deficit certificate

In the proof of Proposition 7.6, we used the uniform lower bound (7.4) on ζ , which gives $\Delta = \rho \int \zeta \varphi \geq \rho k_\beta^2 \int \varphi$. For root localization, we need Δ to be at least a constant multiple of the length of the interval on which φ lives, without the factor ρ . In the next lemma, we achieve this by moving mass of the profile toward the middle of the interval. If the variation does not have enough room, or if it would destroy the profile gap, then one of the lower bounds in the conclusion already holds.

Lemma 8.1 (Local deficit certificate). *There exists $c_{\text{cert}} = c_{\text{cert}}(\beta) > 0$ with the following property. Let $I \subseteq (0, 1)$ be an open interval of length ℓ , let γ_0, γ_1 be admissible profiles, and let $\eta > 0$ be such that $F_\rho(\gamma_i) \leq \min F_\rho + \eta$ for $i = 0, 1$. Suppose that there exist $k > 0$, $d_0, d_1 \geq 0$, and $A > 0$ such that, on I , the derivative measures and the profile gap satisfy*

$$d\gamma_0 \leq d_0 dx, \quad k dx \leq d\gamma_1 \leq d_1 dx, \quad \gamma_1 - \gamma_0 \geq A.$$

Then

$$\frac{\eta}{\rho} \geq c_{\text{cert}} \min \left\{ k\ell^3, A\ell^2, A^2\ell^4, \frac{A^2\ell^2}{d_0 + d_1 + k} \right\}. \quad (8.1)$$

Proof. The proof has four steps. In Step 1, we fix a bump function φ on I and an amplitude τ . If τ is too large for the variation below, then one of the first two alternatives in (8.1) holds. Otherwise, we perturb the segment from γ_0 to γ_1 by the derivative of the bump, which moves mass of the profile toward the middle of I (Step 2). Since both endpoints are nearly minimal and F_ρ is convex, the response integrated against the bump is at least of order ℓ along the perturbed segment (Step 3). Then Lemma 7.4 implies a lower bound on the curvature along the perturbed segment. Comparing this curvature with the excesses at its endpoints, we obtain the last two alternatives (Step 4). Throughout, we write $F_{\min} = \min F_\rho$ and $I = (x_0, x_0 + \ell)$.

Step 1: The bump and the room condition. Let $B(y) = 16y^2(1 - y)^2$ for $y \in [0, 1]$, and let $\varphi(x) = B((x - x_0)/\ell)$ for $x \in I$ and $\varphi(x) = 0$ for $x \notin I$. We have

$$B'(y) = 32y(1 - y)(1 - 2y), \quad B''(y) = 32 - 192y + 192y^2,$$

and B and B' vanish at 0 and 1. A direct computation shows that $\int_0^1 B = 8/15$, $\max |B'| = 16\sqrt{3}/9 < 4$, $\max |B''| = 32$, and $\int_0^1 |B'|^2 = 512/105$. These properties of B imply that φ is continuously differentiable on $\mathbb{R}$ with values in $[0, 1]$, that φ' is Lipschitz, that φ and φ' vanish at the endpoints of I , and that

$$\int \varphi = \frac{8\ell}{15}, \quad \|\varphi'\|_\infty \leq \frac{4}{\ell}, \quad \|\varphi''\|_\infty \leq \frac{32}{\ell^2}, \quad \int |\varphi'|^2 = \frac{512}{105\ell}, \quad (8.2)$$

where φ'' is defined away from the endpoints of I . In particular, the distributional derivative of φ' is the function φ'' , which has no atoms at the endpoints of I , and $\int \varphi'' = 0$.

We define $\tau = 30\eta/(\rho\ell)$; the factor 30 is chosen so that $8\eta/(\rho\tau) = 4\ell/15$ is half of $\int \varphi$. Suppose first that the room condition

$$\tau \leq \min \left\{ \frac{k\ell^2}{64}, \frac{A\ell}{8} \right\} \quad (8.3)$$

fails. Then

$$\frac{\eta}{\rho} = \frac{\ell\tau}{30} > \min \left\{ \frac{k\ell^3}{1920}, \frac{A\ell^2}{240} \right\} \geq \frac{1}{1920} \min \{k\ell^3, A\ell^2\},$$

and (8.1) holds whenever $c_{\text{cert}} \leq 1/1920$. We assume from now on that (8.3) holds.

Step 2: The perturbed segment. For $0 \leq \theta \leq 1$, let

$$\gamma_\theta = (1 - \theta)\gamma_0 + \theta\gamma_1, \quad \tilde{\gamma}_\theta = \gamma_\theta - \theta\tau\varphi'.$$

We claim that $\gamma_\theta - u\varphi'$ is admissible for $0 \leq \theta \leq 1$ and $0 \leq u \leq \theta\tau$. The function φ' is continuous and vanishes outside I , and $\bar{I} \subset (0, 1)$. Then $\gamma_\theta - u\varphi'$ is right-continuous and agrees with γ_θ near 0 and near 1. Its derivative measure is $d\gamma_\theta - u\varphi'' dx$. Off I , this is $d\gamma_\theta$. On I , the hypotheses give $d\gamma_\theta \geq \theta k dx$, and by (8.2) and (8.3), we have

$$d(\gamma_\theta - u\varphi') \geq \left(\theta k - \frac{32u}{\ell^2}\right) dx \geq \left(\theta k - \frac{32\theta\tau}{\ell^2}\right) dx \geq \frac{\theta k}{2} dx.$$

The signed measure $\varphi'' dx$ has total mass zero and does not charge 0. We conclude that the derivative measure of $\gamma_\theta - u\varphi'$ is nonnegative, has total mass β , and has the same atom at 0 as $d\gamma_\theta$; that is, $\gamma_\theta - u\varphi'$ is admissible. We have $\tilde{\gamma}_\theta = (1 - \theta)\gamma_0 + \theta(\gamma_1 - \tau\varphi')$, and $\gamma_1 - \tau\varphi'$ is admissible (this is the claim with $\theta = 1$ and $u = \tau$). It follows that $(\tilde{\gamma}_\theta)$ is a segment, with endpoints $\tilde{\gamma}_0 = \gamma_0$ and $\tilde{\gamma}_1 = \gamma_1 - \tau\varphi'$.

Step 3: The response integrated against the bump. Fix $\theta \in (0, 1]$ and define $g_\theta(u) = F_\rho(\gamma_\theta - u\varphi')$ for $0 \leq u \leq \theta\tau$. The profiles $\gamma_\theta - u\varphi'$, $0 \leq u \leq \theta\tau$, form a segment up to a linear change of parameter, so g_θ is convex and continuous by Proposition 7.1. We compute its one-sided derivatives. Let $0 \leq u \leq \theta\tau$, and write $\gamma' = \gamma_\theta - u\varphi'$, $G = G_{\gamma'}$, and $S = S_{\gamma'}$. Increasing u changes the free measure of γ' in the direction $-\varphi'' dx$, and decreasing u changes it in the direction $+\varphi'' dx$; in both cases, the new free measure is that of an admissible profile on the segment. By (7.2), the right derivative of g_θ at $u < \theta\tau$ and the left derivative at $u > 0$ are both equal to $-(\rho/2) \int G\varphi'' dx$. We integrate by parts twice. The function φ' vanishes at the endpoints of I and $G'(x) = x - S(x)$, which gives $\int G\varphi'' dx = -\int (x - S)\varphi' dx$. The function φ also vanishes at the endpoints of I , and S is absolutely continuous by (7.1); integrating by parts once more, we find

$$-\int (x - S)\varphi' dx = \int \varphi d(x - S) = \frac{8\ell}{15} - \int \varphi dS.$$

We obtain

$$g'_\theta(u) = \frac{\rho}{2} \left(\int \varphi dS_{\gamma_\theta - u\varphi'} - \frac{8\ell}{15} \right), \quad (8.4)$$

for the right derivative at $0 \leq u < \theta\tau$ and for the left derivative at $0 < u \leq \theta\tau$.

By the convexity of F_ρ along the segment (γ_θ) , we have $g_\theta(0) = F_\rho(\gamma_\theta) \leq (1 - \theta)F_\rho(\gamma_0) + \theta F_\rho(\gamma_1) \leq F_{\min} + \eta$, and $g_\theta(\theta\tau) \geq F_{\min}$. For a convex function, the left derivative at the right endpoint of an interval is at least the slope of the secant over the interval, which gives

$$g'_\theta(\theta\tau) \geq \frac{g_\theta(\theta\tau) - g_\theta(0)}{\theta\tau} \geq -\frac{\eta}{\theta\tau}.$$

Inserting (8.4) at $u = \theta\tau$, where $\gamma_\theta - \theta\tau\varphi' = \tilde{\gamma}_\theta$, we find, for $1/4 \leq \theta \leq 3/4$,

$$\Delta_\theta = \int \varphi \, dS_{\tilde{\gamma}_\theta} \geq \frac{8\ell}{15} - \frac{2\eta}{\rho\theta\tau} \geq \frac{8\ell}{15} - \frac{8\eta}{\rho\tau} = \frac{4\ell}{15}.$$

By (7.1), we have $dS_{\tilde{\gamma}_\theta} = \rho\zeta_{\tilde{\gamma}_\theta} dx$. In other words, $\Delta_\theta = \rho \int \zeta_{\tilde{\gamma}_\theta} \varphi \, dx$ is the quantity Δ of Lemma 7.4 for the segment $(\tilde{\gamma}_\theta)$ and the cutoff φ .

Step 4: Curvature and the endpoint excesses. By (8.2) and (8.3), the direction of the segment $(\tilde{\gamma}_\theta)$ satisfies

$$\tilde{\gamma}_1 - \tilde{\gamma}_0 = \gamma_1 - \gamma_0 - \tau\varphi' \geq A - \frac{4\tau}{\ell} \geq \frac{A}{2} \quad \text{on } I = \{\varphi > 0\}.$$

On I , the derivative measure of $\tilde{\gamma}_\theta$ is at most $((1-\theta)d_0 + \theta d_1 + 32\theta\tau/\ell^2) dx \leq \bar{d} dx$, where

$$\bar{d} = d_0 + d_1 + \frac{32\tau}{\ell^2} \leq d_0 + d_1 + \frac{k}{2};$$

note that $\bar{d} \geq d_1 \geq k > 0$. Let $\tilde{f}(\theta) = F_\rho(\tilde{\gamma}_\theta)$. By Lemma 7.4, used for the segment $(\tilde{\gamma}_\theta)$ with the cutoff φ , $\sigma = 1$, the level A , and $d = \bar{d}$, together with $\Delta_\theta \geq 4\ell/15$ and $Y_0 = 512/(105\ell)$ from (8.2), we obtain, for $1/4 \leq \theta \leq 3/4$,

$$\tilde{f}''(\theta) \geq \frac{\rho A^2}{72} \min \left\{ \left(\frac{4\ell}{15} \right)^2 \frac{105\ell}{512\beta}, \frac{4\ell}{15\bar{d}} \right\} = \frac{\rho A^2}{72} \min \left\{ \frac{7\ell^3}{480\beta}, \frac{4\ell}{15\bar{d}} \right\}.$$

Because $\beta > 1$, we have $4/15 \geq 7/(480\beta)$, which gives

$$\frac{d^2}{d\theta^2} F_\rho(\tilde{\gamma}_\theta) \geq \frac{7}{34560\beta} \rho A^2 \min \left\{ \ell^3, \frac{\ell}{\bar{d}} \right\} \quad (1/4 \leq \theta \leq 3/4). \quad (8.5)$$

It remains to bound the excesses at the endpoints of the perturbed segment. At $\theta = 0$, we have $\tilde{f}(0) = F_\rho(\gamma_0) \leq F_{\min} + \eta$. At $\theta = 1$, we have $\tilde{f}(1) = g_1(\tau)$, and the convexity of g_1 gives $g_1(\tau) - g_1(0) \leq \tau g_1'(\tau)$, with the left derivative. By (7.1), the function $S_{\tilde{\gamma}_1}$ is nondecreasing with values in $[0, 1]$, and since $0 \leq \varphi \leq 1$, we have $0 \leq \int \varphi \, dS_{\tilde{\gamma}_1} \leq 1$. Then (8.4) gives $g_1'(\tau) \leq \rho/2$. We conclude that $\tilde{f}(1) \leq F_\rho(\gamma_1) + \rho\tau/2 \leq F_{\min} + \eta + 15\eta/\ell$. Using $\ell \leq 1$, we obtain

$$\tilde{f}(0) + \tilde{f}(1) - 2F_{\min} \leq 2\eta + \frac{15\eta}{\ell} \leq \frac{17\eta}{\ell}.$$

We now use the midpoint inequality (7.9) for the segment $(\tilde{\gamma}_\theta)$ with $\Theta = [1/4, 3/4]$, on which the weight has integral $3/16$. Together with (8.5), it yields

$$\frac{3}{16} \cdot \frac{7}{34560\beta} \rho A^2 \min \left\{ \ell^3, \frac{\ell}{\bar{d}} \right\} \leq \frac{17\eta}{\ell},$$

that is,

$$\frac{\eta}{\rho} \geq \frac{7}{3133440\beta} A^2 \min\left\{\ell^4, \frac{\ell^2}{\bar{d}}\right\} \geq \frac{7}{3133440\beta} \min\left\{A^2 \ell^4, \frac{A^2 \ell^2}{d_0 + d_1 + k}\right\},$$

where we used $\bar{d} \leq d_0 + d_1 + k$. Together with Step 1, this proves (8.1) with $c_{\text{cert}} = 7/(3133440\beta)$, which is smaller than $1/1920$ because $\beta > 1$. $\square$

8.2 Root localization

For an admissible profile γ , we define the quantile of its free measure in the energy normalization by

$$Q_\gamma(v) = \inf\{q \in [0, 1] : \nu_\gamma([0, q]) \geq v\}, \quad 0 < v < \beta - \lambda.$$

If a quantile of a minimizer is large compared with the prescribed mass plus its level, then the minimizer is small on an interval where the scalar trial profile has a positive density. We control the density of the partially regularized minimizer there, and then compare the two profiles on that interval using Lemma 8.1. We follow the strategy of the proof of [4, Lemma 7.6], which regularizes only the free mass below the selected quantile and compares the result with the scalar trial profile. We combine this strategy with Lemma 8.1, which gives the dependence on η/ρ in (8.7). The positive density of the trial profile near the origin comes from the next lemma.

Lemma 8.2 (Trial density near the root). *There exists $T_1 = T_1(\beta) \in (0, T_0]$ such that*

$$d\gamma_\beta \geq \frac{k_0}{2} dx \quad \text{as measures on } [0, T_1]. \quad (8.6)$$

Proof. Recall from (7.6) that μ_β has no atom in $[0, q_\beta)$, and that its density ϱ_β is continuous on $[0, q_\beta]$. We have $\gamma_\beta(x) = \beta \int_0^x \varrho_\beta$ for $0 \leq x < q_\beta$, and $\gamma_\beta(x)/x \rightarrow \beta \varrho_\beta(0)$ as $x \downarrow 0$. By (7.5), we have $\gamma_\beta(x)/x \geq k_0$ for $0 < x \leq T_0$, which implies that $\beta \varrho_\beta(0) \geq k_0$. By the continuity of ϱ_β at 0, there exists $T_1 \in (0, \min\{T_0, q_\beta\})$ such that $\beta \varrho_\beta \geq k_0/2$ on $[0, T_1]$. The measure $d\gamma_\beta$ has no atom at 0, because $\gamma_\beta(0) = 0$, and on $[0, T_1]$ it equals $\beta \varrho_\beta dx$; this gives (8.6). The scalar Parisi measure is unique [10, Corollary 1], and q_β and ϱ_β depend only on β . The constants k_0 and T_0 also depend only on β , and we may therefore choose T_1 depending only on β . $\square$

The width T_1 is not explicit, because the continuity of ϱ_β at 0 is a qualitative statement.

Proposition 8.3 (Root localization). *There exists $C = C(\beta) < \infty$ with the following property. Let $N \geq 1$, $0 < \rho \leq 1$, and $0 < \lambda \leq \lambda_0$, let γ_* be a minimizer of F_ρ , and let η be a trial budget for γ_* . Then*

$$Q_{\gamma_*}(v) \leq C \left[\lambda + v + \min\{1, (\eta/\rho)^{1/6}\} \right] \quad (0 < v < \beta - \lambda). \quad (8.7)$$

The constant depends on β only through β , k_0 , C_0 , and T_1 .

Proof. The proof has five steps. After reducing to $0 < \eta < \rho$, we replace the quantile by its truncation $q = \min\{Q_{\gamma_*}(v), T_1\}$ and dispose of the case where q is small compared with $\lambda + v$ (Step 1). Otherwise, we regularize the free mass of γ_* below q , with a width adapted to the moved mass (Step 2), compare the result with γ^{tr} on the middle interval $(q/4, 3q/4)$ (Step 3), and apply Lemma 8.1 (Step 4). The resulting inequality bounds q either by a power of η/ρ or in terms of the density of the regularized profile, which we control with the choice of the width and Young's inequality (Step 5).

Step 1: Reductions. Fix $v \in (0, \beta - \lambda)$, and let $\bar{v} = \lambda + v$. The constant C constructed below is at least one. If $\eta \geq \rho$, the claim follows from $Q_{\gamma_*}(v) \leq 1 = \min\{1, (\eta/\rho)^{1/6}\}$. If $\eta = 0$, every $\eta' \in (0, \rho)$ is also a trial budget for γ_* ; we apply the argument below for positive budgets to η' and let $\eta' \downarrow 0$. We may therefore assume that $0 < \eta < \rho$, and we write $\varepsilon = \eta/\rho \in (0, 1)$.

We set $q = \min\{Q_{\gamma_*}(v), T_1\}$. Then $Q_{\gamma_*}(v) \leq q/T_1$: if $Q_{\gamma_*}(v) \leq T_1$, then $q = Q_{\gamma_*}(v)$ and $T_1 \leq 1$; otherwise $q = T_1$ and $Q_{\gamma_*}(v) \leq 1 = q/T_1$. It suffices to bound q . If $q \leq 16\bar{v}/k_0$, then $Q_{\gamma_*}(v) \leq 16\bar{v}/(k_0 T_1)$, which is of the required form. Suppose instead that

$$q > 16\bar{v}/k_0. \quad (8.8)$$

Step 2: Regularization below q . Let $m = \nu_{\gamma_*}([0, q])$. For every $x < q$, we have $x < Q_{\gamma_*}(v)$, and the definition of the quantile gives $\nu_{\gamma_*}([0, x]) < v$. Since the interval $[0, q]$ is the increasing union of the intervals $[0, x]$, $x < q$, continuity from below implies that $m \leq v$. An atom of ν_{γ_*} at q is not counted in m : it is not moved by the regularization below, and it lies outside the closure of the interval of Step 3.

We take

$$h = \min\{q/8, \sqrt{\varepsilon/m}\} \quad \text{if } m > 0, \quad h = q/8 \quad \text{if } m = 0.$$

Then $h \leq q/8$ and $mh^2 \leq \varepsilon$, and $h \leq 1 - q$ because $q \leq T_1 \leq T_0 \leq 1/4$. Let $\bar{\gamma}$ be the partial regularization of Lemma 7.3(a) with $T = q$. By that lemma, we have $F_\rho(\bar{\gamma}) - F_\rho(\gamma_*) \leq \rho mh^2/12 \leq \eta/12$, and the restriction of $d\bar{\gamma}$ to $(0, q)$ has density at most m/h . For $0 \leq x < q$, the free mass of $\bar{\gamma}$ in $[0, x]$ comes from the moved mass. Using (8.8) in the last step, we find that

$$\bar{\gamma}(x) \leq \lambda + m \leq \bar{v} < \frac{k_0 q}{16} \quad (0 \leq x < q).$$

Step 3: The middle interval. Let $I = (q/4, 3q/4)$. For $x \in I$, (7.5) and (8.8) give $\gamma_\beta(x) \geq k_0 x > k_0 q/4 > 4\bar{v} \geq \lambda$, which implies that $\gamma^{\text{tr}} = \gamma_\beta$ on I . Using $I \subset (0, T_1]$, Lemma 8.2, and the density bound for $d\gamma^{\text{tr}}$ on $(0, T_0]$ stated after (7.5), we find that

$$\frac{k_0}{2} dx \leq d\gamma^{\text{tr}} \leq C_0 dx \quad \text{on } I.$$

By Step 2, we also have $d\bar{\gamma} \leq (m/h) dx$ on I and

$$\gamma^{\text{tr}} - \bar{\gamma} > \frac{k_0 q}{4} - \frac{k_0 q}{16} \geq \frac{k_0 q}{8} \quad \text{on } I.$$

Step 4: The certificate. We apply Lemma 8.1 with the interval $I = (q/4, 3q/4) \subseteq (0, 1)$, the endpoints $\gamma_0 = \bar{\gamma}$ and $\gamma_1 = \gamma^{\text{tr}}$, and the parameters

$$\ell = \frac{q}{2}, \quad A = \frac{k_0 q}{8}, \quad k = \frac{k_0}{2}, \quad d_0 = \frac{m}{h}, \quad d_1 = C_0.$$

Both endpoints have excess at most η over $\min F_\rho = F_\rho(\gamma_*)$: for $\bar{\gamma}$ by Step 2, and for γ^{tr} by (7.7). The four entries of the minimum in (8.1) are

$$\begin{aligned} k\ell^3 &= \frac{k_0 q^3}{16}, & A\ell^2 &= \frac{k_0 q^3}{32}, & A^2\ell^4 &= \frac{k_0^2 q^6}{1024}, \\ \frac{A^2\ell^2}{d_0 + d_1 + k} &= \frac{k_0^2 q^4}{256(m/h + C_0 + k_0/2)}. \end{aligned}$$

Since $q \leq 1$, the first two entries are at least $k_0 q^6/32$. With $c' = c_{\text{cert}} \min\{k_0/32, k_0^2/1024\}$, we find

$$\varepsilon \geq c' \min\left\{q^6, \frac{q^4}{m/h + C_0 + k_0/2}\right\}. \quad (8.9)$$

Step 5: Conclusion. If the minimum in (8.9) is q^6 , then $q \leq (\varepsilon/c')^{1/6}$. Otherwise, we have $q^4 \leq \varepsilon(m/h + C_0 + k_0/2)/c'$. By the choice of h , we have

$$\frac{m}{h} = \max\left\{\frac{8m}{q}, \varepsilon^{-1/2} m^{3/2}\right\} \leq \frac{8m}{q} + \varepsilon^{-1/2} m^{3/2}$$

if $m > 0$, and $m/h = 0$ if $m = 0$. Using $m \leq v \leq \bar{v}$ and (8.8), we have $8m/q < k_0/2$. When the minimum is the second entry, we therefore obtain

$$q^4 \leq \frac{(C_0 + k_0)\varepsilon + \varepsilon^{1/2} m^{3/2}}{c'}, \quad q \leq (c')^{-1/4} ((C_0 + k_0)^{1/4} \varepsilon^{1/4} + \varepsilon^{1/8} m^{3/8}),$$

where we used $(x + y)^{1/4} \leq x^{1/4} + y^{1/4}$ for $x, y \geq 0$. By Young's inequality with exponents $8/5$ and $8/3$, we have

$$\varepsilon^{1/8} m^{3/8} \leq \frac{5}{8} \varepsilon^{1/5} + \frac{3}{8} m \leq \varepsilon^{1/6} + m,$$

where in the last inequality we used that $0 < \varepsilon < 1$. Together with $\varepsilon^{1/4} \leq \varepsilon^{1/6}$ and $m \leq \bar{v}$, this shows that $q \leq (c')^{-1/4} ((C_0 + k_0)^{1/4} + 1)(\bar{v} + \varepsilon^{1/6})$. Combining the three cases (Step 1, and the two cases of (8.9)), we conclude that

$$q \leq C_1(\bar{v} + \varepsilon^{1/6}), \quad C_1 = \max\{16/k_0, (c')^{-1/6}, (c')^{-1/4} ((C_0 + k_0)^{1/4} + 1)\}.$$

Recalling that $Q_{\gamma_*}(v) \leq q/T_1$ and that $\varepsilon^{1/6} = \min\{1, (\eta/\rho)^{1/6}\}$ for $\eta < \rho$, we obtain (8.7) with $C = \max\{1, C_1/T_1\}$. This constant depends only on β (through c_{cert}), k_0 , C_0 , and T_1 . $\square$

We comment on the exponent $1/6$ in (8.7). In the setting of the proof above, the variation in the proof of Lemma 8.1 moves mass toward the middle of I at a cost of at most $\rho\tau/2$, which is of order η/q , and the excesses at the endpoints of the perturbed segment are at most of order η/q . The response integrated against the bump is at least $4\ell/15 = 2q/15$. By (8.5), the second derivative along the perturbed segment is bounded below by a multiple of $\rho \min\{q^5, q^3/\bar{d}\}$, where $\bar{d}$ is the density bound in the proof of Lemma 8.1. Comparing this lower bound with the excesses at the endpoints, we obtain $\eta/\rho \gtrsim_{\beta} \min\{q^6, q^4/\bar{d}\}$, which leads to (8.9). Its first alternative gives the exponent $1/6$. In the other alternative, the choice of the width h controls the density m/h without losing the linear dependence of the bound on v .

Corollary 8.4 (Quantile bound in probability normalization). *Let $N \geq 1$, $0 < \rho \leq 1$, and $0 < \lambda \leq \lambda_0$, and let $s = \lambda/\beta$. Let $\mu = s\delta_0 + \nu$ be a minimizer of $\mathcal{F}_{1-\rho,s}$ in (3.3), so that $\gamma_* = \beta\mu([0, \cdot])$ minimizes F_{ρ} , and let $\eta \geq 0$ be a trial budget for γ_* . Let*

$$Q(v) = \inf\{q \in [0, 1] : \nu([0, q]) \geq v\} \quad (0 < v < 1 - s)$$

be the quantile function (3.6) of ν , and define

$$\omega(\rho, \eta) = \min\{1, (\eta/\rho)^{1/6}\}. \quad (8.10)$$

Then, with $C_Q = \max\{1, \beta C\}$ and C the constant of Proposition 8.3,

$$Q(v) \leq C_Q(s + v + \omega(\rho, \eta)) \quad (0 < v < 1 - s). \quad (8.11)$$

Proof. The free measure of γ_* is $\nu_{\gamma_*} = \beta\nu$, and $\nu_{\gamma_*}([0, q]) \geq \beta v$ if and only if $\nu([0, q]) \geq v$. Hence $Q(v) = Q_{\gamma_*}(\beta v)$, and $0 < \beta v < \beta - \lambda$ for $0 < v < 1 - s$. By Proposition 8.3 and $\lambda = \beta s$, we have

$$Q(v) \leq C[\beta(s + v) + \omega(\rho, \eta)] \leq \beta C(s + v + \omega(\rho, \eta)),$$

where we used $\beta > 1$. We obtain (8.11), because $\beta C \leq C_Q$. □

Scalar response and stability

In this section, we study the scalar Parisi equation of a single spin under an arbitrary profile and the scalar response operator $J_{a,d}^{\text{sc}}$ built from it. The cavity interpolation of Section 10 decouples one spin from the others. At the independent endpoint of that interpolation, the decoupled spin follows the scalar law of Section 9.1 with coefficient $d = \beta^2(1 - t/N)$, and in Section 11, we express the derivative of the cavity covariance at this endpoint through $J_{a,d}^{\text{sc}}$, applied to the covariance of the rest self-overlap. The contraction argument of Section 12 needs the order bound $J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I$ with $\varepsilon_N \leq \rho/4$ at a minimizing profile. The goal of this section is to prove this bound with $\varepsilon_N \leq C_{\text{st}}[\rho^{-3/4}\eta^{1/4} + \lambda^2 + N^{-1}]$, where η is a trial budget (Proposition 9.9).

The argument has four steps.

1. We show in Lemma 9.6 that, for every profile, $J_{a,d}$ is the derivative of the mean square magnetization $\Gamma_{a,d}$ when the covariance levels are transported by $q \mapsto q + \theta\varphi(q)$.
2. The scalar Parisi functional is convex. We compare its second derivative along a transport of levels with the lower bound given by convexity and the first variation, and we find that $J_{a,d}^{\text{sc}}$ is self-adjoint and that it is dominated by multiplication by the multiplier $k_{a,d}(q) = d \mathbb{E}^0 h_q^2$, for every profile (Lemma 9.7).
3. At the scalar Parisi measure, we have $k = 1$ on $\text{supp } \mu_\beta = [0, q_\beta]$ and $k(q) \leq e^{-\beta^2(q-q_\beta)/2}$ above it (Lemma 9.8). On the support, this follows from an identity of Talagrand. Above the support, we use the log-concavity of a transformed density of the scalar diffusion and its heat evolution, both from [48].
4. The multiplier depends Lipschitz continuously on the profile in L^1 and on d (Lemma 9.3). By Proposition 7.6, a finite-dimensional minimizer is close in L^1 to the scalar trial profile, which is close to $\beta\alpha_\beta$. Combining these facts with the first three steps, we obtain Proposition 9.9.

The bound is marginal: at the Parisi measure, $k = 1$ on the support, where the multiplier gives no contraction. The contraction in Section 12 comes instead from the factor $t = 1 - \rho$, and it needs only $\varepsilon_N \leq \rho/4$. In the physics literature, De Dominicis and Kondor [35, pp. 606–608] found that Parisi’s solution with replica symmetry breaking is marginally stable in a truncated replica model near the critical temperature. The scalar identity on the support that expresses this marginality appears in [32, Eqs. (22)–(23)].

Sommers and Dupont gave an earlier variational formulation of the coupled equations for the local magnetization and the distribution of frozen fields [77, Section 2, Eqs. (1)–(7), and Eq. (33)]; see also [32, Section III, (14)–(18), (22)–(23)]. Those equations use the physical field and the order-parameter function $q(x)$. Here we use the covariance level q and the integrated formulation in terms of probability measures described below.

We use the following results from earlier sections: the generalized law and Lemma 4.2 of Section 4 for the basic properties of the scalar law, Lemma 4.3 for the differentiation in Lemma 9.6, and, from Section 7, the trial budget (7.7), the root growth (7.5), the description (7.6) of μ_β , and Proposition 7.6. The results of this section are used later as follows. The scalar law, described by (9.1) and (9.9), and the kernel bounds of Lemma 9.4 enter the cavity expansion in Sections 10 and 11. The operator (9.19) and its kernel form (9.20) are used in Lemma 11.2, Lemma 9.7 in Proposition 12.4, and Proposition 9.9 in the proof of Theorem 1.1 in Section 13.

Standing assumptions and notation. Throughout this section, $\beta > 1$ is fixed and $d \in [\beta^2/2, \beta^2]$. A *profile* is the distribution function $a(q) = \mu([0, q])$ of a Borel probability measure μ on $[0, 1]$; it is nondecreasing and right-continuous, and $a(1) = 1$. We write $\mu = s\delta_0 + \nu$ with $s \geq 0$, where the measure $\nu \geq 0$ may charge 0. The scalar law depends only on a . The decomposition enters only through ν , which appears in the response operator. A *step profile* is a profile with finitely many values. Integrals over $[0, q]$ include an atom at q , and integrals over $(q, 1]$ exclude it. Unless stated otherwise, C_β denotes a constant depending only on β , whose value may change from line to line. It does not depend on $d \in [\beta^2/2, \beta^2]$, on s , or on the profile, including the positions and masses of its atoms. No profile is assumed to minimize a functional until Section 9.6.

9.1 The scalar law

The scalar Parisi equation and its diffusion are

$$\begin{aligned} \partial_q u &= -\frac{d}{2}(\partial_{xx} u + a(q)(\partial_x u)^2), & u(1, x) &= \log(2 \cosh x), \\ dY_q &= d a(q) \partial_x u(q, Y_q) dq + \sqrt{d} dB_q, & Y_0 &= 0, \end{aligned} \quad (9.1)$$

where B is a standard Brownian motion. We write $u = u_{a,d}$ when the parameters need to be shown. Subscripts x denote spatial derivatives, and $u_j = \partial_x^j u$. When $a = \alpha_\beta$ and $d = \beta^2$, the value $u(0, 0)$ is the first term of the scalar Parisi functional (1.1) at its minimizer.

Solutions and regularity. For a step profile, (9.1) has a classical solution on each interval where a is constant. For a general profile, u is the uniform limit of the solutions for step profiles whose measures converge weakly to μ [48, Section 2.1]. By [48, Proposition 2.1],

applied with β^2 replaced by d , the function $u(q, \cdot)$ is even, its spatial derivatives of every positive order are bounded and jointly continuous on $[0, 1] \times \mathbb{R}$, and the spatial derivatives of the approximating solutions converge uniformly [48, (2.2)]. The terminal value in [48] is $\log \cosh x$; it differs from ours by the constant $\log 2$, which shifts u by $\log 2$ and does not change its spatial derivatives. The distribution functions of the approximating measures converge to a at every continuity point of a , that is, at all but countably many points. Passing to the limit in the equations integrated in the level variable, we find that

$$u(q, x) - u(r, x) = \frac{d}{2} \int_q^r (u_{xx} + a u_x^2)(v, x) dv, \quad 0 \leq q \leq r \leq 1, x \in \mathbb{R}. \quad (9.2)$$

In particular, u is Lipschitz in q , since the integrand is bounded. For the weak space–time interpretation of the Parisi equation for general measures, see [43, Theorem 4]. At an atom of μ , the solution need not be differentiable in q . We read (9.1) and every other parabolic equation in this section in this integrated form, including the linearized equations derived below, and we never differentiate a .

Linear backward equations. Several proofs below use linear equations of the form

$$\partial_q w + \frac{c(q)}{2} w_{xx} + b w_x + V w = -F, \quad (9.3)$$

where c is measurable with values in $[\beta^2/4, 2\beta^2]$, and b , V , and F are bounded functions of (q, x) that are measurable in q and have bounded continuous spatial derivatives. We use (9.3) only for functions w that are Lipschitz in q and have bounded continuous spatial derivatives of every order. We read it in integrated form, as in (9.2). Fix (q, x) , and let Z solve $dZ_r = b(r, Z_r) dr + \sqrt{c(r)} dB_r$ for $r \geq q$, with $Z_q = x$. Itô's formula applies to $r \mapsto e^{\int_q^r V(v, Z_v) dv} w(r, Z_r)$, with the derivative of w in q taken almost everywhere; to see this, one may mollify w in the level variable and pass to the limit, using that the derivatives are bounded. Taking expectations, we obtain the Feynman–Kac representation

$$w(q, x) = \mathbb{E} \left[e^{\int_q^1 V(v, Z_v) dv} w(1, Z_1) + \int_q^1 e^{\int_q^r V(v, Z_v) dv} F(r, Z_r) dr \right]. \quad (9.4)$$

In particular, if $w(1, \cdot) = 0$, then $\sup_x |w(q, x)| \leq e^{\|V\|_\infty} \int_q^1 \sup_x |F(r, x)| dr$.

Notation. We write $\mathbb{P}^0$ and $\mathbb{E}^0$ for probability and expectation under the law of Y , together with a terminal spin $\epsilon \in \{-1, 1\}$, which is sampled given Y_1 with

$$\mathbb{P}^0(\epsilon = \pm 1 \mid Y_1) = \frac{e^{\pm Y_1}}{2 \cosh Y_1} = e^{\pm Y_1 - u(1, Y_1)}.$$

Let $\mathcal{G}_q^0$ be the filtration of Y ; the terminal spin is not $\mathcal{G}_q^0$ -measurable. We define

$$m_q = u_x(q, Y_q), \quad h_q = u_{xx}(q, Y_q), \quad \Gamma_{a,d}(q) = \mathbb{E}^0 m_q^2, \quad k_{a,d}(q) = d \mathbb{E}^0 h_q^2.$$

We call $k_{a,d}$ the *scalar multiplier*. For $0 \leq r \leq \ell \leq 1$, let $P_{r,\ell}^0 \psi(y) = \mathbb{E}^0[\psi(Y_\ell) \mid Y_r = y]$ be the transition operator of the Markov process Y , and set

$$f_p = m_p^2, \quad \bar{f}_p = f_p - \mathbb{E}^0 f_p, \quad g_r^\ell(y) = P_{r,\ell}^0(u_x(\ell, \cdot)^2)(y),$$

so that $g_r^\ell(Y_r) = \mathbb{E}^0[m_\ell^2 \mid \mathcal{G}_r^0]$. When we need the diffusion started from y at level r , we write $Y^{r,y}$ and $\mathbb{E}_{r,y}$.

Lemma 9.1 (Basic properties of the scalar law). *Let a be a profile and $d \in [\beta^2/2, \beta^2]$.*

(a) *On $[0, 1] \times \mathbb{R}$, we have*

$$|\partial_x u| \leq 1, \quad 0 \leq \partial_{xx} u \leq 1 - (\partial_x u)^2 \leq 1. \quad (9.5)$$

(b) *For every $j \geq 1$, there exists a constant $C_j(\beta)$, depending only on β and j , such that $|u_j| \leq C_j(\beta)$ on $[0, 1] \times \mathbb{R}$. Further, $\sup_x |u_{xx}(q, x) - u_{xx}(r, x)| \leq C_\beta |q - r|$ for all $q, r \in [0, 1]$.*

(c) *We have*

$$m_q = \mathbb{E}^0[\epsilon \mid \mathcal{G}_q^0], \quad m_0 = 0, \quad dm_q = \sqrt{d} h_q dB_q, \quad \Gamma'_{a,d}(q) = k_{a,d}(q). \quad (9.6)$$

The function $k_{a,d}$ is continuous on $[0, 1]$, and $\Gamma_{a,d}$ is continuously differentiable on $[0, 1]$, with one-sided derivatives at the endpoints.

Proof. We first identify the scalar law as a generalized law. We consider the generalized law of Definition 4.1 with $N = 1$, $\Lambda = d$, all coupling coefficients equal to zero, root mass $s = 0$, and the probability measure μ in place of ν ; this choice is allowed for every probability measure. The terminal value in (4.1) is then $\log \sum_{\sigma=\pm 1} e^{x\sigma} = \log(2 \cosh x)$, the Gaussian disorder does not enter, and the outer density is one. Hence the generalized law is the law of (Y, ϵ) ; it depends only on a . For one spin, the conditional covariance is $C_q = 1 - m_q^2$, and $H_q = h_q$. Lemma 4.2 implies that $m_q = \mathbb{E}^0[\epsilon \mid \mathcal{G}_q^0]$, that $m_0 = 0$, and that, for every $q \in [0, 1]$, almost surely,

$$a(q)(1 - m_q^2) \leq h_q \leq 1 - m_q^2. \quad (9.7)$$

Differentiating (9.2) in x , we see that u_x satisfies the integrated form of $\partial_q u_x = -(d/2)(u_{xxx} + 2a u_x u_{xx})$. Using Itô's formula along Y , we find that the drift terms cancel and that $dm_q = \sqrt{d} h_q dB_q$. This is the case $N = 1$ of (4.3).

Proof of (a). Fix $q \in (0, 1]$. Since the drift of Y is bounded by d , Girsanov's theorem implies that the law of Y_q is equivalent to the centered Gaussian law with variance dq . In particular, it charges every nonempty open interval. The functions

$$u_{xx}(q, \cdot) - a(q)(1 - u_x(q, \cdot)^2), \quad 1 - u_x(q, \cdot)^2 - u_{xx}(q, \cdot), \quad 1 - u_x(q, \cdot)^2$$

are continuous, and by (9.7) and $|m_q| \leq 1$, each of them is nonnegative at Y_q almost surely. A continuous function with this property is nonnegative at every x . This proves (9.5) for $q > 0$, since $a(q)(1 - u_x^2) \geq 0$. Letting $q \downarrow 0$ and using the joint continuity of u_x and u_{xx} , we see that (9.5) also holds at $q = 0$. (With the right continuity of a , the lower bound $u_{xx} \geq a(q)(1 - u_x^2)$ also holds everywhere.)

Proof of (b). We differentiate the equation and use the Feynman–Kac representation, following the method of [43, Lemma 16]. By (9.5), we have $|u_1| \leq 1$ and $|u_2| \leq 1$. For $j \geq 3$, the Leibniz rule gives

$$\partial_x^j(u_x^2) = 2u_1u_{j+1} + 2ju_2u_j + \sum_{i=2}^{j-2} \binom{j}{i} u_{i+1}u_{j-i+1},$$

with an empty sum for $j = 3$. Differentiating (9.2) j times in x , we see that u_j solves (9.3) with

$$c = d, \quad b = da u_1, \quad V = jda u_2, \quad F = \frac{da}{2} \sum_{i=2}^{j-2} \binom{j}{i} u_{i+1}u_{j-i+1},$$

and terminal value $\partial_x^j \log(2 \cosh x)$. Recall that the derivatives of u are bounded by [48, Proposition 2.1]; in particular, the hypotheses of (9.4) hold, with $Z = Y$. By (9.5), V takes values in $[0, j\beta^2]$; note that the exponent in (9.4) is nonnegative. Then (9.4) implies that

$$\|u_j\|_\infty \leq e^{j\beta^2} \left(\|\partial_x^j \log(2 \cosh x)\|_\infty + \frac{\beta^2}{2} \sum_{i=2}^{j-2} \binom{j}{i} \|u_{i+1}\|_\infty \|u_{j-i+1}\|_\infty \right).$$

Every derivative on the right has order between 3 and $j - 1$. By induction on j , we conclude that $\|u_j\|_\infty \leq C_j(\beta)$ for all $j \geq 1$, uniformly over profiles and $d \in [\beta^2/2, \beta^2]$. Differentiating (9.2) twice in x , we find

$$u_{xx}(q, x) - u_{xx}(r, x) = \frac{d}{2} \int_q^r \left[u_4 + 2a(u_2^2 + u_1u_3) \right] (v, x) dv,$$

and the integrand is bounded by a constant depending only on β . This proves the Lipschitz bound in (b).

Proof of (c). The first three identities in (9.6) were shown above. By Itô's isometry and $m_0 = 0$, we have

$$\Gamma_{a,d}(q) = \mathbb{E}^0 m_q^2 = d \int_0^q \mathbb{E}^0 h_r^2 dr = \int_0^q k_{a,d}(r) dr.$$

By (b), we have $|h_q - h_r| \leq C_\beta |q - r| + C_3(\beta) |Y_q - Y_r|$. Since Y has continuous paths and $|h| \leq 1$, dominated convergence implies that $k_{a,d}$ is continuous. It follows that $\Gamma_{a,d}$ is continuously differentiable on $[0, 1]$, with $\Gamma'_{a,d} = k_{a,d}$. $\square$

We also use the exact path likelihood of the scalar law, for every profile, including profiles with atoms. It is the change of measure of [44, Lemma 8.5 and Corollary 8.7]; we include the derivation in order to record the conventions at atoms. It extends the likelihood used for step profiles in the proof of Lemma 4.2 to every profile, without approximation.

Lemma 9.2 (Exact path likelihood). *Let a be a profile and $d \in [\beta^2/2, \beta^2]$. Under a reference law, let $X_q = \sqrt{d} B_q$, where B is a standard Brownian motion. For every $q \in [0, 1]$, the law of $(Y_r)_{0 \leq r \leq q}$ has density*

$$\exp\left\{a(q)u(q, X_q) - \int_{[0, q]} u(r, X_r) \mu(dr)\right\} \quad (9.8)$$

with respect to the law of $(X_r)_{0 \leq r \leq q}$. The joint law of (Y, ϵ) has density

$$\exp\left\{\epsilon X_1 - su(0, 0) - \int_{[0, 1]} u(q, X_q) \nu(dq)\right\} \quad (9.9)$$

with respect to the product of the law of X and counting measure on $\{-1, 1\}$. The same holds if the coefficient d in (9.1) is replaced by a bounded measurable function $c(q) > 0$ almost everywhere, and $X_q = \int_0^q \sqrt{c(r)} dB_r$.

Proof. We write $Y_r = u(r, X_r)$. The drift $da u_x$ of Y is bounded, and by (9.5), we have

$$\exp\left\{\frac{d}{2} \int_0^q a(r)^2 u_x(r, X_r)^2 dr\right\} \leq e^{\beta^2/2},$$

which is Novikov's condition. By Girsanov's theorem, the density of the law of $(Y_r)_{r \leq q}$ with respect to the law of $(X_r)_{r \leq q}$ is

$$\exp\left\{\int_0^q a(r)u_x(r, X_r) dX_r - \frac{d}{2} \int_0^q a(r)^2 u_x(r, X_r)^2 dr\right\}.$$

Under the reference law, using Itô's formula and (9.2), we have

$$dY_r = u_x(r, X_r) dX_r - \frac{d}{2} a(r) u_x(r, X_r)^2 dr.$$

The exponent therefore equals $\int_0^q a(r) dY_r$. The function a is right-continuous and of bounded variation, with $da = \mu$ on $(0, 1]$, and Y is a continuous semimartingale. Integrating by parts, we have

$$a(q)Y_q - a(0)Y_0 = \int_{(0, q]} a(r-) dY_r + \int_{(0, q]} Y_r \mu(dr).$$

The functions $a(r-)$ and $a(r)$ differ only at the countably many atoms of μ , which affects neither the stochastic integral nor the Lebesgue integral in dY_r . Since $X_0 = 0$

and $a(0) = \mu(\{0\})$, we have $a(0)Y_0 = \int_{\{0\}} u(r, X_r) \mu(dr)$. Inserting these facts into the previous display, we obtain

$$\int_0^q a(r) dY_r = a(q)u(q, X_q) - \int_{[0,q]} u(r, X_r) \mu(dr).$$

This proves (9.8). At an atom of μ at q , the jump of the first term is canceled by the same atom in the integral, and at $q = 0$ the density equals one.

Given the path, the terminal spin has probability $e^{\epsilon Y_1 - u(1, Y_1)}$ relative to counting measure. Multiplying (9.8) at $q = 1$ by this factor, and using $a(1) = 1$ and $\mu = s\delta_0 + \nu$, we obtain (9.9). No further normalization is needed, since both factors are probability densities.

For a coefficient $c(q)$ that is positive almost everywhere, we first choose the everywhere positive representative $c^+(q) = c(q)$ when $c(q) > 0$ and $c^+(q) = 1$ otherwise. It is measurable and bounded and equals c almost everywhere. This change preserves every prefix clock, the integrated equation and drift, and the stochastic integral defining X . Thus the law and the likelihood are independent of this choice of representative. Let $C(q) = \int_0^q c(r) dr$. The substitution $\tau = C(q)/C(1)$ turns the equation with coefficient $c(q)$ into (9.1) with coefficient $C(1)$ and the profile of the image of μ under $q \mapsto C(q)/C(1)$. This substitution is made in the integrated equation, using $d\tau = c(q) dq/C(1)$. Since $c > 0$ almost everywhere, $q \mapsto C(q)/C(1)$ is an increasing homeomorphism of $[0, 1]$; no positive lower bound on c is required. By [48, Proposition 2.1], applied with coefficient $C(1)$, the solution satisfies $|u_x| \leq 1$ and has bounded spatial derivatives. The argument above then applies with $d dr$ replaced by $c(r) dr$; in Novikov's condition, β^2 is replaced by an upper bound for c . $\square$

9.2 Continuity and kernel bounds

The proof of the next lemma follows the comparison argument of [43, Lemma 14], which bounds the differences of the solutions and of their first spatial derivatives by the L^1 distance of the profiles. Here we also allow d to vary and compare the second derivatives along the diffusion.

Lemma 9.3 (Continuity of the scalar multiplier). *There exists $C_\beta < \infty$ such that, for any two profiles $a, \tilde{a}$ and any $d, \tilde{d} \in [\beta^2/2, \beta^2]$,*

$$\sup_{0 \leq q \leq 1} |k_{a,d}(q) - k_{\tilde{a},\tilde{d}}(q)| \leq C_\beta (\|a - \tilde{a}\|_{L^1(0,1)} + |d - \tilde{d}|). \quad (9.10)$$

Proof. Let u and $\tilde{u}$ solve (9.1) for the two parameter pairs, and let Y and $\tilde{Y}$ be the corresponding diffusions. We set $\hat{u} = u - \tilde{u}$ and $\Delta = \|a - \tilde{a}\|_{L^1(0,1)} + |d - \tilde{d}|$. The proof has three steps. We first bound $\hat{u}_x$ and $\hat{u}_{xx}$ by $C_\beta \Delta$, then couple the two diffusions, and finally compare the multipliers.

Step 1: the difference of the solutions. Subtracting the two equations and writing $u_x^2 - \tilde{u}_x^2 = (u_x + \tilde{u}_x)\hat{u}_x$, we find that $\hat{u}$ solves (9.3) with

$$c = d, \quad b = \omega, \quad V = 0, \quad F = \Xi, \quad \hat{u}(1, \cdot) = 0,$$

where

$$\omega = \frac{d a(q)}{2} (u_x + \tilde{u}_x), \quad \Xi = \frac{d - \tilde{d}}{2} \tilde{u}_{xx} + \delta(q) \tilde{u}_x^2, \quad \delta(q) = \frac{d a(q) - \tilde{d} \tilde{a}(q)}{2}.$$

Using Lemma 9.1(a) and (b), we see that $0 \leq \omega_x \leq \beta^2$ and $|\omega_{xx}| \leq C_\beta$. Further,

$$\Xi_x = \frac{d - \tilde{d}}{2} \tilde{u}_3 + 2\delta \tilde{u}_1 \tilde{u}_2, \quad \Xi_{xx} = \frac{d - \tilde{d}}{2} \tilde{u}_4 + 2\delta(\tilde{u}_2^2 + \tilde{u}_1 \tilde{u}_3),$$

and $2|\delta(q)| \leq |d - \tilde{d}| + \beta^2|a(q) - \tilde{a}(q)|$. Using Lemma 9.1(b) again, we find

$$\int_0^1 \sup_x (|\Xi_x| + |\Xi_{xx}|)(q, x) dq \leq C_\beta \int_0^1 (|d - \tilde{d}| + |a(q) - \tilde{a}(q)|) dq \leq C_\beta \Delta.$$

Differentiating the equation for $\hat{u}$ once and twice in x , we see that $w_1 = \hat{u}_x$ and $w_2 = \hat{u}_{xx}$ solve (9.3) with $c = d$, $b = \omega$, zero terminal values, and

$$V = \omega_x, \quad F = \Xi_x \quad \text{and} \quad V = 2\omega_x, \quad F = \Xi_{xx} + \omega_{xx}w_1,$$

respectively. Then (9.4) implies that

$$\begin{aligned} \|w_1\|_\infty &\leq e^{\beta^2} \int_0^1 \|\Xi_x(q, \cdot)\|_\infty dq \leq C_\beta \Delta, \\ \|w_2\|_\infty &\leq e^{2\beta^2} \int_0^1 (\|\Xi_{xx}(q, \cdot)\|_\infty + \|\omega_{xx}\|_\infty \|w_1\|_\infty) dq \leq C_\beta \Delta. \end{aligned}$$

This shows that

$$\sup_{q,x} (|u_x - \tilde{u}_x| + |u_{xx} - \tilde{u}_{xx}|) \leq C_\beta \Delta. \quad (9.11)$$

Step 2: coupling the diffusions. We drive Y and $\tilde{Y}$ by the same Brownian motion and set $D(q) = \sup_{r \leq q} |Y_r - \tilde{Y}_r|$. By (9.5), the drift $d a(r)u_x(r, \cdot)$ is β^2 -Lipschitz in space, and at a common point it differs from $\tilde{d} \tilde{a}(r)\tilde{u}_x(r, \cdot)$ by at most $|d - \tilde{d}| + \beta^2|a(r) - \tilde{a}(r)| + \beta^2 C_\beta \Delta$, where we used (9.11). Subtracting the integral equations for Y and $\tilde{Y}$, we obtain

$$D(q) \leq \beta^2 \int_0^q D(r) dr + C_\beta \Delta + |\sqrt{d} - \sqrt{\tilde{d}}| \sup_{r \leq 1} |B_r|.$$

Since $d, \tilde{d} \geq \beta^2/2$, we have $|\sqrt{d} - \sqrt{\tilde{d}}| \leq |d - \tilde{d}|/(\sqrt{2}\beta)$. By Gronwall's inequality and $\mathbb{E} \sup_{r \leq 1} |B_r| \leq 2$, we conclude that

$$\mathbb{E} \sup_{q \leq 1} |Y_q - \tilde{Y}_q| \leq C_\beta (\Delta + |\sqrt{d} - \sqrt{\tilde{d}}| \mathbb{E} \sup_{q \leq 1} |B_q|) \leq C_\beta \Delta. \quad (9.12)$$

The same argument, applied to the two diffusions started from a common point y at a common level r , yields the same bound, uniformly in (r, y) .

Step 3: the multipliers. We write $h = u_{xx}(q, Y_q)$ and $\tilde{h} = \tilde{u}_{xx}(q, \tilde{Y}_q)$. Note that $|h^2 - \tilde{h}^2| \leq 2|h - \tilde{h}|$, because $0 \leq h, \tilde{h} \leq 1$. Then

$$\begin{aligned} |k_{a,d}(q) - k_{\tilde{a},\tilde{d}}(q)| &\leq |d - \tilde{d}| + 2\beta^2 \mathbb{E}|h - \tilde{h}| \\ &\leq |d - \tilde{d}| + 2\beta^2 \left(\|u_{xx} - \tilde{u}_{xx}\|_\infty + \|u_3\|_\infty \mathbb{E}|Y_q - \tilde{Y}_q| \right). \end{aligned}$$

Inserting (9.11), (9.12), and the bound $\|u_3\|_\infty \leq C_3(\beta)$ into the last line, we obtain (9.10). $\square$

The next lemma collects the bounds on the kernels of the response operator that the cavity estimates of Section 11 use. Each bound retains a factor Γ at the relevant levels.

Lemma 9.4 (Scalar kernels and their level weight). *For every profile a and $d \in [\beta^2/2, \beta^2]$, write $\Gamma = \Gamma_{a,d}$. Then Γ is nondecreasing and*

$$0 \leq \Gamma(p) \leq \min\{1, dp\} \leq \min\{1, \beta^2 p\}. \quad (9.13)$$

For all $p, r, \ell \in [0, 1]$,

$$|\mathbb{E}^0[m_p h_p m_r]| \leq \sqrt{\Gamma(p)\Gamma(r)}, \quad |\mathbb{E}^0[\tilde{f}_p \epsilon m_r]| \leq \sqrt{\Gamma(p)\Gamma(r)}, \quad (9.14)$$

$$|\mathbb{E}^0[\tilde{f}_p m_r m_\ell]| \leq \sqrt{\Gamma(p)\Gamma(r \wedge \ell)}, \quad |\mathbb{E}^0[\tilde{f}_p m_p m_r]| \leq \sqrt{\Gamma(p)\Gamma(p \wedge r)}, \quad (9.15)$$

$$0 \leq \text{Var}^0(f_p) \leq \Gamma(p), \quad 0 \leq \mathbb{E}^0[m_p^2 h_p] \leq \Gamma(p). \quad (9.16)$$

For $\ell \geq p$,

$$|\text{Cov}^0(f_p, g_p^\ell(Y_p))| \leq \sqrt{\Gamma(p)\Gamma(\ell)}. \quad (9.17)$$

Proof. By Lemma 9.1(c) and (9.5), we have $\Gamma(p) = d \int_0^p \mathbb{E}^0 h_r^2 dr$, where $0 \leq h_r \leq 1$. This implies that Γ is nondecreasing and that $0 \leq \Gamma(p) \leq dp$. We also have $\Gamma \leq 1$, because $|m_p| \leq 1$. This proves (9.13). Using $|m_p| \leq 1$ again, we have

$$\text{Var}^0(f_p) = \mathbb{E}^0 \tilde{f}_p^2 \leq \mathbb{E}^0 m_p^4 \leq \mathbb{E}^0 m_p^2 = \Gamma(p).$$

The bounds below follow from the Cauchy–Schwarz inequality, together with $|h_p| \leq 1$, $|\epsilon| = 1$, and $|m| \leq 1$:

$$\begin{aligned} |\mathbb{E}^0[m_p h_p m_r]|^2 &\leq \mathbb{E}^0[m_p^2 h_p^2] \mathbb{E}^0 m_r^2 \leq \Gamma(p)\Gamma(r), \\ |\mathbb{E}^0[\tilde{f}_p \epsilon m_r]|^2 &\leq \mathbb{E}^0 \tilde{f}_p^2 \mathbb{E}^0 m_r^2 \leq \Gamma(p)\Gamma(r), \\ |\mathbb{E}^0[\tilde{f}_p m_r m_\ell]|^2 &\leq \mathbb{E}^0 \tilde{f}_p^2 \mathbb{E}^0[m_r^2 m_\ell^2] \leq \Gamma(p)\Gamma(r \wedge \ell). \end{aligned}$$

In the last line, we used $m_r^2 m_\ell^2 \leq m_{r \wedge \ell}^2$, which holds because the factor at the smaller level is one of the two factors and the other is at most one. This proves (9.14) and the first

bound in (9.15); the second bound in (9.15) is the case $\ell = p$. The bounds (9.16) follow from the estimate on $\text{Var}^0(f_p)$ and from $0 \leq h_p \leq 1$, which gives $0 \leq \mathbb{E}^0[m_p^2 h_p] \leq \mathbb{E}^0 m_p^2$. Finally, let $\ell \geq p$. We have $0 \leq g_p^\ell \leq 1$ and $\mathbb{E}^0 g_p^\ell(Y_p) = \mathbb{E}^0 m_\ell^2 = \Gamma(\ell)$, which imply that

$$\text{Var}^0(g_p^\ell(Y_p)) \leq \mathbb{E}^0(g_p^\ell(Y_p))^2 \leq \mathbb{E}^0 g_p^\ell(Y_p) = \Gamma(\ell),$$

and (9.17) follows from the Cauchy–Schwarz inequality and the bound on $\text{Var}^0(f_p)$. $\square$

9.3 The response operator

For every profile, the operator defined below is the response of Γ to a transport of covariance levels (Lemma 9.6). In [56, Lemma 5.2, (H2)], a scalar response is written as a diagonal multiplier minus a positive semidefinite Hessian kernel. Here we identify the second variation of the scalar Parisi functional along transports of levels, and we compare it using the convexity of this functional along segments of distribution functions [10]. In Section 9.4, this shows that the scalar response operator is dominated by multiplication by $k_{a,d}$ at every profile. The transports fix the levels 0 and 1, and the operator acts on the *scalar transport space*

$$\mathcal{H}_v = \{v \in L^2(v) : v(0) = 0 \text{ if } v(\{0\}) > 0, v(1) = 0 \text{ if } v(\{1\}) > 0\}.$$

This is a closed subspace of $L^2(v)$, and the orthogonal projection onto it is multiplication by $\mathbf{1}_{(0,1)}$.

The operator involves the derivative of g_r^ℓ in its argument. Since the noise of Y is additive, the diffusion $Y^{r,y}$ started from y at level r solves an ordinary differential equation driven by the Brownian path, and it is differentiable in y . Its derivative is $\xi_\ell = \exp\{d \int_r^\ell a(v) u_{xx}(v, Y_v^{r,y}) dv\}$, which lies in $[1, e^{\beta^2}]$ by (9.5). Differentiating $g_r^\ell(y) = \mathbb{E}[u_x(\ell, Y_\ell^{r,y})^2]$ under the expectation, we find

$$\partial_y g_r^\ell(y) = \mathbb{E}[2u_x(\ell, Y_\ell^{r,y}) u_{xx}(\ell, Y_\ell^{r,y}) \xi_\ell], \quad |\partial_y g_r^\ell| \leq 2e^{\beta^2}. \quad (9.18)$$

Definition 9.5 (Response operator). Let a be a profile with $\mu = s\delta_0 + \nu$, and let $d \in [\beta^2/2, \beta^2]$. For a bounded Borel function φ on $[0, 1]$ and $p \in [0, 1]$, we define

$$\begin{aligned} \frac{1}{d}(J_{a,d}\varphi)(p) &= \mathbb{E}^0[h_p^2 + 2(1 - a(p))m_p^2 h_p] \varphi(p) - \int_{(p,1]} \mathbb{E}^0[m_p \partial_y g_p^\ell(Y_p)] \varphi(\ell) \nu(d\ell) \\ &\quad - 2 \int \mathbb{E}^0[m_p h_p m_r] \varphi(p \wedge r) \nu(dr) - \int \mathbb{E}^0[\bar{f}_p \epsilon m_r] \varphi(r) \nu(dr) \\ &\quad + \frac{1}{2} \int a(r) \mathbb{E}^0[\bar{f}_p m_r^2] \varphi(r) \nu(dr) + \frac{1}{2} \iint \mathbb{E}^0[\bar{f}_p m_r m_\ell] \varphi(r \wedge \ell) \nu(dr) \nu(d\ell) \\ &\quad + \frac{1}{2} \int \nu(dr) \int_{(r,1]} \mathbb{E}^0[\bar{f}_p g_r^\ell(Y_r)] \varphi(\ell) \nu(d\ell). \end{aligned} \quad (9.19)$$

Integrals without a domain are over $[0, 1]$. The same formula defines $J_{a,d}v$ for $v \in L^2(\nu)$. The scalar response operator $J_{a,d}^{\text{sc}}$ is the compression of $J_{a,d}$ to $\mathcal{H}_\nu$, that is, $J_{a,d}^{\text{sc}}v = \mathbf{1}_{(0,1)}J_{a,d}v$ for $v \in \mathcal{H}_\nu$.

At an atom of ν at p , the right-continuous value $a(p)$ includes the atom, and the strict future interval $(p, 1]$ excludes it. For $v \in L^2(\nu)$, the formula involves v only through $v(p)$ and through values at points that are integrated against ν , because $p \wedge r \in \{p, r\}$ and $r \wedge \ell \in \{r, \ell\}$. We show in the next paragraph that it defines $J_{a,d}v$ as an element of $L^2(\nu)$.

Kernel form and boundedness. We now rewrite (9.19). For $r > p$, conditioning on $\mathcal{G}_p^0$ and using (9.6), we have $\mathbb{E}^0[m_p h_p m_r] = \mathbb{E}^0[m_p^2 h_p]$, and $\nu((p, 1]) = 1 - a(p)$. Splitting the third term of (9.19) at $r = p$, we see that its part over $(p, 1]$ equals $-2(1 - a(p))\mathbb{E}^0[m_p^2 h_p]\varphi(p)$, which cancels the second part of the first coefficient. In the double integral, we split $\varphi(r \wedge \ell)$ into the regions $r < \ell$, $\ell < r$, and the diagonal $r = \ell$, which carries mass only at atoms of ν ; the first two regions give equal contributions by symmetry. In the last term, we exchange the order of integration. All integrands are bounded, and Fubini's theorem justifies these steps. The result is

$$(J_{a,d}v)(p) = k_{a,d}(p)v(p) + d \int \mathcal{K}_{a,d}(p, r)v(r) \nu(dr), \quad (9.20)$$

for every bounded Borel v , where

$$\begin{aligned} \mathcal{K}_{a,d}(p, r) = & -\mathbf{1}_{\{r > p\}}\mathbb{E}^0[m_p \partial_y g_p^r(Y_p)] - 2\mathbf{1}_{\{r \leq p\}}\mathbb{E}^0[m_p h_p m_r] - \mathbb{E}^0[\bar{f}_p \epsilon m_r] \\ & + \frac{1}{2}a(r)\mathbb{E}^0[\bar{f}_p m_r^2] + \int_{(r, 1]} \mathbb{E}^0[\bar{f}_p m_r m_\ell] \nu(d\ell) + \frac{1}{2}\nu(\{r\})\mathbb{E}^0[\bar{f}_p m_r^2] \\ & + \frac{1}{2} \int_{[0, r)} \mathbb{E}^0[\bar{f}_p g_\ell^r(Y_\ell)] \nu(d\ell). \end{aligned} \quad (9.21)$$

The term with $\nu(\{r\})$ is the diagonal part of the double integral. The kernel $\mathcal{K}_{a,d}$ is a bounded Borel function. Indeed, the first two terms are never both nonzero and are bounded by $2e^{\beta^2}$ by (9.18); since $|\bar{f}_p| \leq 1$, $|m| \leq 1$, and $0 \leq g \leq 1$, the third and fourth terms are bounded by 1 and 1/2, and the last three terms together by $\nu([0, 1]) \leq 1$. We conclude that $|\mathcal{K}_{a,d}| \leq 2e^{\beta^2} + 3$. By the Cauchy-Schwarz inequality, $\|\int \mathcal{K}_{a,d}(\cdot, r)v(r) \nu(dr)\|_{L^2(\nu)} \leq (2e^{\beta^2} + 3)\nu([0, 1])\|v\|_{L^2(\nu)}$. It follows that (9.20) extends to $v \in L^2(\nu)$, where it agrees with (9.19). Using $0 \leq k_{a,d} \leq d \leq \beta^2$, we find $\|J_{a,d}\|_{L^2(\nu) \rightarrow L^2(\nu)} \leq \beta^2(4 + 2e^{\beta^2})$. Compression does not increase the norm, which gives

$$\|J_{a,d}^{\text{sc}}\|_{\mathcal{H}_\nu \rightarrow \mathcal{H}_\nu} \leq C_\beta. \quad (9.22)$$

Transports of levels. Let φ be the restriction to $[0, 1]$ of a smooth function, with $\varphi(0) = \varphi(1) = 0$, and define $T_\theta(q) = q + \theta\varphi(q)$. We fix $\theta_0 \in (0, 1]$ such that $\theta_0\|\varphi'\|_\infty \leq 1/2$. For $|\theta| \leq \theta_0$, the map T_θ is an increasing bijection of $[0, 1]$ with $T'_\theta \in [1/2, 3/2]$. We replace the

covariance $d(q \wedge r)$ of X in (9.9) by $d T_\theta(q \wedge r)$ while keeping ν . Equivalently, we replace the coefficient d in (9.1) by

$$c_\theta(q) = d T'_\theta(q) = d(1 + \theta \varphi'(q)) \in \left[\frac{d}{2}, \frac{3d}{2}\right] \subset \left[\frac{\beta^2}{4}, 2\beta^2\right],$$

and keep a . We write u_θ and Y^θ for the resulting solution and diffusion, $u_{\theta,x} = \partial_x u_\theta$ and similarly for higher derivatives, and

$$\tilde{\Gamma}_\theta(p) = \mathbb{E}[\partial_x u_\theta(p, Y_p^\theta)^2]$$

for the resulting mean square magnetization. Let $a_\theta = a \circ T_\theta^{-1}$ be the distribution function of the image measure $T_{\theta\#}\mu$. Under the substitution $x' = T_\theta(q)$ in the level variable, the equation for u_θ becomes (9.1) with coefficient d and profile a_θ , and the diffusion Y^θ becomes a time change of the diffusion for a_θ . In formulas,

$$u_\theta(q, x) = u_{a_\theta, d}(T_\theta(q), x), \quad \tilde{\Gamma}_\theta(p) = \Gamma_{a_\theta, d}(T_\theta(p)), \quad (9.23)$$

that is, (u_θ, Y^θ) is the scalar law of $T_{\theta\#}\mu$ with coefficient d , observed at the transported levels. In particular, $\tilde{\Gamma}_0 = \Gamma_{a, d}$, and by Lemma 9.1(b), the spatial derivatives of u_θ of each order are bounded uniformly in $|\theta| \leq \theta_0$ and over profiles.

Lemma 9.6 (Response identity). *Let a be a profile, let $d \in [\beta^2/2, \beta^2]$, and let φ and $\tilde{\Gamma}_\theta$ be as above. Then, for every $p \in [0, 1]$, the function $\theta \mapsto \tilde{\Gamma}_\theta(p)$ is continuously differentiable near $\theta = 0$, and*

$$\frac{d}{d\theta} \tilde{\Gamma}_\theta(p) \Big|_{\theta=0} = (J_{a, d} \varphi)(p).$$

Proof. The proof has four steps. In Step 1, we show that u_θ and its spatial derivatives are continuously differentiable in θ ; the argument works with the integrated equations and never differentiates a . In Step 2, we compute the parameter derivative of u_θ and prove (9.28). In Step 3, we write $\tilde{\Gamma}_\theta(p)$ as a normalized expectation on a fixed Gaussian path space and differentiate it with Lemma 4.3. In Step 4, we match the resulting terms with those of (9.19). Throughout the proof, $|\theta| \leq \theta_0$, and constants may depend on β and φ .

Step 1: differentiability in the parameter. Fix θ , and for small $\delta \neq 0$ let $v = u_{\theta+\delta}$, $u = u_\theta$, and $D_\delta = (v - u)/\delta$. Since $c_{\theta+\delta} - c_\theta = \delta d\varphi'$, subtracting the two equations and dividing by δ , we find

$$\partial_q D_\delta + \frac{c_\theta}{2} (D_\delta)_{xx} + \frac{c_\theta a}{2} (v_x + u_x) (D_\delta)_x = -\frac{d\varphi'}{2} (v_{xx} + a v_x^2), \quad D_\delta(1, \cdot) = 0. \quad (9.24)$$

This is (9.3) with $c = c_\theta$, $V = 0$, and a drift and a source whose spatial derivatives of each order are bounded uniformly in δ and θ . Differentiating (9.24) j times in x , we see

that $\partial_x^j D_\delta$ solves an equation of the same form, with zeroth-order coefficient j times the spatial derivative of the drift, and with a source that involves only derivatives of D_δ of order less than j and spatial derivatives of the drift and of the source of (9.24). By (9.4) and induction on j , we have $\sup_{q,x} |\partial_x^j D_\delta| \leq C_j$ for every $j \geq 0$, uniformly in δ and θ . In particular, $v - u = \delta D_\delta$ and all its spatial derivatives are $O(|\delta|)$.

Next, we write the drift in (9.24) as $c_\theta a u_x + (c_\theta a/2)(v_x - u_x)$ and move the second part to the right side. For two increments δ and δ' , the difference $D_\delta - D_{\delta'}$ then solves (9.3) with drift $c_\theta a u_x$, $V = 0$, zero terminal value, and the source

$$\frac{d\varphi'}{2} \left[(v - v')_{xx} + a(v_x^2 - v_x'^2) \right] + \frac{c_\theta a}{2} \left[(v_x - u_x)(D_\delta)_x - (v'_x - u_x)(D_{\delta'})_x \right],$$

where $v' = u_{\theta+\delta'}$. By the bounds just proved, every spatial derivative of this source is bounded by $C_j(|\delta| + |\delta'|)$. By the same induction, we have $\sup_{q,x} |\partial_x^j (D_\delta - D_{\delta'})| \leq C_j(|\delta| + |\delta'|)$. It follows that, as $\delta \rightarrow 0$, the functions D_δ converge uniformly, together with all their spatial derivatives, to a function $\dot{u}_\theta$. Passing to the limit in the integrated form of (9.24), we find that $\dot{u}_\theta$ solves

$$\partial_q \dot{u}_\theta + \frac{c_\theta}{2} (\dot{u}_\theta)_{xx} + c_\theta a u_{\theta,x} (\dot{u}_\theta)_x = -\frac{d\varphi'}{2} (u_{\theta,xx} + a u_{\theta,x}^2), \quad \dot{u}_\theta(1, \cdot) = 0. \quad (9.25)$$

We conclude that u_θ and each of its spatial derivatives are differentiable in θ , with $\partial_\theta \partial_x^j u_\theta = \partial_x^j \dot{u}_\theta$, and that these derivatives are bounded uniformly in (q, x) and θ . To prove that these derivatives depend continuously on θ , we set $b_\theta = c_\theta a u_{\theta,x}$, $F_\theta = (d\varphi'/2)(u_{\theta,xx} + a u_{\theta,x}^2)$, and $R = \dot{u}_\theta - \dot{u}_{\theta'}$. Subtracting (9.25) at the two parameters, we obtain

$$\left(\partial_q + \frac{c_\theta}{2} \partial_{xx} + b_\theta \partial_x \right) R = -(F_\theta - F_{\theta'}) - \frac{c_\theta - c_{\theta'}}{2} (\dot{u}_{\theta'})_{xx} - (b_\theta - b_{\theta'}) (\dot{u}_{\theta'})_x, \quad R(1, \cdot) = 0.$$

Since $c_\theta - c_{\theta'} = (\theta - \theta') d\varphi'$, and since $u_\theta - u_{\theta'}$ and all its spatial derivatives are $O(|\theta - \theta'|)$ by the bound on $\partial_\theta \partial_x^j u_\theta$, every fixed-order spatial derivative of the right side is $O(|\theta - \theta'|)$. Using the same backward estimates, we have $\sup_{q,x} |\partial_x^j R| \leq C_j |\theta - \theta'|$, so $\partial_x^j \dot{u}_\theta$ depends continuously on θ , uniformly in (q, x) . The same argument applies without change to transports $q \mapsto q + \theta_1 \varphi_1(q) + \theta_2 \varphi_2(q)$ with several parameters, and gives continuous partial derivatives in each parameter.

Step 2: the variation of the solution. We write $\dot{u} = \dot{u}_0$ for the derivative at $\theta = 0$. By (9.25) and (9.4), with $Z = Y$, we have

$$\dot{u}(q, y) = \frac{d}{2} \int_q^1 \varphi'(r) \Psi_q(r; y) dr, \quad \Psi_q(r; y) = \mathbb{E}_{q,y} [h_r + a(r) m_r^2], \quad (9.26)$$

where $\mathbb{E}_{q,y}$ refers to the diffusion started from y at level q . To evaluate $\Psi_q(r; y)$, we use Itô's formula along Y . Differentiating (9.2) twice in x and using Itô's formula for $u_{xx}(r, Y_r)$, we find

$$dh_r = -d a(r) h_r^2 dr + \sqrt{d} u_{xxx}(r, Y_r) dB_r,$$

and by (9.6), $d(m_r^2) = dh_r^2 dr + 2\sqrt{d}m_r h_r dB_r$. The product rule for the continuous semimartingale m_r^2 and the deterministic function a of bounded variation gives $d(a(r)m_r^2) = a(r-)d(m_r^2) + m_r^2\mu(dr)$ on $(0, 1]$, and $a(r-) = a(r)$ except at countably many r . Adding the two differentials, we see that the drift terms $\mp da(r)h_r^2 dr$ cancel, and

$$d(h_r + a(r)m_r^2) = \sqrt{d}[u_{xxx}(r, Y_r) + 2a(r)m_r h_r] dB_r + m_r^2 \mu(dr).$$

Taking conditional expectations given $Y_q = y$, and recalling that $\mu = \nu$ on $(q, r] \subset (0, 1]$, we obtain

$$\Psi_q(r; y) - \Psi_q(q; y) = \int_{(q, r]} g_q^v(y) \nu(dv), \quad q \leq r \leq 1, \quad (9.27)$$

where $\Psi_q(q; y) = u_{xx}(q, y) + a(q)u_x(q, y)^2$.

We insert (9.27) into (9.26) and exchange the order of integration over $q < v \leq r \leq 1$. Note that $\int_q^1 \varphi'(r) dr = -\varphi(q)$ and $\int_v^1 \varphi'(r) dr = -\varphi(v)$, since $\varphi(1) = 0$. We obtain

$$\dot{u}(q, y) = -\frac{d}{2} \left[\varphi(q)(u_{xx}(q, y) + a(q)u_x(q, y)^2) + \int_{(q, 1]} \varphi(r)g_q^r(y) \nu(dr) \right]. \quad (9.28)$$

An atom at q belongs to $a(q)$ and not to the integral over $(q, 1]$. An atom at 1 belongs to the integral, but its factor $\varphi(1)$ vanishes.

The same computation holds at every base point θ , with Y^θ in place of Y . The diffusion coefficient d is then replaced by $c_\theta(r)$; the drift terms $\mp c_\theta(r)a(r)h_r^2 dr$ still cancel, and the source in (9.25) still carries the factor $d\varphi'$, which is the derivative of c_θ in θ . At $(q, y) = (0, 0)$, we have $\varphi(0) = 0$ and $\mathbb{E}[\partial_x u_\theta(r, Y_r^\theta)^2 \mid Y_0^\theta = 0] = \tilde{\Gamma}_\theta(r)$. This gives

$$\frac{d}{d\theta} u_\theta(0, 0) = -\frac{d}{2} \int \varphi(r) \tilde{\Gamma}_\theta(r) \nu(dr), \quad |\theta| < \theta_0. \quad (9.29)$$

Step 3: Gaussian differentiation at the fixed profile. Let B be a standard Brownian motion, and under the reference law let $X_q^\theta = B_{dT_\theta(q)}$. This Gaussian path has covariance $K_\theta(q, r) = dT_\theta(q \wedge r)$ and the law of $\int_0^q \sqrt{c_\theta(r)} dB_r$. Then Lemma 9.2, applied with the coefficient c_θ , implies that the joint law of (Y^θ, ϵ) has density (9.9) with u_θ in place of u , relative to the law of X^θ and counting measure. The factor $e^{-su_\theta(0, 0)}$ does not depend on the path and cancels from normalized expectations. With

$$\ell_\theta(X, \epsilon) = \epsilon X_1 - \int u_\theta(r, X_r) \nu(dr), \quad \Phi_\theta(X) = u_{\theta, x}(p, X_p)^2,$$

we have $\tilde{\Gamma}_\theta(p) = \mathbb{E}_\theta \Phi_\theta$, where $\mathbb{E}_\theta$ is the normalized expectation with weight $W_\theta = e^{\ell_\theta}$ in the notation of Lemma 4.3. At $\theta = 0$, the path X under $\mathbb{E}_0$ has the law of Y . As in Section 9.1, we write $m_r = u_x(r, X_r)$ and $h_r = u_{xx}(r, X_r)$, and we also write $u_{3, r} = u_{xxx}(r, X_r)$.

We check that the assumptions (A1)–(A4) of Lemma 4.3 hold, with $z = \theta$ in the compact interval $I = [-\theta_0/2, \theta_0/2]$. The Gaussian coordinates consist of one field path on one edge, with the deterministic clock $dT_\theta(q)$, which is continuous in (q, θ) , continuously differentiable in θ , and strictly increasing in q ; there are no finite coordinates, and the terminal spin ϵ is the counting variable. This is (A1). The kernels K_θ and $\dot{K}_\theta(q, r) = d\varphi(q \wedge r)$ are jointly continuous in (θ, q, r) , which is (A2). In the notation of Lemma 4.3, where M and Q denote derivative measures, the derivative measures of ℓ_θ and Φ_θ are

$$\begin{aligned} S &= \epsilon \delta_1 - u_{\theta,x}(r, X_r) \nu(dr), & M &= -\iota_\#(u_{\theta,xx}(r, X_r) \nu(dr)), \\ P &= 2u_{\theta,x}u_{\theta,xx}(p, X_p) \delta_p, & Q &= 2(u_{\theta,xx}^2 + u_{\theta,x}u_{\theta,xxx})(p, X_p) \delta_{(p,p)}, \end{aligned}$$

where $\iota(r) = (r, r)$. The measure M is carried by the diagonal, with an atom at (r, r) for each atom r of ν , as in Remark 4.4. Using Lemma 9.1(b), we see that these measures have total variation bounded uniformly in θ and X , and by Step 1, they are continuous in (θ, X) in total variation, where X carries the supremum norm. The maps $X \mapsto \ell_\theta(X, \epsilon)$ and $X \mapsto \Phi_\theta(X)$ are twice continuously differentiable, since u_θ has bounded continuous spatial derivatives. This is (A4). The observable Φ_θ takes values in $[0, 1]$, and the explicit parameter derivatives

$$\dot{\Phi}_\theta = 2u_{\theta,x}(p, X_p) \partial_x \dot{u}_\theta(p, X_p), \quad \dot{\ell}_\theta = - \int \dot{u}_\theta(r, X_r) \nu(dr)$$

are bounded and jointly continuous in (θ, X) by Step 1. Since $|u_{\theta,x}| \leq 1$ and $u_\theta(r, 0)$ is bounded, we have $|\ell_\theta| \leq C + 2 \sup_r |X_r|$, which is the linear growth required in (A3). We conclude that Lemma 4.3 applies to $\mathbb{E}_\theta \Phi_\theta$. By Remark 4.4, the right side of (4.5) is continuous in θ , which proves that $\theta \mapsto \tilde{\Gamma}_\theta(p)$ is continuously differentiable on $(-\theta_0/2, \theta_0/2)$. Each term on the right side is bounded by a constant depending only on β and φ ; in particular, the derivative is bounded uniformly in $p \in [0, 1]$ and θ . We use this uniform bound in the proof of Lemma 9.7.

Step 4: the terms of the identity. We evaluate the right side of (4.5) at $\theta = 0$ and divide it by d . For a signed measure R on $[0, 1]^2$, we have $d^{-1} \langle \dot{K}_0, R \rangle = \iint \varphi(q \wedge r) R(dq, dr)$. Note that, under $\mathbb{E}_0$, the centered observable is $\Phi_0 - \mathbb{E}_0 \Phi_0 = \bar{f}_p$. We group the contributions as follows.

1. The second derivative of the observable contributes

$$I_1 = \frac{1}{2d} \mathbb{E}_0 \langle \dot{K}_0, Q \rangle = \varphi(p) \mathbb{E}^0 [h_p^2 + m_p u_{3,p}].$$

2. The explicit derivative of the observable is $\dot{\Phi}_0 = 2m_p \partial_y \dot{u}(p, X_p)$. Differentiating (9.28) in y under the integral, which is allowed by (9.18), we find

$$I_2 = \frac{1}{d} \mathbb{E}_0 \dot{\Phi}_0 = -\varphi(p) \mathbb{E}^0 [m_p u_{3,p} + 2a(p) m_p^2 h_p] - \int_{(p,1]} \varphi(\ell) \mathbb{E}^0 [m_p \partial_y g_p^\ell(Y_p)] \nu(d\ell).$$

3. Since $\dot{K}_0$ is symmetric and $\dot{K}_0(1, p) = d\varphi(p)$, the score–observable terms give

$$\begin{aligned} I_3 &= \frac{1}{2d} \mathbb{E}_0 \langle \dot{K}_0, S \otimes P + P \otimes S \rangle = \frac{1}{d} \mathbb{E}_0 \langle \dot{K}_0, S \otimes P \rangle \\ &= 2\varphi(p) \mathbb{E}^0[m_p h_p \epsilon] - 2 \int \varphi(p \wedge r) \mathbb{E}^0[m_p h_p m_r] \nu(dr). \end{aligned}$$

4. The product of the scores contributes the following; the terminal–terminal term vanishes, since $\varphi(1) = 0$:

$$\begin{aligned} I_4 &= \frac{1}{2d} \mathbb{E}_0 [\bar{f}_p \langle \dot{K}_0, S \otimes S \rangle] \\ &= - \int \varphi(r) \mathbb{E}^0[\bar{f}_p \epsilon m_r] \nu(dr) + \frac{1}{2} \iint \varphi(r \wedge \ell) \mathbb{E}^0[\bar{f}_p m_r m_\ell] \nu(dr) \nu(d\ell). \end{aligned}$$

5. The Hessian of the logarithmic weight contributes

$$I_5 = \frac{1}{2d} \mathbb{E}_0 [\bar{f}_p \langle \dot{K}_0, M \rangle] = -\frac{1}{2} \int \varphi(r) \mathbb{E}^0[\bar{f}_p h_r] \nu(dr).$$

6. The explicit derivative of the weight is $\dot{\ell}_0 = - \int \dot{u}(r, X_r) \nu(dr)$. Evaluating it with (9.28), we find

$$\begin{aligned} I_6 &= \frac{1}{d} \text{Cov}_0(\Phi_0, \dot{\ell}_0) = \frac{1}{2} \int \varphi(r) \mathbb{E}^0[\bar{f}_p (h_r + a(r) m_r^2)] \nu(dr) \\ &\quad + \frac{1}{2} \int \nu(dr) \int_{(r,1]} \varphi(\ell) \mathbb{E}^0[\bar{f}_p g_r^\ell(Y_r)] \nu(d\ell). \end{aligned}$$

By (4.5), we have

$$\frac{1}{d} \frac{d}{d\theta} \tilde{\Gamma}_\theta(p) \Big|_{\theta=0} = I_1 + \cdots + I_6.$$

Conditioning on $\mathcal{G}_p^0$ and using (9.6), we have $\mathbb{E}^0[m_p h_p \epsilon] = \mathbb{E}^0[m_p^2 h_p]$. The following table records where each part of $I_1, \dots, I_6$ goes; “line i ” refers to the i -th line of (9.19).

Group	Part	Contribution
I_1	$\varphi(p) \mathbb{E}^0 h_p^2$	line 1, diagonal coefficient
	$\varphi(p) \mathbb{E}^0[m_p u_{3,p}]$	cancels with I_2
I_2	$-\varphi(p) \mathbb{E}^0[m_p u_{3,p}]$	cancels with I_1
	$-2a(p) \varphi(p) \mathbb{E}^0[m_p^2 h_p]$	line 1, diagonal coefficient
	integral over $(p, 1]$	line 1, second term
I_3	$2\varphi(p) \mathbb{E}^0[m_p h_p \epsilon] = 2\varphi(p) \mathbb{E}^0[m_p^2 h_p]$	line 1, diagonal coefficient
	integral of $\mathbb{E}^0[m_p h_p m_r]$	line 2, first term
I_4	integral of $\mathbb{E}^0[\bar{f}_p \epsilon m_r]$	line 2, second term
	double integral	line 3, second term
I_5	$-(1/2) \int \varphi(r) \mathbb{E}^0[\bar{f}_p h_r] \nu(dr)$	cancels with I_6
I_6	$(1/2) \int \varphi(r) \mathbb{E}^0[\bar{f}_p h_r] \nu(dr)$	cancels with I_5
	$(1/2) \int \varphi(r) a(r) \mathbb{E}^0[\bar{f}_p m_r^2] \nu(dr)$	line 3, first term
	iterated integral	line 4

In summary, the terms with $m_p u_{3,p}$ cancel, the covariance terms with h_r cancel, and the diagonal coefficient is $\mathbb{E}^0[h_p^2 - 2a(p)m_p^2 h_p + 2m_p^2 h_p] = \mathbb{E}^0[h_p^2 + 2(1 - a(p))m_p^2 h_p]$. The sum $I_1 + \dots + I_6$ is the right side of (9.19), including its signs, its factors $1/2$, and its strict future intervals. This completes the proof. $\square$

9.4 Domination by the scalar multiplier

We now show that the scalar response operator is dominated by its multiplier. We use the scalar Parisi functional with coefficient d ,

$$\mathcal{P}_d(a) = u_{a,d}(0,0) - \frac{d}{2} \int_0^1 q a(q) dq,$$

defined on the convex set of profiles. For $d = \beta^2$, it is the functional $\mathcal{P}_\beta$ of (1.1).

Lemma 9.7 (Scalar multiplier domination). *For every profile a and $d \in [\beta^2/2, \beta^2]$, the operator $J_{a,d}^{\text{sc}}$ is self-adjoint on $\mathcal{H}_\nu$, with norm at most C_β , and*

$$J_{a,d}^{\text{sc}} \preceq M_{k_{a,d}}, \quad (M_{k_{a,d}} v)(q) = k_{a,d}(q) v(q). \quad (9.30)$$

Proof. The bound on the norm is (9.22). The proof of the rest has five steps. We first recall that $\mathcal{P}_d$ is convex and compute its first variation. Then we follow $\mathcal{P}_d$ along a transport of levels and compute its second derivative with Lemma 9.6. Next, we bound the same second derivative from below, using convexity with the first variation frozen at a ; the comparison yields the inequality (9.30) for smooth directions. Two-parameter transports then give the symmetry, and a density argument extends both conclusions to $\mathcal{H}_\nu$. No minimizing property of a is used. We abbreviate $\Gamma = \Gamma_{a,d}$ and $k = k_{a,d}$, and $\langle \cdot, \cdot \rangle$ denotes the inner product of $L^2(\nu)$.

Step 1: convexity and the first variation. We first note that $\mathcal{P}_d$ is convex on profiles. Indeed, by [10, Theorem 2], applied with $\xi(q) = dq^2/2$ and zero external field, the value at $(0,0)$ of the solution with terminal value $\log \cosh x$ is a strictly convex function of the measure, and hence of its distribution function. This value differs from $u_{a,d}(0,0)$ by $\log 2$, and the penalty term of $\mathcal{P}_d$ is affine in a .

The first variation of $\mathcal{P}_d$ along segments is the scalar case of [4, (2.8)] (see also [44, Lemma 3.2] and [25, Theorem 2]): for profiles a and $\tilde{a}$,

$$\lim_{\vartheta \downarrow 0} \frac{\mathcal{P}_d(a + \vartheta(\tilde{a} - a)) - \mathcal{P}_d(a)}{\vartheta} = \frac{d}{2} \int_0^1 (\Gamma(q) - q)(\tilde{a}(q) - a(q)) dq. \quad (9.31)$$

The statement in [4] assumes a positive root mass. We note that the argument of Step 1 of the proof of Lemma 9.6, with the coefficient fixed and the profile varied, gives a direct

proof of (9.31) for all profiles. The difference quotients of $u_{a+\vartheta(\tilde{a}-a),d}$ converge, with all spatial derivatives, to the solution $\dot{u}$ of

$$\partial_q \dot{u} + \frac{d}{2} \dot{u}_{xx} + d a u_x \dot{u}_x = -\frac{d}{2} (\tilde{a} - a) u_x^2, \quad \dot{u}(1, \cdot) = 0,$$

and (9.4) gives

$$\dot{u}(0,0) = \frac{d}{2} \int_0^1 (\tilde{a}(q) - a(q)) \mathbb{E}^0 m_q^2 dq = \frac{d}{2} \int_0^1 (\tilde{a}(q) - a(q)) \Gamma(q) dq.$$

The penalty contributes $-(d/2) \int_0^1 q(\tilde{a}(q) - a(q)) dq$. Since $\mathcal{P}_d$ is convex, the difference quotient in (9.31) is nondecreasing in $\vartheta \in (0, 1]$. Comparing its value at $\vartheta = 1$ with its limit as $\vartheta \downarrow 0$, we obtain

$$\mathcal{P}_d(\tilde{a}) - \mathcal{P}_d(a) \geq \frac{d}{2} \int_0^1 (\Gamma(q) - q)(\tilde{a}(q) - a(q)) dq. \quad (9.32)$$

Step 2: the second derivative along a transport. Let φ , T_θ , θ_0 , and $a_\theta = a \circ T_\theta^{-1}$ be as in Section 9.3, and let $\Pi(\theta) = \mathcal{P}_d(a_\theta)$ for $|\theta| < \theta_0/2$. By (9.23) at $q = 0$, we have $u_{a_\theta,d}(0,0) = u_\theta(0,0)$. Using Fubini's theorem, we have

$$\int_0^1 x a_\theta(x) dx = \int \left(\int_{T_\theta(q)}^1 x dx \right) \mu(dq) = \frac{1}{2} - \frac{1}{2} \int T_\theta(q)^2 \mu(dq),$$

which gives

$$\Pi(\theta) = u_\theta(0,0) - \frac{d}{4} + \frac{d}{4} \int T_\theta^2 d\mu.$$

The derivative of the last term is $(d/2) \int T_\theta \varphi d\mu$, in which the atom $s\delta_0$ does not contribute, since $\varphi(0) = 0$. Combining this with (9.29), we obtain

$$\Pi'(\theta) = \frac{d}{2} \int \varphi(T_\theta - \tilde{\Gamma}_\theta) d\nu, \quad \Pi''(0) = \frac{d}{2} \langle \varphi, (I - J_{a,d}) \varphi \rangle. \quad (9.33)$$

For the second identity, we used Lemma 9.6 and differentiated under the integral over ν . This is justified by dominated convergence, since by Step 3 of the proof of Lemma 9.6, $\partial_\theta \tilde{\Gamma}_\theta(p)$ is continuous in θ and bounded uniformly in p and θ . By the same argument, Π'' is continuous near 0. At $\theta = 0$, the first identity in (9.33) is the first variation along transports in [9, Section 3, (19)], where it is credited to [79, Lemma 3.7, Eq. (3.34)].

Step 3: the lower bound from convexity. Let $G(y) = \int_y^1 (\Gamma(q) - q) dq$. Recall that $a_\theta(q) = \mu(\{y : T_\theta(y) \leq q\})$. By Fubini's theorem, we have $\int_0^1 (\Gamma(q) - q) a_\theta(q) dq = \int G(T_\theta y) \mu(dy)$. Applying (9.32) with $\tilde{a} = a_\theta$, we obtain

$$\Pi(\theta) - \Pi(0) \geq \frac{d}{2} \int (G(T_\theta y) - G(y)) \mu(dy) = \frac{d}{2} \int (G(T_\theta y) - G(y)) \nu(dy),$$

where the atom $s\delta_0$ does not contribute because $T_\theta(0) = 0$. Then Lemma 9.1(c) implies that G is twice continuously differentiable on $[0, 1]$, with

$$G'(y) = y - \Gamma(y), \quad G''(y) = 1 - k(y).$$

The function G'' is uniformly continuous and φ is bounded, so by Taylor's formula, we have

$$G(T_\theta y) - G(y) = \theta \varphi(y) G'(y) + \frac{\theta^2}{2} \varphi(y)^2 (1 - k(y)) + o(\theta^2),$$

uniformly in $y \in [0, 1]$. By (9.33) at $\theta = 0$, we have $\Pi'(0) = (d/2) \int \varphi G' d\nu$. Integrating the expansion against ν and inserting it into the previous inequality, we obtain

$$\Pi(\theta) - \Pi(0) - \theta \Pi'(0) \geq \frac{\theta^2}{2} \cdot \frac{d}{2} \int (1 - k) \varphi^2 d\nu + o(\theta^2).$$

On the other hand, Π is twice continuously differentiable near 0, so the left side equals $(\theta^2/2)\Pi''(0) + o(\theta^2)$. Dividing by $\theta^2/2$ and letting $\theta \rightarrow 0$, we find

$$\Pi''(0) \geq \frac{d}{2} \int (1 - k) \varphi^2 d\nu.$$

Comparing this with (9.33), we conclude that

$$\langle \varphi, J_{a,d} \varphi \rangle \leq \langle \varphi, k \varphi \rangle. \quad (9.34)$$

Note that Π need not be convex, since $\theta \mapsto a_\theta$ is not affine. The lower bound accounts for this through the term $1 - k = G''$.

Step 4: symmetry. We now take two functions φ_1 and φ_2 as above, and we consider the transport $T(q) = q + \theta_1 \varphi_1(q) + \theta_2 \varphi_2(q)$, which fixes both endpoints. Let $\tilde{\Gamma}_{\theta_1, \theta_2}$ and $\Pi(\theta_1, \theta_2)$ be defined as before, for small $|\theta_1|, |\theta_2|$. By Step 1 of the proof of Lemma 9.6 for several parameters, and by Lemma 4.3 applied in each parameter with the other one fixed, the function $\tilde{\Gamma}_{\theta_1, \theta_2}(p)$ has continuous first partial derivatives, bounded uniformly in p . The computation of Step 2 of that proof at every base point, with the coefficient derivative $d\varphi'_i$, yields

$$\partial_{\theta_i} \Pi = \frac{d}{2} \int \varphi_i (T - \tilde{\Gamma}_{\theta_1, \theta_2}) d\nu, \quad i = 1, 2.$$

It follows that Π has continuous second partial derivatives, and using Lemma 9.6 in the direction φ_2 (with $\theta_1 = 0$), we find that

$$\partial_{\theta_2} \partial_{\theta_1} \Pi(0, 0) = \frac{d}{2} \langle \varphi_1, (I - J_{a,d}) \varphi_2 \rangle.$$

By the equality of mixed partial derivatives, we conclude that this bilinear form is symmetric, that is, $\langle \varphi_1, J_{a,d} \varphi_2 \rangle = \langle \varphi_2, J_{a,d} \varphi_1 \rangle$.

Step 5: extension to the transport space. In Steps 3 and 4, we proved the domination inequality and the symmetry for restrictions to $[0, 1]$ of smooth functions vanishing at both endpoints. Such functions lie in $\mathcal{H}_\nu$, and they are dense in it. Indeed, let $v \in \mathcal{H}_\nu$, and let ψ_n be continuous functions with $\|\psi_n - v\| \rightarrow 0$, which exist because continuous functions are dense in $L^2(\nu)$. For $n \geq 3$, let χ_n be continuous with values in $[0, 1]$, vanishing at 0 and 1, and equal to one on $[1/n, 1 - 1/n]$. Then $\|\chi_n \psi_n - v\| \leq \|\psi_n - v\| + \|(1 - \chi_n)v\|$, and $\|(1 - \chi_n)v\| \rightarrow \|\mathbf{1}_{\{0,1\}}v\| = 0$ by dominated convergence, because v vanishes ν -almost everywhere on $\{0, 1\}$. Finally, let f be continuous with $f(0) = f(1) = 0$. If P is a polynomial with $|P - f| \leq \tau$ on $[0, 1]$, then the polynomial $P(x) - P(0)(1 - x) - P(1)x$ vanishes at both endpoints and is within 3τ of f ; by the Weierstrass theorem, f is a uniform limit of such polynomials. For $v \in \mathcal{H}_\nu$, we have $\langle v, J_{a,d}^{\text{sc}} v \rangle = \langle v, J_{a,d} v \rangle$, and both sides of (9.34) and of the symmetry identity are continuous on $L^2(\nu)$, since $J_{a,d}$ is bounded on $L^2(\nu)$ (see the proof of (9.22)) and $0 \leq k \leq \beta^2$. Both conclusions extend to $\mathcal{H}_\nu$ by continuity. Multiplication by k maps $\mathcal{H}_\nu$ into itself, and (9.30) follows. $\square$

9.5 The reference multiplier

A finite-dimensional minimizer may place mass outside the scalar support. For this reason, we control the multiplier at the scalar Parisi measure on all of $[0, 1]$. Recall from (7.6) that $\text{supp } \mu_\beta = [0, q_\beta]$ with $0 < q_\beta < 1$ [48, Theorem 1.1], and that μ_β has an atom at q_β . In particular, $\alpha_\beta = 1$ on $[q_\beta, 1]$.

Lemma 9.8 (Reference multiplier on the whole interval). *At the scalar Parisi measure,*

$$k_{\alpha_\beta, \beta^2}(q) = 1 \quad (0 \leq q \leq q_\beta), \quad k_{\alpha_\beta, \beta^2}(q) \leq e^{-\beta^2(q - q_\beta)/2} \quad (q_\beta \leq q \leq 1). \quad (9.35)$$

In particular, $\sup_{q \in [0, 1]} k_{\alpha_\beta, \beta^2}(q) \leq 1$.

Proof. We abbreviate $d = \beta^2$, $a = \alpha_\beta$, $k = k_{\alpha_\beta, d}$, and $\Gamma = \Gamma_{\alpha_\beta, d}$, and we write u and Y for the corresponding solution and diffusion. The proof has four steps. We first prove that $k = 1$ on the support. We then recall that a transformed density of Y_{q_β} is even and log-concave, show that this density evolves by the heat equation above q_β , and finally use a rearrangement inequality to bound k there.

Step 1: the support. By an identity of Talagrand [79], in the form stated in [9, Theorem 5] and [48, Lemma 2.6] (see also [44, Proposition 1.1] and [25, Proposition 1]), we have $\Gamma(q) = q$ for every $q \in \text{supp } \mu_\beta = [0, q_\beta]$. The function Γ of [48] coincides with ours, since u_x does not depend on the constant $\log 2$ in the terminal value. By Lemma 9.1(c), we have $\Gamma' = k$, which implies that $k = 1$ on $(0, q_\beta)$. Since k is continuous, we conclude that $k = 1$ on $[0, q_\beta]$.

Step 2: log-concavity of the transformed density. For $q \in (0, 1]$, let π_q be the density of Y_q , and set

$$z_q(x) = \pi_q(x) e^{-a(q)u(q,x)}, \quad V_q(x) = -\log z_q(x) = a(q)u(q,x) - \log \pi_q(x).$$

By (9.8), $z_q(x) dx$ is the distribution of the endpoint X_q under the finite measure obtained by weighting the law of the Gaussian path, without normalization, by $\exp\{-\int_{[0,q]} u(r, X_r) \mu_\beta(dr)\}$. By [48, Proposition 4.1], which applies to every probability measure on $[0, 1]$, including measures with atoms, the density π_q is smooth and positive, and the potential $a(q)u^L(q, x) - \log \pi_q(x)$ is smooth, even, and has second derivative at least $1/(dq)$; here $u^L = u - \log 2$ is the solution with the terminal value of [48], and $a(q)$ is the right-continuous value. The two potentials differ by the constant $a(q) \log 2$, which implies that

$$V_q''(x) \geq \frac{1}{dq} > 0, \quad V_q(-x) = V_q(x), \quad x \in \mathbb{R}.$$

Note that $V_q'(0) = 0$ by evenness. Then $V_q'(x) \geq x/(dq) > 0$ for $x > 0$. This implies that z_q is even, log-concave, and decreasing on $(0, \infty)$. We use this at $q = q_\beta$, where the right-continuous value is $a(q_\beta) = 1$, including the atom of μ_β at q_β ; in particular, $z_{q_\beta} = \pi_{q_\beta} e^{-u(q_\beta, \cdot)}$.

Step 3: heat evolution above the support. On $[q_\beta, 1]$, we have $a = 1$, and the solution of (9.1) is

$$u(q, x) = \log(2 \cosh x) + \frac{d(1-q)}{2}, \quad q_\beta \leq q \leq 1,$$

by [48, Lemma 2.4(i)]; this formula can also be checked by direct substitution, since $u_x = \tanh x$, $u_{xx} = \text{sech}^2 x$, $u_{xx} + u_x^2 = 1$, and $\partial_q u = -d/2$. We write G_v for the centered Gaussian density with variance v . By [48, Lemma 2.4(ii)], applied with mass $m = 1$ on $[q_\beta, 1)$, we have, for $q_\beta < q \leq 1$,

$$\pi_q(x) = e^{u(q,x)} \int G_{d(q-q_\beta)}(x-y) e^{-u(q_\beta,y)} \pi_{q_\beta}(y) dy;$$

the constants $\log 2$ in $u(q, x)$ and $u(q_\beta, y)$ cancel. This means that

$$z_q = G_{d(q-q_\beta)} * z_{q_\beta}, \quad q_\beta < q \leq 1. \quad (9.36)$$

Equivalently, z_q solves the heat equation $\partial_q z_q = (d/2) \partial_{xx} z_q$. To see this directly, we substitute $\pi_q = e^u z_q$ into the forward equation of Y on the plateau, which gives

$$\partial_q z_q + (\partial_q u) z_q = e^{-u} \left(\frac{d}{2} \partial_{xx} \pi_q - d \partial_x (u_x \pi_q) \right) = \frac{d}{2} \partial_{xx} z_q - \frac{d}{2} (u_{xx} + u_x^2) z_q;$$

the zeroth-order terms cancel because $u_{xx} + u_x^2 = 1$ and $\partial_q u = -d/2$.

On the plateau, we have $u_{xx} = \text{sech}^2$ and $\pi_q = e^{u(q, \cdot)} z_q = 2 \cosh(\cdot) e^{d(1-q)/2} z_q$. The multiplier is therefore

$$k(q) = d \int \pi_q(x) \text{sech}^4 x \, dx = 2d e^{d(1-q)/2} \int_{\mathbb{R}} z_q(x) \text{sech}^3 x \, dx, \quad q_\beta \leq q \leq 1. \quad (9.37)$$

Step 4: the rearrangement inequality. Let f and g be nonnegative, even, continuous functions that are nonincreasing on $[0, \infty)$. We claim that

$$\int f(x-y)g(x) \, dx \leq \int f(x)g(x) \, dx, \quad y \in \mathbb{R}. \quad (9.38)$$

This is the centered rearrangement inequality in one dimension, which also follows from Anderson's inequality [3, Theorem 1] by the layer-cake representation. We include a proof based on intersections of intervals. By the layer-cake representation and Tonelli's theorem,

$$\int f(x-y)g(x) \, dx = \int_0^\infty \int_0^\infty |(\{f > t\} + y) \cap \{g > t'\}| \, dt \, dt',$$

where $|\cdot|$ is Lebesgue measure. The superlevel sets $\{f > t\}$ and $\{g > t'\}$ are open intervals centered at 0. For two centered intervals I and I' , we have $|(I+y) \cap I'| \leq \min\{|I|, |I'|\} = |I \cap I'|$, and (9.38) follows.

We use (9.38) with $f = z_{q_\beta}$ and $g = \text{sech}^3$, which satisfy its hypotheses by Step 2. All the integrals below are finite, since $\text{sech} \leq 1$ and $\int z_{q_\beta} = \mathbb{E}[e^{-u(q_\beta, Y_{q_\beta})}] \leq 1/2$. By (9.36) and Tonelli's theorem, for $q_\beta < q \leq 1$,

$$\int z_q(x) \text{sech}^3 x \, dx = \int G_{d(q-q_\beta)}(y) \left(\int z_{q_\beta}(x-y) \text{sech}^3 x \, dx \right) dy \leq \int z_{q_\beta}(x) \text{sech}^3 x \, dx.$$

Inserting this into (9.37), and using (9.37) at $q = q_\beta$ and $k(q_\beta) = 1$ from Step 1, we find

$$k(q) \leq 2d e^{d(1-q)/2} \int z_{q_\beta}(x) \text{sech}^3 x \, dx = e^{-d(q-q_\beta)/2} k(q_\beta) = e^{-d(q-q_\beta)/2}, \quad q_\beta \leq q \leq 1.$$

Together with Step 1, this proves (9.35). $\square$

9.6 Stability at a minimizing profile

The isolation estimate of Proposition 7.6 compares a minimizer with the scalar trial profile $\gamma_\lambda^{\text{tr}} = \max\{\lambda, \gamma_\beta\}$, where $\gamma_\beta = \beta \alpha_\beta$. We first bound the distance between the trial profile and γ_β . Recall from (7.5) that $\gamma_\beta(q) \geq k_0 q$ for $0 \leq q \leq T_0$, with constants $k_0, T_0 > 0$ depending only on β . We claim that, for every $\lambda > 0$,

$$\|\gamma_\lambda^{\text{tr}} - \gamma_\beta\|_{L^1(0,1)} \leq C_\beta \lambda^2. \quad (9.39)$$

Indeed, $\gamma_\lambda^{\text{tr}} - \gamma_\beta = (\lambda - \gamma_\beta)_+$. Suppose first that $\lambda \leq k_0 T_0$. For $q \geq T_0$, since γ_β is nondecreasing, we have $\gamma_\beta(q) \geq \gamma_\beta(T_0) \geq k_0 T_0 \geq \lambda$, and the difference vanishes there. This gives

$$\|\gamma_\lambda^{\text{tr}} - \gamma_\beta\|_{L^1(0,1)} = \int_0^{T_0} (\lambda - \gamma_\beta(q))_+ dq \leq \int_0^{\lambda/k_0} (\lambda - k_0 q) dq = \frac{\lambda^2}{2k_0}.$$

If $\lambda > k_0 T_0$, the difference is at most λ everywhere, and $\lambda \leq \lambda^2/(k_0 T_0)$. Since $T_0 \leq 1/4 < 2$, both cases give (9.39) with $C_\beta = 1/(k_0 T_0)$.

Proposition 9.9 (Scalar stability at the minimizing profile). *There exists a constant $C_{\text{st}} \geq 1$, depending only on β (in particular, not on the minimizer or its atoms), with the following property. Let $N \geq 2$, $0 < \rho \leq 1$, $t = 1 - \rho$, and $0 < \lambda \leq \lambda_0$, and set $s = \lambda/\beta$. Let $\mu_N = s\delta_0 + \nu_N$ be a minimizer of $\mathcal{F}_{t,s}$ in (3.3), let $a_N(q) = \mu_N([0, q])$ be its profile, and let η be a trial budget for the minimizer βa_N of F_ρ , in the sense of (7.7). Let $d_N = \beta^2(1 - t/N)$. Then*

$$\|a_N - \alpha_\beta\|_{L^1(0,1)} \leq \min\{1, C_{\text{st}}[\rho^{-3/4}\eta^{1/4} + \lambda^2]\}. \quad (9.40)$$

Further, let $\varepsilon_N = \max\{0, \sup_{q \in [0,1]} k_{a_N, d_N}(q) - 1\}$. Then, on the scalar transport space $\mathcal{H}_{\nu_N}$ of μ_N ,

$$J_{a_N, d_N}^{\text{sc}} \preceq (1 + \varepsilon_N)I, \quad \varepsilon_N \leq C_{\text{st}}[\rho^{-3/4}\eta^{1/4} + \lambda^2 + N^{-1}]. \quad (9.41)$$

Proof. We use the energy normalization of Section 3, in which $\gamma = \beta a$ and $F_\rho(\gamma) = \beta^{-1}\mathcal{F}_{t,s}(\mu)$. The map $\mu \mapsto \beta a$ is a bijection onto the admissible profiles of Section 7, and multiplication by β^{-1} preserves minimizers. Hence $\gamma_N = \beta a_N$ is a minimizer of F_ρ , and η is a trial budget for it. The hypotheses of Proposition 7.6 hold, and (7.13) implies that $\|\gamma_N - \gamma_\lambda^{\text{tr}}\|_{L^1} \leq C\rho^{-3/4}\eta^{1/4}$, with a constant C depending only on β . Using the triangle inequality and (9.39), we have

$$\|a_N - \alpha_\beta\|_{L^1} \leq \beta^{-1}(\|\gamma_N - \gamma_\lambda^{\text{tr}}\|_{L^1} + \|\gamma_\lambda^{\text{tr}} - \gamma_\beta\|_{L^1}) \leq C_\beta[\rho^{-3/4}\eta^{1/4} + \lambda^2].$$

The distance is also at most one, since both functions take values in $[0, 1]$. This proves (9.40).

Since $N \geq 2$ and $0 \leq t < 1$, we have $d_N \in [\beta^2/2, \beta^2]$ and $|d_N - \beta^2| = \beta^2 t/N \leq \beta^2/N$. Then Lemma 9.3, applied with $(a, d) = (a_N, d_N)$ and $(\tilde{a}, \tilde{d}) = (\alpha_\beta, \beta^2)$, and Lemma 9.8 imply that, for every $q \in [0, 1]$, whether or not q lies in $\text{supp } \mu_\beta$,

$$k_{a_N, d_N}(q) \leq k_{\alpha_\beta, \beta^2}(q) + C_\beta(\|a_N - \alpha_\beta\|_{L^1} + N^{-1}) \leq 1 + C_\beta(\|a_N - \alpha_\beta\|_{L^1} + N^{-1}).$$

Together with (9.40), this yields the bound on ε_N in (9.41). Finally, since $k_{a_N, d_N} \leq 1 + \varepsilon_N$ pointwise, Lemma 9.7, applied to the profile a_N with $\nu = \nu_N$, implies that $J_{a_N, d_N}^{\text{sc}} \preceq M_{k_{a_N, d_N}} \preceq (1 + \varepsilon_N)I$ on $\mathcal{H}_{\nu_N}$. Note that only the multipliers of the two profiles are compared; the operator inequality is proved on the transport space of the minimizer itself. We choose $C_{\text{st}} \geq 1$ larger than the constants in the two bounds; it depends only on β . $\square$

PART IV

The cavity contraction

The cavity interpolation and its tree laws

This section and the next two prove Proposition 12.4. Suppose that at a minimizing profile, the scalar response operator $J_{a,d}^{\text{sc}}$ of Section 9 satisfies $J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I$ with $\varepsilon_N \leq \rho/4$. The proposition states that the self-overlap variance $V = \int \text{Var}(S_q) \nu(\mathrm{d}q)$ then obeys $(3\rho/4)V \leq C_{\text{cav}}[N^{-1} + A + J + H_3]$, where C_{cav} depends only on β . In Corollary 12.5, we deduce from it the bound on the defect $\mathcal{D}_N(\rho)$ that is used in Section 13.

The argument is a cavity interpolation on one site. We single out the last spin and move the variance of its couplings to the other $n = N - 1$ spins into its auxiliary field. At the end of the interpolation, the last spin is independent of the others. By exchangeability, V equals n/N times the integrated covariance between the squared conditional magnetization of the last spin and the self-overlap of the remaining spins, up to an error of order N^{-1} . This covariance vanishes at the independent endpoint, and its value for the full system is its derivative there plus a Taylor remainder. In Section 11, we show that this derivative is a multiple of the scalar response operator applied to the covariance of the self-overlap of the remaining spins, plus a nonpositive term and explicit errors (Lemma 11.2). We also bound the errors by A and J there (Lemma 11.3). In Section 12, we compare the variance of this self-overlap at the independent endpoint with V , with coefficient one (Lemma 12.1). We then express V through the expansion (Lemma 12.2), bound the cubic moment that controls the remainder by A , J , and H_3 (Lemma 12.3), and close the resulting inequality.

In this section, we set up the interpolation and the calculus that the next two sections use. In Section 10.1, we define the cavity law $\mathbb{P}_u$, $0 \leq u \leq 1$, identify its two endpoints, and record a variance bound (10.4) for the remaining spins. In Section 10.2, we represent products of conditional expectations by laws of finitely many descendants on marked trees. We show that the solution of the Parisi equation depends smoothly on the incident row and on the variance of the last field (Lemma 10.2). We compute the density of the tree laws (Lemma 10.3) and the responses (10.11) and (10.13) to these two parameters (Lemma 10.4). From them, we derive an exact formula for the derivative of a tree moment in u (Lemma 10.5) and a comparison of tree moments at different values of u (Corollary 10.6). In Section 10.3, we bound integrated products of three overlap defects by a cubic moment $\mathcal{M}_3(u)$ (Lemma 10.7), and deduce a bound on the Taylor remainder (Lemma 10.8).

We use the generalized law of Section 4, with its variance bound (Lemma 4.2) and its heat identity (Lemma 4.3). We also use the scalar law (9.1) with Lemma 9.4, and the differentiation and approximation arguments of [4, Sections 2.1, 2.3, and 2.6]. The results of this section are used as follows. The identity (10.11) and the density (10.9) with one leaf enter the proof of Lemma 11.2. The bound (10.4) and Corollary 10.6 enter the proofs of Lemmas 11.3 and 12.3. The bound of Lemma 10.7 enters the proof of Lemma 12.1, and Lemma 10.8 enters the proofs of Lemma 12.2 and Proposition 12.4.

The cavity method originates in the physics literature [51, 52]. We use its rigorous form on one site, the framework of differentiation, comparison, and contraction in [81, Section 1.6] (see Lemmas 1.6.3–1.6.4 and Propositions 1.6.5–1.6.6 there). An expansion on one site for hierarchical Gaussian laws appears in [56, Lemma 2.2]. There, a site is chosen by exchangeability and deleted, and its Gaussian kernels are interpolated to those of an independent scalar model on one site. The linear cavity kernel is half of a connected correlation of four spins of the removed site [56, (A8)], and a linear equation with this kernel is inverted under a gap condition on the scalar response in [56, Proposition 5.1]. The tree laws of Section 10.2 and their calculus are adapted from [54, Section 2]; we give the precise references there.

Standing assumptions. Throughout Part IV, we fix $N \geq 2$ and write $n = N - 1$. We also fix $0 < \rho \leq 1$, $t = 1 - \rho$, $c = \beta^2 \rho$, and $0 < s = \lambda/\beta \leq 1/2$. In this section, the profile $\mu = s\delta_0 + \nu$ is deterministic and otherwise arbitrary. The measure ν may have atoms anywhere in $[0, 1]$, including 0 and 1. No result of this section uses a minimizing property. This matters because at the independent endpoint, the remaining n spins form a system whose law is built from the profile of the full system, and this profile need not minimize any functional of the n -spin system. Constants C are absolute, and constants C_β depend only on β . Neither depends on N , ρ , s , the profile, or the positions and masses of its atoms. Note that the interpolation parameter u is not the solution u of (9.1); in this part of the paper, we write U^0 for the latter.

10.1 Removing the incident row

We write $\sigma = (\tau, \epsilon) \in \{-1, 1\}^n \times \{-1, 1\}$, so that $\epsilon = \sigma_N$ is the last spin. The couplings g_{iN} , $i \leq n$, between ϵ and τ form the *incident row*. Let $u \in [0, 1]$ be the interpolation parameter; t stays fixed. We set

$$\alpha = \beta^2 t \frac{n}{N}, \quad d = c + \alpha = \beta^2 \left(1 - \frac{t}{N}\right), \quad b(u) = d - \alpha u, \quad 0 \leq u \leq 1. \quad (10.1)$$

The second identity holds because $c + \alpha = \beta^2(\rho + t - t/N) = \beta^2(1 - t/N)$. Since $0 \leq t < 1$ and $N \geq 2$, we have $d \in [\beta^2/2, \beta^2]$. Further, $b(u) \geq b(1) = c > 0$ for every u .

For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$, let

$$\mathcal{E}(\sigma; x, y) = \mathcal{E}^-(\tau) + \sqrt{\frac{\alpha u}{n}} \epsilon \sum_{i=1}^n g_{iN} \tau_i + x \cdot \tau + y \epsilon,$$

$$\mathcal{E}^-(\tau) = \frac{\beta \sqrt{t}}{\sqrt{N}} \sum_{1 \leq i < j \leq n} g_{ij} \tau_i \tau_j,$$

where the couplings are independent standard Gaussian variables under the reference law. Let $U(q, x, y)$ solve

$$\partial_q U = -\frac{1}{2} \left\{ c(\Delta_x U + a(q)|\nabla_x U|^2) + b(u)(\partial_{yy} U + a(q)(\partial_y U)^2) \right\},$$

$$U(1, x, y) = \log \sum_{\sigma} e^{\mathcal{E}(\sigma; x, y)}. \quad (10.2)$$

At $u = 1$, each coupling of the row has the variance $\alpha/n = \beta^2 t/N$ of the full system, and the last field has the variance coefficient $b(1) = c$ of the full system. As u decreases to 0, the variance $\alpha u/n$ of each coupling in the row is removed, and the coefficient of the last field increases to $b(0) = d = c + \alpha$.

The completion of Section 3 adds to the terminal exponent the Gaussian term $G = \beta \sqrt{t} \sum_{i \leq N} z_i / \sqrt{2N}$, which does not depend on the spins. It changes U into $U + G$ and multiplies the tilt below by e^{sG} . Since G is independent of the couplings g_{ij} , this factor becomes, after normalization, an independent tilt of the variables z_i , and these variables do not enter the field or the spins. We omit it.

We construct the law $\mathbb{P}_u$, with expectation $\mathbb{E}_u$, as follows. We tilt the reference law of the couplings by $e^{sU(0,0,0)} / Z_s(u)$, where $Z_s(u) = \mathbb{E}_{\text{ref}} e^{sU(0,0,0)}$. Given the couplings, the field path $(X_q, Y_q) \in \mathbb{R}^n \times \mathbb{R}$ solves

$$\begin{aligned} dX_q &= c a(q) \nabla_x U(q, X_q, Y_q) dq + \sqrt{c} dB_q, & X_0 &= 0, \\ dY_q &= b(u) a(q) \partial_y U(q, X_q, Y_q) dq + \sqrt{b(u)} dB'_q, & Y_0 &= 0, \end{aligned}$$

where (B, B') is a standard Brownian motion in $\mathbb{R}^{n+1}$. The diffusion matrix is $\text{diag}(cI_n, b(u))$, and the drift is $a(q)$ times this matrix applied to $\nabla U(q, X_q, Y_q)$. Given the path, we draw σ with probabilities proportional to $e^{\mathcal{E}(\sigma; X_1, Y_1)}$. Let $\mathcal{G}_q$ be the σ -field generated by the couplings and the path up to level q .

The two endpoints. At $u = 1$, the law $\mathbb{P}_1$ is the tilted law of Section 3.3 for the full system, with the completion omitted as above. At $u = 0$, the row is absent, and the terminal value separates as $U(1, x, y) = U^-(1, x) + \log(2 \cosh y)$, where $U^-(1, x) = \log \sum_{\tau} e^{\mathcal{E}^-(\tau) + x \cdot \tau}$. Let U^- be the solution for the n -spin system with Hamiltonian $\mathcal{E}^-$, field coefficient c , and profile μ , and let U^0 solve (9.1) with coefficient d and profile μ .

We claim that $U = U^- + U^0$ at $u = 0$. For a step profile, each Hopf–Cole step of (10.2) factors, because the increments of x and of y are independent and $b(0) = d$. Applying the estimate [4, (2.19)] to each of the three equations, we extend the identity to every profile through the approximation of [4, Section 2.6]. The outer normalizer factors as $Z_s(0) = e^{sU^0(0,0)} \mathbb{E}_{\text{ref}} e^{sU^-(0,0)}$, and the deterministic factor $e^{sU^0(0,0)}$ cancels from the tilt. The drift of X involves only $\nabla_x U^-$, and that of Y only $\partial_y U^0$. The two Brownian motions are independent, and the terminal Gibbs probability is the product of $e^{\mathcal{E}^-(\tau) + X_1 \cdot \tau - U^-(1, X_1)}$ and $e^{\epsilon Y_1} / (2 \cosh Y_1)$. We conclude that $\mathbb{P}_0$ is the product of the tilted law of the n -spin system, with expectation $\mathbb{E}^-$, and the scalar law of Section 9 with coefficient d and profile μ , with expectation $\mathbb{E}^0$. Since $d \in [\beta^2/2, \beta^2]$, we may use the scalar quantities m_p , h_p , $f_p = m_p^2$, $\bar{f}_p = f_p - \mathbb{E}^0 f_p$, g_r^ℓ , and $\Gamma(p) = \Gamma_{a,d}(p) = \mathbb{E}^0 m_p^2$ of Section 9 for this scalar law, together with the bounds of Lemma 9.4.

The remaining spins. For the first n spins, we define

$$M_q = \mathbb{E}_u[\tau \mid \mathcal{G}_q], \quad A_q^- = \frac{|M_q|^2}{n}, \quad C_q^- = \text{Cov}_u(\tau \mid \mathcal{G}_q), \quad H_q^- = \nabla_x^2 U(q, X_q, Y_q).$$

We let $m_N(q) = \mathbb{E}_u[\epsilon \mid \mathcal{G}_q]$ be the conditional magnetization of the last spin, and we set

$$a_q^-(u) = \frac{\mathbb{E}_u \text{Tr}((C_q^-)^2)}{n^2}, \quad j_q^-(u) = \frac{\mathbb{E}_u[M_q^T C_q^- M_q]}{n^2}. \quad (10.3)$$

The *rest self-overlap* A_q^- is a random variable. It should not be confused with the trace term A of the full system.

We note that $\mathbb{P}_u$ is a generalized law in the sense of Section 4. It has N coordinates, the diffusion matrix $\Lambda = \text{diag}(cI_n, b(u))$, and independent Gaussian couplings with deterministic coefficients: $\beta\sqrt{t/N}$ for the pairs $i < j \leq n$, $(\alpha u/n)^{1/2}$ for the pairs (i, N) , and 0 on the diagonal. The matrix Λ is positive definite because $c > 0$ and $b(u) \geq c$. The coefficient $(\alpha u/n)^{1/2}$ vanishes at $u = 0$, which the generalized law allows. We use Lemma 4.2 for this law, with the first n coordinates selected. It implies that $M_q = \nabla_x U(q, X_q, Y_q)$ and $m_N(q) = \partial_y U(q, X_q, Y_q)$. Since principal submatrices preserve the order $a(q)C_q \preceq H_q \preceq C_q$, it also implies that $a(q)C_q^- \preceq H_q^- \preceq C_q^-$, and in particular that $0 \preceq H_q^- \preceq C_q^-$. Taking $k = \psi$ in the two inequalities in (4.2), we find that for every $u \in [0, 1]$ and every deterministic $\psi \in L^2(\nu)$,

$$\text{Var}_u \left(\int \psi(q) A_q^- \nu(dq) \right) \leq 4 \int \psi(q)^2 j_q^-(u) \nu(dq). \quad (10.4)$$

10.2 Laws of finitely many descendants

Derivatives in u produce products of conditional expectations at several levels. We represent them by additional terminal spins, called leaves, under a joint law in which

deleting a leaf does not change the law of the others. The hierarchical background of such laws is Ruelle's probability cascades [74]; see [59, Chapter 2]. An earlier hierarchical Gaussian change of measure appears in [18, Theorem 4.3, (4.25)–(4.26)]. Normalized finite replica trees of the present kind were developed in [54, Section 2], together with the differentiation of their moments by attaching new leaves and the resulting Taylor and comparison bounds. They are also used in [56, Sections 2.1–2.3]. In [54], the covariance differentiation formula is Proposition 2.5, the attachment rules and their uniform variation bounds are Lemma 2.6, the bounds on the levels at which new pairs split are Lemma 2.8, and the Taylor and comparison bounds are Lemma 2.9. We derive below the versions that we need for the cavity law, with its root tilt and an arbitrary profile, including its atoms.

Marked trees

Definition 10.1 (Marked trees and tree laws). A *marked tree* $\mathcal{T}$ with $k \geq 1$ leaves is a finite rooted tree whose vertices v carry deterministic levels $q_v \in [0, 1]$, with the following properties. The root has level 0, and levels increase strictly along every edge. Every vertex at a level in $(0, 1)$ has at least two outgoing edges. The vertices at level 1 have no outgoing edges; we call them terminal points. Finally, k labeled leaves $a = 1, \dots, k$ are attached to the terminal points, and every terminal point carries at least one leaf.

Given $u \in [0, 1]$, the *tree law* $\mathbb{P}_{\mathcal{T}, u}$, with expectation $\mathbb{E}_{\mathcal{T}, u}$, is constructed as follows. The couplings are drawn once from their tilted law under $\mathbb{P}_u$. Given the couplings, the field starts at the origin at the root and follows the diffusion of Section 10.1 along each edge. At each vertex, it continues along the outgoing edges in conditionally independent ways. Given the couplings and the field, the leaves carry conditionally independent spins $\sigma^a = (\tau^a, \epsilon_a)$. The spin σ^a is drawn with probabilities proportional to $e^{\mathcal{E}(\sigma; X_1^a, Y_1^a)}$, where (X_1^a, Y_1^a) is the field at the terminal point of a . We write $\mathbb{E}_{k, q, u}$ for the tree law of k leaves that all split at the level q .

The field is defined at every point v of $\mathcal{T}$, that is, at every vertex and at every point of an edge. We write (X_v, Y_v) for it and q_v for the level of v . Two points on different branches at the same level are different points, with different fields. If $k \geq 2$ and $0 < q < 1$, the tree of k leaves that split at q has one edge from the root to a vertex at level q , followed by k edges to k terminal points. If $k \geq 2$ and $q = 0$, the root has k outgoing edges, and if $k \geq 2$ and $q = 1$, the k leaves share one terminal point. For $k = 1$, the tree is the single path from level 0 to level 1, independently of q . After deleting leaves, we remove the branches that carry no remaining leaf and suppress the interior vertices with one outgoing edge.

Given the couplings and the field at a vertex, the continuations along different outgoing edges are conditionally independent, and their laws are normalized. Integrating out a leaf yields the tree law of the remaining leaves. We call this the *deletion property*. In particular, every leaf has law $\mathbb{P}_u$. We represent products of conditional expectations

given $\mathcal{G}_q$ by leaves that split at q . For leaves $a \neq b$, let q_{ab} be the level of their last common ancestor, so that $q_{ab} = 1$ if they share a terminal point. We define their *rest overlap*, their *overlap defect*, and the product of the overlap defect with their last spins by

$$R_{ab}^- = \frac{\tau^a \cdot \tau^b}{n}, \quad \Delta_{ab} = R_{ab}^- - q_{ab}, \quad D_{ab} = \epsilon_a \epsilon_b \Delta_{ab}. \quad (10.5)$$

Since $|R_{ab}^-| \leq 1$, we have $|D_{ab}| = |\Delta_{ab}| \leq 2$. For instance, the four spins of $\mathbb{E}_{4,p,u}$ are conditionally independent given $\mathcal{G}_p$, each with the conditional law of σ given $\mathcal{G}_p$. Therefore $\mathbb{E}_u[m_N(p)^2 A_p^-] = \mathbb{E}_{4,p,u}[\epsilon_1 \epsilon_2 R_{34}^-]$.

Regularity in the row and in the last-field variance

To differentiate tree moments, we temporarily replace the row by a deterministic coupling $\epsilon \sum_i \theta_i \tau_i$ with $\theta \in \mathbb{R}^n$, and the coefficient $b(u)$ by a number $b > 0$. We write $\Lambda_b = \text{diag}(cI_n, b)$. We keep the other couplings g_{ij} , $i < j \leq n$, fixed. Under $\mathbb{P}_u$ and under the tree laws, conditionally on the couplings, the field and the spins have the law of this fixed system with $\theta = (\alpha u/n)^{1/2}(g_{1N}, \dots, g_{nN})$ and $b = b(u)$.

Lemma 10.2 (Regularity in the parameters). *Let μ be a deterministic profile, and let $U = U(q, x, y)$ solve (10.2) with the row replaced by $\epsilon \sum_i \theta_i \tau_i$ and $b(u)$ replaced by $b > 0$. Then U has derivatives of every order in (θ, x, y) . Each derivative of positive order is bounded by a constant depending only on N and its order. This bound is uniform in the couplings, θ , b , (q, x, y) , and the profile, and each first derivative is bounded by one in absolute value. The function U and all these derivatives are jointly continuous in $(b, q, \theta, x, y) \in (0, \infty) \times [0, 1] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$. Finally, writing ∇ for the gradient in (x, y) , we have for $0 \leq r \leq h \leq 1$ that*

$$U(r, x, y) - U(h, x, y) = \frac{1}{2} \int_r^h [\text{Tr}(\Lambda_b \nabla^2 U) + a(q) \nabla U^T \Lambda_b \nabla U](q, x, y) dq. \quad (10.6)$$

Proof. The proof has three steps. We first prove the bounds for step profiles, using the density of the future of the field. We then pass to arbitrary profiles with the approximation of [4, Section 2.6], and finally we prove the continuity in q and b .

Step 1: step profiles. Let μ be a step profile. The coefficients of (10.2) do not depend on (x, y) . Translating the starting point of the field by (x', y') has the same effect as adding $x' \cdot \tau + y' \epsilon$ to the terminal exponent. It follows that every derivative in (θ, x, y) is a derivative at $z = 0$ with respect to parameters z_j of the terminal exponent

$$\mathcal{E}_z(\sigma; x, y) = \mathcal{E}(\sigma; x, y) + \sum_j z_j \eta_j(\sigma),$$

where the features η_j range over $\tau_i \epsilon$, τ_i ($i \leq n$), and ϵ . Note that $|\eta_j| \leq 1$. Let U_z be the corresponding solution. We fix a level q , refine the grid of the profile so that it contains

q , and fix the field (x, y) at level q . On a grid interval $[q_\ell, q_{\ell+1})$ with $a = a_\ell > 0$, the Hopf–Cole form of the equation is $e^{a_\ell U_z(q_\ell)} = \mathbb{E} e^{a_\ell U_z(q_{\ell+1})}$, where the expectation is over a Gaussian increment of covariance $(q_{\ell+1} - q_\ell) \Lambda_b$. The transition of the diffusion is this Gaussian transition tilted by $e^{a_\ell (U_z(q_{\ell+1}) - U_z(q_\ell))}$. If $a_\ell = 0$, it is the Gaussian transition. We multiply these factors and the Gibbs probability $e^{\mathcal{E}_z(\sigma) - U_z(1)}$, as in [4, Section 2.1, (2.1)]. The coefficient of U_z at a grid point $q_\ell > q$ is $a_{\ell-1} - a_\ell = -\nu(\{q_\ell\})$, and at level 1 it is $a(1-) - 1 = -\nu(\{1\})$. We find that given the field at level q , the future increments of the field and the spin have the density

$$L_z = \exp \left\{ \mathcal{E}_z(\sigma; Z_1) - a(q) U_z(q, Z_q) - \int_{(q,1]} U_z(r, Z_r) \nu(dr) \right\} \quad (10.7)$$

with respect to driftless Gaussian increments of covariance Λ_b and counting measure on σ . Here Z_r is the field at level r .

We write $\mathbb{E}_z$ for the expectation under this density and $\ell_z = \log L_z$. On a grid interval with $a_\ell > 0$, we differentiate $e^{a_\ell U_z(q_\ell)} = \mathbb{E} e^{a_\ell U_z(q_{\ell+1})}$. This shows that $\partial_{z_j} U_z(q_\ell)$ is the expectation of $\partial_{z_j} U_z(q_{\ell+1})$ under the transition of the diffusion, and the same holds if $a_\ell = 0$. At level 1, $\partial_{z_j} U_z(1)$ is the Gibbs mean of η_j . Iterating, we obtain

$$\partial_{z_j} U_z(q) = \mathbb{E}_z \eta_j, \quad \text{so that} \quad |\partial_{z_j} U_z| \leq 1. \quad (10.8)$$

Since $a(q) + \nu((q, 1]) = 1$, this implies that $|\partial_{z_j} \ell_z| \leq 2$. Further, if all derivatives of U_z of some order $m' \geq 2$ are bounded by $C_{m'}$, then all derivatives of ℓ_z of order m' are bounded by $C_{m'}$. Since the reference law in (10.7) does not depend on z , we may differentiate (10.8) under the expectation. For directions $j, j_1, \dots, j_m$, we have

$$\partial_{z_{j_1}} \cdots \partial_{z_{j_m}} \partial_{z_j} U_z(q) = \mathbb{E}_z \left[\eta_j \sum_{\pi} \prod_{B \in \pi} \partial_B \ell_z \right],$$

where π runs over the partitions of $\{1, \dots, m\}$ and ∂_B is the derivative in the directions j_i , $i \in B$. By induction on m , we find that every derivative of order $m+1$ in the coordinates z_j is bounded by a constant depending only on m . There are $2n+1$ coordinates, and this proves the bounds of the lemma for step profiles.

Step 2: arbitrary profiles. As in [4, Section 2.6.2], we approximate μ by step profiles $\mu_m = s\delta_0 + \nu_m$ that converge weakly to μ and whose cumulative functions a_m converge to a in $L^1(0, 1)$. Let U_m be the corresponding solutions. Following the proof of [4, (2.19)], with cI_N replaced by Λ_b and with the first derivatives bounded by one, we obtain

$$\sup |U_m - U| \leq \frac{\text{Tr } \Lambda_b}{2} \|a_m - a\|_{L^1(0,1)},$$

where the supremum is over q, x, y , the couplings, and θ . For a function f of one real variable and every $h > 0$, we have $\|f'\|_\infty \leq 2\|f\|_\infty/h + (h/2)\|f''\|_\infty$. Using this inequality

in one direction at a time for $U_m - U_{m'}$ and then for its derivatives, together with the bounds of Step 1, we find that every derivative of U_m in (θ, x, y) converges uniformly. This implies that U has derivatives of every order in (θ, x, y) , that they are the uniform limits of those of U_m , and that they satisfy the same bounds. Each U_m satisfies (10.6) with a_m in place of a . As $a_m \rightarrow a$ in L^1 , the identity passes to the limit.

Step 3: continuity. For a step profile, the derivatives of U in (θ, x, y) satisfy, on each grid interval, the equations obtained by differentiating (10.2). By Step 1, their right sides are bounded by constants depending only on N, c, b , and the order. More precisely, for every multi-index κ in (θ, x, y) , including $\kappa = 0$, differentiation and the product rule give

$$\sup |\partial_q D^\kappa U_m| \leq C_{N,|\kappa|} (nc + b)$$

on each grid interval, where $C_{N,|\kappa|}$ depends only on N and $|\kappa|$. The derivatives are continuous across the grid points, so the same Lipschitz bound in q holds on $[0, 1]$, independently of the step approximation, and it passes to the limit. We also find that U and its derivatives in (θ, x, y) are Lipschitz in q , with constants depending only on N , the order, and $nc + b$. For the continuity in b , let U_b and $U_{b'}$ be the solutions for $b, b' > 0$ and the same profile. The difference $U_b - U_{b'}$ vanishes at $q = 1$, because both solutions have the same terminal value. Subtracting their equations, we see that it satisfies a linear backward equation with diffusion matrix $\Lambda_{b'}$, a bounded drift, and the source term $\frac{1}{2}(b - b')[\partial_{yy}U_b + a(\partial_y U_b)^2]$. By Step 1, we have $|\partial_{yy}U_b| \leq 2$ and $|\partial_y U_b| \leq 1$. The backward representation then gives $\sup |U_b - U_{b'}| \leq \frac{3}{2}|b - b'|$, first for step profiles and then, by Step 2, for μ . Using the interpolation inequality of Step 2 again, we find that every derivative in (θ, x, y) is continuous in b , uniformly in the other variables. These derivatives are also Lipschitz in (θ, x, y) and in q , so they are jointly continuous. $\square$

The density of a tree law

Let $\mathcal{T}$ be a marked tree. For a vertex v , let l_v be the number of its outgoing edges if $q_v < 1$, and the number of leaves at v if v is a terminal point. We define the measure ω on the points of $\mathcal{T}$ as follows. On the interior of an edge whose endpoints have levels $r < h$, it is the copy of the restriction of ν to (r, h) . A vertex v carries the atom

$$\omega_v = \omega(\{v\}) = \nu(\{q_v\}) + (l_v - 1)a(q_v).$$

For example, the root carries $\nu(\{0\}) + (l - 1)a(0)$ if it has l outgoing edges, and a terminal point with l leaves carries $\nu(\{1\}) + l - 1$. Below, the notation $\sum_v \omega_v F(v)$ means $\int F d\omega$; for an atomic profile, it is an ordinary sum. The *driftless Gaussian field* on $\mathcal{T}$ starts at the origin at the root and has independent increments along the edges. On each edge, its increments have covariance Λ times the length of the level interval.

Lemma 10.3 (Density of a tree law). *Let μ be a deterministic profile, $u \in [0, 1]$, and $\mathcal{T}$ a marked tree with k leaves. Consider the product of the original Gaussian law of the couplings, the driftless Gaussian field on $\mathcal{T}$, and counting measure on the spins. With respect to this reference measure, the joint law of the couplings, the field on $\mathcal{T}$, and the spins of the leaves under $\mathbb{P}_{\mathcal{T}, u}$ has density $Z_s(u)^{-1} \mathcal{D}_{\mathcal{T}}$, where*

$$\mathcal{D}_{\mathcal{T}} = \exp \left\{ \sum_{a=1}^k \mathcal{E}(\sigma^a; X_1^a, Y_1^a) - \int_{\mathcal{T}} U(q_v, X_v, Y_v) \omega(\mathrm{d}v) \right\}, \quad \omega \geq 0, \quad \omega(\mathcal{T}) = k - s. \quad (10.9)$$

For fixed couplings, the conditional density of the field and the spins is $e^{-sU(0,0,0)} \mathcal{D}_{\mathcal{T}}$.

Proof. The proof has three steps: we compute the likelihood of one edge, collect the coefficients of U at the vertices, and compute the mass of ω . We fix the couplings. By Lemma 10.2, the first derivatives of U in (x, y) are bounded by one. The drift $a(q)\Lambda \nabla U$ is bounded, with $a(q)^2 |\Lambda^{1/2} \nabla U|^2 \leq nc + b(u)$, and Novikov's condition holds.

Step 1: one edge. Let Z be the driftless field on an edge with levels $r < h$, so that $\mathrm{d}Z_q = \Lambda^{1/2} \mathrm{d}B_q$. By Girsanov's formula, the law of the diffusion on this edge, started from Z_r , has density

$$\exp \left\{ \int_r^h a(q) \nabla U(q, Z_q) \cdot \mathrm{d}Z_q - \frac{1}{2} \int_r^h a(q)^2 |\Lambda^{1/2} \nabla U(q, Z_q)|^2 \mathrm{d}q \right\}$$

with respect to the law of Z . By Lemma 10.2, U is smooth in (x, y) with bounded derivatives and satisfies (10.6), whose integrand is bounded; in particular, U is Lipschitz in q . We may therefore apply Itô's formula (for instance after smoothing U in q), and by (10.6), we have

$$\mathrm{d}U(q, Z_q) = \nabla U(q, Z_q) \cdot \mathrm{d}Z_q - \frac{a(q)}{2} |\Lambda^{1/2} \nabla U(q, Z_q)|^2 \mathrm{d}q.$$

In the last identity, we used that the time derivative of U is given by (10.2) for almost every q . We conclude that the exponent of the density is $\int_r^h a(q) \mathrm{d}U(q, Z_q)$. The process $U_q = U(q, Z_q)$ is continuous, and a is right-continuous with Stieltjes measure ν on $(r, h]$. We integrate by parts. The countably many jumps of a do not affect the integral against U_q , and we obtain $\int_r^h a \mathrm{d}U_q = a(h)U_h - a(r)U_r - \int_{(r,h]} U_q \nu(\mathrm{d}q)$. As $a(h) - \nu(\{h\}) = a(h-)$, the edge likelihood is

$$\exp \left\{ a(h-)U(h, X_h, Y_h) - a(r)U(r, X_r, Y_r) - \int_{(r,h]} U(q, X_q, Y_q) \nu(\mathrm{d}q) \right\}, \quad (10.10)$$

where we wrote the field as (X, Y) . We write the values at the endpoints separately, so that this formula also covers atoms of ν at r and h .

Step 2: the vertices. We multiply the root tilt $e^{sU(0,0,0)} / Z_s(u)$, the edge likelihoods (10.10), and the Gibbs probabilities $e^{\mathcal{E}(\sigma^a; X_1^a, Y_1^a) - U(1, X_1^a, Y_1^a)}$ of the leaves. The open edges contribute $-\int U dv$ along each edge. At a vertex v with $0 < q_v < 1$, the incoming edge contributes $a(q_v -)U(v)$, and the l_v outgoing edges contribute $-l_v a(q_v)U(v)$. The coefficient of $U(v)$ is

$$a(q_v -) - l_v a(q_v) = -[\nu(\{q_v\}) + (l_v - 1)a(q_v)] = -\omega_v.$$

At the root, the tilt and the l outgoing edges give $s - la(0) = -[\nu(\{0\}) + (l - 1)a(0)]$, since $a(0) = s + \nu(\{0\})$. At a terminal point with l leaves, the incoming edge and the l Gibbs probabilities give $a(1 -) - l = -[\nu(\{1\}) + l - 1]$, since $a(1) = 1$. All these coefficients are nonpositive, which shows that $\omega \geq 0$, and the product is $Z_s(u)^{-1} \mathcal{D}_{\mathcal{T}}$ with $\mathcal{D}_{\mathcal{T}}$ as in (10.9). The edge likelihoods and the Gibbs probabilities are normalized. Integrating $\mathcal{D}_{\mathcal{T}}$ over the driftless field and the spins thus gives $e^{sU(0,0,0)}$, and this proves the claim about fixed couplings. Averaging over the original law of the couplings, we obtain $Z_s(u)$; in particular, the normalizer is $Z_s(u)$ for every tree.

Step 3: the mass. For a single leaf, ω is a copy of ν on the path from the root to the terminal point, and $\omega(\mathcal{T}) = \nu([0, 1]) = 1 - s$. Every marked tree is obtained from a single path by successively attaching $l - 1 \geq 1$ new continuations at a point at some level q , each running to level 1. Such a step adds $l - 1$ leaves. It adds $(l - 1)a(q)$ to the weight of the point, which is or becomes a vertex, and $(l - 1)\nu((q, 1])$ on the new continuations. Since $a(q) + \nu((q, 1]) = 1$, the mass of ω increases by $l - 1$, and $\omega(\mathcal{T}) = 1 - s + (k - 1) = k - s$. $\square$

We also see from the density (10.9) that an *extended* tree, obtained by attaching new leaves to $\mathcal{T}$, carries the weights of its own shape, and that by deletion it induces $\mathbb{P}_{\mathcal{T}, u}$ on the original leaves. The same proof applies when the row is a deterministic coupling $\epsilon \sum_i \theta_i \tau_i$ and $b(u)$ is replaced by a number $b > 0$, as in Lemma 10.2. We then take the density with respect to the Gaussian law of the couplings g_{ij} , $i < j \leq n$, and the normalizer is the expectation of $e^{sU(0,0,0)}$ over these couplings.

Responses to the row and to the last-field variance

Let the field (X, Y) and the spin σ have the law of the fixed system of Lemma 10.2, and let $\mathcal{G}_q$ be generated by the field up to level q . We set

$$\pi_i(q) = \mathbb{E}[\tau_i \epsilon \mid \mathcal{G}_q], \quad m_N(q) = \mathbb{E}[\epsilon \mid \mathcal{G}_q],$$

and we evaluate the derivatives of U at (q, X_q, Y_q) , writing for instance $\partial_{\theta_i} U(q)$.

Lemma 10.4 (Parameter responses). *Let μ be a deterministic profile, and let U , π_i , and m_N be as above, for fixed couplings, $\theta \in \mathbb{R}^n$, and $b > 0$. Then for every $q \in [0, 1]$ and $i \leq n$,*

$$\partial_{\theta_i} U(q) = \pi_i(q), \quad \partial_{\theta_i}^2 U(q) = 1 - a(q)\pi_i(q)^2 - \int_{(q, 1]} \mathbb{E}[\pi_i(r)^2 \mid \mathcal{G}_q] \nu(dr), \quad (10.11)$$

$$\partial_y U(q) = m_N(q), \quad \partial_{yy} U(q) = 1 - a(q)m_N(q)^2 - \int_{(q,1]} \mathbb{E}[m_N(r)^2 \mid \mathcal{G}_q] \nu(dr). \quad (10.12)$$

Further, the derivative of U in b exists and, for each fixed q , is jointly continuous in (b, θ, x, y) . It satisfies

$$\partial_b U(q) = \frac{1}{2} \left[(1-q) - \int_{(q,1]} (r-q) \mathbb{E}[m_N(r)^2 \mid \mathcal{G}_q] \nu(dr) \right]. \quad (10.13)$$

In particular, $0 \leq \partial_b U(q) \leq (1-q)/2$.

For a single row, (10.11) and (10.12) are the conditional form of [4, (2.10)]. The identity (10.13) is the conditional, one-coordinate form of the field-variance computation following [4, (2.11)]. Under $\mathbb{P}_u$ and the tree laws, the identities hold with $\mathcal{G}_q$ generated by the couplings and the field, with $\theta = (\alpha u/n)^{1/2}(g_{1N}, \dots, g_{nN})$ and $b = b(u)$.

Proof. The proof has four steps. We prove (10.11) and (10.12) for step profiles with the density (10.7), and we pass to arbitrary profiles with the exact density of Lemma 10.3. We then treat the derivative in b in the same two stages.

Step 1: step profiles. Let μ be a step profile, and let η be one of the features $\tau_i \epsilon$ and ϵ , with parameter z . By (10.8), we have $\partial_z U(r) = \mathbb{E}[\eta \mid \mathcal{G}_r]$ for every r , and we write $\omega(r)$ for this conditional expectation. Differentiating (10.8) once more under the density (10.7), we find

$$\partial_z^2 U(q) = \mathbb{E}[\eta \partial_z \ell_z \mid \mathcal{G}_q] = \mathbb{E}[\eta^2 \mid \mathcal{G}_q] - a(q)\omega(q)^2 - \int_{(q,1]} \mathbb{E}[\eta \partial_z U(r) \mid \mathcal{G}_q] \nu(dr).$$

For $r \geq q$, by the tower property, we have $\mathbb{E}[\eta \partial_z U(r) \mid \mathcal{G}_q] = \mathbb{E}[\omega(r) \mathbb{E}[\eta \mid \mathcal{G}_r] \mid \mathcal{G}_q] = \mathbb{E}[\omega(r)^2 \mid \mathcal{G}_q]$. Since $\eta^2 = 1$, this proves (10.11) for $\eta = \tau_i \epsilon$ and (10.12) for $\eta = \epsilon$. Note that $\mu([0, q]) = a(q)$ and $\mu = \nu$ on $(q, 1]$. We record the result in the form

$$\partial_z^2 U(q) = 1 - \int_{[0,1]} \mathbb{E}[\omega(q \vee r)^2 \mid \mathcal{G}_q] \mu(dr). \quad (10.14)$$

Step 2: arbitrary profiles. Let $\mu_m \Rightarrow \mu$ be the step profiles of the proof of Lemma 10.2. Applying the proof of Lemma 10.3 to the single edge from q to 1 followed by the Gibbs probability, we see that (10.7) with $z = 0$ is the density of the future given the field at q for every profile. Its exponent equals $\mathcal{E}(\sigma; Z_1) - \int_{[0,1]} U(q \vee r, Z_{q \vee r}) \mu(dr)$, and the integrand is continuous in r . The reference law does not depend on the profile, $U_m \rightarrow U$ uniformly, and $\mu_m \Rightarrow \mu$. Then the densities for μ_m converge pointwise to the density for μ . Writing L_m for the density for μ_m , we have

$$\log L_m = \mathcal{E}(\sigma; Z_1) - U_m(1, Z_1) + \int_{[0,1]} [U_m(1, Z_1) - U_m(q \vee r, Z_{q \vee r})] \mu_m(dr),$$

since $\mu_m([0, 1]) = 1$. The first term is nonpositive, because $U_m(1, \cdot)$ is the logarithm of a sum that contains $e^{\mathcal{E}(\sigma; \cdot)}$. Using this, the bounds on the first derivatives in Lemma 10.2, and the Lipschitz bound in q from Step 3 of its proof, we find

$$\log L_m \leq C_N \left(nc + b + \sup_{q \leq r \leq 1} |Z_r| \right),$$

where C_N depends only on N . The same bound holds for the density for μ , and the densities have a common Gaussian integrable envelope. They therefore converge in L^1 of the reference law, and expectations of bounded functions converge. The left side of (10.8) converges by Lemma 10.2, so $\partial_z U(r) = \mathbb{E}[\eta \mid \mathcal{G}_r]$ also for μ . Next, we consider the integrand $r \mapsto \mathbb{E}[\omega(q \vee r)^2 \mid \mathcal{G}_q]$ in (10.14). Its value for μ_m differs from its value for μ by at most the L^1 distance of the two densities plus twice the supremum of the difference of the first derivatives. It therefore converges uniformly in r . The limit is continuous in r , because the first derivatives are jointly continuous by Lemma 10.2. Together with the weak convergence of μ_m , this proves (10.14), and hence (10.11) and (10.12), for μ . In these identities, an atom of ν at q is counted in $a(q)$, and an atom at 1 is counted in the integral.

Step 3: the derivative in b for step profiles. Let μ be a step profile and $q_\ell \leq q < q_{\ell+1}$, with $a = a_\ell$ on $[q_\ell, q_{\ell+1})$. Suppose first that $a_\ell > 0$, and let $\phi = e^{a_\ell U}$. On this interval, ϕ solves the backward heat equation $\partial_q \phi + \frac{1}{2} \text{Tr}(\Lambda_b \nabla^2 \phi) = 0$. In particular, $\phi(q)$ is the Gaussian average of $\phi(q_{\ell+1})$ over an increment of covariance $(q_{\ell+1} - q)\Lambda_b$. We differentiate in b , using the Gaussian heat identity in the last coordinate, and divide by $a_\ell \phi(q)$. We obtain

$$\partial_b U(q) = \mathbb{E}[\partial_b U(q_{\ell+1}) \mid \mathcal{G}_q] + \frac{q_{\ell+1} - q}{2} \mathbb{E}[(\partial_{yy} U + a_\ell (\partial_y U)^2)(q_{\ell+1}) \mid \mathcal{G}_q].$$

The function $\partial_{yy} \phi = a_\ell (\partial_{yy} U + a_\ell (\partial_y U)^2) \phi$ also solves the backward heat equation. Further, the transition of the field from q to $r \leq q_{\ell+1}$ is the Gaussian transition tilted by $\phi(r)/\phi(q)$. It follows that $\mathbb{E}[(\partial_{yy} U + a_\ell (\partial_y U)^2)(r) \mid \mathcal{G}_q]$ does not depend on $r \in [q, q_{\ell+1}]$, and that the last term equals $\frac{1}{2} \int_q^{q_{\ell+1}} \mathbb{E}[(\partial_{yy} U + a(\partial_y U)^2)(r) \mid \mathcal{G}_q] dr$. The case $a_\ell = 0$ is the same with $\phi = U$. Iterating over the grid intervals and using $\partial_b U(1) = 0$, we find

$$\partial_b U(q) = \frac{1}{2} \int_q^1 \mathbb{E}[\partial_{yy} U(r) + a(r) m_N(r)^2 \mid \mathcal{G}_q] dr, \quad (10.15)$$

where we used that $\partial_y U(r) = m_N(r)$.

Step 4: the derivative in b for arbitrary profiles. By (10.12), the integrand in (10.15) lies in $[0, 1]$. Let $R(b)$ be the right side of (10.15) for μ . For the step profiles μ_m , the right side of (10.15) converges to $R(b)$ at each fixed b . Indeed, the densities converge as in Step 2, the derivatives converge uniformly, and $a_m \rightarrow a$ in $L^1(0, 1)$. These expressions are bounded

by 1/2. By dominated convergence in the identity $U_m(b_2) - U_m(b_1) = \int_{b_1}^{b_2} \partial_b U_m db$, we obtain $U(b_2) - U(b_1) = \int_{b_1}^{b_2} R(b) db$. We write the last-field increments of the reference law as $\sqrt{b}$ times a fixed Brownian motion. Using Lemma 10.2 and dominated convergence, we then see that, for each fixed q , R is jointly continuous in (b, θ, x, y) . It follows that $\partial_b U$ exists, equals R , and satisfies (10.15) for μ . Inserting (10.12) at level r and using the tower property, we have

$$\partial_b U(q) = \frac{1}{2} \int_q^1 \left(1 - \int_{(r,1]} \mathbb{E}[m_N(\ell)^2 \mid \mathcal{G}_q] \nu(d\ell) \right) dr.$$

By Fubini's theorem, we have $\int_q^1 dr \int_{(r,1]} F(\ell) \nu(d\ell) = \int_{(q,1]} (\ell - q) F(\ell) \nu(d\ell)$ for bounded $F \geq 0$, and (10.13) follows. Finally, the integral in (10.13) lies between 0 and $(1 - q)\nu((q, 1]) \leq 1 - q$, which gives $0 \leq \partial_b U(q) \leq (1 - q)/2$. $\square$

Differentiation in the interpolation parameter

For a point v of $\mathcal{T}$, we write $\hat{v}$ and $\check{v}$ for new leaves whose paths leave the path through v at the level q_v . If v is an interior point of an edge, the extended tree has a new vertex at v . If v is a vertex, the new edge starts at v , and if v is a terminal point, the new leaf is attached there. Distinct new leaves are conditionally independent, also when they are attached at the same point. For $r > q_v$, we write $\hat{r}$ and $\check{r}$ for two new leaves that follow a common new branch from v to the level r and split there. Given the couplings and the field of $\mathcal{T}$, a new branch from v is a continuation of the diffusion from the field at v , and it is independent of everything else. Let F be a bounded function of the spins of $\mathcal{T}$, and let G, G' be bounded functions of one spin. Then

$$\begin{aligned} \mathbb{E}[F G(\sigma^{\hat{v}})] &= \mathbb{E}_{\mathcal{T}, u} [F \mathbb{E}[G(\sigma) \mid \mathcal{G}_{q_v}](v)], \\ \mathbb{E}[F G(\sigma^{\hat{r}}) G'(\sigma^{\check{r}})] &= \mathbb{E}_{\mathcal{T}, u} \left[F \mathbb{E}[\mathbb{E}[G(\sigma) \mid \mathcal{G}_r] \mathbb{E}[G'(\sigma) \mid \mathcal{G}_r] \mid \mathcal{G}_{q_v}](v) \right]. \end{aligned} \quad (10.16)$$

Here the left sides are taken under the extended trees. A conditional expectation under $\mathbb{P}_u$ given $\mathcal{G}_{q_v}$ is a function of the couplings and of the field at level q_v , and (v) means that it is evaluated at the couplings and at the field at v . In the same way, two new leaves $\hat{v}$ and $\check{v}$ attached at v and w represent the product of the two conditional expectations, evaluated at v and at w .

Let φ_{ab} be a quantity attached to each pair of distinct leaves of an extended tree. For a bounded function F of the spins of $\mathcal{T}$, we define the five-pattern sum

$$\begin{aligned} \mathfrak{S}_\varphi(F) &= \sum_{a \neq b} \mathbb{E}(F \epsilon_a \epsilon_b \varphi_{ab}) - 2 \sum_{a, v} \omega_v \mathbb{E}(F \epsilon_a \epsilon_{\hat{v}} \varphi_{a\hat{v}}) + \sum_{v, w} \omega_v \omega_w \mathbb{E}(F \epsilon_{\hat{v}} \epsilon_{\check{w}} \varphi_{\hat{v}\check{w}}) \\ &\quad + \sum_v \omega_v a(q_v) \mathbb{E}(F \epsilon_{\hat{v}} \epsilon_{\check{v}} \varphi_{\hat{v}\check{v}}) + \sum_v \omega_v \int_{(q_v, 1]} \mathbb{E}(F \epsilon_{\hat{r}} \epsilon_{\check{r}} \varphi_{\hat{r}\check{r}}) \nu(dr). \end{aligned}$$

The sums run over leaves a, b of $\mathcal{T}$ and points v, w of $\mathcal{T}$, and the weights ω are those of $\mathcal{T}$. We take each expectation under the extended tree that contains the leaves involved. We use this sum with $\varphi_{ab} = R_{ab}^-$, with $\varphi_{ab} = q_{ab}$, and with $\varphi_{ab} = \Delta_{ab}$. Then $\mathfrak{S}_\Delta = \mathfrak{S}_{R^-} - \mathfrak{S}_q$, and $\epsilon_a \epsilon_b \Delta_{ab} = D_{ab}$. The following table lists the five patterns. Its last column gives the level q_{ab} of the pair.

pair of leaves	new leaves	coefficient	level of the pair
two leaves $a \neq b$ of $\mathcal{T}$	none	1	q_{ab}
a leaf a and a leaf attached at v	$\hat{v}$	$-2\omega_v$	$q_{a\hat{v}}$
leaves attached at v and at w	$\hat{v}, \hat{w}$	$\omega_v \omega_w$	$q_{\hat{v}\hat{w}}$
two leaves attached at v	$\hat{v}, \check{v}$	$a(q_v)\omega_v$	q_v
two leaves splitting at $r > q_v$	$\hat{r}, \check{r}$	$\omega_v \nu(dr)$	r

Lemma 10.5 (Differentiation under the tree law). *Let μ be a deterministic profile, $\mathcal{T}$ a marked tree with k leaves, and F a bounded function of the spins of its leaves. Then $u \mapsto \mathbb{E}_{\mathcal{T},u} F$ belongs to $C^2([0,1])$, with one-sided derivatives at the endpoints, and*

$$\begin{aligned} \frac{d}{du} \mathbb{E}_{\mathcal{T},u} F &= \frac{\alpha}{2} \left[\sum_{a \neq b} \mathbb{E}(FD_{ab}) - 2 \sum_{a,v} \omega_v \mathbb{E}(FD_{a\hat{v}}) + \sum_{v,w} \omega_v \omega_w \mathbb{E}(FD_{\hat{v}\hat{w}}) \right. \\ &\quad + \sum_v \omega_v a(q_v) \mathbb{E}(FD_{\hat{v}\check{v}}) + \sum_v \omega_v \int_{(q_v,1]} \mathbb{E}(FD_{\hat{r}\check{r}}) \nu(dr) \\ &\quad \left. + s \mathbb{E}_{\mathcal{T},u} F \int \mathbb{E}_{2,r,u} D_{12} \nu(dr) \right]. \end{aligned} \quad (10.17)$$

The sums run over leaves a, b and points v, w of $\mathcal{T}$, with the weighted sums interpreted as integrals against ω as above. The first five expectations are taken under extended trees, and the last term is a product of expectations over independent copies. The absolute coefficients in the bracket sum to at most $4k^2 + 2k + 1$. The second derivative is obtained by applying the same formula to each tree moment on its right side, and the product rule to independent products of moments.

In the notation above, (10.17) reads $\frac{d}{du} \mathbb{E}_{\mathcal{T},u} F = \frac{\alpha}{2} [\mathfrak{S}_\Delta(F) + s \mathbb{E}_{\mathcal{T},u} F \int \mathbb{E}_{2,r,u} D_{12} \nu(dr)]$.

Proof. If $\alpha = 0$, the law does not depend on u , and both sides vanish. Let $\alpha > 0$. The proof has five steps. We check that Lemma 4.3 applies, and we compute the contributions of the row and of the last field to the derivative of the unnormalized moment. We then compute the derivative of the normalizer. Finally, we bound the coefficients and prove the C^2 regularity.

Step 1: the heat identity. By Lemma 10.3, we have $\mathbb{E}_{\mathcal{T},u} F = Z_s(u)^{-1} \mathbb{E}_{\text{ref}}[F \mathcal{D}_{\mathcal{T}}]$, where $\mathbb{E}_{\text{ref}}$ denotes integration against the reference measure of that lemma. We write the row as $\theta = (\alpha u/n)^{1/2} (g_{iN})_{i \leq n}$, and we use Lemma 4.3 on $I = [0,1]$ with the following data. The finite coordinates are the entries of θ , with variances $\alpha u/n$. The field paths are the

$n + 1$ coordinates of the driftless field on the edges of $\mathcal{T}$. For an edge e with endpoint levels $r_e < h_e$, we use the local coordinate $v \in [0, 1]$, and the clocks are

$$T_{e,i,u}(v) = c(h_e - r_e)v \quad (i \leq n), \quad T_{e,N,u}(v) = b(u)(h_e - r_e)v.$$

The couplings g_{ij} , $i < j \leq n$, do not depend on u , and the spins are the counting variable. The row, the Brownian motions of the different edges and components, and the couplings g_{ij} are mutually independent. We take $\ell = \log \mathcal{D}_{\mathcal{T}}$ and $\Phi = F$. The elements of the space $\mathcal{C}$ of Lemma 4.3 need not agree at the common endpoints of edges. To define ℓ on the whole space, we evaluate the field at each nonroot vertex at the endpoint of its incoming edge. At the root, we choose a fixed outgoing edge e_0 and evaluate the field at the initial point of e_0 ; its coordinates vanish on every path of the driftless field. This agrees with (10.9) on the paths of the driftless field. The derivative measures below, including the atom of ω at the root, are those of this extension. We check the hypotheses (A1)–(A4) of Lemma 4.3.

- (A1): the variances $\alpha u/n$ are positive for $u > 0$. The clocks are affine in u , and they are strictly increasing in v , because $c > 0$, $b(u) \geq c$, and every edge has positive length.
- (A2): the covariance of the last component at two field points v, w is $b(u)h(v, w)$, where $h(v, w)$ is the level of the last common point of the paths from the root to v and to w . For the other components, it is $c h(v, w)$. Hence K is affine in u , with $\dot{K} = (\alpha/n)I_n$ on the row block, $\dot{K} = -\alpha h$ on the last-field block, and $\dot{K} = 0$ elsewhere. These kernels are continuous.
- (A3): $\Phi = F$ is bounded and $\dot{\Phi} = 0$. The explicit derivative of ℓ comes from $b(u)$ in U , so that $\dot{\ell} = -b'(u) \sum_v \omega_v \partial_b U(v) = \alpha \sum_v \omega_v \partial_b U(v)$. The levels q_v and the finite measure ω are fixed. Using the fixed-level continuity and uniform bound of Lemma 10.4, together with dominated convergence against ω , we see that $\dot{\ell}$ is continuous in (u, G) ; it is bounded by $\alpha k/2$. If we treat each rest coupling g_{ij} as an additional source parameter with bounded feature $\beta \sqrt{t/N} \tau_i \tau_j$ in Step 1 of the proof of Lemma 10.2, we obtain joint continuity of U and its spatial derivatives also in the rest couplings. The fixed-level response formula of Lemma 10.4 then gives joint continuity of $\partial_b U$ in those couplings by dominated convergence, using the uniform bound of that lemma. Thus ℓ and $\dot{\ell}$ are jointly continuous in (u, G, g) , as required by (A3). The first derivatives of U in the row and in the field are bounded by one. Its derivative in g_{ij} , $i < j \leq n$, is the conditional mean of $\beta \sqrt{t/N} \tau_i \tau_j$, by the argument of Step 1 of the proof of Lemma 10.2. Integrating (10.2) in q and using that the first spatial derivatives of U are bounded by one and the second by two (Step 1 of that proof), we have $|U(q, 0, 0)| \leq N \log 2 + \frac{3}{2}(nc + b(u))$ when all couplings, the row, and the field vanish. Since $\mathcal{E}$ is linear in the couplings, the row, and the field, we conclude that $|\ell| \leq C(1 + |g| + |\theta| + \|G^{\text{path}}\|_{\infty})$, where C depends only on N, k , and β . Here g collects the couplings g_{ij} , $i < j \leq n$.

- (A4): the score of ℓ consists of the atoms $\tau_i^a \epsilon_a$ (row), τ^a and ϵ_a (field at the terminal point of a), and of $-\nabla U(v) \omega(\mathrm{d}v)$, where ∇ is the gradient in the row and the field. Since $\mathcal{E}$ is linear in these variables, the Hessian of ℓ is $-\nabla^2 U(v) \omega(\mathrm{d}v)$, placed at the pairs of indices of the point v . It has a row–row block $-\partial_{\theta_i \theta_j} U(v) \omega(\mathrm{d}v)$, mixed row–field blocks such as $-\partial_{\theta_i y} U(v) \omega(\mathrm{d}v)$ and $-\partial_{\theta_i x_j} U(v) \omega(\mathrm{d}v)$ with their transposes, and field–field blocks such as $-\partial_{yy} U(v) \omega(\mathrm{d}v)$, placed at the diagonal pairs of index points of v . By Lemma 10.2, these derivatives of U are bounded and jointly continuous in (b, q, θ, x, y) . As the points (q_v, X_v, Y_v) stay in a compact set when G varies in a bounded set, the measures are continuous in total variation in (u, G) . Their total variations are at most $C_N k$. The measures for Φ vanish.

The constants in these hypotheses depend on N and k . We use them only to justify the differentiation, and the identity below has coefficients that do not depend on N . Since $\Phi = F$ depends neither on the Gaussian coordinates nor on u , (4.5) and (4.6) imply that

$$\frac{\mathrm{d}}{\mathrm{d}u} \mathbb{E}_{\mathcal{T}, u} F = \frac{\partial_u \mathbb{E}_{\text{ref}}[F \mathcal{D}_{\mathcal{T}}]}{Z_s(u)} - \mathbb{E}_{\mathcal{T}, u} F \frac{\mathrm{d}}{\mathrm{d}u} \log Z_s(u), \quad (10.18)$$

where the derivative of the unnormalized moment is

$$\frac{\partial_u \mathbb{E}_{\text{ref}}[F \mathcal{D}_{\mathcal{T}}]}{Z_s(u)} = \mathbb{E}_{\mathcal{T}, u} \left[F \left(\dot{\ell} + \frac{\alpha}{2n} \sum_{i=1}^n ((\partial_{\theta_i} \ell)^2 + \partial_{\theta_i}^2 \ell) - \frac{\alpha}{2} \langle h, S_y \otimes S_y + M_{yy} \rangle \right) \right].$$

Here S_y and M_{yy} are the score and Hessian measures of ℓ in the last-field coordinates.

Step 2: the row. By (10.9) and (10.11), we have

$$\partial_{\theta_i} \ell = \sum_a \tau_i^a \epsilon_a - \sum_v \omega_v \pi_i(v), \quad \partial_{\theta_i}^2 \ell = - \sum_v \omega_v \partial_{\theta_i}^2 U(v),$$

where $\pi_i(v)$ and $\partial_{\theta_i}^2 U(v)$ are evaluated at the couplings and at the field at v . We expand the square and insert (10.11). The diagonal terms $(\tau_i^a \epsilon_a)^2 = 1$ contribute k , and the constant 1 in (10.11) contributes $-\omega(\mathcal{T})$. By Lemma 10.3, these constants add up to s . The other terms of $(\partial_{\theta_i} \ell)^2 + \partial_{\theta_i}^2 \ell$ are

$$\begin{aligned} & \sum_{a \neq b} \tau_i^a \epsilon_a \tau_i^b \epsilon_b - 2 \sum_{a, v} \omega_v \tau_i^a \epsilon_a \pi_i(v) + \sum_{v, w} \omega_v \omega_w \pi_i(v) \pi_i(w) \\ & + \sum_v \omega_v a(q_v) \pi_i(v)^2 + \sum_v \omega_v \int_{(q_v, 1]} \mathbb{E}[\pi_i(r)^2 \mid \mathcal{G}_{q_v}](v) \nu(\mathrm{d}r). \end{aligned}$$

By (10.16), inside $\mathbb{E}_{\mathcal{T}, u}[F \cdot]$ we may replace $\pi_i(v)$ by $\tau_i^{\hat{v}} \epsilon_{\hat{v}}$, and $\pi_i(v) \pi_i(w)$ by the product for $\hat{v}$ and $\hat{w}$. We may also replace $\pi_i(v)^2$ by the product for $\hat{v}$ and $\check{v}$, and $\mathbb{E}[\pi_i(r)^2 \mid \mathcal{G}_{q_v}](v)$ by the product for $\hat{r}$ and $\check{r}$. Averaging over i then replaces $\tau_i^a \tau_i^b$ by R_{ab}^- . We obtain

$$\frac{\alpha}{2n} \sum_{i=1}^n \mathbb{E}_{\mathcal{T}, u} [F ((\partial_{\theta_i} \ell)^2 + \partial_{\theta_i}^2 \ell)] = \frac{\alpha}{2} [s \mathbb{E}_{\mathcal{T}, u} F + \mathfrak{S}_{R^-}(F)]. \quad (10.19)$$

Step 3: the last field. By (10.9) and (10.12), the score and Hessian measures of ℓ in the last-field coordinates are

$$S_y = \sum_a \epsilon_a \delta_a - m_N(v) \omega(\mathrm{d}v), \quad M_{yy} = -\mathrm{diag}_\#(\partial_{yy}U(v) \omega(\mathrm{d}v)),$$

where δ_a is the unit mass at the terminal point of a . Here $\mathrm{diag}_\#$ is the image of a measure under $v \mapsto (v, v)$, so that atoms of ω stay at exact diagonal pairs. Let a leaf stand for its terminal point. Then the kernel h satisfies $h(a, a) = 1$, $h(a, b) = q_{ab}$, $h(a, v) = q_{a\hat{v}}$, $h(v, w) = q_{\hat{v}\hat{w}}$, and $h(v, v) = q_v$. Expanding $\langle h, S_y \otimes S_y \rangle$ as in Step 2 and replacing $m_N(v)$ by $\epsilon_{\hat{v}}$ with (10.16), we find

$$\begin{aligned} -\frac{\alpha}{2} \mathbb{E}_{\mathcal{T}, u} [F \langle h, S_y \otimes S_y \rangle] &= -\frac{\alpha}{2} \left[k \mathbb{E}_{\mathcal{T}, u} F + \sum_{a \neq b} q_{ab} \mathbb{E}(F \epsilon_a \epsilon_b) \right. \\ &\quad \left. - 2 \sum_{a, v} \omega_v q_{a\hat{v}} \mathbb{E}(F \epsilon_a \epsilon_{\hat{v}}) + \sum_{v, w} \omega_v \omega_w q_{\hat{v}\hat{w}} \mathbb{E}(F \epsilon_{\hat{v}} \epsilon_{\hat{w}}) \right]. \end{aligned}$$

The bracket consists of the constant $k \mathbb{E}_{\mathcal{T}, u} F$ and the first three patterns of $\mathfrak{S}_q(F)$. The Hessian measure contributes $-\frac{\alpha}{2} \langle h, M_{yy} \rangle = \frac{\alpha}{2} \sum_v \omega_v q_v \partial_{yy}U(v)$, and we recall that $\dot{\ell} = \alpha \sum_v \omega_v \partial_b U(v)$. Together, these two terms give $\frac{\alpha}{2} \sum_v \omega_v [q_v \partial_{yy}U(v) + 2\partial_b U(v)]$. By (10.12) and (10.13), we have

$$q_v \partial_{yy}U(v) + 2\partial_b U(v) = 1 - q_v a(q_v) m_N(v)^2 - \int_{(q_v, 1]} r \mathbb{E}[m_N(r)^2 \mid \mathcal{G}_{q_v}](v) \nu(\mathrm{d}r).$$

In the integral, we combined $q_v \int_{(q_v, 1]}$ from (10.12) with $\int_{(q_v, 1]} (r - q_v)$ from (10.13). After the replacements (10.16), the constant gives $\frac{\alpha}{2} \omega(\mathcal{T}) \mathbb{E}_{\mathcal{T}, u} F = \frac{\alpha}{2} (k - s) \mathbb{E}_{\mathcal{T}, u} F$. The other two terms are the last two patterns of $-\frac{\alpha}{2} \mathfrak{S}_q(F)$, with the levels $q_{\hat{v}\hat{v}} = q_v$ and $q_{\hat{v}\hat{r}} = r$. Altogether, the last field and the explicit derivative contribute $\frac{\alpha}{2} [-k + (k - s)] \mathbb{E}_{\mathcal{T}, u} F - \frac{\alpha}{2} \mathfrak{S}_q(F)$. Adding (10.19), we see that the constants $\alpha s/2$, $-\alpha k/2$, and $\alpha(k - s)/2$ cancel, and since $\mathfrak{S}_{R^-} - \mathfrak{S}_q = \mathfrak{S}_\Delta$, we conclude that

$$\frac{\partial_u \mathbb{E}_{\mathrm{ref}}[F \mathcal{D}_{\mathcal{T}}]}{Z_s(u)} = \frac{\alpha}{2} \mathfrak{S}_\Delta(F). \quad (10.20)$$

In words, the row inserts a rest overlap, the last field inserts the level of the corresponding pair, and their difference is an overlap defect.

Step 4: the normalizer. We compute the derivative of $Z_s(u) = \mathbb{E}_{\mathrm{ref}} e^{sU(0,0,0)}$ with (4.6); now the only active coordinates are those of the row. The row contributes $\frac{\alpha s}{2n} \sum_i \mathbb{E}_u [\partial_{\theta_i}^2 U(0) + s(\partial_{\theta_i} U(0))^2]$. As $a(0) - s = \nu(\{0\})$, we obtain from (10.11) that

$$\partial_{\theta_i}^2 U(0) + s(\partial_{\theta_i} U(0))^2 = 1 - \int_{[0, 1]} \mathbb{E}[\pi_i(r)^2 \mid \mathcal{G}_0] \nu(\mathrm{d}r).$$

Here, as in [4, (2.11)], the term $s(\partial_{\theta_i} U)^2$ removes the prescribed root mass s , and a possible atom $\nu(\{0\})$ remains in the integral. Since $\mathbb{E}_u \mathbb{E}[\pi_i(r)^2 \mid \mathcal{G}_0] = \mathbb{E}_{2,r,u}[\tau_i^1 \epsilon_1 \tau_i^2 \epsilon_2]$, the row contributes $\frac{\alpha s}{2} [1 - \int \mathbb{E}_{2,r,u}[\epsilon_1 \epsilon_2 R_{12}^-] \nu(dr)]$. The coefficient $b(u)$ contributes $sb'(u) \mathbb{E}_u \partial_b U(0) = -\alpha s \mathbb{E}_u \partial_b U(0)$. By (10.13), this equals $-\frac{\alpha s}{2} [1 - \int r \mathbb{E}_{2,r,u}[\epsilon_1 \epsilon_2] \nu(dr)]$. Adding the two contributions, we find

$$\frac{d}{du} \log Z_s(u) = -\frac{\alpha s}{2} \int \mathbb{E}_{2,r,u} D_{12} \nu(dr). \quad (10.21)$$

Inserting (10.20) and (10.21) into (10.18), we obtain (10.17). In its last term, $\mathbb{E}_{2,r,u}$ is the expectation under an independent copy of the complete tree law, including its couplings.

Step 5: coefficients and regularity. Recall that $\omega(\mathcal{T}) = k - s \leq k$, $a \leq 1$, and $\nu([0, 1]) \leq 1$. The absolute coefficients of the six terms in (10.17) are at most $k(k-1)$, $2k^2$, k^2 , k , k , and 1. Their sum is $4k^2 + k + 1 \leq 4k^2 + 2k + 1$.

We now show that each extended-tree moment on the right side of (10.17) is continuous in u . We write the row as $(\alpha u/n)^{1/2} g$, and the last-field coordinates as $\sqrt{b(u)}$ times a fixed Brownian field on the tree. By Lemma 10.2, the density is then continuous in u for fixed Gaussian variables. Its linear growth gives a Gaussian envelope that is uniform in u . By dominated convergence, the moments and their normalizers are continuous. The weights and the measures ω and ν over which the right side is integrated do not depend on u , and the integrands are bounded by $2\|F\|_\infty$. Bounded convergence then shows that the right side is continuous. The derivative exists at every $u \in [0, 1]$, one-sided at the endpoints, so $u \mapsto \mathbb{E}_{\mathcal{T},u} F$ is C^1 . Next, we apply (10.17) to each bounded spin observable FD on the extended trees; these observables do not depend on u . The numbers of leaves grow by at most two, and the coefficients are bounded in terms of k . Each of these moments therefore has a derivative bounded by $C_k \alpha \|F\|_\infty$, uniformly in the points v, w and the level r . This lets us differentiate under the integrals against ω and ν by dominated convergence, and we use the product rule for the last term. The resulting right side is continuous by the same argument. This proves the C^2 regularity and the rule for the second derivative. $\square$

Corollary 10.6 (Comparison of nonnegative observables). *For every deterministic profile, every marked tree $\mathcal{T}$ with k leaves, and every nonnegative bounded function F of the spins of its leaves,*

$$\mathbb{E}_{\mathcal{T},u} F \leq K_{\beta,k} \mathbb{E}_{\mathcal{T},v} F \quad (0 \leq u, v \leq 1), \quad K_{\beta,k} = \exp(\beta^2(4k^2 + 2k + 1)). \quad (10.22)$$

Proof. Since $|D_{ab}| \leq 2$ and $F \geq 0$, each expectation in (10.17) is at most twice the expectation of F under the corresponding extended tree. By deletion, the latter equals $\mathbb{E}_{\mathcal{T},u} F$, and the same bound holds for the product in the last term. Using Lemma 10.5 and $\alpha \leq \beta^2$, we find

$$|\partial_u \mathbb{E}_{\mathcal{T},u} F| \leq \beta^2(4k^2 + 2k + 1) \mathbb{E}_{\mathcal{T},u} F, \quad 0 \leq u \leq 1.$$

By Gronwall's inequality on the interval between u and v , we obtain $\mathbb{E}_{\mathcal{T},u} F \leq e^{\beta^2(4k^2+2k+1)|u-v|} \mathbb{E}_{\mathcal{T},v} F$. In particular, (10.22) holds, also in the case $\mathbb{E}_{\mathcal{T},v} F = 0$, in which the moment vanishes for all u . The comparisons in [54, Lemmas 2.9 and 4.3] are proved by the same differential inequality. The classical analog on one site is [81, Lemma 1.6.4]. $\square$

10.3 Products of three overlap defects

In the Taylor remainder of the cavity covariance, and in the variance comparison of Lemma 12.1, we differentiate tree moments of observables that already contain overlap defects. Each differentiation adds one more defect. We show that integrated moments of products of three defects are controlled by the integrated third absolute moment

$$\mathcal{M}_3(u) = \int \mathbb{E}_{2,q,u} |R_{12}^- - q|^3 \nu(\mathrm{d}q) \quad (10.23)$$

of the overlap defect of two descendants. In Lemma 12.3, we bound $\mathcal{M}_3$ by A , J , and H_3 .

Terms and marks. Fix integers $k_0 \geq 1$ and $j \geq 0$. For each $p \in [0, 1]$, let $\mathcal{T}_p$ be the tree of k_0 leaves that split at p , and let F_p be a bounded function of their spins, jointly measurable in p and the spins. We consider the integrated tree moment $\int \nu(\mathrm{d}p) \mathbb{E}_{\mathcal{T}_p,u} F_p$. We also consider the terms produced from it by j applications of (10.17), with the product rule for products of moments over independent trees. Each application multiplies by $\alpha/2$ and replaces a moment by the six terms of (10.17). Each term produced after j applications is therefore $(\alpha/2)^j$ times an integral against a signed measure on finitely many *marks*, which are the following variables. The marks are p , the points v and w at which the differentiations attached leaves, the levels r at which new branches split, and the split levels r of the independent trees produced by the last term of (10.17). The integrand is a product of expectations over independent trees, whose shapes are determined by the marks. The signed measure collects the coefficients 1 , $-2\omega_v$, $\omega_v\omega_w$, $a(q_v)\omega_v$, $\omega_v\nu(\mathrm{d}r)$, and $s\nu(\mathrm{d}r)$ of the attachments, and the multiplicities of the product rule. We let $\mathfrak{m}$ be its total variation. The factor $(\alpha/2)^j$ is not part of $\mathfrak{m}$. The level of a mark is p , q_v , q_w , or r .

Lemma 10.7 (Three-defect bound). *For every deterministic profile, every such term containing three defects $\Delta_1, \Delta_2, \Delta_3$, of pairs of leaves in the same or in independent trees, satisfies*

$$\int \mathbb{E} |\Delta_1 \Delta_2 \Delta_3| \, \mathrm{d}\mathfrak{m} \leq C \mathcal{M}_3(u). \quad (10.24)$$

Here $\mathbb{E}$ is the expectation under the product of the tree laws of the term, and C depends only on the numbers of leaves and of differentiations.

The bounds on the marks in Step 1 of the proof are a counterpart, for an arbitrary profile, of the bounds on split levels in [54, Lemma 2.8].

Proof. The proof has three steps. We show that the level of every mark has a marginal under $\mathfrak{m}$ dominated by a multiple of ν , and that the level of every pair of leaves is the level of a mark. We then use Hölder's inequality.

Step 1: marginals of the marks. Each differentiation adds at most two leaves to one tree, or creates an independent tree with two leaves. After i differentiations, every tree of a term has at most $K_i = k_0 + 2i$ leaves, and a product has at most $k_0 + i$ factors. Let $\mathfrak{m}_i$ be the measure $\mathfrak{m}$ of a term after $i \leq j$ differentiations. We show by induction on i that $\mathfrak{m}_i$ has total mass at most C_i , and that the level of each of its marks has a marginal at most $C_i \nu$. Here C_i depends only on k_0 and i . For $i = 0$, the only mark is p , and $\mathfrak{m}_0 = \nu(dp)$.

For the induction step, let $K = K_i$. We first bound the image of the weights ω of a tree with at most K leaves under the level map $v \mapsto q_v$. At most K points of the tree lie at any given level. The copies of ν on the edges and the parts $\nu(\{q_v\})$ of the vertex weights contribute at most $K\nu$. The parts $(l_v - 1)a(q_v) \leq l_v - 1$ contribute at most $\sum_v (l_v - 1)\delta_{q_v}$, where v runs over the vertices with $l_v \geq 2$, and $\sum_v (l_v - 1) \leq K$. Note that every such q_v is the level of an earlier mark. We now consider one more differentiation. By Lemma 10.5, it multiplies $\mathfrak{m}_i$ by a kernel of total mass at most $(k_0 + i)(4K^2 + 2K + 1)$; here $k_0 + i$ bounds the number of factors to which the product rule applies. The total mass and the marginals of the old marks thus grow by at most the factor $(k_0 + i)(4K^2 + 2K + 1)$. A new point v or w chosen according to ω has a level whose conditional law is at most $K[\nu + \sum_{r'} \delta_{r'}]$, where r' runs over the levels of the old marks. A new level r chosen according to $\nu(dr)$ on $(q_v, 1]$, or according to $s\nu(dr)$, has conditional law at most ν . Factors $a(q_v) \leq 1$ and the restriction $r > q_v$ only decrease these measures. For every Borel set B , the marginal of the level r_{new} of a new mark therefore satisfies

$$\mathfrak{m}_{i+1}(r_{\text{new}} \in B) \leq C \left[\mathfrak{m}_i(1) \nu(B) + \sum_{r'} \mathfrak{m}_i(r' \in B) \right] \leq C_{i+1} \nu(B),$$

where C depends only on k_0 and i . This completes the induction.

Step 2: levels of pairs. In a tree of such a term, every vertex with at least two outgoing edges or leaves lies at the level of a mark. This level is the original split level p , a level q_v at which leaves were attached, a level r at which a new branch split, or the split level of an independent tree. For two leaves a, b of one tree, q_{ab} is the level of such a vertex. In particular, $\chi(q_{ab}) \leq \sum_{r'} \chi(r')$ for every $\chi \geq 0$, where r' runs over the levels of the marks. By Step 1, this implies that

$$\int \chi(q_{ab}) \, d\mathfrak{m} \leq C \int \chi \, d\nu.$$

Step 3: Hölder's inequality. As in [54, Remark 2.4], we apply Hölder's inequality unconditionally under the positive tree laws, and we bound each factor after deleting all

leaves except one pair. On the product of the probability space of the term and the mark space with $\mathfrak{m}$, we have

$$\int \mathbb{E}|\Delta_1\Delta_2\Delta_3| \, d\mathfrak{m} \leq \prod_{i=1}^3 \left(\int \mathbb{E}|\Delta_i|^3 \, d\mathfrak{m} \right)^{1/3}.$$

If we delete all leaves except the two leaves of Δ_i , two descendants remain, splitting at their common level q_i . By deletion, we have $\mathbb{E}|\Delta_i|^3 = \mathbb{E}_{2,q_i,u}|R_{12}^- - q_i|^3$; for a product over independent trees, the other trees integrate to one. We use Step 2 with $\chi(q) = \mathbb{E}_{2,q,u}|R_{12}^- - q|^3$, and we find that each factor is at most $(C\mathcal{M}_3(u))^{1/3}$. This proves (10.24). $\square$

The same proof applies to products of finitely many independent initial tree moments whose leaves split at p , integrated against $\nu(dp)$. The constant then depends on the total number of their initial leaves and on the number of differentiations. Indeed, the initial mark measure is still $\nu(dp)$, the number of factors to which the product rule applies is at most the total number of initial leaves plus the number of differentiations, and Steps 2 and 3 are unchanged.

The cavity covariance. In particular, Lemma 10.7 applies to the covariance between the squared conditional magnetization of the last spin and the rest self-overlap, and to its Taylor remainder,

$$\mathcal{C}_{pp}(u) = \text{Cov}_u(m_N(p)^2, A_p^-), \quad \mathcal{R}_{\text{cav}} = \int \nu(dp) \int_0^1 (1-u) \mathcal{C}_{pp}''(u) \, du.$$

Representing conditional expectations by leaves, we can write $\mathcal{C}_{pp} = \Psi_p - \Psi_p^\epsilon \Psi_p^\tau$, where

$$\Psi_p(u) = \mathbb{E}_{4,p,u}[\epsilon_1 \epsilon_2 (R_{34}^- - p)], \quad \Psi_p^\epsilon(u) = \mathbb{E}_{2,p,u} \epsilon_1 \epsilon_2, \quad \Psi_p^\tau(u) = \mathbb{E}_{2,p,u} (R_{12}^- - p).$$

Indeed, by conditional independence given $\mathcal{G}_p$, we have $\Psi_p = \mathbb{E}_u[m_N(p)^2(A_p^- - p)]$, $\Psi_p^\epsilon = \mathbb{E}_u m_N(p)^2$, and $\Psi_p^\tau = \mathbb{E}_u(A_p^- - p)$. The constant p does not change the covariance. Note that the observables in these tree moments, including the center p , do not depend on u . They are bounded by 2, and the trees have at most four leaves. By Lemma 10.5, each of the three moments belongs to $C^2([0,1])$, with first and second derivatives bounded by $C\alpha$ and $C\alpha^2$. Then $\mathcal{C}_{pp} \in C^2([0,1])$ for every deterministic profile, and $|\mathcal{C}_{pp}''|$ is bounded by a constant depending only on β , uniformly in p . In particular, $\mathcal{R}_{\text{cav}}$ is well defined, and by Taylor's formula with integral remainder, we have

$$\mathcal{R}_{\text{cav}} = \int [\mathcal{C}_{pp}(1) - \mathcal{C}_{pp}(0) - \mathcal{C}_{pp}'(0)] \nu(dp).$$

Lemma 10.8 (Taylor remainder). *For every deterministic profile and every $u \in [0,1]$,*

$$\int |\mathcal{C}_{pp}''(u)| \nu(dp) \leq C_\beta \mathcal{M}_3(u), \quad |\mathcal{R}_{\text{cav}}| \leq C_\beta \sup_{0 \leq w \leq 1} \mathcal{M}_3(w). \quad (10.25)$$

Proof. We have

$$\mathcal{C}_{pp}'' = \Psi_p'' - (\Psi_p^\epsilon)'' \Psi_p^\tau - 2(\Psi_p^\epsilon)'(\Psi_p^\tau)' - \Psi_p^\epsilon(\Psi_p^\tau)''.$$

Each application of (10.17) adds a factor $\alpha/2 \leq \beta^2/2$ and one overlap defect, either inside a factor D or in the independent tree of its last term. The spin factors have absolute value one. The moments Ψ_p and Ψ_p^τ contain one defect, and Ψ_p^ϵ contains none. We count the defects before differentiation and those added by the two differentiations.

term	defects before	added	total
Ψ_p''	1	2	3
$(\Psi_p^\epsilon)'' \Psi_p^\tau$	$0 + 1$	$2 + 0$	3
$(\Psi_p^\epsilon)'(\Psi_p^\tau)'$	$0 + 1$	$1 + 1$	3
$\Psi_p^\epsilon(\Psi_p^\tau)''$	$0 + 1$	$0 + 2$	3

Products of moments, and the last term of (10.17), are products over independent trees. Every term of $\mathcal{C}_{pp}''$ is therefore $(\alpha/2)^2$ times a term of the kind treated in Lemma 10.7, or in its extension to products of independent initial moments stated after its proof, with at most four initial leaves and $j = 2$. It is multiplied by a bounded multiplicity and by factors of absolute value at most one. Integrating p against ν and using Lemma 10.7 and this extension, we obtain the first bound in (10.25) with $C_\beta = C\beta^4$. We integrate it against $(1 - u) du$ and use Tonelli's theorem. This gives $|\mathcal{R}_{\text{cav}}| \leq C\beta^4 \int_0^1 (1 - u) \mathcal{M}_3(u) du \leq C\beta^4 \sup_w \mathcal{M}_3(w)$, which is the second bound. $\square$

The derivative at the independent endpoint

In Section 10, we constructed the cavity interpolation $\mathbb{P}_u$, $0 \leq u \leq 1$. At $u = 1$, it is the tilted law of the full system with N spins. At $u = 0$, it is a product law: the last spin ϵ is independent of the remaining $n = N - 1$ spins. We also introduced the covariance

$$\mathcal{C}_{pp}(u) = \text{Cov}_u(m_N(p)^2, A_p^-)$$

between the squared conditional magnetization of the last spin and the rest self-overlap. By Lemma 10.5, it is twice continuously differentiable in u (see Section 10.3), and Lemma 10.8 bounds its Taylor remainder. The covariance vanishes at $u = 0$. In this section, we compute its derivative there.

The main result is Lemma 11.2. It states that after integration against ν , the derivative $\mathcal{C}'_{pp}(0)$ equals α/d times the trace of the scalar response operator $J_{a,d}^{\text{sc}}$ of Section 9 against the covariance operator Σ^- of the rest self-overlap, minus a nonnegative term $\alpha\mathcal{P}/2$, plus explicit error terms. The identity before integration is Lemma 11.1. In Lemma 11.3, we bound the integrated error terms by the trace term A and the mean-covariance term J of the full system, up to an additive N^{-2} . The three lemmas hold for every deterministic profile with $\nu(\{1\}) = 0$; no minimizing property is used. In Section 12, we insert Lemma 11.2 into the proof of the trace identity of Lemma 12.2 and use Lemma 11.3 to bound its error term in the proof of Proposition 12.4.

The computation uses the exact density (10.9) of the cavity law and the parameter responses (10.11) from Section 10, the Gaussian heat identity of Lemma 4.3, and the explicit form (9.19) of the scalar response operator. The error bounds also use the kernel bounds of Lemma 9.4, the posterior variance bound (10.4), and the comparison inequality of Corollary 10.6.

Standing assumptions. We keep the assumptions of Section 10: $N \geq 2$, $n = N - 1$, $0 < \rho \leq 1$, $t = 1 - \rho$, $c = \beta^2 \rho$, and $0 < s = \lambda/\beta \leq 1/2$. The parameters α , d , and $b(u)$ are as in (10.1). In particular, $0 \leq \alpha \leq \beta^2$ and $d \in [\beta^2/2, \beta^2]$. The profile $\mu = s\delta_0 + \nu$ is deterministic and otherwise arbitrary, with cumulative function $a(q) = \mu([0, q])$. From now on, we also assume that $\nu(\{1\}) = 0$. This holds at every minimizing profile by (3.9); Sections 10.2 and 10.3 did not use it. Under this assumption, we show below that the covariance operator of the rest self-overlap has its range in the scalar transport space $\mathcal{H}_\nu$ of Section 9, on which $J_{a,d}^{\text{sc}}$ acts.

11.1 Linear response at the independent endpoint

The product law at the independent endpoint

In this section, we take all expectations at $u = 0$ unless a subscript says otherwise. Recall from Section 10.1 that $\mathbb{P}_0$ is a product law. Its first factor is the tilted law of the n -spin system with Hamiltonian $\mathcal{E}^-$, field coefficient c , and profile μ , and the second is the scalar law of Section 9 with coefficient d . We write $\mathbb{E}^-$, Var^- , and Cov^- for the first factor (the rest system) and $\mathbb{E}^0$, Var^0 , and Cov^0 for the second; $\mathcal{G}_q^-$ and $\mathcal{G}_q^0$ denote their filtrations. At $u = 0$, the solution of (10.2) splits as $U(q, x, y) = U^-(q, x) + U^0(q, y)$. We define

$$\begin{aligned} M(r, x) &= \nabla_x U^-(r, x), & M_i(r, x) &= \partial_{x_i} U^-(r, x), \\ A^-(r, x) &= \frac{|M(r, x)|^2}{n}, & m(r, y) &= \partial_y U^0(r, y), \end{aligned}$$

so that $M_r = M(r, X_r)$, $A_r^- = A^-(r, X_r)$, $H_r^- = \nabla_x^2 U^-(r, X_r)$, $m_r = m(r, Y_r)$, and $h_r = \partial_y m(r, Y_r)$. We use the scalar quantities $f_p = m_p^2$, $\bar{f}_p = f_p - \mathbb{E}^0 f_p$, $g_r^\ell(y) = P_{r,\ell}^0(m(\ell, \cdot)^2)(y)$, and $\Gamma = \Gamma_{a,d}$, with $\Gamma(p) = \mathbb{E}^0 m_p^2$, of Section 9. Recall that $g_r^\ell(Y_r) = \mathbb{E}^0[m_\ell^2 \mid \mathcal{G}_r^0]$ and $m_r = \mathbb{E}^0[\epsilon \mid \mathcal{G}_r^0]$. For $r \leq \ell$, let $P_{r,\ell}^-$ be the transition operator of the rest diffusion. Its state includes the rest couplings, so that $(P_{r,\ell}^- A_\ell^-)(X_r) = \mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-]$. For notational convenience, we abbreviate $a_r^- = a_r^-(0)$ and $j_r^- = j_r^-(0)$ for the quantities of (10.3) at $u = 0$.

Note that the rest Hamiltonian $\mathcal{E}^-$ is invariant under the global spin reversal $\tau \mapsto -\tau$. Then $U^-(0, \cdot)$ is even, which implies that $M_0 = 0$ and

$$A_0^- = 0. \tag{11.1}$$

Since the spatial derivatives of U^- are bounded and jointly continuous, $r \mapsto A_r^-$ is continuous with values in $[0, 1]$.

The rest self-overlap. We define

$$\begin{aligned} \Sigma^-(p, q) &= \text{Cov}^-(A_p^-, A_q^-), & \bar{A}_p^- &= A_p^- - \mathbb{E}^- A_p^-, \\ \Theta(r; \ell, p) &= \text{Cov}^-(P_{r,\ell}^- A_\ell^-, A_p^-) \quad (r \leq \ell), \end{aligned}$$

where we recall that $P_{r,\ell}^- A_\ell^-$ stands for $\mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-]$. Note that Σ^- is continuous and bounded by one, and that $\Sigma^-(0, p) = 0$ by (11.1). If $r \geq p$, then $\bar{A}_p^-$ is $\mathcal{G}_r^-$ -measurable, and by the tower property, we have

$$\Theta(r; \ell, p) = \mathbb{E}^-[\mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-] \bar{A}_p^-] = \mathbb{E}^-[A_\ell^- \bar{A}_p^-] = \Sigma^-(\ell, p) \quad (r \geq p). \tag{11.2}$$

Fluctuations along the path. Let τ be the terminal rest spin. For levels $r, \ell \in [0, 1]$, we define

$$x_r = \frac{\tau \cdot M_r}{n} - A_r^-, \quad y_{r\ell} = \frac{M_r \cdot M_\ell}{n} - A_{r\wedge\ell}^-,$$

random variables of the rest system; the subscripts distinguish them from the field variables x and y . The vector $M_q = \mathbb{E}^-[\tau \mid \mathcal{G}_q^-]$ is a martingale, and by the conditional law of total covariance, we have, for $r \leq \ell$,

$$C_r^- = \mathbb{E}^-[C_\ell^- \mid \mathcal{G}_r^-] + \mathbb{E}^-[(M_\ell - M_r)(M_\ell - M_r)^T \mid \mathcal{G}_r^-]. \quad (11.3)$$

In particular, $\mathbb{E}^-[C_\ell^- \mid \mathcal{G}_r^-] \preceq C_r^-$ and $\text{Cov}^-(M_\ell \mid \mathcal{G}_r^-) \preceq C_r^-$. We have $x_r = M_r \cdot (\tau - M_r)/n$, and $y_{r\ell} = M_r \cdot (M_\ell - M_r)/n$ if $r \leq \ell$. By conditioning on $\mathcal{G}_r^-$, we deduce from (11.3) that

$$\mathbb{E}^-[x_r \mid \mathcal{G}_r^-] = 0, \quad \mathbb{E}^-[x_r^2 \mid \mathcal{G}_r^-] = j_r^-, \quad \mathbb{E}^-[y_{r\ell}] = 0, \quad \mathbb{E}^-[y_{r\ell}^2 \mid \mathcal{G}_r^-] \leq j_{r\wedge\ell}^-. \quad (11.4)$$

Indeed, $\mathbb{E}^-[x_r^2 \mid \mathcal{G}_r^-] = M_r^T C_r^- M_r / n^2$, and for $r \leq \ell$, the vector $M_\ell - M_r$ has conditional mean zero and conditional covariance at most C_r^- given $\mathcal{G}_r^-$, so that $\mathbb{E}^-[y_{r\ell}^2 \mid \mathcal{G}_r^-] \leq M_r^T C_r^- M_r / n^2$. The case $r > \ell$ follows by symmetry.

The error terms. We set $v_p^0 = \text{Var}^0(f_p)$ and

$$a_p^H = \frac{\mathbb{E}^-[\text{Tr}((H_p^-)^2)]}{n^2}, \quad b_p^H = \frac{\mathbb{E}^-[M_p^T H_p^- M_p]}{n^2}, \quad b_{r,p}^H = \frac{\mathbb{E}^-[M_r^T H_p^- M_p]}{n^2},$$

$$z_{p,\ell} = \frac{1}{n} \mathbb{E}^-[M_p \cdot \nabla(P_{p,\ell}^- A_\ell^-)(X_p)] \quad (\ell \geq p),$$

where $\nabla(P_{p,\ell}^- A_\ell^-)$ is the gradient of $x \mapsto (P_{p,\ell}^- A_\ell^-)(x)$, with the rest couplings fixed. The error terms are

$$T_p^{(1)} = 2\mathbb{E}^0[m_p h_p \epsilon] \mathbb{E}^-[x_p \bar{A}_p^-] - 2 \int \mathbb{E}^0[m_p h_p m_r] \mathbb{E}^-[y_{pr} \bar{A}_p^-] \nu(\mathrm{d}r), \quad (11.5)$$

$$T_p^{(2)} = - \int \mathbb{E}^0[\bar{f}_p \epsilon m_r] \mathbb{E}^-[x_r \bar{A}_p^-] \nu(\mathrm{d}r) \\ + \frac{1}{2} \iint \mathbb{E}^0[\bar{f}_p m_r m_\ell] \mathbb{E}^-[y_{r\ell} \bar{A}_p^-] \nu(\mathrm{d}r) \nu(\mathrm{d}\ell), \quad (11.6)$$

$$T_p^{(3)} = v_p^0 a_p^H + [4\mathbb{E}^0(m_p^2 h_p) + 2(1 - a(p))v_p^0] b_p^H \\ - \int_{(p,1]} \text{Cov}^0(f_p, g_p^\ell(Y_p)) z_{p,\ell} \nu(\mathrm{d}\ell) - 2 \int \mathbb{E}^0[\bar{f}_p m_p m_r] b_{r,p}^H \nu(\mathrm{d}r), \quad (11.7)$$

and $T_p = T_p^{(1)} + T_p^{(2)} + T_p^{(3)}$. When we integrate against ν without indicating a domain, we integrate over $[0, 1]$, including any atom of ν at 0. The terms $T_p^{(1)}$ and $T_p^{(2)}$ contain the fluctuations x_r and $y_{r\ell}$, and $T_p^{(3)}$ contains the derivatives of the rest observable A_p^- in the coupling vector of the incident row.

The conditional future correction. For $p, \ell > r$, we define

$$\Lambda_r(p, \ell) = \text{Cov}^0(g_r^p(Y_r), g_r^\ell(Y_r)), \quad \Sigma_r^-(p, \ell) = \mathbb{E}^- \text{Cov}^-(A_p^-, A_\ell^- \mid \mathcal{G}_r^-),$$

and

$$\mathcal{P} = \int \nu(\mathrm{d}r) \iint_{p, \ell > r} \Lambda_r(p, \ell) \Sigma_r^-(p, \ell) \nu(\mathrm{d}p) \nu(\mathrm{d}\ell). \quad (11.8)$$

Note that Λ_r is a covariance of the scalar law and that Σ_r^- is an averaged conditional covariance of the rest system. We show in the proof of Lemma 11.2 that $\mathcal{P} \geq 0$ and that $-\alpha\mathcal{P}/2$ is the contribution of the difference between Θ and Σ^- .

The scalar response operator applied to Σ^- . For a bounded Borel function φ on $[0, 1]$ and $p \in [0, 1]$, the seven terms on the right side of (9.19), in the order written there, are

$$(J1) \quad \mathbb{E}^0[h_p^2 + 2(1 - a(p))m_p^2 h_p] \varphi(p),$$

$$(J2) \quad - \int_{(p, 1]} \mathbb{E}^0[m_p \partial_y g_p^\ell(Y_p)] \varphi(\ell) \nu(\mathrm{d}\ell),$$

$$(J3) \quad - 2 \int \mathbb{E}^0[m_p h_p m_r] \varphi(p \wedge r) \nu(\mathrm{d}r),$$

$$(J4) \quad - \int \mathbb{E}^0[\bar{f}_p \epsilon m_r] \varphi(r) \nu(\mathrm{d}r),$$

$$(J5) \quad \frac{1}{2} \int a(r) \mathbb{E}^0[\bar{f}_p m_r^2] \varphi(r) \nu(\mathrm{d}r),$$

$$(J6) \quad \frac{1}{2} \iint \mathbb{E}^0[\bar{f}_p m_r m_\ell] \varphi(r \wedge \ell) \nu(\mathrm{d}r) \nu(\mathrm{d}\ell),$$

$$(J7) \quad \frac{1}{2} \int \nu(\mathrm{d}r) \int_{(r, 1]} \mathbb{E}^0[\bar{f}_p g_r^\ell(Y_r)] \varphi(\ell) \nu(\mathrm{d}\ell),$$

and their sum is $d^{-1}(J_{a,d}\varphi)(p)$. All kernels in (J1)–(J7) are bounded by constants depending only on β (Section 9). For $p \in [0, 1]$, we denote by $(J_{a,d}^{\text{sc}}\Sigma^-)(p, p)$ the value at p of the function $J_{a,d}\Sigma^-(\cdot, p)$, that is, of $J_{a,d}$ applied to the first variable of Σ^- . This notation agrees with the compression $J_{a,d}^{\text{sc}}$. Indeed, let φ be bounded with $\varphi(0) = 0$. Each of the terms (J2)–(J7) contains a factor m_p or $\bar{f}_p$, and $m_0 = \bar{f}_0 = 0$. At $p = 0$, these terms vanish, and (J1) equals $h_0^2\varphi(0) = 0$. As $\nu(\{1\}) = 0$, the function $J_{a,d}\varphi$ therefore lies in $\mathcal{H}_\nu$, and it coincides with $J_{a,d}^{\text{sc}}\varphi$ when $\varphi \in \mathcal{H}_\nu$. By (11.1), this applies to $\varphi = \Sigma^-(\cdot, p)$ for every p .

The endpoint derivative

Lemma 11.1 (Endpoint derivative at a fixed level). *Let $\mu = s\delta_0 + \nu$ be a deterministic profile with $\nu(\{1\}) = 0$, and let $p \in [0, 1]$. If $\alpha > 0$, then*

$$\begin{aligned} \frac{C'_{pp}(0)}{\alpha} &= \frac{1}{d} (J_{a,d}^{\text{sc}}\Sigma^-)(p, p) + T_p \\ &\quad + \frac{1}{2} \int \nu(\mathrm{d}r) \int_{(r, 1]} \text{Cov}^0(f_p, g_r^\ell(Y_r)) [\Theta(r; \ell, p) - \Sigma^-(\ell, p)] \nu(\mathrm{d}\ell). \end{aligned} \quad (11.9)$$

If $\alpha = 0$, then $C'_{pp}(0) = 0$.

The first term on the right side of (11.9) is the scalar response applied to $\Sigma^-(\cdot, p)$ in place of φ , and T_p collects the error terms. The last line has no scalar counterpart. It comes from the conditional expectation $\mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-]$ in the second derivative of the density, which replaces $\Sigma^-(\ell, p)$ by $\Theta(r; \ell, p)$ in (J7). After integration over p , it becomes the nonpositive term $-\mathcal{P}/2$.

Lemma 11.2 (Integrated linear response). *Let $\mu = s\delta_0 + \nu$ be a deterministic profile with $\nu(\{1\}) = 0$, and let $V^- = \int \text{Var}^-(A_p^-) \nu(dp)$. Then*

$$\int C'_{pp}(0) \nu(dp) = \alpha \left[\frac{1}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-) - \frac{1}{2} \mathcal{P} + \int T_p \nu(dp) \right], \quad \mathcal{P} \geq 0, \quad (11.10)$$

where $\mathcal{P}$ is given by (11.8), $\text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-) = \int (J_{a,d}^{\text{sc}} \Sigma^-)(p, p) \nu(dp)$, and T_p is given by (11.5)–(11.7). Further, Σ^- is a positive trace-class covariance operator on $L^2(\nu)$ with trace V^- and range in $\mathcal{H}_\nu$. With $\bar{A}^-$ the random function $p \mapsto \bar{A}_p^-$, the number $\text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-)$ equals $\mathbb{E}^- \langle \bar{A}^-, J_{a,d}^{\text{sc}} \bar{A}^- \rangle_{L^2(\nu)}$, the trace of the operator $J_{a,d}^{\text{sc}} \Sigma^-$ on $\mathcal{H}_\nu$.

Proof of the endpoint derivative at a fixed level

Proof of Lemma 11.1. If $\alpha = 0$, then the row variance $\alpha u/n$ vanishes and $b(u) = d$ for every u , so that $\mathbb{P}_u = \mathbb{P}_0$ for every u and C_{pp} vanishes identically. From now on, we assume that $\alpha > 0$ and fix $p \in [0, 1]$. The proof has six steps. In Step 1, we show that $C'_{pp}(0)$ is a Laplacian, in the coupling vector of the incident row, of a covariance under laws with deterministic rows. In Step 2, we write this Laplacian as a sum of four groups of terms, using the exact density of these laws. In Steps 3–5, we compute the four groups. In Step 6, we match the result with (J1)–(J7). At fixed N , the spin observables and the required positive-order derivatives of U are bounded. The reference-density integrands have the integrable Gaussian envelopes established below, so we may exchange expectations and finite level integrals by Fubini's theorem.

Step 1: reduction to fixed couplings. For $w \geq 0$ and $b \in [c, d]$, let $\mathbb{P}_{w,b}$ be the law of Section 10.1 with an incident row of variance w per coordinate in place of $\alpha u/n$, and with last-field coefficient b in place of $b(u)$. In this notation, $\mathbb{P}_u = \mathbb{P}_{\alpha u/n, b(u)}$, and $b(u) \in [c, d]$. We denote expectation and covariance under these laws by $\mathbb{E}_{w,b}$ and $\text{Cov}_{w,b}$, and define

$$C(w, b) = \text{Cov}_{w,b}(m_N(p)^2, A_p^-), \quad \text{so that} \quad C_{pp}(u) = C(\alpha u/n, b(u)).$$

At $w = 0$, the row vanishes, and U splits into a function of (q, x) and a function of (q, y) , as at $u = 0$ but with d replaced by b . Then $\mathbb{P}_{0,b}$ is a product law, under which $m_N(p)$ and A_p^- are independent, and therefore

$$C(0, b) = 0 \quad \text{for every } b \in [c, d]. \quad (11.11)$$

Applying Lemma 4.3 with the row variance w as its parameter, together with Lemmas 10.2 and 10.4, we see that the function $w \mapsto C(w, b)$ is continuous on $[0, \alpha/n]$ and differentiable on $(0, \alpha/n]$, with a right derivative at $w = 0$. The derivative $\partial_w C$ is jointly continuous in (w, b) . To see this, we couple the row as $\sqrt{w}g$ and the last field as $\sqrt{b}B$ and use the Gaussian envelope of the density, as in the continuity part of the proof of Lemma 10.5. Let $0 < u \leq 1$. By (11.11) and the mean value theorem in w , there exists $w_u \in (0, \alpha u/n)$ such that

$$C_{pp}(u) = C(\alpha u/n, b(u)) - C(0, b(u)) = \frac{\alpha u}{n} \partial_w C(w_u, b(u)).$$

Dividing by u and letting $u \downarrow 0$, we find

$$C'_{pp}(0) = \frac{\alpha}{n} \partial_w C(0, d). \quad (11.12)$$

In particular, the variation of the last-field coefficient $b(u)$ does not contribute to $C'_{pp}(0)$.

We now compute $\partial_w C(0, d)$. For $\theta \in \mathbb{R}^n$, let $\mathbb{P}^\theta$ be the law of Section 10.1 with last-field coefficient d and with the incident row replaced by a deterministic coupling: the term $\sqrt{\alpha u/n} \epsilon \sum_i g_{iN} \tau_i$ in $\mathcal{E}(\sigma; x, y)$ becomes $\epsilon \theta \cdot \tau$. We write $U(q, x, y; \theta)$ for the corresponding solution of (10.2), $\mathbb{E}^\theta$ and Cov^θ for expectation and covariance under $\mathbb{P}^\theta$, and

$$Z(\theta) = \mathbb{E}_{\text{ref}} e^{sU(0,0,0;\theta)}$$

for its outer normalizer, which is an expectation over the couplings g_{ij} , $1 \leq i < j \leq n$, only. Under $\mathbb{P}^\theta$, observables such as $m_N(p) = \partial_y U(p, X_p, Y_p; \theta)$ and $M_p = \nabla_x U(p, X_p, Y_p; \theta)$ are computed from $U(\cdot; \theta)$. Under $\mathbb{P}_{w,d}$, we may integrate first over the couplings g_{ij} and then over $\theta = \sqrt{w}(g_{1N}, \dots, g_{nN})$. Given θ , the conditional law is $\mathbb{P}^\theta$. Because of the outer tilt $e^{sU(0,0,0)}$, the vector θ has the density $Z(\theta)/\mathbb{E}Z(\theta)$ with respect to $N(0, wI_n)$. We obtain, for every bounded observable G ,

$$\mathbb{E}_{w,d} G = \frac{\mathbb{E}[Z(\theta) \mathbb{E}^\theta G]}{\mathbb{E}Z(\theta)}, \quad \theta \sim N(0, wI_n). \quad (11.13)$$

This mixture is not the Gaussian average of the laws $\mathbb{P}^\theta$, because θ is weighted by $Z(\theta)$. We now check that this weight does not change the derivative at $w = 0$.

Let $f = m_N(p)^2$. For an observable G , we write $\Phi_G(\theta) = \mathbb{E}^\theta G$. Replacing the summation variable ϵ by $-\epsilon$ in the terminal value of (10.2), we find $U(1, x, y; \theta) = U(1, x, -y; -\theta)$. The invariance of (10.2) under $y \mapsto -y$ then implies that $U(q, x, y; \theta) = U(q, x, -y; -\theta)$ for every q . Hence the reflection $(\epsilon, Y) \mapsto (-\epsilon, -Y)$ of the last spin and its field maps $\mathbb{P}^\theta$ to $\mathbb{P}^{-\theta}$, preserves $U(0, 0, 0; \theta)$ and M_p , and changes the sign of $m_N(p)$. It follows that the functions

$$Z, \quad \Phi_f, \quad \Phi_{A_p^-}, \quad \Phi_{fA_p^-}$$

are even in θ , so that their gradients vanish at $\theta = 0$. At fixed N , they are twice continuously differentiable, with second derivatives bounded by $Ce^{C|\theta|}$ for a constant C that may depend on N . Indeed, by (10.11), we have $|\partial_{\theta_i} U| \leq 1$ and $0 \leq \partial_{\theta_i}^2 U \leq 1$. By Lemma 10.2, all derivatives of U of positive order in (θ, x, y) are bounded at fixed N . Further, the density of $\mathbb{P}^\theta$ and its first two derivatives in θ are bounded by the integrable envelope

$$C \exp \left\{ C \left(1 + \sum_{1 \leq i < j \leq n} |g_{ij}| + |\theta| + \sup_{q \leq 1} |X_q| + \sup_{q \leq 1} |Y_q| \right) \right\}$$

under the Gaussian reference law of the couplings and the paths. Constants that we use only to justify differentiation may depend on N ; they do not enter the identity we prove.

Let $F \in C^2(\mathbb{R}^n)$ satisfy $|\nabla^2 F(\theta)| \leq Ce^{C|\theta|}$, and let g be a standard Gaussian vector in $\mathbb{R}^n$. For $w > 0$, Gaussian integration by parts shows that the derivative of $w \mapsto \mathbb{E} F(\sqrt{w} g)$ is $\mathbb{E} \Delta F(\sqrt{w} g)/2$, which tends to $\Delta F(0)/2$ as $w \downarrow 0$ by dominated convergence. By the mean value theorem, the right derivative at $w = 0$ is $\Delta F(0)/2$. This is the heat equation for the Gaussian semigroup. We use this fact for $F = Z\Phi_G$ and for $F = Z$ in (11.13). By the quotient rule, we have

$$\begin{aligned} \partial_w \mathbb{E}_{w,d} G|_{w=0} &= \frac{\Delta(Z\Phi_G)(0) - \Phi_G(0)\Delta Z(0)}{2Z(0)} \\ &= \frac{1}{2}\Delta\Phi_G(0) + \frac{\nabla Z(0) \cdot \nabla\Phi_G(0)}{Z(0)} = \frac{1}{2}\Delta\Phi_G(0), \end{aligned}$$

where in the second equality we used $\Delta(Z\Phi_G) = Z\Delta\Phi_G + 2\nabla Z \cdot \nabla\Phi_G + \Phi_G\Delta Z$, and in the third that $\nabla Z(0) = 0$. We use this for $G \in \{f, A_p^-, fA_p^-\}$, together with the product rule. The cross term $\nabla\Phi_f \cdot \nabla\Phi_{A_p^-}$ in the expansion of $\Delta(\Phi_f\Phi_{A_p^-})/2$ vanishes at $\theta = 0$, because $\nabla\Phi_f(0) = 0$. We conclude that

$$\partial_w C(0, d) = \frac{1}{2}\Delta\Phi_{fA_p^-}(0) - \frac{1}{2}\Delta\Phi_f(0)\Phi_{A_p^-}(0) - \frac{1}{2}\Phi_f(0)\Delta\Phi_{A_p^-}(0) = \frac{1}{2}\Delta_\theta \text{Cov}^\theta(f, A_p^-)|_{\theta=0}.$$

Inserting this into (11.12), we obtain

$$C'_{pp}(0) = \frac{\alpha}{2n} \sum_{i=1}^n \partial_{\theta_i}^2 \text{Cov}^\theta(f, A_p^-)|_{\theta=0}. \quad (11.14)$$

Step 2: derivatives in one coordinate of the row. For $\theta = 0$, the law $\mathbb{P}^\theta$ is the product law $\mathbb{P}_0$, and $U(q, x, y; 0) = U^-(q, x) + U^0(q, y)$. In this proof, $\mathbb{E}$ and Cov without superscript denote expectation and covariance under $\mathbb{P}_0$. Let

$$\dot{U}_i(r, x, y) = \partial_{\theta_i} U(r, x, y; \theta)|_{\theta=0}, \quad \ddot{U}_i(r, x, y) = \partial_{\theta_i}^2 U(r, x, y; \theta)|_{\theta=0}.$$

Under $\mathbb{P}_0$, the conditional law of (τ, ϵ) given $\mathcal{G}_r$ is the product of the conditional laws of τ given $\mathcal{G}_r^-$ and of ϵ given $\mathcal{G}_r^0$. The conditional feature $\pi_i(r) = \mathbb{E}[\tau_i \epsilon \mid \mathcal{G}_r]$ of (10.11) therefore

equals $M_i(r, X_r)m_r$, and for $\ell \geq r$, we have $\mathbb{E}[\pi_i(\ell)^2 \mid \mathcal{G}_r] = (P_{r,\ell}^- M_i(\ell, \cdot)^2)(X_r) g_r^\ell(Y_r)$. Inserting these identities into (10.11), we obtain

$$\dot{U}_i(r, x, y) = M_i(r, x) m(r, y), \quad (11.15)$$

$$\ddot{U}_i(r, x, y) = 1 - a(r) M_i(r, x)^2 m(r, y)^2 - \int_{(r,1]} (P_{r,\ell}^- M_i(\ell, \cdot)^2)(x) g_r^\ell(y) \nu(d\ell). \quad (11.16)$$

Since $\sum_i M_i(\ell, \cdot)^2 = nA^-(\ell, \cdot)$, the average of (11.16) over i is

$$\ddot{U}_{\text{av}}(r, x, y) = \frac{1}{n} \sum_{i=1}^n \ddot{U}_i(r, x, y) = 1 - a(r) A^-(r, x) m(r, y)^2 - \int_{(r,1]} (P_{r,\ell}^- A_\ell^-)(x) g_r^\ell(y) \nu(d\ell). \quad (11.17)$$

By Lemma 10.3 with one leaf, applied with the row fixed at θ , the law $\mathbb{P}^\theta$ has the density

$$\omega_\theta = \frac{1}{Z(\theta)} \exp \left\{ \mathcal{E}^-(\tau) + \epsilon \theta \cdot \tau + X_1 \cdot \tau + Y_1 \epsilon - \int_{[0,1]} U(q, X_q, Y_q; \theta) \nu(dq) \right\}$$

with respect to the Gaussian reference law of the couplings g_{ij} , the Gaussian reference paths X and Y with covariances $c(q \wedge q') I_n$ and $d(q \wedge q')$, and counting measure on $\sigma = (\tau, \epsilon)$. Recall from Step 1 that $\nabla Z(0) = 0$. Then the θ_i -score of ω_θ at $\theta = 0$ is

$$\mathfrak{s}_i = \partial_{\theta_i} \log \omega_\theta|_{\theta=0} = \tau_i \epsilon - \int M_i(r) m_r \nu(dr), \quad M_i(r) = M_i(r, X_r),$$

where we used (11.15), and the second derivative satisfies

$$\frac{\partial_{\theta_i}^2 \omega_\theta}{\omega_\theta} \Big|_{\theta=0} = \mathfrak{s}_i^2 + \partial_{\theta_i}^2 \log \omega_\theta|_{\theta=0} = \mathfrak{s}_i^2 - \int \ddot{U}_i(r, X_r, Y_r) \nu(dr) - \partial_{\theta_i}^2 \log Z(0). \quad (11.18)$$

This quantity has mean zero under $\mathbb{P}_0$, because ω_θ integrates to one for every θ .

An observable G_θ under $\mathbb{P}^\theta$ depends on θ explicitly, through $U(\cdot; \theta)$, and through the law. We write G_i and G_{ii} for its first and second explicit θ_i -derivatives at $\theta = 0$, with the couplings, the paths, and the spins fixed. Differentiating under the reference integral, which the envelope of Step 1 permits, we have

$$\partial_{\theta_i}^2 \mathbb{E}^\theta G_\theta|_{\theta=0} = \mathbb{E} \left[G_{ii} + 2G_i \mathfrak{s}_i + G \frac{\partial_{\theta_i}^2 \omega_\theta}{\omega_\theta} \Big|_{\theta=0} \right].$$

This formula holds for $G = fA_p^-$, f , and A_p^- . At $\theta = 0$, we have $f = f_p$. We denote the explicit derivatives of f by $f_{p,i}$ and $f_{p,ii}$, and those of A_p^- by $A_{p,i}^-$ and $A_{p,ii}^-$. The first derivatives of $\mathbb{E}^\theta f$ and $\mathbb{E}^\theta A_p^-$ vanish at $\theta = 0$ by Step 1. Expanding $\text{Cov}^\theta(f, A_p^-) = \mathbb{E}^\theta[fA_p^-] - \mathbb{E}^\theta f \mathbb{E}^\theta A_p^-$ with the product rule, and using that (11.18) has mean zero, we find that $\partial_{\theta_i}^2 \text{Cov}^\theta(f, A_p^-)$ at $\theta = 0$ equals

$$\mathbb{E}[f_{p,ii} \bar{A}_p^- + \bar{f}_p A_{p,ii}^- + 2f_{p,i} A_{p,i}^-] + 2\mathbb{E}[\mathfrak{s}_i(f_{p,i} \bar{A}_p^- + \bar{f}_p A_{p,i}^-)] + \mathbb{E}\left[\bar{f}_p \bar{A}_p^- \frac{\partial_{\theta_i}^2 \omega_\theta}{\omega_\theta} \Big|_{\theta=0}\right].$$

Since $\mathbb{P}_0$ is a product law, we have $\mathbb{E}[\bar{f}_p \bar{A}_p^-] = 0$, and the constant $\partial_{\theta_i}^2 \log Z(0)$ in (11.18) drops out. Summing over i , we see from (11.14) that

$$\begin{aligned} \frac{C'_{pp}(0)}{\alpha} &= \frac{1}{2n} \sum_i \left\{ \mathbb{E}[f_{p,ii} \bar{A}_p^- + \bar{f}_p A_{p,ii}^- + 2f_{p,i} A_{p,i}^-] + 2\mathbb{E}[\mathfrak{s}_i(f_{p,i} \bar{A}_p^- + \bar{f}_p A_{p,i}^-)] \right. \\ &\quad \left. + \mathbb{E}[\bar{f}_p \bar{A}_p^- (\mathfrak{s}_i^2 - \int \ddot{U}_i d\nu)] \right\}. \end{aligned} \quad (11.19)$$

We sort the right side into four groups, each including the factor $1/(2n)$ and the sum over i :

- (a) the terms $f_{p,ii} \bar{A}_p^-$;
- (b) the terms $2\mathfrak{s}_i f_{p,i} \bar{A}_p^-$;
- (c) the terms $\bar{f}_p \bar{A}_p^- (\mathfrak{s}_i^2 - \int \ddot{U}_i d\nu)$;
- (d) the three terms containing a derivative of A_p^- , namely $\bar{f}_p A_{p,ii}^-$, $2f_{p,i} A_{p,i}^-$, and $2\mathfrak{s}_i \bar{f}_p A_{p,i}^-$.

We now compute the explicit derivatives. We have $f = (\partial_y U(p, X_p, Y_p; \theta))^2$ and $A_p^- = |\nabla_x U(p, X_p, Y_p; \theta)|^2/n$, and by (11.15), $\partial_y \dot{U}_i(p) = M_i(p)h_p$ and $\nabla_x \dot{U}_i(p) = m_p H_p^- e_i$. Differentiating once and twice in θ_i , we obtain

$$\begin{aligned} f_{p,i} &= 2m_p h_p M_i(p), & f_{p,ii} &= 2M_i(p)^2 h_p^2 + 2m_p \partial_y \ddot{U}_i(p), \\ A_{p,i}^- &= \frac{2m_p}{n} (H_p^- M_p)_i, & A_{p,ii}^- &= \frac{2}{n} \{m_p^2 |H_p^- e_i|^2 + M_p \cdot \nabla_x \ddot{U}_i(p)\}, \end{aligned} \quad (11.20)$$

where $\ddot{U}_i(p)$ and its derivatives are evaluated at (p, X_p, Y_p) and e_i is the i th coordinate vector of $\mathbb{R}^n$.

Step 3: the groups (a) and (b). The identities (11.20) and (11.17), together with $\partial_y(m^2) = 2mh$, yield

$$\begin{aligned} \frac{1}{2n} \sum_i f_{p,ii} &= A_p^- h_p^2 + m_p \partial_y \ddot{U}_{av}(p) \\ &= (h_p^2 - 2a(p)m_p^2 h_p) A_p^- - \int_{(p,1]} m_p \partial_y g_p^\ell(Y_p) (P_{p,\ell}^- A_\ell^-)(X_p) \nu(d\ell). \end{aligned}$$

We multiply by $\bar{A}_p^-$ and take the expectation. The scalar and rest factors are independent, $\mathbb{E}^-[A_p^- \bar{A}_p^-] = \Sigma^-(p, p)$, and $\mathbb{E}^-[(P_{p,\ell}^- A_\ell^-) \bar{A}_p^-] = \Theta(p; \ell, p) = \Sigma^-(\ell, p)$ by (11.2). We conclude that group (a) equals

$$(\mathbb{E}^0 h_p^2 - 2a(p)\mathbb{E}^0[m_p^2 h_p])\Sigma^-(p, p) - \int_{(p,1]} \mathbb{E}^0[m_p \partial_y g_p^\ell(Y_p)] \Sigma^-(\ell, p) \nu(d\ell). \quad (11.21)$$

For group (b), we insert the definitions of $\mathfrak{s}_i$, x_p , and $y_{pr} = y_{rp}$ and find

$$\begin{aligned} \frac{1}{n} \sum_i \mathfrak{s}_i f_{p,i} &= 2m_p h_p \left[\epsilon \frac{\tau \cdot M_p}{n} - \int m_r \frac{M_r \cdot M_p}{n} \nu(dr) \right] \\ &= 2m_p h_p \left[\epsilon (A_p^- + x_p) - \int m_r (A_{p \wedge r}^- + y_{pr}) \nu(dr) \right]. \end{aligned}$$

We multiply by $\bar{A}_p^-$ and take the expectation. Since $\mathbb{E}^0[\epsilon \mid \mathcal{G}_p^0] = m_p$, we have $\mathbb{E}^0[m_p h_p \epsilon] = \mathbb{E}^0[m_p^2 h_p]$. The terms containing A^- produce covariances of the rest self-overlap, and the terms containing x_p and y_{pr} make up $T_p^{(1)}$ of (11.5). Therefore, group (b) equals

$$2\mathbb{E}^0[m_p^2 h_p] \Sigma^-(p, p) - 2 \int \mathbb{E}^0[m_p h_p m_r] \Sigma^-(p \wedge r, p) \nu(dr) + T_p^{(1)}. \quad (11.22)$$

The coefficients of $\Sigma^-(p, p)$ in (11.21) and (11.22) add up to $\mathbb{E}^0[h_p^2 + 2(1 - a(p))m_p^2 h_p]$, the coefficient in (J1).

Step 4: the group (c). We average the last factor of (11.19) over i . Using $\tau_i^2 = \epsilon^2 = 1$, the definitions of x_r and $y_{r\ell}$, the identity (11.17), and $\nu([0, 1]) = 1 - s$, we have

$$\begin{aligned} \frac{1}{n} \sum_i \left(s_i^2 - \int \ddot{U}_i d\nu \right) &= s - 2 \int \epsilon m_r (A_r^- + x_r) \nu(dr) + \iint m_r m_\ell (A_{r \wedge \ell}^- + y_{r\ell}) \nu(dr) \nu(d\ell) \\ &\quad + \int a(r) m_r^2 A_r^- \nu(dr) + \int \nu(dr) \int_{(r, 1]} g_r^\ell(Y_r) (P_{r, \ell}^- A_\ell^-)(X_r) \nu(d\ell). \end{aligned} \quad (11.23)$$

Here the constant s is $1 - \nu([0, 1])$: the squares $\tau_i^2 \epsilon^2$ in s_i^2 contribute 1, and the constant in (11.17) integrates to $\nu([0, 1])$. We multiply (11.23) by $\bar{f}_p \bar{A}_p^- / 2$ and take the expectation. The constant contributes $s \mathbb{E}[\bar{f}_p \bar{A}_p^-] / 2 = 0$. Each term containing A^- factorizes into a scalar expectation times $\mathbb{E}^-[A_q^- \bar{A}_p^-] = \Sigma^-(q, p)$, with $q = r$ or $q = r \wedge \ell$, or, in the last integral, times $\mathbb{E}^-[(P_{r, \ell}^- A_\ell^-) \bar{A}_p^-] = \Theta(r; \ell, p)$. The terms containing x_r and $y_{r\ell}$ make up $T_p^{(2)}$ of (11.6). Collecting these terms, we find that group (c) equals

$$\begin{aligned} & - \int \mathbb{E}^0[\bar{f}_p \epsilon m_r] \Sigma^-(r, p) \nu(dr) + \frac{1}{2} \iint \mathbb{E}^0[\bar{f}_p m_r m_\ell] \Sigma^-(r \wedge \ell, p) \nu(dr) \nu(d\ell) \\ & + \frac{1}{2} \int a(r) \mathbb{E}^0[\bar{f}_p m_r^2] \Sigma^-(r, p) \nu(dr) + \frac{1}{2} \int \nu(dr) \int_{(r, 1]} \mathbb{E}^0[\bar{f}_p g_r^\ell(Y_r)] \Theta(r; \ell, p) \nu(d\ell) + T_p^{(2)}. \end{aligned} \quad (11.24)$$

Step 5: the group (d). We compute the three terms in turn. First, by (11.20), $\sum_i |H_p^- e_i|^2 = \text{Tr}((H_p^-)^2)$, and (11.17), we have

$$\frac{1}{2n} \sum_i A_{p, ii}^- = \frac{m_p^2}{n^2} \text{Tr}((H_p^-)^2) + \frac{1}{n} M_p \cdot \nabla_x \ddot{U}_{\text{av}}(p).$$

We differentiate (11.17) in x , using $\nabla_x A^-(p, x) = 2H^-(p, x)M(p, x)/n$, where $H^-(p, x) = \nabla_x^2 U^-(p, x)$, and obtain

$$M_p \cdot \nabla_x \ddot{U}_{\text{av}}(p) = -\frac{2a(p)m_p^2}{n} M_p^T H_p^- M_p - \int_{(p, 1]} g_p^\ell(Y_p) M_p \cdot \nabla(P_{p, \ell}^- A_\ell^-)(X_p) \nu(d\ell).$$

We multiply by $\bar{f}_p$ and take the expectation. Using $\mathbb{E}^0[\bar{f}_p m_p^2] = \text{Var}^0(f_p) = v_p^0$ and $\mathbb{E}^0[\bar{f}_p g_p^\ell(Y_p)] = \text{Cov}^0(f_p, g_p^\ell(Y_p))$, we find

$$\frac{1}{2n} \sum_i \mathbb{E}[\bar{f}_p A_{p,ii}^-] = v_p^0 a_p^H - 2a(p) v_p^0 b_p^H - \int_{(p,1]} \text{Cov}^0(f_p, g_p^\ell(Y_p)) z_{p,\ell} \nu(d\ell).$$

Second, (11.20) gives $n^{-1} \sum_i f_{p,i} A_{p,i}^- = 4n^{-2} m_p^2 h_p M_p^T H_p^- M_p$, and therefore

$$\frac{1}{2n} \sum_i \mathbb{E}[2f_{p,i} A_{p,i}^-] = 4\mathbb{E}^0[m_p^2 h_p] b_p^H.$$

Third, by the definition of $\mathfrak{s}_i$, we have

$$\frac{1}{n} \sum_i \mathfrak{s}_i A_{p,i}^- = \frac{2m_p}{n^2} \left[\epsilon \tau^T H_p^- M_p - \int m_r M_r^T H_p^- M_p \nu(dr) \right].$$

We multiply by $\bar{f}_p$ and take the expectation. Conditioning ϵ on $\mathcal{G}_p^0$ and τ on $\mathcal{G}_p^-$, we have $\mathbb{E}^0[\bar{f}_p m_p \epsilon] = \mathbb{E}^0[\bar{f}_p m_p^2] = v_p^0$ and $\mathbb{E}^0[\tau^T H_p^- M_p] = \mathbb{E}^0[M_p^T H_p^- M_p]$. We conclude that

$$\frac{1}{2n} \sum_i \mathbb{E}[2\mathfrak{s}_i \bar{f}_p A_{p,i}^-] = 2v_p^0 b_p^H - 2 \int \mathbb{E}^0[\bar{f}_p m_p m_r] b_{r,p}^H \nu(dr).$$

Adding the three terms, we find that group (d) equals

$$\begin{aligned} & v_p^0 a_p^H + [4\mathbb{E}^0(m_p^2 h_p) + 2(1 - a(p))v_p^0] b_p^H \\ & - \int_{(p,1]} \text{Cov}^0(f_p, g_p^\ell(Y_p)) z_{p,\ell} \nu(d\ell) - 2 \int \mathbb{E}^0[\bar{f}_p m_p m_r] b_{r,p}^H \nu(dr), \end{aligned} \quad (11.25)$$

which is $T_p^{(3)}$ of (11.7).

Step 6: comparison with the scalar response operator. By (11.19), $C'_{pp}(0)/\alpha$ is the sum of (11.21), (11.22), (11.24), and (11.25). We take $\varphi = \Sigma^-(\cdot, p)$ in (J1)–(J7) and compare; Table 11.1 lists the origin of every term.

In detail, the integral in (11.21) is (J2), and the integral in (11.22) is (J3), with $\varphi(p \wedge r) = \Sigma^-(p \wedge r, p)$. The first three terms of (11.24) are (J4), (J6), and (J5). In the fourth term of (11.24), we write $\Theta(r; \ell, p) = \Sigma^-(\ell, p) + [\Theta(r; \ell, p) - \Sigma^-(\ell, p)]$ and use $\mathbb{E}^0[\bar{f}_p g_r^\ell(Y_r)] = \text{Cov}^0(f_p, g_r^\ell(Y_r))$. The first part is (J7), and the second part is the last line of (11.9). The remaining terms are $T_p^{(1)}$, $T_p^{(2)}$, and $T_p^{(3)}$. The terms (J1)–(J7) add up to $d^{-1}(J_{a,d} \Sigma^-(\cdot, p))(p) = d^{-1}(J_{a,d}^{\text{sc}} \Sigma^-)(p, p)$, and (11.9) follows.

Atoms of ν need no separate treatment, because the computation uses only the exact density (10.9) and the exact parameter responses (10.11). At an atom, the formulas use the right-continuous value $a(p)$ together with the strict future intervals $(p, 1]$ and $(r, 1]$, as (9.19) does, and the double integrals over $[0, 1]^2$ include their diagonal atoms. This completes the proof. $\square$

Term	Origin in (11.19)	Group
(J1)	coefficients of $\Sigma^-(p, p)$ in (11.21) and (11.22)	(a), (b)
(J2)	future part of $\partial_y \ddot{U}_i(p)$ in $f_{p,ii}$	(a)
(J3)	integral part of $\mathfrak{s}_i$, times $f_{p,i}$	(b)
(J4)	cross term of $\mathfrak{s}_i^2$	(c)
(J5)	present-level part of $-\int \ddot{U}_i d\nu$	(c)
(J6)	square of the integral part of $\mathfrak{s}_i$	(c)
(J7)	future part of $-\int \ddot{U}_i d\nu$, with $\Sigma^-(\ell, p)$ for $\Theta(r; \ell, p)$	(c)
last line of (11.9)	future part of $-\int \ddot{U}_i d\nu$, the difference $\Theta - \Sigma^-$	(c)
$T_p^{(1)}$	fluctuations x_p, y_{pr} in the score term	(b)
$T_p^{(2)}$	fluctuations $x_r, y_{r\ell}$ in $\mathfrak{s}_i^2$	(c)
$T_p^{(3)}$	all derivatives of A_p^-	(d)

Table 11.1. The terms of (11.9), with $\varphi = \Sigma^-(\cdot, p)$ in (J1)–(J7), and the groups of (11.19) from which they arise.

Proof of the integrated linear response

Proof of Lemma 11.2. The proof has three steps. First, we show that the last line of (11.9), integrated over p , equals $-\mathcal{P}/2$, and that $\mathcal{P} \geq 0$. Second, we prove the assertions about Σ^- and the trace. Third, we integrate (11.9).

Step 1: the conditional future correction. If $r \geq p$, the bracket in (11.9) vanishes by (11.2). As the inner integral runs over $\ell > r$, only levels with $p, \ell > r$ contribute. For such levels, we have $g_r^p(Y_r) = \mathbb{E}^0[f_p \mid \mathcal{G}_r^0]$, which implies that

$$\text{Cov}^0(f_p, g_r^\ell(Y_r)) = \text{Cov}^0(g_r^p(Y_r), g_r^\ell(Y_r)) = \Lambda_r(p, \ell).$$

By the law of total covariance given $\mathcal{G}_r^-$, we have

$$\begin{aligned} \Sigma^-(\ell, p) &= \mathbb{E}^- \text{Cov}^-(A_\ell^-, A_p^- \mid \mathcal{G}_r^-) + \text{Cov}^-(\mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-], \mathbb{E}^-[A_p^- \mid \mathcal{G}_r^-]) \\ &= \Sigma_r^-(p, \ell) + \Theta(r; \ell, p), \end{aligned}$$

where in the second equality we used that the covariance of the $\mathcal{G}_r^-$ -measurable variable $\mathbb{E}^-[A_\ell^- \mid \mathcal{G}_r^-]$ with A_p^- equals its covariance with $\mathbb{E}^-[A_p^- \mid \mathcal{G}_r^-]$. We conclude that $\Theta(r; \ell, p) - \Sigma^-(\ell, p) = -\Sigma_r^-(p, \ell)$. Integrating the last line of (11.9) against $\nu(dp)$, and exchanging the order of integration, we obtain

$$\frac{1}{2} \int \nu(dp) \int \nu(dr) \int_{(r,1]} \text{Cov}^0(f_p, g_r^\ell(Y_r)) [\Theta(r; \ell, p) - \Sigma^-(\ell, p)] \nu(d\ell) = -\frac{1}{2} \mathcal{P}. \quad (11.26)$$

To show that $\mathcal{P} \geq 0$, we fix $r \in [0, 1]$ and define, under the product of the scalar and rest laws,

$$\Xi_r = \int_{(r,1]} (g_r^p(Y_r) - \mathbb{E}^0 g_r^p(Y_r)) (A_p^- - \mathbb{E}^-[A_p^- \mid \mathcal{G}_r^-]) \nu(dp).$$

We expand the square of Ξ_r and use the independence of the two factors. Since $\mathbb{E}^-[(A_p^- - \mathbb{E}^-[A_p^- | \mathcal{G}_r^-])(A_\ell^- - \mathbb{E}^-[A_\ell^- | \mathcal{G}_r^-])] = \Sigma_r^-(p, \ell)$, we have

$$\mathbb{E} \Xi_r^2 = \iint_{p, \ell > r} \Lambda_r(p, \ell) \Sigma_r^-(p, \ell) \nu(dp) \nu(d\ell), \quad \text{and hence} \quad \mathcal{P} = \int \mathbb{E} \Xi_r^2 \nu(dr) \geq 0.$$

Individual entries of Λ_r and Σ_r^- need not have a sign; only the integrated form is a square.

Step 2: the covariance operator and the trace. Let $B = \bar{A}^-$ be the random function $p \mapsto \bar{A}_p^-$. The function $p \mapsto A_p^-$ is continuous with values in $[0, 1]$, and $\nu([0, 1]) \leq 1$. Then B is a random element of $L^2(\nu)$ with $\|B\|_{L^2(\nu)} \leq 1$. Its covariance operator $\Sigma^- = \mathbb{E}^-[B \otimes B]$, where $(B \otimes B)w = \langle B, w \rangle B$, is a Bochner integral in the space of trace-class operators, since $\mathbb{E}^-\|B \otimes B\|_1 = \mathbb{E}^-\|B\|^2 < \infty$. Hence Σ^- is positive and trace class, its integral kernel is $\Sigma^-(p, q)$, and its trace is $\mathbb{E}^-\|B\|_{L^2(\nu)}^2 = \int \text{Var}^-(A_p^-) \nu(dp) = V^-$. By (11.1), $B(0) = 0$; together with $\nu(\{1\}) = 0$, this places B in $\mathcal{H}_\nu$ for every realization. Since $\mathcal{H}_\nu$ is closed, the vector $\Sigma^-w = \mathbb{E}^-[\langle B, w \rangle B]$ lies in $\mathcal{H}_\nu$ for every $w \in L^2(\nu)$, that is, the range of Σ^- is contained in $\mathcal{H}_\nu$.

By (9.22), $J_{a,d}^{\text{sc}}$ is bounded on $\mathcal{H}_\nu$. The rank-one operator $J_{a,d}^{\text{sc}}(B \otimes B)$ has trace $\langle B, J_{a,d}^{\text{sc}} B \rangle$. As the trace is continuous in the trace norm and $\Sigma^- = \mathbb{E}^-[B \otimes B]$, the trace of $J_{a,d}^{\text{sc}} \Sigma^-$ on $\mathcal{H}_\nu$ is $\mathbb{E}^-\langle B, J_{a,d}^{\text{sc}} B \rangle$. Since $B \in \mathcal{H}_\nu$ and $J_{a,d}^{\text{sc}}$ is the compression of $J_{a,d}$, we have $\langle B, J_{a,d}^{\text{sc}} B \rangle = \langle B, J_{a,d} B \rangle = \int B(p)(J_{a,d} B)(p) \nu(dp)$. We expand $(J_{a,d} B)(p)$ as the sum (J1)–(J7) with $\varphi = B$. Each term is a kernel bounded by a constant depending only on β , integrated against ν , and evaluated on B at one of the levels $p, r, \ell, p \wedge r$, or $r \wedge \ell$. As $|B| \leq 1$ and ν is finite, Fubini's theorem lets us take $\mathbb{E}^-$ inside all integrals, and $\mathbb{E}^-[B(p)B(q)] = \Sigma^-(q, p)$. We obtain

$$\mathbb{E}^-\langle B, J_{a,d}^{\text{sc}} B \rangle = \int (J_{a,d} \Sigma^-(\cdot, p))(p) \nu(dp) = \text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-).$$

Step 3: integration. Let $\alpha > 0$. All terms of (11.9) are bounded measurable functions of p . We integrate (11.9) against $\nu(dp)$, use (11.26) and the definition of Tr_ν , and multiply by α , which gives (11.10). If $\alpha = 0$, then $C'_{pp}(0) = 0$ by Lemma 11.1, and both sides of (11.10) vanish. $\square$

11.2 The integrated response errors

If we bounded the error terms T_p of Lemma 11.2 one level at a time, we would meet covariances such as $\mathbb{E}^-[x_r \bar{A}_p^-]$, whose Cauchy–Schwarz bound involves $\text{Var}^-(A_p^-)$. Its integral V^- is what the cavity argument is meant to control, so such a bound is of no use. Instead, we first integrate in p against the deterministic scalar kernels. The error terms then become covariances of centered variables of the rest system with ν -averages of A^- , to which the variance bound (10.4) applies. The resulting error is linear in A and J .

As in Section 11.1, $\mathbb{E}^-$ and $\mathbb{E}^0$ refer to $u = 0$, $a_p^- = a_p^-(0)$, $j_p^- = j_p^-(0)$, and $\Gamma = \Gamma_{a,d}$ is as in Lemma 9.4. We define

$$\mathfrak{a} = \int \Gamma(p) a_p^- \nu(\mathrm{d}p), \quad \mathfrak{j} = \int \Gamma(p) j_p^- \nu(\mathrm{d}p).$$

The quantities A and J below are those of Section 3 for the full system with N spins and the same profile μ .

Lemma 11.3 (Signed response estimate). *Let $\mu = s\delta_0 + \nu$ be a deterministic profile with $\nu(\{1\}) = 0$. Then*

$$\left| \int T_p \nu(\mathrm{d}p) \right| \leq \mathfrak{a} + 24\mathfrak{j} \quad (11.27)$$

and, with $K_{\beta,3}$ from Corollary 10.6,

$$\alpha \left| \int T_p \nu(\mathrm{d}p) \right| \leq 192\beta^2 K_{\beta,3} [A + J + N^{-2}]. \quad (11.28)$$

Proof. The proof has eight steps. In Step 1, we record two elementary inequalities. In Steps 2–6, we split $\int T_p \nu(\mathrm{d}p)$ into eight pieces and bound each of them: the remaining term of $T^{(1)}$, the two terms of $T^{(2)}$, the first line of (11.7), the explicit and score parts of its future-gradient term, and its last integral over $r \geq p$ and over $r < p$. In Step 7, we add the bounds (Table 11.2) and obtain (11.27). In Step 8, we compare $\mathfrak{a} + \mathfrak{j}$ with the full system and deduce (11.28). Bounded observables and integrable Gaussian envelopes justify the exchanges of expectations and finite level integrals below; in Step 1, integrability follows from $\mathbb{E}^- \xi^2 < \infty$.

Step 1: two elementary inequalities. Let ξ be a random variable of the rest system with $\mathbb{E}^- \xi = 0$ and $\mathbb{E}^- \xi^2 < \infty$, let $r \in [0, 1]$, and let ψ be a deterministic Borel function on $[0, 1]$ such that $\psi(q)^2 \leq \Gamma(r)\Gamma(q)$ for all q . Note that $\mathbb{E}^- [\xi Y] = \mathbb{E}^- [\xi(Y - \mathbb{E}^- Y)]$ for every square-integrable Y , because ξ is centered. Using the Cauchy–Schwarz inequality and (10.4) at $u = 0$, we find

$$\left| \mathbb{E}^- \left[\xi \int \psi A^- \mathrm{d}\nu \right] \right| \leq (\mathbb{E}^- \xi^2)^{1/2} \mathrm{Var}^- \left(\int \psi A^- \mathrm{d}\nu \right)^{1/2} \leq 2(\Gamma(r)\mathfrak{j} \mathbb{E}^- \xi^2)^{1/2}, \quad (11.29)$$

where in the last inequality we used $4 \int \psi^2 j^- \mathrm{d}\nu \leq 4\Gamma(r) \int \Gamma j^- \mathrm{d}\nu = 4\Gamma(r)\mathfrak{j}$. Next, for every Borel function $\chi \geq 0$, we have

$$\int \sqrt{\Gamma j^-} \mathrm{d}\nu \leq \sqrt{\mathfrak{j}}, \quad \iint \chi(r \wedge \ell) \nu(\mathrm{d}r) \nu(\mathrm{d}\ell) \leq 2 \int \chi \mathrm{d}\nu, \quad (11.30)$$

where the first bound follows from the Cauchy–Schwarz inequality and $\nu([0, 1]) \leq 1$, and the second from $\chi(r \wedge \ell) \leq \chi(r) + \chi(\ell)$.

Step 2: the term $T^{(1)}$. The first term of (11.5) vanishes, since $\mathbb{E}^-[x_p | \mathcal{G}_p^-] = 0$ by (11.4) and $\bar{A}_p^-$ is $\mathcal{G}_p^-$ -measurable. In the second term, if $r \geq p$, then $y_{pr} = M_p \cdot (M_r - M_p)/n$ has conditional mean zero given $\mathcal{G}_p^-$, and the term vanishes. If $r < p$, then $\mathbb{E}^-[x_r | \mathcal{G}_p^-] = M_p \cdot M_r/n - A_r^- = y_{pr}$, so that $\mathbb{E}^-[y_{pr}\bar{A}_p^-] = \mathbb{E}^-[x_r\bar{A}_p^-] = \mathbb{E}^-[x_r A_p^-]$, since x_r is centered. Exchanging the integrals over p and r , we obtain

$$\int T_p^{(1)} \nu(dp) = -2 \int \mathbb{E}^-\left[x_r \int_{(r,1]} \mathbb{E}^0[m_p h_p m_r] A_p^- \nu(dp)\right] \nu(dr).$$

For fixed r , the kernel $\psi(p) = \mathbf{1}_{\{p>r\}} \mathbb{E}^0[m_p h_p m_r]$ satisfies $\psi^2 \leq \Gamma(r)\Gamma$ by (9.14), and $\mathbb{E}^-x_r^2 = j_r^-$ by (11.4). By (11.29), the inner expectation is at most $2(\Gamma(r)j_r^-)^{1/2}$ in absolute value. Integrating over r and using (11.30), we conclude that

$$\left| \int T_p^{(1)} \nu(dp) \right| \leq 4\sqrt{j} \int \sqrt{\Gamma j^-} d\nu \leq 4j.$$

Step 3: the term $T^{(2)}$. Since x_r is centered, the first term of (11.6), integrated in p , equals

$$- \int \mathbb{E}^-\left[x_r \int \mathbb{E}^0[\bar{f}_p \epsilon m_r] A_p^- \nu(dp)\right] \nu(dr).$$

The kernel $\psi(p) = \mathbb{E}^0[\bar{f}_p \epsilon m_r]$ satisfies $\psi^2 \leq \Gamma(r)\Gamma$ by (9.14). Using (11.29) and (11.30) as in Step 2, but without the factor 2, we bound this term by $2\sqrt{j} \int \sqrt{\Gamma j^-} d\nu \leq 2j$. In the second term of (11.6), the variable $y_{r\ell}$ is centered with $\mathbb{E}^-y_{r\ell}^2 \leq j_{r\wedge\ell}^-$ by (11.4), and the kernel $\psi(p) = \mathbb{E}^0[\bar{f}_p m_r m_\ell]$ satisfies $\psi^2 \leq \Gamma(r \wedge \ell)\Gamma$ by (9.15). Using (11.29) at the level $r \wedge \ell$ and both parts of (11.30), with $\chi = \sqrt{\Gamma j^-}$, we bound this term by

$$\frac{1}{2} \iint 2\sqrt{j} \chi(r \wedge \ell) \nu(dr) \nu(d\ell) \leq 2\sqrt{j} \int \chi d\nu \leq 2j.$$

Step 4: the first line of $T^{(3)}$. By Lemma 4.2, as recalled in Section 10.1, we have $0 \preceq H_p^- \preceq C_p^-$. For positive semidefinite matrices $0 \preceq H \preceq C$, we have $\text{Tr}(H^2) \leq \text{Tr}(HC) \leq \text{Tr}(C^2)$, because $\text{Tr}(H(C-H)) = \text{Tr}(H^{1/2}(C-H)H^{1/2}) \geq 0$ and $\text{Tr}((C-H)C) = \text{Tr}(C^{1/2}(C-H)C^{1/2}) \geq 0$. Further, $M_p^T H_p^- M_p \leq M_p^T C_p^- M_p$. We conclude that

$$0 \leq a_p^H \leq a_p^-, \quad 0 \leq b_p^H \leq j_p^-.$$

In addition, (9.16) implies that $0 \leq v_p^0 \leq \Gamma(p)$ and $0 \leq \mathbb{E}^0[m_p^2 h_p] \leq \Gamma(p)$, and we have $0 \leq 1 - a(p) \leq 1$. The first line of (11.7) is therefore nonnegative and at most $\Gamma(p)a_p^- + (4+2)\Gamma(p)j_p^-$. Integrating against ν , we find that it contributes at most $a + 6j$.

Step 5: the future gradient in $T^{(3)}$. We first compute $z_{p,\ell}$ for $\ell \geq p$. Given $\mathcal{G}_p^-$, the future of the rest system consists of the increments $(X_r - X_p)_{r \geq p}$ and the terminal spin τ . By

the edge likelihood (10.10) on $[p, 1]$ and the terminal Gibbs probability, as in (10.9), their conditional law has the density

$$\exp\left\{\mathcal{E}^-(\tau) + X_1 \cdot \tau - a(p)U^-(p, X_p) - \int_{(p,1]} U^-(r, X_r) \nu(\mathrm{d}r)\right\}$$

with respect to driftless Gaussian increments and counting measure on τ . If we vary $X_p = x$ with the increments fixed, every X_r with $r \geq p$ moves by the same vector. The gradient in x of the logarithm of this density is therefore the score $\tau - a(p)M_p - \int_{(p,1]} M_r \nu(\mathrm{d}r)$, while the gradient in x of $A_\ell^- = A^-(\ell, X_\ell)$ is $2H_\ell^- M_\ell/n$. We may differentiate $(P_{p,\ell}^- A_\ell^-)(x)$ under the integral, because at fixed N the density and its x -gradient have an integrable Gaussian envelope. This gives

$$\nabla(P_{p,\ell}^- A_\ell^-)(X_p) = \mathbb{E}^- \left[\frac{2}{n} H_\ell^- M_\ell + A_\ell^- \left(\tau - a(p)M_p - \int_{(p,1]} M_r \nu(\mathrm{d}r) \right) \mid \mathcal{G}_p^- \right].$$

Taking the inner product with M_p/n and using $a(p) + \nu((p, 1]) = 1$, we obtain

$$\begin{aligned} z_{p,\ell} &= \frac{2}{n^2} \mathbb{E}^- [M_p^T H_\ell^- M_\ell] + \mathbb{E}^- [\zeta_p A_\ell^-], \\ \zeta_p &= \frac{M_p \cdot (\tau - M_p)}{n} - \int_{(p,1]} \frac{M_p \cdot (M_r - M_p)}{n} \nu(\mathrm{d}r). \end{aligned} \quad (11.31)$$

We bound the explicit term. For $\ell \geq p$, (11.3) implies that $\mathbb{E}^- [C_\ell^- \mid \mathcal{G}_p^-] \preceq C_p^-$. Recall also that $H_\ell^- \preceq C_\ell^-$. Using the Cauchy–Schwarz inequality for the positive semidefinite form of H_ℓ^- and then for $\mathbb{E}^-$, we find

$$\frac{2}{n^2} |\mathbb{E}^- [M_p^T H_\ell^- M_\ell]| \leq \frac{2}{n^2} (\mathbb{E}^- [M_p^T C_\ell^- M_p] \mathbb{E}^- [M_\ell^T C_\ell^- M_\ell])^{1/2} \leq 2\sqrt{j_p^- j_\ell^-},$$

where in the last inequality we conditioned on $\mathcal{G}_p^-$ in the first factor. Next, both parts of ζ_p have conditional mean zero given $\mathcal{G}_p^-$, since $\mathbb{E}^- [\tau \mid \mathcal{G}_p^-] = M_p$ and $\mathbb{E}^- [M_r \mid \mathcal{G}_p^-] = M_p$ for $r \geq p$. The first part has conditional second moment $M_p^T C_p^- M_p/n^2$. By (11.3), each $M_p \cdot (M_r - M_p)/n$ has conditional second moment at most $M_p^T C_p^- M_p/n^2$. As $\nu((p, 1]) \leq 1$, the conditional Minkowski inequality bounds the conditional L^2 norm of the second part by $(M_p^T C_p^- M_p)^{1/2}/n$. We conclude that

$$\mathbb{E}^- \zeta_p = 0, \quad \mathbb{E}^- \zeta_p^2 \leq 4j_p^-.$$

Let $\mathcal{K}(p, \ell) = \mathbf{1}_{\{\ell > p\}} \text{Cov}^0(f_p, g_p^\ell(Y_p))$. By (9.17), we have $\mathcal{K}(p, \ell)^2 \leq \Gamma(p)\Gamma(\ell)$. The future-gradient term of (11.7), integrated in p , is $-\iint \mathcal{K}(p, \ell) z_{p,\ell} \nu(\mathrm{d}p) \nu(\mathrm{d}\ell)$. Inserting (11.31), we split it into an explicit part and a score part. Using (11.30), we find that the explicit part is at most

$$2 \iint \sqrt{\Gamma(p)j_p^-} \sqrt{\Gamma(\ell)j_\ell^-} \nu(\mathrm{d}p) \nu(\mathrm{d}\ell) = 2 \left(\int \sqrt{\Gamma j^-} \mathrm{d}\nu \right)^2 \leq 2j.$$

The score part is $-\int \mathbb{E}^-[\zeta_p \int \mathcal{K}(p, \ell) A_\ell^- \nu(d\ell)] \nu(dp)$. Using (11.29) with $\xi = \zeta_p$, $\psi = \mathcal{K}(p, \cdot)$, and the level p , and then (11.30), we bound it by $\int 2(\Gamma(p)j \cdot 4j_p^-)^{1/2} \nu(dp) \leq 4\sqrt{j} \int \sqrt{\Gamma j^-} d\nu \leq 4j$.

Step 6: the last integral in $T^{(3)}$. If $r \geq p$, conditioning on $\mathcal{G}_p^-$ gives $\mathbb{E}^-[M_r^T H_p^- M_p] = \mathbb{E}^-[M_p^T H_p^- M_p]$, so that $b_{r,p}^H = b_p^H \in [0, j_p^-]$. By (9.15), the kernel satisfies $|\mathbb{E}^0[\bar{f}_p m_p m_r]| \leq \sqrt{\Gamma(p)\Gamma(p \wedge r)} = \Gamma(p)$. This part contributes at most $2 \int \Gamma(p) j_p^- \nu([p, 1]) \nu(dp) \leq 2j$. If $r < p$, the Cauchy–Schwarz inequalities of Step 5 imply that

$$|b_{r,p}^H| \leq n^{-2} (\mathbb{E}^-[M_r^T C_p^- M_r] \mathbb{E}^-[M_p^T C_p^- M_p])^{1/2} \leq \sqrt{j_r^- j_p^-},$$

where we used $\mathbb{E}^-[C_p^- | \mathcal{G}_r^-] \preceq C_r^-$ in the first factor. The kernel is at most $\sqrt{\Gamma(r)\Gamma(p)}$ by (9.15), and by (11.30), this part also contributes at most $2(\int \sqrt{\Gamma j^-} d\nu)^2 \leq 2j$.

Step 7: the sum. Table 11.2 collects the bounds of Steps 2–6, with their running total.

Piece of $\int T_p \nu(dp)$	Step	Bound	Running total
$T^{(1)}$, second term	2	4j	4j
$T^{(2)}$, first term	3	2j	6j
$T^{(2)}$, second term	3	2j	8j
$T^{(3)}$, first line	4	$\alpha + 6j$	$\alpha + 14j$
$T^{(3)}$, future gradient, explicit part	5	2j	$\alpha + 16j$
$T^{(3)}$, future gradient, score part	5	4j	$\alpha + 20j$
$T^{(3)}$, last integral, $r \geq p$	6	2j	$\alpha + 22j$
$T^{(3)}$, last integral, $r < p$	6	2j	$\alpha + 24j$

Table 11.2. Bounds for the pieces of the integrated error terms in the proof of Lemma 11.3.

Summing the bounds, we obtain $|\int T_p \nu(dp)| \leq \alpha + (4 + 2 + 2 + 6 + 2 + 4 + 2 + 2)j = \alpha + 24j$, which is (11.27).

Step 8: transfer to the full system. Fix $q \in [0, 1]$, and consider three descendants splitting at q , with rest spins τ^1, τ^2, τ^3 . Given $\mathcal{G}_q$, they are conditionally independent with conditional mean M_q and conditional covariance C_q^- . The difference $D = \tau^1 - \tau^2$ has conditional mean zero and conditional second-moment matrix $2C_q^-$, and τ^3 is conditionally independent of it, with second-moment matrix $C_q^- + M_q M_q^T$. Then $\mathbb{E}[(D \cdot \tau^3)^2 | \mathcal{G}_q] = \text{Tr}(2C_q^- (C_q^- + M_q M_q^T)) = 2 \text{Tr}((C_q^-)^2) + 2M_q^T C_q^- M_q$, and

$$a_q^-(u) + j_q^-(u) = \frac{1}{2} \mathbb{E}_{3,q,u} \left[\left(\frac{(\tau^1 - \tau^2) \cdot \tau^3}{n} \right)^2 \right] \quad (0 \leq u \leq 1).$$

The observable on the right is nonnegative and bounded, and it does not depend on u . By Corollary 10.6 with $k = 3$, its mean at $u = 0$ is at most $K_{\beta,3}$ times its mean at $u = 1$.

At $u = 1$, we use the notation of Section 3 for the full system: $m_q \in \mathbb{R}^N$ is the full conditional mean, C_q the full conditional covariance, and a_q^{th} and j_q are as defined there. Then C_q^- is the principal $n \times n$ block of C_q , and $(M_q, 0) = m_q - m_N(q)e_N$, where e_N is the last coordinate vector, so that $M_q^T C_q^- M_q = (M_q, 0)^T C_q (M_q, 0)$. The sum of the squared entries of a principal block is at most that of the whole matrix, which implies that $\text{Tr}((C_q^-)^2) \leq \text{Tr}(C_q^2)$. Since $C_q \succeq 0$, we have

$$(M_q, 0)^T C_q (M_q, 0) \leq 2m_q^T C_q m_q + 2m_N(q)^2 (C_q)_{NN},$$

and $m_N(q)^2 (C_q)_{NN} \leq 1$. Dividing by n^2 and taking expectations, we find

$$a_q^-(1) \leq \left(\frac{N}{n}\right)^2 a_q^{\text{th}}, \quad j_q^-(1) \leq 2\left(\frac{N}{n}\right)^2 j_q + \frac{2}{n^2}.$$

We multiply the comparison by $\Gamma(q) \in [0, 1]$, integrate against ν , and use these bounds, $\nu([0, 1]) \leq 1$, and $n \geq N/2$. This yields

$$\mathfrak{a} + \mathfrak{j} \leq K_{\beta,3} \left[\left(\frac{N}{n}\right)^2 A + 2\left(\frac{N}{n}\right)^2 J + \frac{2}{n^2} \right] \leq 8K_{\beta,3} [A + J + N^{-2}].$$

Since $\alpha \leq \beta^2$, (11.27) implies that $\alpha |\int T_p \nu(\mathrm{d}p)| \leq 24\beta^2(\mathfrak{a} + \mathfrak{j})$. Together with the previous bound, this proves (11.28), with $192 = 24 \cdot 8$. $\square$

Closing the cavity inequality

In this section, we complete the cavity argument of Sections 10 and 11. The main result is Proposition 12.4: if the scalar response operator at a minimizing profile satisfies $J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I$ with $0 \leq \varepsilon_N \leq \rho/4$, then the self-overlap variance satisfies $(3\rho/4)V \leq C_{\text{cav}}[N^{-1} + A + J + H_3]$. In Corollary 12.5, we absorb the cubic moment H_3 and bound the defect $\mathcal{D}_N(\rho)$ by $C_\beta \log^2(2 + N)/(N\lambda\rho^2)$ for $1/N \leq \rho \leq 1$, wherever the defect is defined and some minimizer satisfies the stability assumption. In Section 13, we use this corollary to bound the contribution of the stable interval in the proof of Theorem 1.1, and we verify the stability assumption there with Proposition 9.9.

We rely on the following earlier results. Section 10 provides the laws of finitely many descendants with Lemma 10.5, the comparison of nonnegative observables in Corollary 10.6, the three-defect bound of Lemma 10.7, and the Taylor remainder of Lemma 10.8. Section 11 provides the integrated linear response of Lemma 11.2 and the signed response estimate of Lemma 11.3, and Lemma 9.7 states that $J_{a,d}^{\text{sc}}$ is bounded and self-adjoint. We also use the contact identities (3.9), the conditional decomposition in the proof of Lemma 3.5, the trace bound (3.10), the bound on J in Proposition 6.4, the cubic moment bound of Corollary 5.3, and the defect identity (3.15).

Standing assumptions and notation. We keep the setting of Section 10: $N \geq 2$, $n = N - 1$, $0 < \rho \leq 1$, $t = 1 - \rho$, $c = \beta^2\rho$, and $0 < s = \lambda/\beta \leq 1/2$. The parameters α , d , and $b(u)$ are given by (10.1). The profile $\mu = s\delta_0 + \nu$ is deterministic, with cumulative function a , and each statement says whether μ must minimize $\mathcal{F}_{t,s}$ or satisfy $\nu(\{1\}) = 0$. Constants C_β depend only on β ; they do not depend on N , ρ , s , the profile, or the positions and masses of its atoms.

We write $\mathbb{E}_1$, Var_1 , and Cov_1 for expectation, variance, and covariance under $\mathbb{P}_u$ at $u = 1$. Recall from Section 10.1 that $\mathbb{P}_1$ is the tilted law of Section 3.3 for the full system, so that $\mathbb{E}_1 = \mathbb{E}$. At $u = 0$, we write $\mathbb{E}^-$ and Var^- for the rest system, as in Section 11. We also recall the rest self-overlap $A_p^- = |M_p|^2/n$, the conditional magnetization $m_N(p)$ of the last spin, the descendant laws $\mathbb{E}_{k,q,u}$, the overlap defects Δ_{ab} and D_{ab} of (10.5), and the integrated third moment $\mathcal{M}_3(u)$ of (10.23). In Section 10.3, we defined the covariance $\mathcal{C}_{pp}(u) = \text{Cov}_u(m_N(p)^2, A_p^-)$ and its Taylor remainder $\mathcal{R}_{\text{cav}}$. In Section 11.1, we defined the covariance Σ^- of the rest self-overlap, the centered rest self-overlap $\bar{A}_p^-$, the error terms T_p of (11.5)–(11.7), and the quantity $\mathcal{P}$ of (11.8). Lemma 11.2 uses the rest variance

$V^- = \int \text{Var}^-(A_p^-) \nu(dp)$ and shows that $\mathcal{P} \geq 0$.

Exchangeability. The tilted law of Section 3.3 is invariant under a simultaneous permutation of the N coordinates of the couplings, the auxiliary field, and the terminal spin. Indeed, let π be a permutation of $\{1, \dots, N\}$, and let W^π and x^π denote the matrix and the vector with permuted coordinates. The reference law of W is invariant under $W \mapsto W^\pi$, and since the terminal condition in (3.2) and the Laplacian are invariant, uniqueness gives $U(q, x^\pi; W^\pi) = U(q, x; W)$. The tilt, the drift in (3.5), and the Gibbs law of the terminal spin transform accordingly. In particular, the vector $(m_1(p), \dots, m_N(p))$ of the conditional magnetizations $m_i(p) = \mathbb{E}[\sigma_i \mid \mathcal{G}_p]$ is exchangeable. As $S_p = N^{-1} \sum_i m_i(p)^2$ is a symmetric function of this vector, the pairs $(S_p, m_i(p)^2)$, $1 \leq i \leq N$, all have the same law. For every deterministic profile and every $p \in [0, 1]$, this implies that

$$\mathbb{E}_1 m_N(p)^2 = \mathbb{E} S_p, \quad \text{Cov}_1(S_p, m_N(p)^2) = \frac{1}{N} \sum_{i=1}^N \text{Cov}(S_p, m_i(p)^2) = \text{Var}(S_p). \quad (12.1)$$

At $u = 1$, the vector M_p consists of the first n coordinates of m_p , and we have

$$NS_p = m_N(p)^2 + nA_p^-. \quad (12.2)$$

Exchangeability of the sites is used in the same way in the cavity method of [81, Section 1.6].

12.1 The mechanism of the contraction

Before giving the proofs, we explain how the inequality closes. In Lemma 12.2 below, we prove the identity

$$V = \frac{V^{\text{site}}}{N} + \frac{n}{N} \left[\frac{\alpha}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-) - \frac{\alpha}{2} \mathcal{P} + \alpha \int T_p \nu(dp) + \mathcal{R}_{\text{cav}} \right],$$

where $0 \leq V^{\text{site}} \leq 1/4$. The rest variance $V^- = \text{Tr} \Sigma^-$ enters this identity through the trace pairing $\text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-)$. The other terms in the bracket are either nonpositive, as $-\alpha \mathcal{P}/2$ is, or errors controlled by A , J , H_3 , and N^{-1} (Lemmas 10.8, 11.3, and 12.3). Under the stability assumption $J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I$, the first term in the bracket is at most $(\alpha/d)(1 + \varepsilon_N)V^-$.

The contraction comes from the factor α/d . We consider two descendants a, b that split at the level q_{ab} , and fix their rest spins τ^a, τ^b . Under the reference Gaussian law, the fields acting on ϵ_a and ϵ_b are the incident row plus the last coordinate of the auxiliary field. The row contributes covariance $(\alpha u/n) \tau^a \cdot \tau^b = \alpha u R_{ab}^-$, and the last field contributes $b(u)q_{ab}$. The total covariance is

$$b(u)q_{ab} + \alpha u R_{ab}^- = d q_{ab} + \alpha u (R_{ab}^- - q_{ab}).$$

Only the part $\alpha u(R_{ab}^- - q_{ab})$ depends on the fluctuations of the rest overlap. Its coefficient αu is at most α , while the coefficient of q_{ab} is d . By (10.1), the ratio of the two satisfies

$$\frac{\alpha}{d} = \frac{\beta^2 t n / N}{\beta^2 (N - t) / N} = t \frac{N - 1}{N - t} \leq t = 1 - \rho. \quad (12.3)$$

The remaining part of the field has deterministic covariance; at $u = 1$, the last field has variance coefficient $b(1) = c = \beta^2 \rho$. Inserting this bound on the first term into the identity above, we find $V \leq t(1 + \varepsilon_N)V^- + \text{errors}$.

To close the inequality, we must compare V^- with V , and the comparison must have coefficient exactly one. Indeed, suppose that we only knew $V^- \leq KV + \text{errors}$ for a constant $K > 1$, which is the kind of bound that a comparison of the two endpoints through Corollary 10.6 gives, with $K = K_{\beta,k}$. We would then have $V \leq tK(1 + \varepsilon_N)V + \text{errors}$, which bounds V only if $tK(1 + \varepsilon_N) < 1$. This fails whenever $\rho \leq 1 - 1/K$. With coefficient one, Lemma 12.1 implies that $V^- \leq V + \text{errors}$, and hence $V \leq (1 - \rho)(1 + \varepsilon_N)V + \text{errors}$. As $(1 - \rho)(1 + \rho/4) \leq 1 - 3\rho/4$, a margin of order ρ remains once $\varepsilon_N \leq \rho/4$.

The comparison with coefficient one has two parts. At the independent endpoint $u = 0$, the rest variance is at most the second moment $\mathbb{E}^-(A_p^- - p)^2$, which is the mean of a product of two overlap defects on four descendants that split at p . The u -derivative of this mean consists of terms with three overlap defects, and Lemma 10.7 bounds them by the third moment $\mathcal{M}_3$. At the full endpoint $u = 1$, exchangeability and the contact identity $\mathbb{E}S_p = p$ show that the value of this mean is $\text{Var}(S_p)$, up to an error $1/(4n^2)$. Finally, Lemma 12.3 bounds $\mathcal{M}_3$ by A , J , H_3 , and N^{-3} , and Corollary 5.3 bounds H_3 by a small multiple of V plus a thermal term.

The rest of this section carries out this plan. In Section 12.2, we prove Lemma 12.1 and the trace identity of Lemma 12.2. In Section 12.3, we bound $\mathcal{M}_3$ in Lemma 12.3. In Section 12.4, we prove Proposition 12.4 and Corollary 12.5.

12.2 Comparison with coefficient one and the trace identity

Lemma 12.1 (Four-descendant variance comparison). *Let $\mu = s\delta_0 + v$ minimize $\mathcal{F}_{t,s}$. Then*

$$V^- \leq V + C_\beta \sup_{0 \leq u \leq 1} \mathcal{M}_3(u) + \frac{1}{4n^2}. \quad (12.4)$$

Proof. For each level p , we represent $\mathbb{E}_u(A_p^- - p)^2$ as the mean of a product of two overlap defects on four descendants. The proof has two steps. In the first step, we compare the values of this mean at $u = 0$ and $u = 1$ using Lemmas 10.5 and 10.7. In the second step, we evaluate it at $u = 1$ using exchangeability and the contact identity. Only the second step uses that μ is a minimizer.

Step 1: The interpolation. Fix $p \in [0, 1]$ and consider four descendants that split at p . Each of the pairs $(1, 2)$ and $(3, 4)$ splits at p , and we have $\Delta_{12} = R_{12}^- - p$ and $\Delta_{34} = R_{34}^- - p$. We set

$$F_p = \Delta_{12}\Delta_{34} = (R_{12}^- - p)(R_{34}^- - p).$$

Given $\mathcal{G}_p$, the four spins are conditionally independent, each with the conditional law of σ given $\mathcal{G}_p$ under $\mathbb{P}_u$. In particular, $\mathbb{E}_{4,p,u}[R_{12}^- \mid \mathcal{G}_p] = |M_p|^2/n = A_p^-$, and the same holds for R_{34}^- . As the two pairs are conditionally independent given $\mathcal{G}_p$, we have $\mathbb{E}_{4,p,u}[F_p \mid \mathcal{G}_p] = (A_p^- - p)^2$. We define

$$\mathcal{V}_p(u) = \mathbb{E}_{4,p,u}F_p = \mathbb{E}_u(A_p^- - p)^2, \quad \mathcal{V}(u) = \int \mathcal{V}_p(u) \nu(dp).$$

Since $|F_p| \leq 4$, Lemma 10.5 implies that $\mathcal{V}_p \in C^2([0, 1])$, with $\mathcal{V}'_p(u)$ given by (10.17). The right side of (10.17) is $\alpha/2$ times a signed combination of terms of two kinds. A term of the first kind is the expectation, under a tree that extends the four descendants, of F_p times one further overlap defect and a spin factor $\epsilon_x \epsilon_y$ of modulus one. The term of the second kind is $s \mathcal{V}_p(u) \int \mathbb{E}_{2,r,u} D_{12} \nu(dr)$, the product of $\mathcal{V}_p(u)$ with an expectation over an independent pair of descendants. Each term contains three overlap defects, two of which come from F_p . The absolute values of the coefficients, together with $\nu(dp)$, form a mark measure of the kind considered in Section 10.3. Using Lemma 10.7, with four initial leaves and one differentiation, and $\alpha \leq \beta^2$, we find

$$\int |\mathcal{V}'_p(u)| \nu(dp) \leq C_\beta \mathcal{M}_3(u) \quad (0 \leq u \leq 1). \quad (12.5)$$

For each p , we have $\mathcal{V}_p(0) = \mathcal{V}_p(1) - \int_0^1 \mathcal{V}'_p(u) du$. At fixed N , the derivative $\mathcal{V}'_p(u)$ is bounded uniformly in p and u , because every term in (10.17) is. We integrate the identity against $\nu(dp)$, bound $|\int_0^1 \mathcal{V}'_p(u) du|$ by $\int_0^1 |\mathcal{V}'_p(u)| du$, and exchange the integrals in p and u by Tonelli's theorem. Using (12.5), we conclude that

$$\mathcal{V}(0) \leq \mathcal{V}(1) + C_\beta \sup_{0 \leq u \leq 1} \mathcal{M}_3(u). \quad (12.6)$$

At $u = 0$, the variance of A_p^- under the rest law is at most its second moment about p , which gives $V^- \leq \mathcal{V}(0)$. Note that this bound does not require $\mathbb{E}^- A_p^- = p$. We do not know whether this identity holds, because the profile μ need not minimize the functional of the rest system.

Step 2: The full endpoint. We write $f_p^{\text{site}} = m_N(p)^2$. By (12.2), we have $nA_p^- = NS_p - f_p^{\text{site}}$. Let $p \in \text{supp } \nu$. By (12.1) and the contact identity $\mathbb{E}S_p = p$ of (3.9), we have

$\mathbb{E}_1 f_p^{\text{site}} = p$, and hence $\mathbb{E}_1 A_p^- = (Np - p)/n = p$. Using (12.1) again, we find

$$\begin{aligned} \text{Var}_1(A_p^-) &= \frac{N^2 \text{Var}(S_p) - 2N \text{Cov}_1(S_p, f_p^{\text{site}}) + \text{Var}_1(f_p^{\text{site}})}{n^2} \\ &= \frac{N^2 - 2N}{n^2} \text{Var}(S_p) + \frac{\text{Var}_1(f_p^{\text{site}})}{n^2} \leq \text{Var}(S_p) + \frac{1}{4n^2}. \end{aligned}$$

In the last step, we used $N^2 - 2N = n^2 - 1$ and $\text{Var}_1(f_p^{\text{site}}) \leq 1/4$, which holds because $0 \leq f_p^{\text{site}} \leq 1$. As $\mathbb{E}_1 A_p^- = p$, the left side equals $\mathcal{V}_p(1)$. Integrating against ν , which is carried by $\text{supp } \nu$ and has mass at most one, we find $\mathcal{V}(1) \leq V + 1/(4n^2)$. Together with (12.6) and $V^- \leq \mathcal{V}(0)$, this proves (12.4). $\square$

The next lemma expresses V itself through the cavity interpolation. It uses exchangeability but not the minimizing property.

Lemma 12.2 (Trace identity for the full variance). *Let the profile be deterministic with $\nu(\{1\}) = 0$, and let*

$$V^{\text{site}} = \int \text{Var}_1(m_N(p)^2) \nu(dp).$$

Then $0 \leq V^{\text{site}} \leq 1/4$ and

$$V = \frac{V^{\text{site}}}{N} + \frac{n}{N} \left[\frac{\alpha}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}} \Sigma^-) - \frac{\alpha}{2} \mathcal{P} + \alpha \int T_p \nu(dp) + \mathcal{R}_{\text{cav}} \right]. \quad (12.7)$$

Proof. Each variance in the definition of V^{site} lies in $[0, 1/4]$, because $0 \leq m_N(p)^2 \leq 1$, and $\nu([0, 1]) \leq 1$; this proves the bounds on V^{site} .

Fix $p \in [0, 1]$. By (12.1) and (12.2), we have

$$\text{Var}(S_p) = \text{Cov}_1(S_p, m_N(p)^2) = \frac{1}{N} \text{Var}_1(m_N(p)^2) + \frac{n}{N} \mathcal{C}_{pp}(1). \quad (12.8)$$

At $u = 0$, the law $\mathbb{P}_0$ is the product of the law of the rest system and the scalar law. The variable $m_N(p)$ depends only on the scalar component and A_p^- only on the rest component; hence $\mathcal{C}_{pp}(0) = 0$. Recall from Section 10.3 that $\mathcal{C}_{pp} \in C^2([0, 1])$. By Taylor's formula with integral remainder, we have

$$\mathcal{C}_{pp}(1) = \mathcal{C}'_{pp}(0) + \int_0^1 (1-u) \mathcal{C}''_{pp}(u) du.$$

At fixed N , all terms in this identity are bounded uniformly in p , and we may integrate it against $\nu(dp)$. The integral of the remainder is $\mathcal{R}_{\text{cav}}$ by definition. As $\nu(\{1\}) = 0$, Lemma 11.2 applies, and (11.10) gives the integral of $\mathcal{C}'_{pp}(0)$. Integrating (12.8) against $\nu(dp)$, recalling that $V = \int \text{Var}(S_p) \nu(dp)$, and inserting these two integrals, we obtain (12.7). $\square$

12.3 The third moment of the rest overlap

The remaining input to Lemma 10.8 and Lemma 12.1 is the integrated third moment $\mathcal{M}_3$ of the rest overlap defect. We bound it by the thermal quantities A and J and the cubic moment H_3 of the full system.

Lemma 12.3 (Third moment of the rest overlap). *For every deterministic profile,*

$$\sup_{0 \leq u \leq 1} \mathcal{M}_3(u) \leq K_{\beta,2} [64(A + J) + 16H_3 + 32N^{-3}],$$

where $K_{\beta,2} = e^{21\beta^2}$ is the constant of Corollary 10.6 for two leaves, and A , J , and H_3 are computed for the full system with the same profile.

Proof. The proof has three steps: an exact formula for the conditional second moment of $R_q - S_q$ in the full system, the passage from the full overlap to the rest overlap at $u = 1$, and the comparison along the interpolation.

Step 1: The full overlap. Fix $q \in [0, 1]$, and let σ^1, σ^2 be conditionally independent given $\mathcal{G}_q$, each with the conditional law of σ , so that $R_q = \sigma^1 \cdot \sigma^2 / N$. With $\delta^a = \sigma^a - m_q$, we have

$$R_q - S_q = \frac{\delta^1 \cdot \delta^2 + m_q \cdot \delta^1 + m_q \cdot \delta^2}{N}.$$

Given $\mathcal{G}_q$, the vectors δ^1 and δ^2 are independent and centered, with covariance C_q . The three summands are pairwise orthogonal in conditional L^2 ; for instance, $\mathbb{E}[(\delta^1 \cdot \delta^2)(m_q \cdot \delta^1) \mid \mathcal{G}_q] = 0$, because δ^2 is centered and independent of δ^1 . We find

$$\mathbb{E}[(R_q - S_q)^2 \mid \mathcal{G}_q] = \frac{\text{Tr } C_q^2 + 2m_q^T C_q m_q}{N^2}.$$

This is the conditional orthogonality in the proof of Lemma 3.5 (see also [4, Section 2.5, before (2.17)]); it does not use the contact identity and holds for every deterministic profile. Taking expectations, we have $\mathbb{E}(R_q - S_q)^2 = a_q^{\text{th}} + 2j_q$. As $|R_q| \leq 1$ and $0 \leq S_q \leq 1$, we have $|R_q - S_q| \leq 2$ and hence $|R_q - S_q|^3 \leq 2(R_q - S_q)^2$. Integrating against ν , we conclude that

$$\int \mathbb{E}|R_q - S_q|^3 \nu(dq) \leq 2(A + 2J). \quad (12.9)$$

Step 2: The rest overlap at $u = 1$. Since $\mathbb{P}_1$ is the full tilted law, two descendants that split at q have, under $\mathbb{E}_{2,q,1}$, the joint law of the pair (σ^1, σ^2) of Step 1. Their full and rest overlaps satisfy $NR_q = nR_{12}^- + \epsilon_1\epsilon_2$, which implies that

$$|R_{12}^- - R_q| = \frac{|R_{12}^- - \epsilon_1\epsilon_2|}{N} \leq \frac{2}{N}.$$

By the convexity of $x \mapsto |x|^3$, we have $|x + y|^3 \leq 4(|x|^3 + |y|^3)$. Applying this first to $(R_{12}^- - R_q) + (R_q - q)$ and then to $(R_q - S_q) + (S_q - q)$, we find

$$|R_{12}^- - q|^3 \leq 32N^{-3} + 16|R_q - S_q|^3 + 16|S_q - q|^3.$$

We take expectations and integrate against ν , whose mass is at most one. Using (12.9) and $H_3 = \int \mathbb{E}|S_q - q|^3 \nu(\mathrm{d}q)$, we obtain

$$\mathcal{M}_3(1) \leq 32N^{-3} + 32A + 64J + 16H_3 \leq 64(A + J) + 16H_3 + 32N^{-3}.$$

Step 3: Comparison. For each q , the function $|R_{12}^- - q|^3$ is nonnegative, bounded by 8, and depends only on the spins of two leaves that split at q . By Corollary 10.6 with $k = 2$, we have $\mathbb{E}_{2,q,u}|R_{12}^- - q|^3 \leq K_{\beta,2}\mathbb{E}_{2,q,1}|R_{12}^- - q|^3$ for every $u \in [0, 1]$. Integrating in q against ν , we obtain $\mathcal{M}_3(u) \leq K_{\beta,2}\mathcal{M}_3(1)$, and the lemma follows from Step 2. $\square$

12.4 Closing the variance inequality

We now combine the trace identity, the error bounds, and the comparison with coefficient one. The stability assumption below concerns the scalar response operator of Section 9 with the cavity coefficient d of (10.1), acting on the scalar transport space $\mathcal{H}_\nu$ of the minimizing profile.

Proposition 12.4 (Self-overlap contraction). *Let $N \geq 2$, $0 < \rho \leq 1$, and $0 < s \leq 1/2$, and let $\mu = s\delta_0 + \nu$ minimize $\mathcal{F}_{t,s}$, with cumulative function a . Suppose that the scalar response with coefficient $d = \beta^2(1 - t/N)$ satisfies*

$$J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I \text{ on } \mathcal{H}_\nu, \quad 0 \leq \varepsilon_N \leq \rho/4. \quad (12.10)$$

Then, with a constant C_{cav} depending only on β ,

$$\frac{3\rho}{4} V \leq C_{\text{cav}} [N^{-1} + A + J + H_3]. \quad (12.11)$$

Proof. The proof has three steps. We first collect the bounds on the error terms in (12.7) and on V^- . We then bound the trace pairing using (12.10). Finally, we insert the comparison (12.4) into (12.7) and absorb the resulting multiple of V .

Step 1: The errors. As μ is a minimizer, (3.9) implies that $\text{supp } \nu \subset [0, 1)$. Hence $\nu(\{1\}) = 0$, and Lemmas 11.2, 11.3, and 12.2 apply. Let

$$\mathfrak{e} = H_3 + A + J + N^{-3}.$$

We claim that

$$\alpha \left| \int T_p \nu(\mathrm{d}p) \right| \leq C_\beta [A + J + N^{-2}], \quad |\mathcal{R}_{\text{cav}}| \leq C_\beta \mathfrak{e}, \quad V^- \leq V + C_\beta \mathfrak{e} + \frac{1}{4n^2}. \quad (12.12)$$

The first bound is (11.28), with $C_\beta = 192\beta^2 K_{\beta,3}$. For the second, by Lemma 10.8, we have $|\mathcal{R}_{\text{cav}}| \leq C_\beta \sup_w \mathcal{M}_3(w)$, and Lemma 12.3 bounds this supremum by $64K_{\beta,2}\mathfrak{e}$. For the third, we insert the same bound on $\mathcal{M}_3$ into (12.4).

Step 2: The trace pairing. Let $\bar{A}^-$ be the random function $p \mapsto \bar{A}_p^-$. By Lemma 11.2, the trace pairing is $\text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-) = \mathbb{E}^- \langle \bar{A}^-, J_{a,d}^{\text{sc}}\bar{A}^- \rangle_{L^2(\nu)}$. The function $\bar{A}^-$ lies in $\mathcal{H}_\nu$ for every realization, since $\bar{A}_0^- = 0$ by (11.1) and $\nu(\{1\}) = 0$. By Fubini's theorem, we also have $\mathbb{E}^- \|\bar{A}^-\|_{L^2(\nu)}^2 = \int \text{Var}^-(A_p^-) \nu(dp) = V^-$. By Lemma 9.7, $J_{a,d}^{\text{sc}}$ is a bounded self-adjoint operator on $\mathcal{H}_\nu$, and (12.10) states that $\langle v, J_{a,d}^{\text{sc}}v \rangle_{L^2(\nu)} \leq (1 + \varepsilon_N) \|v\|_{L^2(\nu)}^2$ for every $v \in \mathcal{H}_\nu$. Taking $v = \bar{A}^-$ and averaging over the rest system, we obtain

$$\text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-) \leq (1 + \varepsilon_N)V^-.$$

Equivalently, we may test (12.10) on the eigenvectors of Σ^- with nonzero eigenvalues, which lie in $\mathcal{H}_\nu$. We multiply the last display by α/d , which lies in $[0, t]$ by (12.3). Using $(1 + \varepsilon_N)V^- \geq 0$, we find

$$\frac{\alpha}{d} \text{Tr}_\nu(J_{a,d}^{\text{sc}}\Sigma^-) \leq t(1 + \varepsilon_N)V^-, \quad (12.13)$$

whatever the sign of the trace.

Step 3: Absorption. The term $-\alpha\mathcal{P}/2$ in (12.7) is nonpositive, because $\mathcal{P} \geq 0$ by Lemma 11.2. By (12.12) and (12.13), the bracket in (12.7) is at most

$$t(1 + \varepsilon_N)V^- + C_\beta[A + J + N^{-2} + \mathfrak{e}],$$

and this upper bound is nonnegative. As $0 < n/N \leq 1$, the product of the bracket with n/N is also at most this upper bound. Using $V^{\text{site}} \leq 1/4$ in (12.7), we have

$$V \leq \frac{1}{4N} + t(1 + \varepsilon_N)V^- + C_\beta[A + J + N^{-2} + \mathfrak{e}].$$

The stability assumption (12.10) implies that

$$t(1 + \varepsilon_N) \leq (1 - \rho) \left(1 + \frac{\rho}{4}\right) = 1 - \frac{3\rho}{4} - \frac{\rho^2}{4} \leq 1 - \frac{3\rho}{4} \leq 1.$$

Inserting the bound on V^- from (12.12), using $t(1 + \varepsilon_N)V \leq (1 - 3\rho/4)V$, and bounding $t(1 + \varepsilon_N)$ by one in the error terms, we obtain

$$V \leq \left(1 - \frac{3\rho}{4}\right)V + \frac{1}{4N} + \frac{1}{4n^2} + C_\beta[A + J + N^{-2} + \mathfrak{e}].$$

We move $(1 - 3\rho/4)V$ to the left side and note that each of $(4n^2)^{-1}$, N^{-2} , and N^{-3} is at most N^{-1} ; for the first, we used $n \geq N/2$. This proves (12.11), with a constant C_{cav} depending only on β . $\square$

By Corollary 5.3, H_3 is at most an arbitrarily small multiple of V plus a thermal term. Absorbing this multiple in (12.11), we obtain the bound on the defect that we use in Section 13.

Corollary 12.5 (Defect on the stable interval). *Let $N \geq 2$, $1/N \leq \rho \leq 1$, and $0 < s = \lambda/\beta \leq 1/2$.*

- (a) *Every minimizer $\mu = s\delta_0 + \nu$ of $\mathcal{F}_{t,s}$ at these parameters that satisfies (12.10) for some ε_N , with a its cumulative function, obeys*

$$\beta(A + 2J + V) \leq C_\beta \frac{\log^2(2 + N)}{N\lambda\rho^2}, \quad (12.14)$$

where A , J , and V are computed at μ , and C_β depends only on β .

- (b) *If the defect $\mathcal{D}_N(\rho)$ is defined at ρ by (3.15) and some minimizer at ρ satisfies (12.10), then $\mathcal{D}_N(\rho)$ satisfies the bound (12.14).*
- (c) *If $1/N \leq \rho_0 \leq 1$ and, for every $\rho \in [\rho_0, 1]$, every minimizer at ρ satisfies (12.10), then*

$$\int_{\rho_0}^1 \mathcal{D}_N(\rho) d\rho \leq C_\beta \frac{\log^2(2 + N)}{N\lambda\rho_0}. \quad (12.15)$$

Proof. We first prove (a). Fix C_{cav} from (12.11), and let $\kappa_\beta = \min\{1/4, (4C_{\text{cav}})^{-1}\}$. As $\beta > 1$ and $\rho \geq 1/N$, the parameter $M = Nc = \beta^2 N\rho$ satisfies $1 < \beta^2 \leq M \leq \beta^2 N$, and $\kappa_\beta \rho \in (0, 1/4]$. Hence the hypotheses of Corollary 5.3 hold with $z = \kappa_\beta \rho$, and (5.6) gives

$$H_3 \leq \kappa_\beta \rho V + \frac{425 \log(2M)}{Ms} \log \frac{2}{\kappa_\beta \rho}.$$

We insert this bound into (12.11). The resulting term $C_{\text{cav}} \kappa_\beta \rho V$ is at most $\rho V/4$, and moving it to the left side leaves $(\rho/2)V$ there. We find

$$\rho V \leq 2C_{\text{cav}} \left[N^{-1} + A + J + \frac{425 \log(2M)}{Ms} \log \frac{2}{\kappa_\beta \rho} \right].$$

By (3.10) and (6.7), we have $A \leq (Ms)^{-1}$ and $J \leq 341(1 + \log_+ M)^2/(Ms)$. As $1 \leq M \leq \beta^2 N$ and $\rho \geq 1/N$, we also have

$$\log(2M) \leq \log(2\beta^2) + \log N, \quad 1 + \log_+ M \leq 1 + \log \beta^2 + \log N,$$

and $\log(2/(\kappa_\beta \rho)) \leq \log(2/\kappa_\beta) + \log N$. Each of these three quantities is at most $C_\beta \log(2 + N)$, because $\log(2 + N) \geq \log 4 > 1$. Recalling that $Ms = \beta^2 N\rho \cdot \lambda/\beta = \beta N\rho\lambda$, we find

$$\beta V \leq \frac{C_\beta}{N\rho} + \frac{C_\beta \log^2(2 + N)}{N\lambda\rho^2}, \quad \beta(A + 2J) \leq \frac{C_\beta \log^2(2 + N)}{N\lambda\rho}.$$

Finally, $\rho \leq 1$, $\lambda = \beta s \leq \beta/2$, and $\log(2 + N) \geq 1$ give

$$\frac{1}{N\rho} = \frac{\lambda\rho}{N\lambda\rho^2} \leq \frac{\beta}{2} \cdot \frac{\log^2(2 + N)}{N\lambda\rho^2}, \quad \frac{\log^2(2 + N)}{N\lambda\rho} \leq \frac{\log^2(2 + N)}{N\lambda\rho^2},$$

which proves (12.14).

For (b), recall from (3.15) that, at every ρ where $\mathcal{D}_N(\rho)$ is defined, $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ for every minimizer at ρ . We evaluate the right side at a minimizer that satisfies (12.10), and (a) gives the bound for $\mathcal{D}_N(\rho)$.

For (c), recall from Section 3 that $\mathcal{D}_N$ is a nonnegative measurable function defined at almost every level, and that a minimizer exists at every ρ . Under the hypothesis of (c), part (b) applies at almost every $\rho \in [\rho_0, 1]$, and gives $\mathcal{D}_N(\rho) \leq C_\beta \log^2(2 + N)/(N\lambda\rho^2)$ there. This bound holds at almost every level whichever minimizer is used, so no measurable choice of minimizers is needed to integrate it. As $\int_{\rho_0}^1 \rho^{-2} d\rho = \rho_0^{-1} - 1 \leq \rho_0^{-1}$, integrating yields (12.15). $\square$

PART V

Proof of the main theorem

Proof of the main theorem

In this section, we prove Theorem 1.1. By (3.15), four times the fractional-moment error $e_{N,\beta}(\lambda)$ is the integral of the defect $\mathcal{D}_N(\rho)$ over $\rho \in (0, 1]$. We split this integral at a cutoff ρ_0 . Above the cutoff, we verify the stability assumption of the cavity argument with Proposition 9.9 and bound the defect with Corollary 12.5. Below the cutoff, we bound the defect in Proposition 13.1, using the quantile-shell estimate of Proposition 5.2 and the root localization of Corollary 8.4. The cutoff is $\rho_0 = \min\{1, K\eta^{1/7}\}$, where K depends only on β and $\eta = e_{N,\beta}(\lambda) + L_s\lambda^5$ is the trial budget. This budget contains the unknown error, and the two estimates combine into an implicit inequality for $e_{N,\beta}(\lambda)$. In Section 13.2, we choose the tilt $\lambda = N^{-7/50}$, normalize this inequality by $N^{-7/10}$, and close it. Then Proposition 3.9 converts the bound on $e_{N,\beta}(N^{-7/50})$ into the bound on the mean bias. In Remark 13.2, we explain how the exponent 7/10 arises.

We use results from Section 3 (the defect identity (3.15), the bound (3.16), the trial budget of Lemma 3.8, the trace bound (3.10), the decomposition of Lemma 3.5, and the reduction of the mean bias in Proposition 3.9), from Sections 5 and 6 (Propositions 5.2 and 6.4), from Section 8 (Corollary 8.4), from Section 9 (Proposition 9.9), and from Section 12 (Corollary 12.5). Before the proof of Theorem 1.1, we list these inputs with the parameters at which they are used.

Standing assumptions and notation. Throughout this section, $\beta > 1$ is fixed, and C_β denotes a constant depending only on β whose value may change from line to line. The constants $\lambda_0 \leq \min\{1, \beta/4\}$ and $L_s \geq 1$ are those of Section 3; in particular, (3.17) holds for $0 < \lambda \leq 2\lambda_0$. Recall that $c = \beta^2\rho$, $M = Nc$, $s = \lambda/\beta$, and $\mu = s\delta_0 + \nu$, and that $\log_+ x = \log(\max\{1, x\})$ for $x \geq 0$. As in Section 3, a *trial budget at ρ* is a number $\eta \geq 0$ that bounds the left side of (3.18). Since every minimizer γ_* of F_ρ attains $\min_\gamma F_\rho(\gamma)$, such an η is a trial budget for every minimizer at ρ in the sense of (7.7).

13.1 The interval near $\rho = 0$

We estimate the three terms of $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ separately. By the trace bound (3.10) and (6.7), we have at every minimizer

$$A \leq \frac{1}{Ms}, \quad J \leq \frac{341(1 + \log_+ M)^2}{Ms}. \quad (13.1)$$

To estimate the self-overlap variance V , we combine Proposition 5.2 on dyadic quantile shells with the root localization of Corollary 8.4. This follows Steps 2–5 of the proof of [4, Proposition 7.7], with Proposition 5.2 in place of [4, Lemma 3.3] and Corollary 8.4 in place of [4, Lemma 7.6].

Proposition 13.1 (The interval near $\rho = 0$). *Let $N \geq 2$, $0 < \lambda \leq \lambda_0$, and $0 < \rho \leq 1$, and let $\eta > 0$ be a trial budget at ρ , that is, a bound for the left side of (3.18). Let $\omega(\rho, \eta) = \min\{1, (\eta/\rho)^{1/6}\}$ be as in (8.10). Then every minimizing profile μ at these parameters satisfies*

$$\beta(A + 2J + V) \leq C_\beta \left[\frac{\log^2(2 + N)}{N\rho\lambda} + (N\rho)^{-2/3} + (N\rho)^{-2/3}\lambda^{-1/3}\omega(\rho, \eta)^{2/3} \right], \quad (13.2)$$

$$\beta(A + 2J + V) \leq C_\beta (N\rho)^{-1/2}, \quad (13.3)$$

where A , J , and V are computed at μ . In particular, the same bounds hold for $\mathcal{D}_N(\rho)$ whenever $\mathcal{D}_N(\rho)$ is defined. If $0 < \rho_0 \leq 1$ and η is a trial budget at every $\rho \in (0, \rho_0]$, then

$$\int_0^{\rho_0} \mathcal{D}_N(\rho) \, d\rho \leq C_\beta \left[N^{-1} + \frac{\log^3(2 + N)}{N\lambda} + N^{-2/3}\rho_0^{1/3} + N^{-2/3}\lambda^{-1/3}\eta^{1/9}\rho_0^{2/9} \right]. \quad (13.4)$$

The constants depend only on β ; in particular, they do not depend on N , ρ , λ , η , the minimizer, or the positions and masses of the atoms of ν .

Proof. The proof has five steps. We partition the quantile variable into dyadic shells and bound the diameter of the image of each shell by the root localization. We then apply Proposition 5.2 on each shell and sum over the shells to bound V . Finally, we deduce the pointwise bounds and integrate them in ρ .

Step 1: Shells and localization. Fix a minimizing profile $\mu = s\delta_0 + \nu$, and let Q be the quantile function of ν from (3.6). Recall that Q pushes Lebesgue measure on $(0, 1 - s)$ forward to ν and that $a(Q(v)) \geq s + v$ for $0 < v < 1 - s$. In particular,

$$V = \int \text{Var}(S_q) \nu(dq) = \int_0^{1-s} \text{Var}(S_{Q(v)}) \, dv. \quad (13.5)$$

For the dyadic levels $u = 2^j s < 1$, $j \geq 0$, of which there are finitely many, we consider the shells

$$I_u = \{v \in (0, 1 - s) : u \leq s + v < 2u\}, \quad h_u = |I_u|.$$

(Here u denotes a level; it is unrelated to the parameter of the cavity interpolation in Section 10.) For each $v \in (0, 1 - s)$, we have $s < s + v < 1$, and there exists exactly one $j \geq 0$ such that $2^j s \leq s + v < 2^{j+1} s$; for this j , we have $u = 2^j s < 1$. Hence the shells partition $(0, 1 - s)$. Each I_u is an interval of positive length $h_u \leq u$, and $a(Q(v)) \geq s + v \geq u$ for $v \in I_u$. In other words, I_u has threshold u in the sense of

Section 5.1. A shell may contain only part of the quantile interval of an atom of v , which Proposition 5.2 allows.

We write $\omega = \omega(\rho, \eta)$, and we let $\ell(I_u) = \text{ess sup}_{v, w \in I_u} |Q(v) - Q(w)|$ be the diameter of the image of I_u , as in Section 5.3. Since μ is a minimizer, $0 < \lambda \leq \lambda_0$, and η is a trial budget for $\gamma_* = \beta\mu([0, \cdot])$, Corollary 8.4 implies that $0 \leq Q(v) \leq C_Q(s + v + \omega)$ for $0 < v < 1 - s$, with C_Q depending only on β . For $v \in I_u$, we have $s + v < 2u$. As $Q \geq 0$, this gives

$$\ell(I_u) \leq C_Q(2u + \omega) \leq 2C_Q(u + \omega).$$

Step 2: The variance on one shell. We use Proposition 5.2 on I_u , with threshold u , length h_u , diameter $\ell(I_u)$, and $\epsilon_u = (Mu)^{-1}$. The first term in (5.3) is at most $\epsilon_u(1 + \log_+(h_u/\epsilon_u)) \leq (Mu)^{-1}(1 + \log_+(Mu^2))$, because $h_u/\epsilon_u = Muh_u \leq Mu^2$. For the second term, we use $h_u \leq u$ and $(u + \omega)^{2/3} \leq u^{2/3} + \omega^{2/3}$ and find

$$\begin{aligned} (\epsilon_u \ell(I_u))^{2/3} h_u^{1/3} &\leq (2C_Q)^{2/3} (Mu)^{-2/3} (u + \omega)^{2/3} u^{1/3} \\ &\leq (2C_Q)^{2/3} M^{-2/3} (u^{1/3} + u^{-1/3} \omega^{2/3}). \end{aligned}$$

We conclude that

$$\int_{I_u} \text{Var}(S_{Q(v)}) dv \leq C_\beta \left[\frac{1 + \log_+(Mu^2)}{Mu} + M^{-2/3} (u^{1/3} + u^{-1/3} \omega^{2/3}) \right]. \quad (13.6)$$

The term $u^{-1/3} \omega^{2/3}$ accounts for the range of levels that the localization permits beyond the shell's own mass.

Step 3: Summation over the shells. For $u = 2^j s$, we have $\log_+(Mu^2) \leq \log_+(Ms^2) + 2j \log 2$. Using $\sum_{j \geq 0} 2^{-j} = \sum_{j \geq 0} j 2^{-j} = 2$, we see that the first terms of (13.6) sum to at most

$$\frac{2(1 + \log_+(Ms^2)) + 4 \log 2}{Ms} \leq \frac{5(1 + \log_+(Ms^2))}{Ms},$$

where we used $4 \log 2 < 3$. Let $c_3 = (1 - 2^{-1/3})^{-1}$. Summing downward from the largest level, which is less than one, we have $\sum_{u < 1} u^{1/3} \leq c_3$, and summing upward from s , we have $\sum_{u < 1} u^{-1/3} \leq c_3 s^{-1/3}$. As the shells partition $(0, 1 - s)$, summing (13.6) and using (13.5) yields

$$V \leq C_\beta \left[\frac{1 + \log_+(Ms^2)}{Ms} + M^{-2/3} (1 + s^{-1/3} \omega^{2/3}) \right]. \quad (13.7)$$

Step 4: The pointwise bounds. We insert (13.1) and (13.7) into $\beta(A + 2J + V)$ and use

$$Ms = \beta N \rho \lambda, \quad \beta M^{-2/3} = \beta^{-1/3} (N \rho)^{-2/3}, \quad \beta M^{-2/3} s^{-1/3} = (N \rho)^{-2/3} \lambda^{-1/3}.$$

As $M \leq \beta^2 N$ and $Ms^2 = N \rho \lambda^2 \leq N$, both $1 + \log_+ M$ and $1 + \log_+(Ms^2)$ are at most $C_\beta \log(2 + N)$. Using also $\log(2 + N) \geq 1$, we find that βA , $2\beta J$, and β times the first

term of (13.7) are at most $C_\beta \log^2(2+N)/(N\rho\lambda)$, while the other two terms of (13.7) give the last two terms of (13.2). This proves (13.2). No lower bound on M is needed here.

For (13.3), recall the first bound in (3.16): at every minimizer, $\int \mathbb{E}(R_q - q)^2 \nu(dq) \leq C \min\{1, M^{-1/2}\}$ with a numerical constant C . The measure ν is carried by the contact levels, where (3.12) gives $\mathbb{E}(R_q - q)^2 = a_q^{\text{th}} + 2j_q + \text{Var}(S_q)$. Integrating against ν , we find $\beta(A + 2J + V) = \beta \int \mathbb{E}(R_q - q)^2 \nu(dq) \leq C\beta M^{-1/2} = C(N\rho)^{-1/2}$, which is (13.3).

At every ρ at which $\mathcal{D}_N(\rho)$ is defined, a minimizer exists, and $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ for every minimizer at ρ by (3.15). This implies that the two bounds also hold for $\mathcal{D}_N(\rho)$.

Step 5: Integration. Suppose that η is a trial budget at every $\rho \in (0, \rho_0]$. By Step 4, the bounds (13.2) and (13.3) hold for $\mathcal{D}_N(\rho)$ at almost every $\rho \in (0, \rho_0]$. As they hold for every minimizer, no measurable choice of minimizers is needed to integrate them. We first use (13.3), whose singularity at $\rho = 0$ is integrable:

$$\int_0^{\min\{\rho_0, 1/N\}} \mathcal{D}_N(\rho) d\rho \leq C_\beta N^{-1/2} \int_0^{1/N} \rho^{-1/2} d\rho = 2C_\beta N^{-1}.$$

If $\rho_0 \leq 1/N$, this proves (13.4). If $\rho_0 > 1/N$, we integrate (13.2) over $[1/N, \rho_0]$. The first term contributes at most

$$\frac{\log^2(2+N)}{N\lambda} \int_{1/N}^{\rho_0} \frac{d\rho}{\rho} = \frac{\log^2(2+N) \log(N\rho_0)}{N\lambda} \leq \frac{\log^3(2+N)}{N\lambda},$$

and the second at most $N^{-2/3} \int_0^{\rho_0} \rho^{-2/3} d\rho = 3N^{-2/3} \rho_0^{1/3}$. For the third term, we use $\omega(\rho, \eta)^{2/3} = \min\{1, (\eta/\rho)^{1/9}\} \leq (\eta/\rho)^{1/9}$, which holds for all $\rho > 0$. With this bound, the third term contributes at most

$$N^{-2/3} \lambda^{-1/3} \eta^{1/9} \int_0^{\rho_0} \rho^{-2/3-1/9} d\rho = \frac{9}{2} N^{-2/3} \lambda^{-1/3} \eta^{1/9} \rho_0^{2/9}.$$

Adding the contributions, we obtain (13.4). The constants in Steps 1–5 depend only on β , through C_Q , through the absolute constants of Proposition 5.2, Proposition 6.4, and (3.16), and through powers of β . $\square$

13.2 Closure at the critical tilt

We now fix the tilt. We take $\lambda = N^{-7/50}$, so that the scalar cost λ^5 in the trial budget equals $N^{-7/10}$. The cutoff ρ_0 is defined through the trial budget, which contains the unknown error $e_{N,\beta}(\lambda)$, and we do not assume any preliminary bound on this error. After division by $N^{-7/10}$, the terms that depend on the budget become sublinear powers of the single variable $x = \eta/N^{-7/10}$, and the resulting inequality bounds x . We follow the closure in the proof of [4, Proposition 7.7]: there, the trial budget [4, (7.44)] also contains the unknown error, and the resulting inequality [4, (7.52)] is closed in Step 6 because the error enters only through sublinear powers. Here the cutoff produces the leading term $N^{-2/3} \eta^{1/21}$.

The inputs. The proof of Theorem 1.1 uses the following results, at the parameters listed. In each case, the hypotheses are verified in the step indicated.

Result	Conclusion used	Parameters in the proof
Lemma 3.8	$\eta = e + L_s \lambda^5$ is a trial budget at every $\rho \in (0, 1]$	$N \geq N_1, \lambda = N^{-7/50} \leq \lambda_0$ (Step 1)
(3.15)	$4e = \int_0^1 \mathcal{D}_N(\rho) d\rho$, and $\mathcal{D}_N(\rho) = \beta(A + 2J + V)$ at every minimizer, for almost every ρ	$s = \lambda/\beta \leq 1/4$ (Steps 1 and 4)
Proposition 9.9	$J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I$ with $\varepsilon_N \leq C_{\text{st}}[\rho^{-3/4}\eta^{1/4} + \lambda^2 + N^{-1}]$, at every minimizer	$N \geq 2, \rho \in [\rho_0, 1]$, trial budget η (Step 3)
Corollary 12.5(c)	$\int_{\rho_0}^1 \mathcal{D}_N \leq C_\beta \log^2(2 + N)/(N\lambda\rho_0)$	$1/N \leq \rho_0 \leq 1, s \leq 1/2$, stability on $[\rho_0, 1]$ (Step 3)
Proposition 13.1	the integrated bound (13.4)	$0 < \rho_0 \leq 1$, trial budget $\eta > 0$ on $(0, \rho_0]$ (Step 4)
Proposition 3.9	$(P_\beta - p_{N,\beta})/\beta \leq 32L_s\lambda^5 + 2e$	$N \geq N_1, 0 < \lambda \leq \lambda_0$ (Step 6)

Proof of Theorem 1.1. The proof has six steps. We fix the parameters and the common trial budget, define the cutoff, bound the integral of the defect above and below the cutoff, close the resulting inequality after normalization, and convert the bound on the fractional-moment error into the bound on the mean bias.

Step 1: Parameters and the trial budget. Let $N_1 = \max\{2, \lceil \lambda_0^{-50/7} \rceil\}$, which depends only on β . For every integer $N \geq N_1$, we have $N^{-7/50} \leq \lambda_0$, and we set

$$\lambda = N^{-7/50}, \quad r = N^{-7/10} = \lambda^5, \quad e = e_{N,\beta}(\lambda), \quad \eta = e + L_s r. \quad (13.8)$$

Then $0 < \lambda \leq \lambda_0 \leq \min\{1, \beta/4\}$ and $s = \lambda/\beta \leq 1/4$. By Lemma 3.8, η is a trial budget at every $\rho \in (0, 1]$. The number e is finite. Indeed, the constrained value $E_\beta(\lambda)$ is finite, and $F_{N,\beta}^c/\beta$ is a Lipschitz function of finitely many Gaussian variables (Section 3), which has finite exponential moments of every order; these moments enter (3.1). As $e \geq 0$ by (3.15) and $L_s \geq 1$, we have

$$\infty > \eta \geq L_s \lambda^5 \geq \lambda^5 > 0. \quad (13.9)$$

Step 2: The cutoff. Let $C_{\text{st}} \geq 1$ be the constant of Proposition 9.9. We define

$$K = 12C_{\text{st}}, \quad \rho_0 = \min\{1, K\eta^{1/7}\}. \quad (13.10)$$

In both cases $\rho_0 < 1$ and $\rho_0 = 1$, we have

$$\rho_0 \leq K\eta^{1/7}, \quad \rho_0 \geq \min\{1, K\lambda^{5/7}\} \geq \lambda^{5/7} = N^{-1/10} \geq N^{-1},$$

where we used (13.9), $K \geq 1$, and $\lambda \leq 1$.

Step 3: The stable interval. Suppose first that $\rho_0 < 1$, and let $\rho \in [\rho_0, 1]$. Then $\rho \geq K\eta^{1/7} \geq K\lambda^{5/7}$. Since $N^{-1} \leq \lambda \leq 1$, this implies

$$\lambda^2 \leq \lambda^{5/7} \leq \frac{\rho}{K}, \quad N^{-1} \leq \lambda^{5/7} \leq \frac{\rho}{K}, \quad \rho^{-3/4} \eta^{1/4} \leq \frac{\rho}{K^{7/4}} \leq \frac{\rho}{K},$$

where the last bound follows from $\rho^{7/4} \geq K^{7/4} \eta^{1/4}$ and $K \geq 1$. Let μ be any minimizer at ρ , with cumulative function a . The hypotheses of Proposition 9.9 hold: $N \geq 2$, $0 < \rho \leq 1$, $0 < \lambda \leq \lambda_0$, and η is a trial budget for μ . Applying the proposition with $d = \beta^2(1 - t/N)$, the cavity coefficient of (10.1), we have

$$J_{a,d}^{\text{sc}} \preceq (1 + \varepsilon_N)I \text{ on } \mathcal{H}_v, \quad 0 \leq \varepsilon_N \leq C_{\text{st}}[\rho^{-3/4} \eta^{1/4} + \lambda^2 + N^{-1}] \leq \frac{3C_{\text{st}}\rho}{K} = \frac{\rho}{4},$$

where $\varepsilon_N \geq 0$ is the number defined in Proposition 9.9. This is the stability condition (12.10), at every minimizer and every $\rho \in [\rho_0, 1]$. As $N \geq 2$, $1/N \leq \rho_0 \leq 1$, and $s \leq 1/2$, Corollary 12.5(c) applies. Using $\rho_0 = K\eta^{1/7} \geq \lambda^{5/7}$ and $\lambda^{12/7} = N^{-6/25}$, we obtain

$$\int_{\rho_0}^1 \mathcal{D}_N(\rho) d\rho \leq C_\beta \frac{\log^2(2+N)}{N\lambda\rho_0} \leq C_\beta \frac{\log^2(2+N)}{N\lambda^{12/7}} = C_\beta N^{-19/25} \log^2(2+N). \quad (13.11)$$

When $\rho_0 = 1$, the integral is zero, so (13.11) holds in both cases.

Step 4: The interval near zero. The hypotheses of Proposition 13.1 hold on $(0, \rho_0]$: we have $N \geq 2$, $0 < \lambda \leq \lambda_0$, and $0 < \rho_0 \leq 1$. By Step 1 and (13.9), $\eta > 0$ is a trial budget at every $\rho \in (0, \rho_0]$. By Step 2, we have $\rho_0 \leq K\eta^{1/7}$ in both cases $\rho_0 < 1$ and $\rho_0 = 1$, which implies that

$$\rho_0^{1/3} \leq K^{1/3} \eta^{1/21}, \quad \eta^{1/9} \rho_0^{2/9} \leq K^{2/9} \eta^{1/9+2/63} = K^{2/9} \eta^{1/7}.$$

We insert these bounds into (13.4), add (13.11), and use $4e = \int_0^1 \mathcal{D}_N(\rho) d\rho$ from (3.15). This gives

$$e \leq C_\beta \left[N^{-1} + N^{-43/50} \log^3(2+N) + N^{-19/25} \log^2(2+N) \right. \\ \left. + N^{-2/3} \eta^{1/21} + N^{-2/3} \lambda^{-1/3} \eta^{1/7} \right],$$

where we used $N^{-1} \lambda^{-1} = N^{-43/50}$. As $1 > 43/50 > 19/25$ and $\log(2+N) \geq 1$, the first three terms are at most $3N^{-19/25} \log^3(2+N)$. We conclude that

$$e \leq C_\beta \left[N^{-19/25} \log^3(2+N) + N^{-2/3} \eta^{1/21} + N^{-2/3} \lambda^{-1/3} \eta^{1/7} \right]. \quad (13.12)$$

Step 5: Normalization and closure. Let

$$x = \frac{\eta}{r} = \frac{e}{r} + L_s,$$

which is finite by (13.9) and satisfies $x \geq L_s \geq 1$. Then $\eta = rx$, and the terms of the bound for e in Step 4 have the following sizes after division by $r = N^{-7/10}$.

Term	Origin	Size after division by r
N^{-1}	$\rho \leq 1/N$, by (13.3)	$N^{-3/10}$
$N^{-43/50} \log^3(2+N)$	first term of (13.2)	$N^{-4/25} \log^3(2+N)$
$N^{-19/25} \log^2(2+N)$	stable interval, (13.11)	$N^{-3/50} \log^2(2+N)$
$N^{-2/3} \eta^{1/21}$	second term of (13.2)	$x^{1/21}$
$N^{-2/3} \lambda^{-1/3} \eta^{1/7}$	third term of (13.2)	$N^{-1/50} x^{1/7}$

The entries in the last column are exact. For the two terms that depend on the budget, we have

$$N^{-2/3} \eta^{1/21} = N^{-2/3} r^{1/21} x^{1/21} = N^{-2/3-1/30} x^{1/21} = r x^{1/21}$$

and

$$N^{-2/3} \lambda^{-1/3} \eta^{1/7} = N^{-2/3+7/150-1/10} x^{1/7} = N^{-18/25} x^{1/7} = r N^{-1/50} x^{1/7},$$

since $2/3 + 1/30 = 7/10$ and $18/25 = 7/10 + 1/50$. For the first term of (13.12), we have $N^{-19/25}/r = N^{-3/50}$. Dividing (13.12) by r and adding L_s to both sides, we find

$$x \leq L_s + C_\beta \left[N^{-3/50} \log^3(2+N) + x^{1/21} + N^{-1/50} x^{1/7} \right] \leq C_\beta x^{1/7}. \quad (13.13)$$

In the last inequality, we used that $\sup_{N \geq 2} N^{-3/50} \log^3(2+N)$ is finite, because $\log(2+N)$ grows more slowly than any power of N . We also used $N^{-1/50} \leq 1$ and the bounds $1 \leq x^{1/7}$ and $x^{1/21} \leq x^{1/7}$, which hold because $x \geq 1$. Dividing (13.13) by $x^{1/7} > 0$, we find $x^{6/7} \leq C_\beta$, that is, $x \leq C_\beta^{7/6}$. Then

$$e \leq \eta = r x \leq C_\beta N^{-7/10}.$$

This argument uses no preliminary bound on e and no smallness of η ; it uses only that x is finite.

Step 6: The mean bias. By Proposition 3.9, which applies since $0 < \lambda \leq \lambda_0$, and by Step 5, we have

$$0 \leq \frac{P_\beta - p_{N,\beta}}{\beta} \leq 32L_s \lambda^5 + 2e \leq (32L_s + 2C_\beta) N^{-7/10}.$$

Multiplying by β , we obtain $P_\beta - p_{N,\beta} \leq C_\beta N^{-7/10}$ for every $N \geq N_1$. For the finitely many integers $2 \leq N < N_1$, the mean bias $P_\beta - p_{N,\beta}$ is finite. The number

$$C_\beta^{\text{small}} = \max_{2 \leq N < N_1} N^{7/10} (P_\beta - p_{N,\beta}),$$

with the maximum interpreted as zero when $N_1 = 2$, is therefore finite and depends only on β . Replacing C_β by $\max\{C_\beta, C_\beta^{\text{small}}\}$, we obtain the bound of Theorem 1.1 for every $N \geq 2$. This completes the proof. $\square$

Remark 13.2. The leading term in (13.12) is $N^{-2/3}\eta^{1/21}$. The exponent $1/21$ is the product of the exponent $1/3$ of ρ_0 in the integrated shell estimate (13.4) and the exponent $1/7$ of η in the cutoff (13.10), which comes from the stability threshold $\rho^{-3/4}\eta^{1/4} \lesssim \rho$. If η is of order N^{-a} , the leading term is of the same order when $a = 2/3 + a/21$, that is, $20a/21 = 2/3$ and $a = 7/10$. The scalar cost λ^5 in the trial budget is then of the same order when $\lambda = N^{-7/50}$. For this choice, the localization correction has the extra factor $N^{-1/50}$ after normalization, and the stable interval and the two elementary terms in (13.4) contribute at most $N^{-3/50} \log^3(2 + N)$, as the table in Step 5 shows. The bound (13.11) on the stable interval has only a squared logarithm; the third logarithm comes from integrating the first term of (13.2) over $[1/N, \rho_0]$. These contributions have strict polynomial slack, and the leading term carries no logarithm. After the closure, $L_s N^{-7/10} \leq \eta \leq C_\beta N^{-7/10}$, so that for all sufficiently large N the cutoff satisfies $\rho_0 < 1$ and is comparable to $N^{-1/10}$.

References

- [1] Michael Aizenman, Joel Louis Lebowitz, and David Ruelle. Some rigorous results on the Sherrington–Kirkpatrick spin glass model. *Communications in Mathematical Physics*, 112(1):3–20, 1987. doi: 10.1007/BF01217677.
- [2] Michael Aizenman, Robert Sims, and Shannon Lee Starr. Extended variational principle for the Sherrington–Kirkpatrick spin-glass model. *Physical Review B*, 68(21):214403, 2003. doi: 10.1103/PhysRevB.68.214403.
- [3] Theodore Wilbur Anderson. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proceedings of the American Mathematical Society*, 6(2):170–176, 1955. doi: 10.1090/S0002-9939-1955-0069229-1.
- [4] P. M. Aronow and Patrick Lopatto. Quantitative Parisi formulas and fluctuations in the Sherrington–Kirkpatrick model, 2026. URL <https://arxiv.org/abs/2609.09103v1>. Statement and equation locators refer to arXiv:2609.09103v1.
- [5] P. M. Aronow and Patrick Lopatto. The Parisi formula for the Edwards–Anderson model in the high-dimensional limit. *Hexagon*, 2026. hexagon:2610.00155v1.
- [6] Timo Aspelmeier. Bond chaos in the Sherrington–Kirkpatrick model. *Journal of Physics A: Mathematical and Theoretical*, 41(20):205005, 2008. doi: 10.1088/1751-8113/41/20/205005. Section locators refer to arXiv:0802.3204v1.
- [7] Timo Aspelmeier. An exact relation between free energy fluctuations and bond chaos in the Sherrington–Kirkpatrick model. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(4):P04018, 2008. doi: 10.1088/1742-5468/2008/04/P04018. Equation locators refer to arXiv:0802.3007v1.
- [8] Timo Aspelmeier, Alain Billoire, Enzo Marinari, and Michael A. Moore. Finite-size corrections in the Sherrington–Kirkpatrick model. *Journal of Physics A: Mathematical and Theoretical*, 41(32):324008, 2008. doi: 10.1088/1751-8113/41/32/324008.
- [9] Antonio Auffinger and Wei-Kuo Chen. On properties of Parisi measures. *Probability Theory and Related Fields*, 161(3–4):817–850, 2015. doi: 10.1007/s00440-014-0563-y.
- [10] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Communications in Mathematical Physics*, 335(3):1429–1444, 2015. doi: 10.1007/s00220-014-2254-z. Statement locators refer to arXiv:1402.5132v3.
- [11] Antonio Auffinger, Wei-Kuo Chen, and Qiang Zeng. The SK model is infinite step replica symmetry breaking at zero temperature. *Communications on Pure and Applied*

- Mathematics*, 73(5):921–943, 2020. doi: 10.1002/cpa.21886.
- [12] Jinho Baik and Ji Oon Lee. Fluctuations of the free energy of the spherical Sherrington–Kirkpatrick model. *Journal of Statistical Physics*, 165(2):185–224, 2016. doi: 10.1007/s10955-016-1610-0. Statement locators refer to arXiv:1505.07349v3.
 - [13] Sergei Natanovich Bernstein. Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités. *Communications de la Société mathématique de Kharkow. Deuxième série*, 13(1):1–2, 1912. URL <https://www.mathnet.ru/eng/khmo107>.
 - [14] Alain Billoire. Numerical estimate of the finite-size corrections to the free energy of the Sherrington–Kirkpatrick model using Guerra–Toninelli interpolation. *Physical Review B*, 73(13):132201, 2006. doi: 10.1103/PhysRevB.73.132201.
 - [15] Stefan Boettcher. Extremal optimization for Sherrington–Kirkpatrick spin glasses. *The European Physical Journal B*, 46(4):501–505, 2005. doi: 10.1140/epjb/e2005-00280-6. Section locators refer to arXiv:cond-mat/0407130v2.
 - [16] Stefan Boettcher. Simulations of ground state fluctuations in mean-field Ising spin glasses. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(7):P07002, 2010. doi: 10.1088/1742-5468/2010/07/P07002. Section locators refer to arXiv:0906.1292v2.
 - [17] Stefan Boettcher. Ground state properties of the diluted Sherrington–Kirkpatrick spin glass. *Physical Review Letters*, 124(17):177202, 2020. doi: 10.1103/PhysRevLett.124.177202. Table locators refer to arXiv:1912.04974v3.
 - [18] Erwin Bolthausen and Alain-Sol Sznitman. On Ruelle’s probability cascades and an abstract cavity method. *Communications in Mathematical Physics*, 197(2):247–276, 1998. doi: 10.1007/s002200050450.
 - [19] Jean-Philippe Bouchaud, Florent Krzakala, and Olivier C. Martin. Energy exponents and corrections to scaling in Ising spin glasses. *Physical Review B*, 68(22):224404, 2003. doi: 10.1103/PhysRevB.68.224404. Section locators refer to arXiv:cond-mat/0212070v2.
 - [20] Anton Bovier and Anton Klimovsky. The Aizenman–Sims–Starr and Guerra’s schemes for the SK model with multidimensional spins. *Electronic Journal of Probability*, 14(8):161–241, 2009. doi: 10.1214/EJP.v14-611. Section locators refer to arXiv:0802.3467v2.
 - [21] Herm Jan Brascamp and Elliott Hershel Lieb. On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366–389, 1976. doi: 10.1016/0022-1236(76)90004-5.
 - [22] Sourav Chatterjee. Disorder chaos and multiple valleys in spin glasses, 2009. Statement locators refer to arXiv:0907.3381v4.
 - [23] Hong-Bin Chen. FRSB in the SK spin glass: convergence to full-interval support at zero temperature, 2026. Statement locators refer to arXiv:2607.18032v1.

- [24] Hong-Bin Chen, Alice Guionnet, Justin Ko, Bertrand Lacroix-A-Chez-Toine, and Jean-Christophe Mourrat. One-sided large deviations for the ground-state energy of spin glasses, 2026. Statement locators refer to arXiv:2603.06368v1.
- [25] Wei-Kuo Chen. Variational representations for the Parisi functional and the two-dimensional Guerra–Talagrand bound. *The Annals of Probability*, 45(6A):3929–3966, 2017. doi: 10.1214/16-AOP1154. Statement locators refer to arXiv:1501.06635v2.
- [26] Wei-Kuo Chen. A Gaussian convexity for logarithmic moment generating functions with applications in spin glasses. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 62(1):195–206, 2026. doi: 10.1214/24-AIHP1527. Statement locators refer to arXiv:2311.08351v2.
- [27] Wei-Kuo Chen and Wai-Kit Lam. Order of fluctuations of the free energy in the SK model at critical temperature. *ALEA, Latin American Journal of Probability and Mathematical Statistics*, 16(1):809–816, 2019. doi: 10.30757/ALEA.v16-29.
- [28] Wei-Kuo Chen, Partha Dey, and Dmitry Panchenko. Fluctuations of the free energy in the mixed p -spin models with external field. *Probability Theory and Related Fields*, 168(1–2):41–53, 2017. doi: 10.1007/s00440-016-0705-5. Statement locators refer to arXiv:1509.07071v2.
- [29] Yifeng Chu and Maxim Raginsky. Upper and lower bounds on expected soft maxima of Gaussian processes, 2026. URL <https://arxiv.org/abs/2606.22611v1>. Section and theorem locators refer to arXiv:2606.22611v1.
- [30] Francis Comets and Jacques Neveu. The Sherrington–Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case. *Communications in Mathematical Physics*, 166(3):549–564, 1995. doi: 10.1007/BF02099887.
- [31] Pierluigi Contucci and Emanuele Mingione. A multi-scale spin-glass mean-field model. *Communications in Mathematical Physics*, 368(3):1323–1344, 2019. doi: 10.1007/s00220-019-03308-8. Statement locators refer to arXiv:1804.00629v3.
- [32] Andrea Crisanti and Tommaso Rizzo. Analysis of the ∞ -replica symmetry breaking solution of the Sherrington–Kirkpatrick model. *Physical Review E*, 65(4):046137, 2002. doi: 10.1103/PhysRevE.65.046137. Equation locators refer to arXiv:cond-mat/0111037v5.
- [33] Andrea Crisanti, Giovanni Paladin, Hans-Jürgen Sommers, and Angelo Vulpiani. Replica trick and fluctuations in disordered systems. *Journal de Physique I*, 2(7):1325–1332, 1992. doi: 10.1051/jp1:1992213.
- [34] Jairo Rolim Lopes de Almeida and David James Thouless. Stability of the Sherrington–Kirkpatrick solution of a spin glass model. *Journal of Physics A: Mathematical and General*, 11(5):983–990, 1978. doi: 10.1088/0305-4470/11/5/028.

- [35] Cirano De Dominicis and Imre Kondor. Eigenvalues of the stability matrix for Parisi solution of the long-range spin-glass. *Physical Review B*, 27(1):606–608, 1983. doi: 10.1103/PhysRevB.27.606.
- [36] Partha Sarathi Dey and Taegu Kang. Fluctuations for the Sherrington–Kirkpatrick spin glass model near the critical temperature, 2026. Statement locators refer to arXiv:2603.05636v1.
- [37] Hang Du and Brice Huang. Fluctuations of the Sherrington–Kirkpatrick free energy at critical temperature, 2026. Statement locators refer to arXiv:2607.02172v2.
- [38] Silvio Franz, Giorgio Parisi, and Federico Ricci-Tersenghi. Overlap locking and nonperturbative effects in spin glasses. *Proceedings of the National Academy of Sciences*, 123(25):e2606171123, 2026. doi: 10.1073/pnas.2606171123. Section, equation, and figure locators refer to arXiv:2602.18399v2.
- [39] Stefano Ghirlanda and Francesco Guerra. General properties of overlap probability distributions in disordered spin systems. towards Parisi ultrametricity. *Journal of Physics A: Mathematical and General*, 31(46):9149–9155, 1998. doi: 10.1088/0305-4470/31/46/006. URL <https://arxiv.org/abs/cond-mat/9807333v1>.
- [40] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Communications in Mathematical Physics*, 233(1):1–12, 2003. doi: 10.1007/s00220-002-0773-5.
- [41] Francesco Guerra and Fabio Lucio Toninelli. The thermodynamic limit in mean field spin glass models. *Communications in Mathematical Physics*, 230(1):71–79, 2002. doi: 10.1007/s00220-002-0699-y.
- [42] Francesco Guerra and Fabio Lucio Toninelli. Quadratic replica coupling in the Sherrington–Kirkpatrick mean field spin glass model. *Journal of Mathematical Physics*, 43(7):3704–3716, 2002. doi: 10.1063/1.1483378. Section locators refer to arXiv:cond-mat/0201091v2.
- [43] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proceedings of the American Mathematical Society*, 144(7):3135–3150, 2016. doi: 10.1090/proc/12968. Statement locators refer to arXiv:1502.04398v4.
- [44] Aukosh Jagannath and Ian Tobasco. Some properties of the phase diagram for mixed p -spin glasses. *Probability Theory and Related Fields*, 167(3–4):615–672, 2017. doi: 10.1007/s00440-015-0691-z. Statement locators refer to arXiv:1504.02731v2.
- [45] Scott Kirkpatrick and David Sherrington. Infinite-ranged models of spin-glasses. *Physical Review B*, 17(11):4384–4403, 1978. doi: 10.1103/PhysRevB.17.4384.
- [46] Imre Kondor. Parisi’s mean-field solution for spin glasses as an analytic continuation in the replica number. *Journal of Physics A: Mathematical and General*, 16(4):L127–L131, 1983. doi: 10.1088/0305-4470/16/4/006.

- [47] Seiichiro Kusuoka and Shuta Nakajima. A quantitative replica-symmetric bound for Sherrington–Kirkpatrick model in the entire de Almeida–Thouless region, 2026. Statement locators refer to arXiv:2608.23413v2.
- [48] Patrick Lopatto. Full replica symmetry breaking in the Sherrington–Kirkpatrick model, 2026. Statement and equation locators refer to arXiv:2607.11756v4.
- [49] Patrick Lopatto. Replica symmetry up to the de Almeida–Thouless line in the Sherrington–Kirkpatrick model, 2026. Statement locators refer to arXiv:2604.11921v2.
- [50] Marc Mézard, Giorgio Parisi, Nicolas Sourlas, Gérard Toulouse, and Miguel Angel Virasoro. Nature of the spin-glass phase. *Physical Review Letters*, 52(13):1156–1159, 1984. doi: 10.1103/PhysRevLett.52.1156.
- [51] Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. SK model: the replica solution without replicas. *Europhysics Letters*, 1(2):77–82, 1986. doi: 10.1209/0295-5075/1/2/006.
- [52] Marc Mézard, Giorgio Parisi, and Miguel Ángel Virasoro. *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific, Singapore, 1987. doi: 10.1142/0271.
- [53] Cécile Monthus and Thomas Garel. Matching between typical fluctuations and large deviations in disordered systems: application to the statistics of the ground state energy in the SK spin-glass model. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(2):P02023, 2010. doi: 10.1088/1742-5468/2010/02/P02023. Section locators refer to arXiv:0912.2875v2.
- [54] OpenAI. The low-temperature Sherrington–Kirkpatrick fluctuation scale. OpenAI mathematics research repository, September 2026. URL <https://github.com/openai/math/tree/adc7f1241b42e322a6451854ab7e4b4c146bf78a/preprints/The-low-temperature-Sherrington-Kirkpatrick-fluctuation-scale-September-24-2026>. September 24, 2026 manuscript; initial repository version at commit adc7f1241b42e322a6451854ab7e4b4c146bf78a.
- [55] OpenAI. Full support of the zero-temperature Sherrington–Kirkpatrick order parameter. OpenAI mathematics research repository, September 2026. URL <https://github.com/openai/math/tree/adc7f1241b42e322a6451854ab7e4b4c146bf78a/preprints/Full-support-of-the-zero-temperature-Sherrington-Kirkpatrick-order-parameter-September-27-2026>. September 27, 2026 manuscript; initial repository version at commit adc7f1241b42e322a6451854ab7e4b4c146bf78a.
- [56] OpenAI. The low-temperature Sherrington–Kirkpatrick free-energy limiting law. OpenAI mathematics research repository, September 2026. URL <https://github.com/openai/math/tree/adc7f1241b42e322a6451854ab7e4b4c146bf78a/preprints/The-low-temperature-Sherrington-Kirkpatrick-free-energy-limiting-law>

- September-24-2026. September 24, 2026 manuscript; initial repository version at commit `adc7f1241b42e322a6451854ab7e4b4c146bf78a`.
- [57] OpenAI. Mathematics research repository. GitHub repository, 2026. URL <https://github.com/openai/math/tree/adc7f1241b42e322a6451854ab7e4b4c146bf78a>. Initial repository version at commit `adc7f1241b42e322a6451854ab7e4b4c146bf78a`.
 - [58] Matteo Palassini. Ground-state energy fluctuations in the Sherrington–Kirkpatrick model. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10005, 2008. doi: 10.1088/1742-5468/2008/10/P10005.
 - [59] Dmitry Panchenko. *The Sherrington–Kirkpatrick Model*. Springer Monographs in Mathematics. Springer, New York, 2013. doi: 10.1007/978-1-4614-6289-7.
 - [60] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Annals of Mathematics*, 177(1):383–393, 2013. doi: 10.4007/annals.2013.177.1.8.
 - [61] Dmitry Panchenko. The Parisi formula for mixed p -spin models. *The Annals of Probability*, 42(3):946–958, 2014. doi: 10.1214/12-AOP800.
 - [62] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Physical Review Letters*, 43(23):1754–1756, 1979. doi: 10.1103/PhysRevLett.43.1754.
 - [63] Giorgio Parisi. The order parameter for spin glasses: a function on the interval 0–1. *Journal of Physics A: Mathematical and General*, 13(3):1101–1112, 1980. doi: 10.1088/0305-4470/13/3/042.
 - [64] Giorgio Parisi. A sequence of approximated solutions to the S-K model for spin glasses. *Journal of Physics A: Mathematical and General*, 13(4):L115–L121, 1980. doi: 10.1088/0305-4470/13/4/009.
 - [65] Giorgio Parisi. Order parameter for spin-glasses. *Physical Review Letters*, 50(24):1946–1948, 1983. doi: 10.1103/PhysRevLett.50.1946.
 - [66] Giorgio Parisi and Paolo Biscari. Finite-volume corrections to the mean-field solution of the SK model. *Journal of Physics A: Mathematical and General*, 25(18):4787–4797, 1992. doi: 10.1088/0305-4470/25/18/013.
 - [67] Giorgio Parisi and Tommaso Rizzo. Large deviations in the free energy of mean-field spin glasses. *Physical Review Letters*, 101(11):117205, 2008. doi: 10.1103/PhysRevLett.101.117205. Page locators refer to arXiv:0706.1180v2.
 - [68] Giorgio Parisi and Tommaso Rizzo. Phase diagram and large deviations in the free energy of mean-field spin glasses. *Physical Review B*, 79(13):134205, 2009. doi: 10.1103/PhysRevB.79.134205.
 - [69] Giorgio Parisi and Tommaso Rizzo. Universality and deviations in disordered systems. *Physical Review B*, 81(9):094201, 2010. doi: 10.1103/PhysRevB.81.094201. Equation locators refer to arXiv:0901.1100v1.

- [70] Giorgio Parisi, Fèlix Ritort, and František Slanina. Several results on the finite-size corrections in the Sherrington–Kirkpatrick spin-glass model. *Journal of Physics A: Mathematical and General*, 26(15):3775–3789, 1993. doi: 10.1088/0305-4470/26/15/026.
- [71] Giorgio Parisi, Fèlix Ritort, and František Slanina. Critical finite-size corrections for the Sherrington–Kirkpatrick spin glass. *Journal of Physics A: Mathematical and General*, 26(2):247–259, 1993. doi: 10.1088/0305-4470/26/2/013.
- [72] Giorgio Parisi, Leopoldo Sarra, and Lorenzo Talamanca. Study of longitudinal fluctuations of the Sherrington–Kirkpatrick model. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(3):033302, 2019. doi: 10.1088/1742-5468/ab0c1b. Section locators refer to arXiv:1812.06334v1.
- [73] Yan Ru Pei. Upper deviations of the Sherrington–Kirkpatrick free energy: the fifth-order law and the onset of constrained Parisi measures, 2026. Statement locators refer to arXiv:2610.06480v2.
- [74] David Ruelle. A mathematical reformulation of Derrida’s REM and GREM. *Communications in Mathematical Physics*, 108(2):225–239, 1987. doi: 10.1007/BF01210613.
- [75] Adrien Schertzer. The order of free energy fluctuations in the critical Sherrington–Kirkpatrick model revisited, 2026. arXiv:2606.21360v1.
- [76] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical Review Letters*, 35(26):1792–1796, 1975. doi: 10.1103/PhysRevLett.35.1792.
- [77] H.-J. Sommers and W. Dupont. Distribution of frozen fields in the mean-field theory of spin glasses. *Journal of Physics C: Solid State Physics*, 17(32):5785–5793, 1984. doi: 10.1088/0022-3719/17/32/012.
- [78] Michel Talagrand. The Parisi formula. *Annals of Mathematics*, 163(1):221–263, 2006. doi: 10.4007/annals.2006.163.221.
- [79] Michel Talagrand. Parisi measures. *Journal of Functional Analysis*, 231(2):269–286, 2006. doi: 10.1016/j.jfa.2005.03.001.
- [80] Michel Talagrand. Large deviations, Guerra’s and A.S.S. schemes, and the Parisi hypothesis. *Journal of Statistical Physics*, 126(4–5):837–894, 2007. doi: 10.1007/s10955-006-9108-9.
- [81] Michel Talagrand. *Mean Field Models for Spin Glasses. Volume I: Basic Examples*, volume 54 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*. Springer, Berlin, Heidelberg, 2011. doi: 10.1007/978-3-642-15202-3. URL <https://michel.talagrand.net/challenge/volume1.pdf>. Section and statement locators refer to the author draft dated September 14, 2010.
- [82] Michel Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*. Springer, Berlin, Heidelberg, 2011. doi: 10.1007/978-3-642-22253-5.

- [83] David James Thouless, Philip Warren Anderson, and Richard G. Palmer. Solution of ‘Solvable model of a spin glass’. *Philosophical Magazine*, 35(3):593–601, 1977. doi: 10.1080/14786437708235992.
- [84] Fabio Lucio Toninelli. About the Almeida–Thouless transition line in the Sherrington–Kirkpatrick mean field spin glass model. *Europhysics Letters*, 60:764–767, 2002. doi: 10.1209/epl/i2002-00667-5.
- [85] Felix Voigtlaender. A general version of Price’s theorem: A tool for bounding the expectation of nonlinear functions of Gaussian random vectors. *Journal of Theoretical Probability*, 34(3):1474–1485, 2021. doi: 10.1007/s10959-020-01017-w.
- [86] Yuxin Zhou. Existence of full replica symmetry breaking for the Sherrington–Kirkpatrick model at low temperature. *Communications on Pure and Applied Mathematics*, 79(7):1746–1770, 2026. doi: 10.1002/cpa.70038.