

FREE ENERGY OF NON-CONVEX MULTI-SPECIES ISING SPIN GLASSES WITH EXTERNAL FIELDS

JEANNE BOURSIER AND JEAN-CHRISTOPHE MOURRAT 

ABSTRACT. We identify the limit free energy of mean-field multi-species spin glasses with ± 1 spins, allowing for the presence of external fields, and without assuming any convexity property of the covariance function of the Hamiltonian. The limit is given by an “uninverted” variational formula over Brownian martingales. The proof consists in first identifying the limit free energy of configurations with a fixed magnetization. Once the magnetization is fixed, we can show the convergence of the enriched free energy to the solution of a Hamilton–Jacobi equation. The magnetization constraint is also used to show the convexity of the initial condition in this equation, and this yields a variational representation of the solution, as a Hopf formula. Convex duality then turns this formula into an optimization over martingales whose initial values are the prescribed magnetizations.

1. INTRODUCTION

In this paper, we identify the limit free energy of mean-field multi-species spin glasses with ± 1 spins. For models with a convex covariance function, the limit free energy is given by the Parisi formula [3, 8, 15, 27, 28, 33]. For models with arbitrary covariance function but no external field, this limit was also identified, and takes the form of a Hopf formula [7]. However, the variational formulas obtained in these works are known not to generalize to non-convex models with external fields [21, Section 6]. In our convention with a minus sign in the definition of the free energy, the Parisi formula (for convex models) and the formula in [7] (for models without external field) represent the limit free energy as a supremum. Here, we obtain instead an “uninverted” variational representation of the limit, which takes the form of an infimum over a space of Brownian martingales. This representation is that introduced in [24] for spin glasses with a single type of spins and extended in [6] to multi-species convex models.

Let us define the model more precisely. Let $\mathcal{S}$ be a finite set of species labels. For every $N \in \mathbb{N}$, let $(I_{N,s})_{s \in \mathcal{S}}$ be a partition of $\{1, \dots, N\}$, and let

$$\lambda_{N,s} := |I_{N,s}|/N, \quad \lambda_N := (\lambda_{N,s})_{s \in \mathcal{S}}.$$

We assume that λ_N converges to a vector λ with strictly positive entries. We also define the multi-species overlap

$$R_{N,s}(\sigma, \sigma') := \frac{1}{N} \sum_{i \in I_{N,s}} \sigma_i \sigma'_i, \quad R_N(\sigma, \sigma') := (R_{N,s}(\sigma, \sigma'))_{s \in \mathcal{S}}.$$

Let $\xi : \mathbb{R}^{\mathcal{S}} \rightarrow \mathbb{R}$ admit the expansion

$$(1.1) \quad \xi(x) = \sum_{k=1}^{\infty} \sum_{s_1, \dots, s_k \in \mathcal{S}} \beta_{s_1, \dots, s_k}^2 \prod_{j=1}^k x_{s_j},$$

which we assume to converge absolutely for every $x \in \mathbb{R}^{\mathcal{S}}$. Let H_N be a centered Gaussian field on $\{-1, 1\}^N$ with covariance

$$(1.2) \quad \mathbb{E}[H_N(\sigma)H_N(\sigma')] = N\xi(R_N(\sigma, \sigma')).$$

As discussed in [23, Section 6], the stated assumption on ξ guarantees the existence of such a Gaussian field, and, in the sense made precise there, functions ξ of this form are essentially the only ones for which a Gaussian field exists satisfying (1.2). We impose no convexity assumption on ξ . We fix a deterministic external field $h = (h_s)_{s \in \mathcal{S}} \in \mathbb{R}^{\mathcal{S}}$, so that every spin of species s feels the field h_s . The aim of this paper is to identify, for each $t \geq 0$, the large- N limit of the free energy

$$(1.3) \quad \bar{F}_N(t, 0) := -\frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \{-1, 1\}^N} 2^{-N} \exp(\sqrt{2t} H_N(\sigma) - Nt\xi(\lambda_N) + \sum_{s \in \mathcal{S}} h_s \sum_{i \in I_{N,s}} \sigma_i).$$

We will later extend $\bar{F}_N$ so that its second argument can be a function taking values in $\mathbb{R}_+^{\mathcal{S}}$; in (1.3) we just take this function to be constant equal to zero.

We next introduce the objects necessary to the description of the limit free energy. Let

$$I(z) := \frac{1+z}{2} \log(1+z) + \frac{1-z}{2} \log(1-z), \quad -1 \leq z \leq 1,$$

with $0 \log 0 = 0$. This is the convex conjugate of $\log \cosh$. Let $T = (T_s)_{s \in \mathcal{S}} \in (0, \infty)^{\mathcal{S}}$, let $(B_s)_{s \in \mathcal{S}}$ be independent standard Brownian motions, and let $\mathcal{F}^s = (\mathcal{F}_v^s)_{v \geq 0}$ be the completed natural filtration of B_s . We denote by $\mathcal{M}_T$ the set of processes $\alpha = (\alpha_s)_{s \in \mathcal{S}}$ such that α_s is an $\mathcal{F}^s$ -martingale on $[0, T_s]$ with values in $[-1, 1]$. In particular, $\alpha_s(0) = \mathbb{E} \alpha_s(T_s)$ is deterministic. For every $\alpha \in \mathcal{M}_T$, we write

$$\mathcal{C}_T(\alpha) := \sum_s \lambda_s \left\{ \mathbb{E}[I(\alpha_s(T_s)) - \sqrt{2} \alpha_s(T_s) B_s(T_s)] + \int_0^{T_s} (1 - \mathbb{E}[\alpha_s(v)^2]) dv \right\},$$

and for every $x \in \mathbb{R}_+^{\mathcal{S}}$, we set

$$(1.4) \quad \mathcal{K}_{t,\alpha}^T(x) := \sup_{z \in [0,1]^{\mathcal{S}}} \left\{ x \cdot (\lambda z) + t\xi(\lambda z) - \sum_s \lambda_s \int_0^{T_s} (z_s - \mathbb{E}[\alpha_s(v)^2])_+ dv \right\},$$

where $\lambda z := (\lambda_s z_s)_s$ and $u_+ := \max(u, 0)$.

Theorem 1.1 (Identification of the limit free energy). *For every $t \geq 0$, $h \in \mathbb{R}^{\mathcal{S}}$, and T satisfying $T_s \geq t \partial_s \xi(\lambda)$ for every $s \in \mathcal{S}$, we have*

$$(1.5) \quad \lim_{N \rightarrow \infty} \bar{F}_N(t, 0) = \inf_{\alpha \in \mathcal{M}_T} \left\{ \mathcal{C}_T(\alpha) - \sum_s \lambda_s h_s \alpha_s(0) + \mathcal{K}_{t,\alpha}^T(0) \right\},$$

and the infimum is attained.

With the help of (9.28), one can check that this formula coincides with that shown in [6, Theorem 1.5] for convex models. In Theorem 10.2, we also give a more abstract representation of the limit free energy that was conjectured in [18] and shown in [6, Theorem 1.4] for convex models.

Our approach to the proof of Theorem 1.1 consists in first identifying the free energy of configurations with a constrained magnetization. For every integer vector $k = (k_s)_{s \in \mathcal{S}}$ with $0 \leq k_s \leq |I_{N,s}|$, we define

$$\Omega_{N,k} := \left\{ \sigma \in \{-1, 1\}^N : |\{i \in I_{N,s} : \sigma_i = 1\}| = k_s \text{ for every } s \right\}, \quad m_{N,k,s} := \frac{2k_s}{|I_{N,s}|} - 1,$$

with $m_{N,k} := (m_{N,k,s})_{s \in \mathcal{S}}$. We call the set $\Omega_{N,k}$ a slice. The free energy restricted to this slice is defined by

$$(1.6) \quad \bar{F}_{N,k}(t, 0) := -\frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \Omega_{N,k}} 2^{-N} \exp(\sqrt{2t} H_N(\sigma) - Nt\xi(\lambda_N)).$$

The contribution of the deterministic field is constant on $\Omega_{N,k}$, and we removed it in the definition of $\overline{F}_{N,k}$. Since there are at most $(N+1)^{|\mathcal{S}|}$ slices, the limit free energies over restricted slices determine that of the unconstrained model.

To state our results concerning these constrained models, we first introduce enriched free energies where the second argument, set at 0 in (1.3) and (1.6), is more general. Let $\mathcal{Q}$ denote the set of right-continuous increasing maps $q : [0, 1) \rightarrow \mathbb{R}_+$, and let $\mathcal{Q}_p := \mathcal{Q} \cap L^p[0, 1)$ for $1 \leq p \leq \infty$. Throughout, the word “increasing” is used in the sense of non-strict inequalities, that is, here we are asking that for every $u \leq v \in [0, 1)$, we have $q(u) \leq q(v)$ (and likewise for “decreasing”).

For every $t \geq 0$, we extend $\overline{F}_{N,k}(t, \cdot)$ to a function on $\mathcal{Q}_1^\mathcal{S}$, the enriched free energy of the slice, and we do the same for $\overline{F}_N(t, \cdot)$. The path q encodes the law of an additional random external field acting on the spins, with an ultrametric structure prescribed by q ; we will comment further on this shortly after stating the next theorem.

The limit of the enriched free energy of a slice is described in terms of the following scalar quantities. For every $T \in (0, +\infty)$ and probability measure ν on $[0, T]$, let $\Phi_\nu = \Phi_\nu(v, x) : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ solve

$$(1.7) \quad -\partial_v \Phi_\nu = \partial_x^2 \Phi_\nu + \nu([0, v])(\partial_x \Phi_\nu)^2, \quad \Phi_\nu(T, x) = \log \cosh x,$$

and let

$$(1.8) \quad \psi_\circ(\nu; y) := T - \Phi_\nu(0, y), \quad g_m(\nu) := \sup_{y \in \mathbb{R}} \{\psi_\circ(\nu; y) + ym\}, \quad m \in [-1, 1].$$

These quantities do not depend on T , as long as $[0, T]$ contains the support of ν . We also evaluate $\psi_\circ$ and g_m at a scalar path q , by evaluating them at the law of $q(U)$, where U is uniform on $[0, 1]$. The initial condition for the slice $m \in [-1, 1]^\mathcal{S}$ is

$$(1.9) \quad \psi_m(q) := \sum_s \lambda_s g_{m_s}(q_s), \quad q \in \mathcal{Q}_1^\mathcal{S}.$$

The extension to $\mathcal{Q}_1^\mathcal{S}$ follows from the Lipschitz estimate proved in Lemma 6.7. For $m = 0$, one can show that $g_0 = \psi_\circ(\cdot; 0)$, and this is the centered initial condition that showed up in [7, 16]. The function g_m is the limit of the free energy, at $t = 0$, of a system with a single species constrained to have magnetization m ; see Proposition 7.5. We define $\mathcal{M}_T^m := \{\alpha \in \mathcal{M}_T : \alpha_s(0) = m_s \text{ for every } s \in \mathcal{S}\}$.

Theorem 1.2 (Limit free energy on a slice). *For every $t \geq 0$ and $q \in \mathcal{Q}_2^\mathcal{S}$, there is a continuous function $m \mapsto f_m(t, q)$ on $[-1, 1]^\mathcal{S}$ such that*

$$(1.10) \quad \lim_{N \rightarrow \infty} \sup_k |\overline{F}_{N,k}(t, q) - f_{m_{N,k}}(t, q)| = 0.$$

For every $m \in [-1, 1]^\mathcal{S}$, the function f_m is the Lipschitz viscosity solution of

$$(1.11) \quad \partial_t f_m - \int_0^1 \xi(\partial_q f_m(r)) dr = 0, \quad f_m(0, \cdot) = \psi_m,$$

on the closed cone $\mathcal{Q}_2^\mathcal{S}$, in the sense specified in Section 8. It has the Hopf representation

$$(1.12) \quad f_m(t, q) = \sup_{p \in \mathcal{Q}_\infty^\mathcal{S}} \inf_{q' \in \mathcal{Q}_\infty^\mathcal{S}} \left\{ \psi_m(q') + \int_0^1 (q(r) - q'(r)) \cdot p(r) dr + t \int_0^1 \xi(p(r)) dr \right\}.$$

For every $q \in \mathcal{Q}_\infty^\mathcal{S}$ and every $T \in (0, \infty)^\mathcal{S}$ with $T_s \geq \|q_s\|_\infty + t \partial_s \xi(\lambda)$ for every s , we have

$$(1.13) \quad f_m(t, q) = \inf_{\alpha \in \mathcal{M}_T^m} \left\{ \mathcal{C}_T(\alpha) + \int_0^1 \mathcal{K}_{t,\alpha}^T(q(r)) dr \right\},$$

and the infimum is attained.

To recover the limit free energy of the full system, we will show that, for every $t \geq 0$ and $q \in \mathcal{Q}_2^\mathcal{S}$,

$$(1.14) \quad \lim_{N \rightarrow \infty} \bar{F}_N(t, q) = \inf_{m \in [-1, 1]^\mathcal{S}} \left\{ f_m(t, q) - \sum_s \lambda_s h_s m_s \right\}.$$

Combining this identity with (1.13) releases the constraint on $\alpha(0)$ and proves Theorem 1.1. It also gives its enriched version, with $\mathcal{K}_{t, \alpha}^T(0)$ replaced by $\int_0^1 \mathcal{K}_{t, \alpha}^T(q(r)) dr$.

Proof ideas. The core of the proof of Theorem 1.2 consists of two facts: the enriched free energy of a slice satisfies the Hamilton–Jacobi equation (1.11) up to an error term that is small, and the initial condition ψ_m is convex as a function of the path variable, so that the solution has the Hopf representation (1.12).

In the usual formulation of the Hamilton–Jacobi approach [5, 7, 8, 9, 13, 21, 22, 23, 25], the enriched free energy is defined by coupling the spins to a Gaussian field indexed by a Poisson–Dirichlet cascade, the path q prescribing the covariance between the fields at two leaves of the cascade as a function of their overlap. The Gibbs measure then features a cascade variable in addition to the spin configuration, and the error term in the PDE vanishes if and only if the overlap of two configurations sampled from the Gibbs measure is determined by the overlap of the cascade variables; see for instance [13, (6.79)]. Showing that the error term is small thus requires the overlap of the configurations to synchronize with the overlap of the cascade variables. This synchronization is derived from the ultrametricity of the Gibbs measure, by the mechanism of [28, 29, 30], and ultrametricity itself is only known after a perturbation of the Hamiltonian, using the approach of [26] based on the Ghirlanda–Guerra identities. What we obtain then is comparable to knowing that a function solves a Hamilton–Jacobi equation only almost everywhere, which does not suffice to identify it in general, and overcoming this difficulty is a major challenge.

In this paper, we use instead an equivalent definition of the enriched free energy that does not resort to any cascade and that features a single random field, which we call the branch. The branch is a stochastic process X on $[0, 1]$, taking values in $\mathbb{R}^N$, and the enriched Hamiltonian features a term $X_1 \cdot \sigma$ in addition to the terms already present in (1.3). The law of the branch corresponds, in the setting of the cascade formulation, to the law of the Gaussian field attached to a leaf that is selected according to the Gibbs weights. This law can be described explicitly as the solution of a stochastic differential equation in $\mathbb{R}^N$; we refer to Section 2 for the precise definition. Appendix A is not necessary for the purposes of showing any result in this paper, but we provide it to explain in more detail this correspondence between the classical definition of the enriched free energy and the one we adopt here.

With this definition, the derivatives of the free energy necessarily take a different form from the classical one. They involve a family of Gibbs measures indexed by the time variable of the branch; this family corresponds to the disintegration of the usual Gibbs measure from the cascade representation according to the value of the overlap between the cascade variables. Controlling the error term in the PDE then amounts to showing that the overlap concentrates under these Gibbs measures. Adapting the approach of [1] for obtaining a quantitative proof of the Parisi formula for the SK model without external field, we show that the fluctuations of the overlap under these Gibbs measures are indeed small, with an explicit bound in terms of N and of a lower bound on the derivative of the path (Theorem 5.1). One key observation is that conditionally on a sampled configuration, the Gaussian data have a log-concave law, and one can therefore use the Brascamp–Lieb inequality to control the fluctuations of functionals of the disorder. No perturbation of the Hamiltonian is involved. This argument bounds

the fluctuations conditionally on the sampled configuration. By working on a slice, we ensure that intra-species permutations of the coordinates act transitively on the model, and thus for observables that are invariant under this transformation, the conditioning can be removed. In the model without external field studied in [1], the spin-flip symmetry is used instead to cover the full configuration space, but the presence of an external field breaks the invariance under such transformations.

The second role played by the restriction of the free energy to slices concerns the initial condition. The Hopf representation of the solution to a Hamilton–Jacobi equation requires a convex initial condition, and the initial condition $\psi_\circ(\cdot; h)$ of the model with a fixed field h is not convex in the path when h is large, as was pointed out in [21, Section 6]. Fixing the magnetization changes this: we prove that g_m is convex as a function of the path variable. That $\psi_\circ(\cdot; 0) = g_0$ satisfies this property has already been shown in [7, 16], and here we extend this property to every m . This yields a variational representation for the limit free energy over each slice, which we turn into the uninverted formula (1.13) using convex-duality arguments, and the full free energy is finally recovered by minimizing over the slices, as in (1.14).

To close this introduction, we also mention that the existence of the limit free energy of a broader class of spin-glass models was very recently shown in [17], also using ideas from [1]. Now that the existence of the limit free energy is established for this broader class of models, we can appeal to [5, 8] to assert that this limit can be represented in terms of a (possibly non-unique) critical point of an explicit functional.

Organization of the paper. Sections 2 to 4 define the enriched free energy and derive the identities for its derivatives, as outlined above, both for the full system and for the one restricted to a slice. This yields that the error term in the Hamilton–Jacobi equation is controlled by an average of the variances of the overlap under the family of Gibbs measures discussed above. Section 5 shows that these variances are small on a slice, uniformly over the slices, once the slopes of the path are bounded below, following the method of [1]. Section 6 presents several properties of the scalar function g_m , including that it can be represented as an infimum over martingales started at m , and that it is a convex function of the path variable. Section 7 identifies the limit of the initial condition $\overline{F}_{N,k}(0, \cdot)$ as $\psi_{m_{N,k}}$, uniformly over the slices. Section 8 uses this to show (1.10)–(1.12) of Theorem 1.2. Section 9 turns the Hopf formula into the martingale formula via a convex-duality argument, yielding (1.13). Section 10 shows (1.14) and completes the proof of Theorem 1.1. Finally, Appendix A relates the enriched free energy used here to its classical definition through Poisson–Dirichlet cascades; the main text does not depend on the results of this appendix.

Statement of AI use. The proofs of our main results were first obtained by prompting ChatGPT 6 Astra. The initial arguments thus obtained were however difficult for us to parse, and our main contribution has been to turn these arguments into a form that we find conceptually appealing. We were assisted by ChatGPT 6 Astra and Claude Fable 5.1 in this process. We have manually checked all statements and proofs and polished the paper, and we take full responsibility for its contents.

Acknowledgements. JCM acknowledges the support of the MSCA grant SLOHD (101203974).

2. THE FREE ENERGY ON ONE BRANCH

In this section, we define the enriched free energy of a slice, and more generally of any weighted set of configurations. Throughout this section and the next three, N is

fixed, and every statement holds for every finite set $\mathcal{S}$, every $N \geq 1$, and every partition $(I_{N,s})_{s \in \mathcal{S}}$ of $\{1, \dots, N\}$ into nonempty sets.

2.1. Setting. Let Ω be a nonempty subset of $\Sigma_N := \{-1, 1\}^N$, and let $w : \Omega \rightarrow (0, \infty)$ be a weight function. Two cases are of primary interest. In the first, Ω is a slice $\Omega_{N,k}$ and $w \equiv 2^{-N}$, so that the slice free energy retains the entropy cost of the constraint, as in (1.6). In the second, Ω is the full cube Σ_N and $w(\sigma) = 2^{-N} e^{h^{(N)} \cdot \sigma}$ is the weight in (1.3), where $h^{(N)} \in \mathbb{R}^N$ denotes the vector with $h_i^{(N)} := h_s$ for $i \in I_{N,s}$, so that $h^{(N)} \cdot \sigma = \sum_s h_s \sum_{i \in I_{N,s}} \sigma_i$. A third case, a coupling of two slices, appears in Section 7. Since the configuration set is finite, the positive semidefinite kernel $(\sigma, \sigma') \mapsto N\xi(R_N(\sigma, \sigma'))$ on Σ_N admits a Gram representation: there exist a finite-dimensional Euclidean space and vectors $(v_\sigma)_{\sigma \in \Sigma_N}$ in it such that

$$(2.1) \quad H_N(\sigma) = g \cdot v_\sigma, \quad v_\sigma \cdot v_{\sigma'} = N\xi(R_N(\sigma, \sigma')), \quad |v_\sigma|^2 = N\xi(\lambda_N),$$

where g is a standard Gaussian vector in this space. We fix one such representation on the full cube and use it for every Ω , so that the disorder is one finite Gaussian vector g , shared by all the slices. We write P_s for the orthogonal projection of $\mathbb{R}^N$ onto the coordinates in $I_{N,s}$, and, for a vector $a = (a_s)_{s \in \mathcal{S}} \in \mathbb{R}^{\mathcal{S}}$, we write $\text{diag}(a)$ for the diagonal $N \times N$ matrix whose i -th entry is a_s for $i \in I_{N,s}$. For every realization of g , every $t \geq 0$ and every $x \in \mathbb{R}^N$, we set

$$(2.2) \quad U(1, x; g) := \log \sum_{\sigma \in \Omega} w(\sigma) \exp(\sqrt{2t} g \cdot v_\sigma + x \cdot \sigma).$$

For every fixed g , the function $U(1, \cdot; g)$ is smooth and convex, its gradient is the mean of σ under the Gibbs measure on Ω with weights w and Hamiltonian $\sqrt{2t} g \cdot v_\sigma + x \cdot \sigma$, its Hessian is the covariance of σ under this measure, and all its derivatives are bounded, being cumulants of the bounded vector σ . In particular $|\partial_{x_i} U(1, \cdot; g)| \leq 1$ for every i .

2.2. The branch equation. We say that a path $q : [0, 1] \rightarrow \mathbb{R}_+^{\mathcal{S}}$ is *regular* if it is continuous and piecewise affine with finitely many pieces, and if

$$(2.3) \quad q_s(0) > 0 \quad \text{and} \quad \dot{q}_s > 0 \text{ on every piece},$$

for every species s ; restricted to $[0, 1)$, such a path belongs to $\mathcal{Q}_\infty^{\mathcal{S}}$. For a regular path, let

$$(2.4) \quad D(r) := \text{diag}(\dot{q}(r)),$$

so that D is piecewise constant and positive definite, and $\text{diag}(q(0))$ is positive definite. The free energy is first defined for regular paths. Its extension to all of $\mathcal{Q}_1^{\mathcal{S}}$ is obtained in Proposition 4.2 below, after the Lipschitz estimate that justifies it. For a fixed realization of g , we consider the backward equation

$$(2.5) \quad \partial_r U + \text{tr}(D(r) \nabla^2 U) + r \nabla U \cdot D(r) \nabla U = 0 \quad \text{on } [0, 1) \times \mathbb{R}^N,$$

with the terminal condition (2.2). All spatial gradients below are with respect to $x \in \mathbb{R}^N$, and we write

$$(2.6) \quad \mathcal{L}[W] := \partial_r W + \text{tr}(D(r) \nabla^2 W) + r \nabla W \cdot D(r) \nabla W$$

for the nonlinear operator in (2.5). This equation is the finite- N analogue of the Parisi equation (1.7), up to a time change discussed further in Section 6.2 for a single spin.

Lemma 2.1 (Well-posedness of the branch equation). *Let q be regular, and fix a realization of g . There is a unique function $U = U(\cdot, \cdot; g)$ on $[0, 1] \times \mathbb{R}^N$, continuous and with $|\partial_{x_i} U| \leq 1$ for every i , which is a classical solution of (2.5) on each piece of q and satisfies the terminal condition (2.2). Its spatial derivatives of all orders are continuous and bounded on $[0, 1] \times \mathbb{R}^N$, and $\partial_r U$ is bounded and continuous away from the breakpoints of q .*

Proof. We argue by backward induction on the pieces of q . On a piece $[s_j, s_{j+1}]$, the matrix D is constant and positive definite, so that (2.5) is a uniformly parabolic semilinear equation with coefficients that are smooth in r and a nonlinearity that is quadratic in the gradient. The terminal datum at s_{j+1} is (2.2) for the last piece, and otherwise the value already constructed at s_{j+1} ; in both cases it has bounded derivatives of all orders. The standard theory of such equations, see for instance [20, Chapter V], provides a unique classical solution with bounded derivatives of all orders on the piece, the gradient bound below supplying the a priori estimate needed for global existence; the solution converges to its terminal datum in every C^k norm as $r \uparrow s_{j+1}$. Hence the spatial derivatives of the glued function are continuous across s_{j+1} . Each partial derivative $w := \partial_{x_i} U$ solves the linear equation $\partial_r w + \text{tr}(D \nabla^2 w) + 2r \nabla U \cdot D \nabla w = 0$, with terminal datum in $[-1, 1]$, so the maximum principle gives $|w| \leq 1$. Uniqueness among continuous solutions with bounded gradient is the case $\phi_1 = \phi_2 = 0$ of Lemma 2.3 below. $\square$

Remark 2.2 (Other terminal data). The proof uses only two properties of the terminal datum (2.2): its derivatives of all orders are bounded, and its first derivatives take values in $[-1, 1]$. Lemma 2.1 therefore holds for every terminal datum with these two properties.

The following comparison principle is used repeatedly: it gives the uniqueness in Lemma 2.1, the stability of the solution with respect to the path (Lemmas 4.3 and 7.3), and the comparison between the free energies of the full model and of the slices (Lemma 7.1).

Lemma 2.3 (Comparison principle). *Let q be regular. Let W_1 and W_2 be continuous functions on $[0, 1] \times \mathbb{R}^N$ with bounded spatial gradients, which on each piece of q are classical solutions of $\mathcal{L}[W_i] = \phi_i$, where ϕ_i is bounded on $[0, 1] \times \mathbb{R}^N$ and continuous on each piece. For every $r \in [0, 1]$, we have*

$$(2.7) \quad \sup_{x \in \mathbb{R}^N} |W_1(r, x) - W_2(r, x)| \leq \sup_{x \in \mathbb{R}^N} |W_1(1, x) - W_2(1, x)| + \int_r^1 \sup_{x \in \mathbb{R}^N} |\phi_1(u, x) - \phi_2(u, x)| du.$$

If moreover $\phi_1 \leq \phi_2$ on $[0, 1] \times \mathbb{R}^N$ and $W_1(1, \cdot) \geq W_2(1, \cdot)$, then $W_1 \geq W_2$ on $[0, 1] \times \mathbb{R}^N$.

Proof. The difference $w := W_1 - W_2$ satisfies, on each piece of q ,

$$\partial_r w + \text{tr}(D \nabla^2 w) + r(\nabla W_1 + \nabla W_2) \cdot D \nabla w = \phi_1 - \phi_2,$$

which is a linear equation with the bounded drift coefficient $rD(\nabla W_1 + \nabla W_2)$. Fix (r, x) . By Girsanov's theorem, there is a weak solution Y on $[r, 1]$ of $dY_u = uD(u)(\nabla W_1 + \nabla W_2)(u, Y_u) du + \sqrt{2D(u)} dB_u$ with $Y_r = x$. Since w grows at most linearly and ∇w is bounded, Itô's formula on each piece and the continuity of w across the breakpoints give

$$w(r, x) = \mathbb{E}[w(1, Y_1)] - \mathbb{E} \int_r^1 (\phi_1 - \phi_2)(u, Y_u) du,$$

the stochastic integral having zero mean. Both assertions follow. $\square$

The arguments used in the proof of Lemma 2.1 also give smooth dependence on parameters: if the terminal condition or the coefficients D and $\text{diag}(q(0))$ depend smoothly on a real parameter, with bounded derivatives, then so does U , and each parameter derivative solves a linear parabolic equation with bounded coefficients, obtained by differentiating (2.5). This is used for the response identity of Section 3 and for the derivatives of Section 4.

2.3. The branch and the definition. We fix $t \geq 0$ and a regular path q . We sample g and $X_0 \sim \mathcal{N}(0, 2 \operatorname{diag}(q(0)))$ independently, and we let B be an independent standard N -dimensional Brownian motion. Since $\nabla^2 U$ is bounded, the stochastic differential equation

$$(2.8) \quad dX_r = 2rD(r)\nabla U(r, X_r; g) dr + \sqrt{2D(r)} dB_r, \quad 0 \leq r \leq 1,$$

has a unique strong solution. We call X the *branch*, and set

$$(2.9) \quad \mathbf{m}_r := \nabla U(r, X_r; g) \in [-1, 1]^N, \quad A_q := \int_0^1 r \mathbf{m}_r \cdot D(r) \mathbf{m}_r dr.$$

Throughout, $\mathbb{E}$ denotes the expectation with respect to all sources of randomness: so far this means the disorder g , the initial field X_0 , and the Brownian motion B .

Definition 2.4 (Enriched free energy on one branch). For $t \geq 0$ and a regular path q , let

$$H_N^{\text{br}}(\sigma) := \sqrt{2t}H_N(\sigma) + X_1 \cdot \sigma - Nt\xi(\lambda_N) - N\lambda_N \cdot q(1) - A_q,$$

and define

$$(2.10) \quad \bar{F}_{N,\Omega}(t, q) := -\frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \Omega} w(\sigma) e^{H_N^{\text{br}}(\sigma)},$$

the dependence on the choice of the weight w being kept implicit. We write $\bar{F}_{N,k} := \bar{F}_{N,\Omega_{N,k}}$ for a slice with the weights 2^{-N} , and $\bar{F}_N := \bar{F}_{N,\Sigma_N}$ for the full cube with the weights $2^{-N} e^{h^{(N)} \cdot \sigma}$.

The three subtracted terms in H_N^{br} do not depend on the configuration σ that is summed over, and will therefore not have any influence on the associated Gibbs measure. The first two are the usual self-overlap corrections: $t\xi(\lambda_N)$ is half the variance of $\sqrt{2t}H_N(\sigma)/\sqrt{N}$, and $\lambda_N \cdot q(1)$ is half the variance per site of the Gaussian part of the field. The term A_q records a relative-entropy cost, see Remark A.9. The definition (2.10) can be rewritten as

$$(2.11) \quad \bar{F}_{N,\Omega}(t, q) = t\xi(\lambda_N) + \lambda_N \cdot q(1) - \frac{1}{N} \mathbb{E}[U(1, X_1; g) - A_q].$$

Lemma 2.5 (Itô's formula along the branch). *For every g , X_0 , and almost every B , we have*

$$(2.12) \quad U(1, X_1; g) - A_q = U(0, X_0; g) + \int_0^1 \mathbf{m}_r \cdot \sqrt{2D(r)} dB_r.$$

Consequently

$$(2.13) \quad \bar{F}_{N,\Omega}(t, q) = t\xi(\lambda_N) + \lambda_N \cdot q(1) - \frac{1}{N} \mathbb{E} U(0, X_0; g).$$

Proof. By Lemma 2.1, Itô's formula applies on each piece of q , and gives

$$\begin{aligned} dU(r, X_r; g) &= (\partial_r U + \operatorname{tr}(D\nabla^2 U) + 2r \nabla U \cdot D\nabla U)(r, X_r; g) dr + \mathbf{m}_r \cdot \sqrt{2D(r)} dB_r \\ &= r \mathbf{m}_r \cdot D(r) \mathbf{m}_r dr + \mathbf{m}_r \cdot \sqrt{2D(r)} dB_r, \end{aligned}$$

where we used (2.5). Integration over $[0, 1]$ gives (2.12). The stochastic integral has zero mean since $|\mathbf{m}_r| \leq \sqrt{N}$, and (2.11) gives (2.13). $\square$

Remark 2.6 (Control characterization). The same calculation, applied to the controlled process $Y_r^v := X_0 + \int_0^r 2uD(u)v_u du + \int_0^r \sqrt{2D(u)} dB_u$ for a progressively measurable v

with values in $[-1, 1]^N$, gives

$$\begin{aligned} U(0, X_0; g) - \mathbb{E}_B \left[U(1, Y_1^v; g) - \int_0^1 r v_r \cdot D(r) v_r dr \right] \\ = \mathbb{E}_B \int_0^1 r |v_r - \nabla U(r, Y_r^v; g)|_{D(r)}^2 dr \geq 0, \end{aligned}$$

where $|w|_D^2 := w \cdot Dw$, with equality exactly for the branch. Thus $U(0, X_0; g)$ is the value of a stochastic control problem, the branch is its optimal state, and $\bar{F}_{N, \Omega}(t, q)$ could equivalently be defined by maximizing the expected terminal log partition function minus the quadratic running cost, separately for each realization of (g, X_0) . This is the finite- N counterpart of the control representation of the Parisi functional [2, 19].

We refer to Appendix A for the relationship between the enriched free energy as defined here and that based on Poisson–Dirichlet cascades.

2.4. Symmetries of the branch equation. Three elementary symmetries of (2.5) will be used: translations of the field and constant shifts, the separation of the equation over blocks of coordinates, and the permutations of the sites within a species. For the last one, let S be a permutation of $\{1, \dots, N\}$ that preserves each $I_{N, s}$, acting on $\mathbb{R}^N$ by permutation of the coordinates. It preserves every species overlap, so the Gram matrix in (2.1) is invariant under S , and there is an orthogonal map O_S of the disorder space with $O_S v_\sigma = v_{S\sigma}$ for every $\sigma \in \Sigma_N$. Such a permutation preserves the slices and the full cube.

Lemma 2.7 (Symmetries). *Let q be regular, and fix a realization of g .*

- (1) *For every $y \in \mathbb{R}^N$ and $c \in \mathbb{R}$, the solution of (2.5) with the terminal datum $x \mapsto U(1, x + y; g) + c$ is $(r, x) \mapsto U(r, x + y; g) + c$.*
- (2) *Let $\{1, \dots, N\} = B_1 \cup \dots \cup B_p$ be a partition into blocks, each contained in a single species, and suppose that the terminal datum is of the form $U(1, x; g) = \sum_\ell U_\ell(1, x_{B_\ell})$, where x_{B_ℓ} denotes the restriction of x to the coordinates in B_ℓ and each $U_\ell(1, \cdot)$ has bounded derivatives of all orders. Then $U(r, x; g) = \sum_\ell U_\ell(r, x_{B_\ell})$, where U_ℓ solves (2.5) in dimension $|B_\ell|$, with D replaced by its restriction to the coordinates in B_ℓ , and with the terminal datum $U_\ell(1, \cdot)$.*
- (3) *Suppose that $S\Omega = \Omega$ and $w \circ S = w$. Then, for every such permutation S ,*

$$(2.14) \quad \begin{aligned} U(r, Sx; O_S g) &= U(r, x; g), & \nabla U(r, Sx; O_S g) &= S \nabla U(r, x; g), \\ \nabla^2 U(r, Sx; O_S g) &= S \nabla^2 U(r, x; g) S^\top. \end{aligned}$$

Proof. In each case, the function on the right-hand side is continuous with bounded gradient, so by the uniqueness in Lemma 2.1 it suffices to check that it solves (2.5) with the stated terminal datum. For (1), the equation only involves derivatives of U and is invariant under translations in x . For (2), since D is diagonal and constant on each species, the function $\sum_\ell U_\ell(r, x_{B_\ell})$ satisfies $\text{tr}(D \nabla^2 \sum_\ell U_\ell) = \sum_\ell \text{tr}(D_{B_\ell} \nabla^2 U_\ell)$, and the quadratic term separates as well, because the gradients of the different blocks are supported on disjoint coordinates. For (3), let $W(r, x) := U(r, Sx; O_S g)$. Then $\nabla W(r, x) = S^\top \nabla U(r, Sx; O_S g)$ and $\nabla^2 W(r, x) = S^\top \nabla^2 U(r, Sx; O_S g) S$. Since D is constant on each species, $S D S^\top = D$, so that $\text{tr}(D \nabla^2 W) = \text{tr}(D \nabla^2 U)$ and $\nabla W \cdot D \nabla W = \nabla U \cdot D \nabla U$, evaluated at $(r, Sx; O_S g)$: the function W solves (2.5). As for the terminal datum, since O_S is orthogonal, $O_S g \cdot v_\sigma = g \cdot O_S^{-1} v_\sigma = g \cdot v_{S^{-1}\sigma}$, and $Sx \cdot \sigma = x \cdot S^{-1}\sigma$. Substituting $\sigma' = S^{-1}\sigma$ in (2.2), which is a bijection of Ω preserving the weights, gives $U(1, Sx; O_S g) = U(1, x; g)$. The identities for the derivatives follow by differentiating the first identity in (2.14). $\square$

3. THE GIBBS MEASURES ALONG THE BRANCH

The quantities that drive the argument are Gibbs averages of the sampled configuration, conditioned on the branch up to an intermediate level. We introduce them now, together with the response identity that expresses the derivatives of U along the branch in terms of these averages. Throughout this section, (Ω, w) is as in Section 2, $t \geq 0$ and q is a regular path.

3.1. The sampled configuration and the Gibbs measure at level r . After sampling g and the branch X , we sample one configuration σ from the normalized Gibbs measure of H_N^{br} , that is,

$$(3.1) \quad \mathbb{P}(\sigma = \tau \mid g, (X_r)_{0 \leq r \leq 1}) = w(\tau) \exp(\sqrt{2t} g \cdot v_\tau + X_1 \cdot \tau - U(1, X_1; g)), \quad \tau \in \Omega.$$

We call σ the *sampled configuration*. Let $\mathcal{G}_r$ be the σ -algebra generated by g , X_0 and $(B_u)_{u \leq r}$; it does not contain the sampled configuration, even for $r = 1$. For $r \in [0, 1]$ and a function b of n configurations $\sigma^1, \dots, \sigma^n$, we write $\langle b \rangle_r$ for the integral of b against the n -fold product of the conditional law of σ given $\mathcal{G}_r$. Thus the replicas at level r are conditionally independent given $\mathcal{G}_r$, each with the conditional law of σ given $\mathcal{G}_r$, and $\langle b \rangle_r = \mathbb{E}[b(\sigma) \mid \mathcal{G}_r]$ when $n = 1$. For a function b of one configuration, we also write

$$\text{Var}_r(b) := \langle b^2 \rangle_r - \langle b \rangle_r^2$$

for its conditional variance given $\mathcal{G}_r$. The tower property gives, for $r \leq u$,

$$(3.2) \quad \langle b \rangle_r = \mathbb{E}[\langle b \rangle_u \mid \mathcal{G}_r],$$

so that $r \mapsto \langle b \rangle_r$ is a martingale. At $r = 1$, $\langle \cdot \rangle_1$ is the Gibbs measure in (3.1).

We note that the conditional laws of σ given $\mathcal{G}_r$ thus form a measure-valued martingale that evokes the stochastic localization scheme of [12, 14]. However, the presence of the coefficient r in (2.5) and (2.8) in place of 1 makes this interpretation difficult to flesh out precisely.

We also point out that, for $r < 1$, the average $\langle \cdot \rangle_r$ integrates over the unobserved continuation $(X_u)_{u > r}$ of the branch, and it is in general not the Gibbs measure obtained by replacing X_1 with X_r in the terminal Hamiltonian. To sample two replicas at level r , one uses two independent continuations of the branch after level r , each followed by an independent draw from the Gibbs measure (3.1) in its terminal field. In the cascade picture of Appendix A, two replicas at level r are two replicas of the cascade Gibbs measure, conditioned on the event that the cascade variables separate at level r ; see Lemma A.4 and Proposition A.6.

Lemma 3.1 below shows that the Gibbs mean $\langle \sigma \rangle_r$ is the drift vector $\mathbf{m}_r = \nabla U(r, X_r; g)$ of (2.9). Anticipating this, we call $\mathbf{m}_r$ the posterior mean, and we define

$$(3.3) \quad C_r := \text{Cov}(\sigma \mid \mathcal{G}_r) = \langle (\sigma - \mathbf{m}_r)(\sigma - \mathbf{m}_r)^\top \rangle_r, \quad K_r := \nabla^2 U(r, X_r; g)$$

for the conditional covariance of the sampled configuration and for the Hessian of U along the branch, which we call the response matrix. We write $C_{r,ss} := P_s C_r P_s$ and $K_{r,ss} := P_s K_r P_s$ for their diagonal blocks. All matrix comparisons below are in the positive semidefinite order $\leq$. Since $\sigma_i^2 = 1$, we have

$$(3.4) \quad \text{tr } C_{r,ss} = |I_{N,s}| - |P_s \mathbf{m}_r|^2, \quad \text{tr } C_r = N - |\mathbf{m}_r|^2 \leq N.$$

Finally, for two replicas σ^1, σ^2 , we write $R_N := R_N(\sigma^1, \sigma^2)$ and $R_{N,s} := R_{N,s}(\sigma^1, \sigma^2)$. We set

$$(3.5) \quad S_{r,s} := \langle R_{N,s} \rangle_r = \frac{|P_s \mathbf{m}_r|^2}{N}, \quad p_{N,s}(r) := \mathbb{E} \langle R_{N,s} \rangle_r = \mathbb{E} S_{r,s},$$

and $S_r, p_N(r)$ denote the corresponding vectors in $\mathbb{R}^{\mathcal{S}}$. The quantity $S_{r,s}$ is the random Gibbs mean of the overlap at level r , and $p_{N,s}(r)$ is its expectation; both lie in $[0, \lambda_{N,s}]$.

3.2. The response identity. For a function $b : \Omega \rightarrow \mathbb{R}$, possibly depending on g , and $y \in \mathbb{R}$, let U^y denote the solution of (2.5) with the terminal condition obtained from (2.2) by adding $y b(\sigma)$ to the Hamiltonian,

$$U^y(1, x; g) := \log \sum_{\sigma \in \Omega} w(\sigma) \exp(\sqrt{2t} g \cdot v_\sigma + x \cdot \sigma + y b(\sigma)),$$

so that $U^0 = U$. As for (2.2), the x -derivatives of this terminal datum are cumulants of σ under a Gibbs measure on Ω , its gradient being the mean, so it satisfies the assumptions of Remark 2.2, and thus Lemma 2.1 applies to U^y ; by the parameter dependence noted after Lemma 2.3, it depends smoothly on y , and its y -derivatives solve the linearized equations written below. We define the response to the observable b at level r by

$$(3.6) \quad \Xi_r(b) := \partial_y^2 U^y(r, X_r; g)|_{y=0}.$$

Lemma 3.1 (Gibbs averages and response along the branch). *Let $b : \Omega \rightarrow \mathbb{R}$. For every $r \in [0, 1]$, we have*

$$(3.7) \quad \partial_y U^y(r, X_r; g)|_{y=0} = \langle b \rangle_r,$$

$$(3.8) \quad \Xi_r(b) = r \text{Var}_r(b) + \int_r^1 \mathbb{E}[\text{Var}_u(b) | \mathcal{G}_r] du.$$

In particular, $\langle \sigma \rangle_r = \nabla U(r, X_r; g) = \mathbf{m}_r$, and the response matrix satisfies

$$(3.9) \quad K_r = r C_r + \int_r^1 \mathbb{E}[C_u | \mathcal{G}_r] du, \quad r C_r \leq K_r \leq C_r \leq N \text{Id}.$$

Moreover,

$$(3.10) \quad d\mathbf{m}_r = K_r \sqrt{2D(r)} dB_r.$$

Proof. Let $\mathcal{A} := \partial_r + \text{tr}(D\nabla^2) + 2r(D\nabla U) \cdot \nabla$ be the backward generator of the branch (2.8). Differentiating (2.5) once and twice in y at $y = 0$, we find that $W := \partial_y U^y|_{y=0}$ and $W_2 := \partial_y^2 U^y|_{y=0}$ solve

$$\mathcal{A}W = 0, \quad \mathcal{A}W_2 + 2r \nabla W \cdot D\nabla W = 0,$$

with terminal conditions $W(1, x) = \partial_y U^y(1, x; g)|_{y=0}$ and $W_2(1, x) = \partial_y^2 U^y(1, x; g)|_{y=0}$. These are the mean and the variance of b under the Gibbs measure with field x , so that $W(1, X_1) = \langle b \rangle_1$ and $W_2(1, X_1) = \text{Var}_1(b)$ by (3.1).

By Itô's formula, $M_r := W(r, X_r)$ is a bounded continuous martingale with terminal value $\langle b \rangle_1 = \mathbb{E}[b(\sigma) | \mathcal{G}_1]$. Hence $M_r = \mathbb{E}[b(\sigma) | \mathcal{G}_r] = \langle b \rangle_r$, which is (3.7), and $d[M]_r = 2\nabla W \cdot D\nabla W(r, X_r) dr$. Set $V_r := \text{Var}_r(b)$. We claim that

$$(3.11) \quad \Xi_r(b) - rV_r + \int_0^r V_u du \quad \text{is a martingale.}$$

Since $\Xi_1(b) = W_2(1, X_1) = V_1$, the terminal value of this martingale is $\int_0^1 V_u du$, and (3.8) follows by conditioning on $\mathcal{G}_r$. To prove the claim, we note that three processes are martingales: first, $\Xi_r(b) + \int_0^r u d[M]_u$, by Itô's formula applied to $W_2(r, X_r)$ and the equation for W_2 ; second, $V_r + [M]_r = \mathbb{E}[b(\sigma)^2 | \mathcal{G}_r] - (M_r^2 - [M]_r)$, by the tower property and the martingale property of $M^2 - [M]$; third, $rY_r - \int_0^r Y_u du$ for every bounded martingale Y , as a direct computation of conditional expectations shows. Applying the third fact to $Y := V + [M]$ and subtracting the result from the first martingale, we find that

$$\Xi_r(b) - rV_r + \int_0^r V_u du + \left(\int_0^r u d[M]_u - r[M]_r + \int_0^r [M]_u du \right)$$

is a martingale, and the expression in parentheses vanishes identically, by integration by parts for the continuous increasing process $[M]$. This proves (3.11).

For $b(\sigma) = c \cdot \sigma$ with $c \in \mathbb{R}^N$, we have $U^y(1, x; g) = U(1, x + yc; g)$, so $U^y(r, x; g) = U(r, x + yc; g)$ by Lemma 2.7(1), and thus $W = c \cdot \nabla U$ and $W_2 = c \cdot \nabla^2 U c$. Now (3.7) gives $c \cdot \nabla U(r, X_r; g) = \langle c \cdot \sigma \rangle_r$, and (3.8) gives $c \cdot K_r c = r c \cdot C_r c + \int_r^1 \mathbb{E}[c \cdot C_u c \mid \mathcal{G}_r] du$; polarization yields (3.9). The lower bound on K_r follows from $C_u \geq 0$. For the upper bound, the conditional variance decomposition of the vector σ reads

$$C_r = \mathbb{E}[C_u \mid \mathcal{G}_r] + \mathbb{E}[(\mathbf{m}_u - \mathbf{m}_r)(\mathbf{m}_u - \mathbf{m}_r)^\top \mid \mathcal{G}_r] \geq \mathbb{E}[C_u \mid \mathcal{G}_r] \quad (r \leq u),$$

and $c \cdot C_r c \leq \langle (c \cdot \sigma)^2 \rangle_r \leq N|c|^2$. Finally, differentiating (2.5) in x shows that each component of ∇U solves $\mathcal{A}(\partial_{x_i} U) = 0$, and Itô's formula gives (3.10). $\square$

We note that at $r = 0$, the identity (3.8) reads

$$(3.12) \quad \Xi_0(b) = \int_0^1 \mathbb{E}[\text{Var}_u(b) \mid \mathcal{G}_0] du,$$

which will yield the formula for the time derivative of the free energy.

Lemma 3.1 and its proof apply verbatim to the branch started from any point (r_0, x_0) in place of $(0, X_0)$. In particular, for every $(r, x) \in [0, 1] \times \mathbb{R}^N$,

$$(3.13) \quad 0 \leq \nabla^2 U(r, x; g) \leq N \text{Id}, \quad \text{tr } \nabla^2 U(r, x; g) \leq N, \quad |\nabla U(r, x; g)|^2 \leq N,$$

all three bounds coming from (3.9) and (3.4), applied at the initial level of the branch started from (r, x) : there $\nabla^2 U(r, x; g) \leq C \leq N \text{Id}$ and $\text{tr } \nabla^2 U(r, x; g) \leq \text{tr } C = N - |\nabla U(r, x; g)|^2 \leq N$, where C is the covariance of σ at that level. In particular, $U(r, \cdot; g)$ is convex for every r .

3.3. The gain in predicting the sampled configuration. We define

$$(3.14) \quad \Gamma(r) := \frac{1}{N} \mathbb{E}|\mathbf{m}_r|^2 = \sum_{s \in \mathcal{S}} p_{N,s}(r) = 1 - \frac{1}{N} \mathbb{E} \text{tr } C_r = 1 - \frac{1}{N} \mathbb{E}|\sigma - \mathbf{m}_r|^2,$$

where we used (3.4). Since $\mathbf{m}_r = \mathbb{E}[\sigma \mid \mathcal{G}_r]$ is the best $\mathcal{G}_r$ -measurable prediction of σ in mean square, and since the trivial prediction 0 has normalized error 1, the number $\Gamma(r)$ is the expected reduction of the normalized squared prediction error afforded by the information available at level r ; equivalently, $1 - \Gamma(r)$ measures how much of the sampled configuration remains unresolved at level r . Since $\mathbf{m}$ is a martingale, its increments over disjoint intervals are orthogonal, and for $u \leq r$,

$$(3.15) \quad p_{N,s}(r) - p_{N,s}(u) = \frac{1}{N} \mathbb{E}|P_s(\mathbf{m}_r - \mathbf{m}_u)|^2, \quad \Gamma(r) - \Gamma(u) = \frac{1}{N} \mathbb{E}|\mathbf{m}_r - \mathbf{m}_u|^2,$$

so that the gain afforded by the levels between u and r is the mean square displacement of the prediction over these levels; in particular, $p_{N,s}$ and Γ are increasing.

Lemma 3.2 (Bounded gain in prediction). *The function Γ is absolutely continuous on $[0, 1]$ and*

$$(3.16) \quad \Gamma'(r) = \frac{2}{N} \mathbb{E} \text{tr}(K_r D(r) K_r), \quad 0 \leq \Gamma(0) \leq \Gamma(1) \leq 1.$$

Each component $p_{N,s}$ is increasing with values in $[0, \lambda_{N,s}]$, and $\sum_s p_{N,s} \leq 1$. If $\dot{q}_s \geq \kappa > 0$ on $[0, 1]$ for every species s , then

$$(3.17) \quad \int_0^1 \mathbb{E} \text{tr } K_r^2 dr \leq \frac{N}{2\kappa}.$$

Proof. By (3.10), Itô's formula gives

$$d|P_s \mathbf{m}_r|^2 = 2P_s \mathbf{m}_r \cdot d\mathbf{m}_r + 2 \operatorname{tr}(P_s K_r D(r) K_r P_s) dr,$$

whose drift is nonnegative. Taking expectations and summing over s gives (3.16) and the monotonicity of $p_{N,s}$; the bounds follow from $|P_s \mathbf{m}_r|^2 \leq |I_{N,s}|$. Under the lower bound $D(r) \geq \kappa \operatorname{Id}$ on the slopes, we have $\operatorname{tr}(K_r D K_r) \geq \kappa \operatorname{tr} K_r^2$, and integrating (3.16) gives (3.17). $\square$

The bound (3.17) turns the lower bound in the slope of the path into a control of the response matrices, and it will play a key role in the derivation of Theorem 5.1.

4. FREE-ENERGY DERIVATIVES

The main result of this section is an explicit representation of the derivatives of the free energy in terms of the Gibbs averages introduced above. From this follow an explicit representation of the remainder in the Hamilton–Jacobi equation for the free energy, and the Lipschitz continuity of the free energy as a function of the path.

4.1. The derivative identities. We call a continuous, piecewise affine $h : [0, 1] \rightarrow \mathbb{R}^{\mathcal{S}}$ an *admissible variation*; for a regular path q , the path $q + \varepsilon h$ is again regular for all ε small enough, and the solution of (2.5) depends smoothly on ε , by the parameter dependence noted after Lemma 2.3. We recall that we denote the overlap between two replicas by $R_N = R_N(\sigma^1, \sigma^2)$.

Theorem 4.1 (Derivative identities). *Let $t \geq 0$ and let q be a regular path. For every admissible variation h , we have*

$$(4.1) \quad \begin{aligned} \frac{d}{d\varepsilon} \bar{F}_{N,\Omega}(t, q + \varepsilon h) \Big|_{\varepsilon=0} &= \int_0^1 h(r) \cdot \mathbb{E} \langle R_N \rangle_r dr, \\ \partial_t \bar{F}_{N,\Omega}(t, q) &= \int_0^1 \mathbb{E} \langle \xi(R_N) \rangle_r dr. \end{aligned}$$

We write $\partial_q \bar{F}_{N,\Omega}(t, q)(r) = p_N(r) = \mathbb{E} \langle R_N \rangle_r$, as in (3.5), for the density of the first variation. Both identities average the Gibbs measures over a level r that is uniformly distributed on $[0, 1]$. In the cascade representation of the free energy, this uniform distribution on $[0, 1]$ arises as the law of the overlap between two cascade variables.

Proof of the time derivative. We write $H_N(\sigma) = g \cdot v_\sigma = \sum_a g_a h_a(\sigma)$, where $h_a(\sigma)$ denotes the a -th coordinate of v_σ in an orthonormal basis of the disorder space, so that $\sum_a h_a(\sigma) h_a(\tau) = N \xi(R_N(\sigma, \tau))$ by (2.1). For a deterministic vector G in the disorder space, let $\Psi(G)$ denote $\mathbb{E} U(0, X_0; g)$ computed with the terminal condition $\log \sum_{\sigma \in \Omega} w(\sigma) e^{G \cdot v_\sigma + x \cdot \sigma}$; then $\mathbb{E} U(0, X_0; g) = \mathbb{E} \Psi(\sqrt{2t}g)$. Since $\sqrt{2t}g$ is a centered Gaussian vector with covariance $2t \operatorname{Id}$, and since Ψ is smooth with bounded second derivatives (by the polarized form of (3.12), $|\partial_{G_a} \partial_{G_b} \Psi| \leq \|h_a\|_\infty \|h_b\|_\infty$), the heat equation gives, for every $t \geq 0$,

$$\partial_t \mathbb{E} U(0, X_0; g) = \mathbb{E} \sum_a \partial_{G_a}^2 \Psi(\sqrt{2t}g) = \sum_a \mathbb{E} \Xi_0(h_a),$$

because a change of G_a perturbs the terminal Hamiltonian by the fixed observable h_a , and the second variation of $U(0, X_0; g)$ in that direction is the response (3.6). By (3.12)

and (2.1),

$$\begin{aligned} \sum_a \mathbb{E} \Xi_0(h_a) &= \int_0^1 \mathbb{E} \sum_a \text{Var}_u(h_a) du \\ &= \int_0^1 \mathbb{E} \left\langle N\xi(R_N(\sigma, \sigma)) - N\xi(R_N(\sigma^1, \sigma^2)) \right\rangle_u du \\ &= N\xi(\lambda_N) - N \int_0^1 \mathbb{E} \langle \xi(R_N) \rangle_u du, \end{aligned}$$

since $R_N(\sigma, \sigma) = \lambda_N$ for every $\sigma \in \Sigma_N$. The self-overlap correction $t\xi(\lambda_N)$ in (2.13) cancels the first term, and the second identity in (4.1) follows. At $t = 0$, the derivative of $t \mapsto \mathbb{E} \Psi(\sqrt{2t}g)$ is a right derivative, and the right-hand side is continuous in t . $\square$

Proof of the path derivative. Let $q^\varepsilon := q + \varepsilon h$, let U^ε be the corresponding solution of (2.5), and let $X_0^\varepsilon \sim \mathcal{N}(0, 2 \text{diag}(q(0)) + 2\varepsilon \text{diag}(h(0)))$. Let $D_h(r) := \text{diag}(\dot{h}(r))$. By (2.13),

$$(4.2) \quad \frac{d}{d\varepsilon} \overline{F}_{N,\Omega}(t, q^\varepsilon) \Big|_{\varepsilon=0} = \lambda_N \cdot h(1) - \frac{1}{N} \frac{d}{d\varepsilon} \mathbb{E} U^\varepsilon(0, X_0; g) \Big|_{\varepsilon=0} - \frac{1}{N} \frac{d}{d\varepsilon} \mathbb{E} U(0, X_0^\varepsilon; g) \Big|_{\varepsilon=0}.$$

The last term is the derivative of a Gaussian expectation with respect to its covariance, and equals

$$(4.3) \quad -\frac{1}{N} \mathbb{E} \text{tr}(\nabla^2 U(0, X_0; g) \text{diag}(h(0))) = -\frac{1}{N} \sum_s h_s(0) \mathbb{E} \text{tr} K_{0,ss}.$$

For the middle term, $W := \partial_\varepsilon U^\varepsilon|_{\varepsilon=0}$ solves the linearized equation

$$\mathcal{A}W = -\text{tr}(D_h \nabla^2 U) - r \nabla U \cdot D_h \nabla U, \quad W(1, \cdot) = 0,$$

where $\mathcal{A}$ is the backward generator of the branch, as in the proof of Lemma 3.1. Itô's formula along the branch gives

$$\mathbb{E} W(0, X_0) = \int_0^1 \mathbb{E} [\text{tr}(D_h(r) K_r) + r \mathbf{m}_r \cdot D_h(r) \mathbf{m}_r] dr.$$

We now express the integrand through the Gibbs measures, for a general diagonal matrix. Let $c \in \mathbb{R}^{\mathcal{J}}$. Since $C_r + \mathbf{m}_r \mathbf{m}_r^\top = \langle \sigma \sigma^\top \rangle_r$ has unit diagonal, the response identity (3.9) gives

$$\begin{aligned} \text{tr}(\text{diag}(c) K_r) + r \mathbf{m}_r \cdot \text{diag}(c) \mathbf{m}_r &= r \text{tr}(\text{diag}(c) \langle \sigma \sigma^\top \rangle_r) + \int_r^1 \text{tr}(\text{diag}(c) \mathbb{E}[C_u | \mathcal{G}_r]) du \\ &= Nc \cdot \left(r \lambda_N + \int_r^1 \mathbb{E}[\lambda_N - S_u | \mathcal{G}_r] du \right), \end{aligned}$$

where we used (3.4) in the last step. Taking expectations and recalling that $\mathbb{E} S_u = p_N(u)$, we obtain, for every $c \in \mathbb{R}^{\mathcal{J}}$ and $r \in [0, 1]$,

$$(4.4) \quad \frac{1}{N} \mathbb{E} [\text{tr}(\text{diag}(c) K_r) + r \mathbf{m}_r \cdot \text{diag}(c) \mathbf{m}_r] = c \cdot \mathbf{a}(r), \quad \mathbf{a}(r) := \lambda_N - \int_r^1 p_N(u) du.$$

We apply this identity twice. With $c = \dot{h}(r)$, so that $\text{diag}(c) = D_h(r)$, it gives the middle term in (4.2):

$$-\frac{1}{N} \frac{d}{d\varepsilon} \mathbb{E} U^\varepsilon(0, X_0; g) \Big|_{\varepsilon=0} = -\frac{1}{N} \mathbb{E} W(0, X_0) = -\int_0^1 \dot{h}(r) \cdot \mathbf{a}(r) dr.$$

With $c = h(0)$ and $r = 0$, the quadratic term in $\mathbf{m}_r$ vanishes, and it gives the root term (4.3):

$$-\frac{1}{N} \mathbb{E} \text{tr}(\nabla^2 U(0, X_0; g) \text{diag}(h(0))) = -\frac{1}{N} \mathbb{E} \text{tr}(\text{diag}(h(0)) K_0) = -h(0) \cdot \mathbf{a}(0).$$

The last two terms in (4.2) together form the derivative of $\mathbb{E} U^\varepsilon(0, X_0^\varepsilon; g)$, in which ε enters both through the equation and through the initial law; we have shown that

$$(4.5) \quad \frac{1}{N} \frac{d}{d\varepsilon} \mathbb{E} U^\varepsilon(0, X_0^\varepsilon; g) \Big|_{\varepsilon=0} = h(0) \cdot \mathbf{a}(0) + \int_0^1 \dot{h}(r) \cdot \mathbf{a}(r) dr.$$

Since h and $\mathbf{a}$ are Lipschitz, an integration by parts gives

$$h(0) \cdot \mathbf{a}(0) + \int_0^1 \dot{h}(r) \cdot \mathbf{a}(r) dr = h(1) \cdot \mathbf{a}(1) - \int_0^1 h(r) \cdot \mathbf{a}'(r) dr,$$

and using also that $\mathbf{a}' = p_N$ and $\mathbf{a}(1) = \lambda_N$, we thus obtain

$$(4.6) \quad \frac{1}{N} \frac{d}{d\varepsilon} \mathbb{E} U^\varepsilon(0, X_0^\varepsilon; g) \Big|_{\varepsilon=0} = \lambda_N \cdot h(1) - \int_0^1 h(r) \cdot p_N(r) dr.$$

Inserting this into (4.2), the self-overlap terms $\lambda_N \cdot h(1)$ cancel, and

$$\frac{d}{d\varepsilon} \bar{F}_{N,\Omega}(t, q^\varepsilon) \Big|_{\varepsilon=0} = \int_0^1 h(r) \cdot p_N(r) dr = \int_0^1 h(r) \cdot \mathbb{E} \langle R_N \rangle_r dr,$$

by the definition (3.5) of p_N . This is the first identity in (4.1). $\square$

4.2. The Hamilton–Jacobi remainder. For $p \in \mathcal{Q}_\infty^\mathcal{S}$, we define $\mathcal{H}(p) := \int_0^1 \xi(p(r)) dr$, and for a regular path q , we set

$$(4.7) \quad \mathcal{E}_N(t, q) := \partial_t \bar{F}_{N,\Omega}(t, q) - \mathcal{H}(\partial_q \bar{F}_{N,\Omega}(t, q)).$$

Theorem 4.1 gives at once

$$(4.8) \quad \mathcal{E}_N(t, q) = \int_0^1 \left\{ \mathbb{E} \langle \xi(R_N) \rangle_r - \xi(\mathbb{E} \langle R_N \rangle_r) \right\} dr.$$

The main result of the next section is that this term is small. We note here that it is controlled by the variance of the overlap. Let $M_\xi := \sup_{x \in [-1,1]^\mathcal{S}} \|\nabla^2 \xi(x)\|$. Since the linear term of the Taylor expansion of ξ around $p_N(r)$ has zero mean under $\mathbb{E} \langle \cdot \rangle_r$, we have

$$(4.9) \quad |\mathcal{E}_N(t, q)| \leq \frac{M_\xi}{2} \mathcal{W}_N(t, q), \quad \text{with } \mathcal{W}_N(t, q) := \int_0^1 \mathbb{E} \langle |R_N - p_N(r)|^2 \rangle_r dr.$$

The conditional variance decomposition separates two sources of fluctuations:

$$(4.10) \quad \mathbb{E} \langle |R_N - p_N(r)|^2 \rangle_r = \mathbb{E} \langle |R_N - S_r|^2 \rangle_r + \mathbb{E} \langle |S_r - p_N(r)|^2 \rangle_r.$$

The first term concerns two replicas within the Gibbs measure at a fixed disorder and branch; the second concerns the fluctuations of the Gibbs mean overlap across the disorder and the branch. Writing $P_s \sigma^a = P_s \mathbf{m}_r + \theta_s^a$ for the two replicas, we have $NR_{N,s} = |P_s \mathbf{m}_r|^2 + P_s \mathbf{m}_r \cdot \theta_s^2 + \theta_s^1 \cdot P_s \mathbf{m}_r + \theta_s^1 \cdot \theta_s^2$, and the three fluctuating terms are uncorrelated under $\langle \cdot \rangle_r$, with variances $\mathbf{m}_r \cdot C_{r,ss} \mathbf{m}_r$ (twice) and $\text{tr}(C_{r,ss}^2)$. With (4.10), this gives

$$(4.11) \quad \mathbb{E} \langle (R_{N,s} - p_{N,s}(r))^2 \rangle_r = \frac{\mathbb{E} \text{tr}(C_{r,ss}^2) + 2 \mathbb{E} [\mathbf{m}_r \cdot C_{r,ss} \mathbf{m}_r]}{N^2} + \text{Var}(S_{r,s}).$$

The first term in (4.11) will be controlled by the response bound $K_r \geq r C_r$ and (3.17). Obtaining a control of the other two terms will be the main focus of the next section.

4.3. Lipschitz continuity and extension to all paths. We set $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^{\mathcal{S}}$.

Proposition 4.2 (Lipschitz bound and extension). *For all regular paths q, q' and all $t, t' \geq 0$, we have*

$$(4.12) \quad |\bar{F}_{N,\Omega}(t, q) - \bar{F}_{N,\Omega}(t', q')| \leq \sum_{s \in \mathcal{S}} \lambda_{N,s} \|q_s - q'_s\|_{L^1} + \xi(\mathbf{1}) |t - t'| \leq \|q - q'\|_{L^1} + \xi(\mathbf{1}) |t - t'|.$$

Consequently, $\bar{F}_{N,\Omega}$ extends uniquely to a function on $[0, \infty) \times \mathcal{Q}_1^{\mathcal{S}}$ satisfying (4.12), and we denote this extension by the same symbol. For a slice, this extension is the enriched slice free energy $\bar{F}_{N,k}(t, q)$ appearing in Theorem 1.2.

Proof. The set of regular paths is convex, so $q + \varepsilon(q' - q)$ is regular for $\varepsilon \in [0, 1]$, with the admissible variation $h = q' - q$. By Theorem 4.1 and $0 \leq p_{N,s}(r) \leq \lambda_{N,s}$, the derivative along this segment is bounded in absolute value by $\sum_s \lambda_{N,s} \|q'_s - q_s\|_{L^1}$, and since $\lambda_{N,s} \geq 0$ and $\sum_s \lambda_{N,s} = 1$, we have $\sum_s \lambda_{N,s} |x_s| \leq \max_s |x_s| \leq |x|$ for every $x \in \mathbb{R}^{\mathcal{S}}$, which gives the second inequality in (4.12) after integration in r . Since $R_N(\sigma^1, \sigma^2) \in [-1, 1]^{\mathcal{S}}$ and the coefficients of ξ are nonnegative, $|\partial_t \bar{F}_{N,\Omega}| \leq \sup_{[-1,1]^{\mathcal{S}}} |\xi| = \xi(\mathbf{1})$. Regular paths are dense in $\mathcal{Q}_1^{\mathcal{S}}$: a path in $\mathcal{Q}_1^{\mathcal{S}}$ is approximated in L^1 by bounded paths, which are approximated by continuous piecewise affine increasing paths, and adding $\eta(1+r)\mathbf{1}$ with small $\eta > 0$ makes them regular. $\square$

The path $q = 0$ is not regular, but substituting it formally in Definition 2.4 gives $X \equiv 0$ and $A_q = 0$, whatever U is, hence the ordinary free energy. The extension indeed takes this value. To see this, we will use the heat interpolation identity

$$(4.13) \quad \mathbb{E}f(V + \sqrt{2\varepsilon}Z) - \mathbb{E}f(V) = \int_0^\varepsilon \mathbb{E} \Delta f(V + \sqrt{2s}Z) ds,$$

valid for $\varepsilon \geq 0$, a random vector V , an independent standard Gaussian vector Z , and a C^2 function f with bounded derivatives. For the functions $U(r, \cdot; g)$, whose Laplacian lies in $[0, N]$ by (3.13), the left-hand side lies in $[0, N\varepsilon]$.

Lemma 4.3 (The value at $q = 0$). *For every $t \geq 0$,*

$$(4.14) \quad \bar{F}_{N,\Omega}(t, 0) = -\frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \Omega} w(\sigma) \exp(\sqrt{2t}H_N(\sigma) - Nt\xi(\lambda_N)),$$

which is the free energy (1.6) for a slice, and the free energy (1.3) for the full cube with the weights $2^{-N} e^{h^{(N)} \cdot \sigma}$.

Proof. For $\varepsilon > 0$, the path $q^\varepsilon(r) := \varepsilon(1+r)\mathbf{1}$ is regular, with $D^\varepsilon = \varepsilon \text{Id}$ and initial field $X_0^\varepsilon = \sqrt{2\varepsilon}Z$, and $\bar{F}_{N,\Omega}(t, q^\varepsilon) \rightarrow \bar{F}_{N,\Omega}(t, 0)$ as $\varepsilon \rightarrow 0$ by (4.12). Let U^ε be the solution of (2.5) for q^ε , with operator $\mathcal{L}^\varepsilon$, and let $W(r, x) := U(1, x; g)$ be the terminal function (2.2), held constant in r . Since $\nabla^2 W$ is a covariance matrix of σ , of trace at most N , and $|\nabla W|^2 \leq N$, we have $\mathcal{L}^\varepsilon[W] = \varepsilon(\text{tr} \nabla^2 W + r|\nabla W|^2) \in [0, 2\varepsilon N]$, while $\mathcal{L}^\varepsilon[U^\varepsilon] = 0$ and both functions have the same terminal datum, so Lemma 2.3 gives $|U^\varepsilon(0, \cdot) - U(1, \cdot; g)| \leq 2\varepsilon N$. By (4.13), $0 \leq \mathbb{E}U(1, \sqrt{2\varepsilon}Z; g) - \mathbb{E}U(1, 0; g) \leq N\varepsilon$. Hence $|\mathbb{E}U^\varepsilon(0, X_0^\varepsilon; g) - \mathbb{E}U(1, 0; g)| \leq 3\varepsilon N$, and (2.13), in which $\lambda_N \cdot q^\varepsilon(1) = 2\varepsilon$, gives

$$\left| \bar{F}_{N,\Omega}(t, q^\varepsilon) - t\xi(\lambda_N) + \frac{1}{N} \mathbb{E}U(1, 0; g) \right| \leq 5\varepsilon.$$

Letting $\varepsilon \rightarrow 0$ yields $\bar{F}_{N,\Omega}(t, 0) = t\xi(\lambda_N) - \frac{1}{N} \mathbb{E}U(1, 0; g)$, which is (4.14), since $U(1, 0; g) = \log \sum_{\sigma \in \Omega} w(\sigma) e^{\sqrt{2t}H_N(\sigma)}$. $\square$

5. CONCENTRATION OF THE OVERLAP ON A SLICE

In this section, $\Omega = \Omega_{N,k}$ is a slice, with the weights $w \equiv 2^{-N}$, and we show that the overlap variance $\mathscr{W}_N(t, q)$ of (4.9) is small, uniformly over the slices and over t , as soon as the slopes of the path are bounded below. By (4.9), this controls the remainder $\mathscr{E}_N(t, q)$ of (4.7), where we recall that $M_\xi = \sup_{x \in [-1, 1]^\mathscr{S}} \|\nabla^2 \xi(x)\|$.

Theorem 5.1 (Quantitative remainder estimate on a slice). *Let $\kappa > 0$, and let q be a regular path with $\dot{q}_s \geq \kappa$ on every piece, for every species s . For every slice $\Omega = \Omega_{N,k}$, every $\delta \in \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ and every $t \geq 0$, we have*

$$(5.1) \quad \mathscr{W}_N(t, q) \leq 13\delta + \frac{5}{2\kappa N \delta^2}.$$

Consequently, for every N , every slice, every $t \geq 0$ and every such path, we have

$$(5.2) \quad \left| \partial_t \bar{F}_{N,k}(t, q) - \mathcal{H}(\partial_q \bar{F}_{N,k}(t, q)) \right| \leq 15M_\xi(\kappa N)^{-1/3}.$$

The argument is adapted from [1], where it is used to prove the Parisi formula for the SK model. As noted after (4.11), the first term there is controlled by the response bound $K_r \geq rC_r$ of (3.9) and by the bounded gain in prediction of Lemma 3.2, which makes the response matrices small in an integrated sense once the slopes are bounded below. The other two terms, which involve the posterior mean $\mathbf{m}_r$, are the main concern of this section. We first show that they are bounded by the variance of a single observable which we call the realized gain in prediction $L_{r,s}$, defined in (5.3). To bound this variance, we reverse the direction of the conditioning: given the sampled configuration, the Gaussian data have a log-concave law, and the Brascamp–Lieb inequality bounds the variance of averages of the realized gain over levels in terms of the response matrices; the symmetry of the slice under permutations of the sites within each species then removes the conditioning on the configuration. This is Proposition 5.4. A last step, based on the martingale property of the posterior mean, passes from averages over levels to the integrated variance of the realized gain; the parameter δ in (5.1) is the length of the intervals over which the realized gain is averaged.

5.1. The realized gain in prediction. For a level r and a species s , we define

$$(5.3) \quad L_{r,s} := \lambda_{N,s} - \frac{1}{N} |P_s(\sigma - \mathbf{m}_r)|^2 = \frac{1}{N} (2\sigma \cdot P_s \mathbf{m}_r - |P_s \mathbf{m}_r|^2),$$

where we used $|P_s \sigma|^2 = |I_{N,s}|$. This is the gain in predicting $P_s \sigma$ by $P_s \mathbf{m}_r$ rather than by 0, realized along the branch rather than averaged, and we call it the realized gain in prediction: summing over the species gives $\sum_s L_{r,s} = 1 - \frac{1}{N} |\sigma - \mathbf{m}_r|^2$, whose expectation is $\Gamma(r)$ by (3.14), and likewise $\mathbb{E} L_{r,s} = p_{N,s}(r)$ by (3.4). The realized gain is thus the quenched, per-species counterpart of the expected gain of Section 3.3. It is determined by the Gibbs measure at level r and the sampled configuration: with σ held fixed and two further replicas τ^1, τ^2 at level r , averaging only over τ^1, τ^2 , we have

$$L_{r,s} = \left\langle 2R_{N,s}(\sigma, \tau^1) - R_{N,s}(\tau^1, \tau^2) \right\rangle_r.$$

The conditional mean of $L_{r,s}$ given $\mathcal{G}_r$ is $S_{r,s}$, and its conditional variance given $\mathcal{G}_r$ equals $4N^{-2} \mathbf{m}_r \cdot C_{r,ss} \mathbf{m}_r$. Hence

$$(5.4) \quad \text{Var}(L_{r,s}) = \frac{4 \mathbb{E}[\mathbf{m}_r \cdot C_{r,ss} \mathbf{m}_r]}{N^2} + \text{Var}(S_{r,s}).$$

Comparing (5.4) with (4.11), we obtain

$$(5.5) \quad \begin{aligned} \mathbb{E} \left((R_{N,s} - p_{N,s}(r))^2 \right)_r &= \frac{\mathbb{E} \operatorname{tr}(C_{r,ss}^2)}{N^2} + \operatorname{Var}(L_{r,s}) - \frac{2 \mathbb{E}[\mathbf{m}_r \cdot C_{r,ss} \mathbf{m}_r]}{N^2} \\ &\leq \frac{\mathbb{E} \operatorname{tr}(C_{r,ss}^2)}{N^2} + \operatorname{Var}(L_{r,s}). \end{aligned}$$

Integrating over r and summing over s leads to

$$(5.6) \quad \mathcal{W}_N(t, q) \leq \frac{1}{N^2} \int_0^1 \mathbb{E} \sum_s \operatorname{tr}(C_{r,ss}^2) dr + \int_0^1 \sum_s \operatorname{Var}(L_{r,s}) dr.$$

5.2. The joint law of the disorder, the branch and the sampled configuration.

Let $Z_r := X_0 + \int_0^r \sqrt{2D(u)} dB_u$ be the Gaussian part of the branch, and let γ denote the law of the pair (g, Z) , that is, the Gaussian disorder together with the Gaussian field path with initial covariance $2 \operatorname{diag}(q(0))$ and increment covariance $2D(r) dr$.

Lemma 5.2 (Joint density). *The law of (g, X, σ) has density*

$$(5.7) \quad 2^{-N} \exp \left(\sqrt{2t} g \cdot v_\sigma + x_1 \cdot \sigma - \int_0^1 U(r, x_r; g) dr \right)$$

at (g, x, σ) with respect to the product of γ and the counting measure on Ω .

Proof. Given g , the branch has drift $2rD\nabla U$ relative to the Gaussian path. Since $|\nabla U| \leq \sqrt{N}$, Girsanov's formula gives the density

$$\exp \left(\int_0^1 r \nabla U \cdot \sqrt{2D} d\tilde{B}_r - \int_0^1 r^2 \nabla U \cdot D \nabla U dr \right)$$

of the branch law with respect to the Gaussian path law. Here $x = (x_r)_{0 \leq r \leq 1}$ is the path variable, ∇U is evaluated at $(r, x_r; g)$, and $\tilde{B}_r := \int_0^r (2D(u))^{-1/2} dx_u$. Since D is piecewise constant and invertible by (2.3), $\tilde{B}$ is a linear function of the path. Under the Gaussian path law, $\tilde{B}$ is a standard N -dimensional Brownian motion and $dx_r = \sqrt{2D(r)} d\tilde{B}_r$, whereas along the branch $\tilde{B}_r = B_r + \int_0^r u \sqrt{2D(u)} \nabla U(u, X_u; g) du$. Under the Gaussian law, Itô's formula and (2.5) give, as in the proof of Lemma 2.5, $dU(r, x_r; g) = -r \nabla U \cdot D \nabla U dr + \nabla U \cdot \sqrt{2D} d\tilde{B}_r$, so that the exponent equals

$$\int_0^1 r dU(r, x_r; g) = U(1, x_1; g) - \int_0^1 U(r, x_r; g) dr.$$

Multiplying by the Gibbs probability (3.1) cancels $U(1, x_1; g)$. $\square$

For a fixed configuration σ , the exponent in (5.7) is the sum of a linear function of the Gaussian data and of $-\int_0^1 U(r, x_r; g) dr$. This second part is concave:

Lemma 5.3 (Convexity in the Gaussian data). *For every $r \in [0, 1]$, the map $(x, g) \mapsto U(r, x; g)$ is smooth, with derivatives of all orders bounded and continuous in (r, x, g) , and it is convex, with*

$$(5.8) \quad 0 \leq \nabla_{(x,g)}^2 U(r, x; g) \leq N(1 + 2t\xi(\lambda_N)) \operatorname{Id}.$$

Proof. The terminal datum (2.2) is smooth in (x, g) with bounded derivatives of all orders, so the regularity follows from Lemma 2.1 and the smooth dependence on parameters noted after Lemma 2.3. Fix a direction $(\hat{x}, \hat{g})$ with $|\hat{x}|^2 + |\hat{g}|^2 = 1$. By Lemma 2.7(1) and the uniqueness in Lemma 2.1, we have $U(r, x + y\hat{x}; g + y\hat{g}) = U^y(r, x; g)$, where U^y solves the equation with the terminal Hamiltonian perturbed by $y b(\sigma)$ with $b(\sigma) := \sqrt{2t} \hat{g} \cdot v_\sigma + \hat{x} \cdot \sigma$. The proof of Lemma 3.1, applied to the branch started from (r, x) , shows that $\partial_y^2 U^y(r, x; g)|_{y=0}$ is a nonnegative combination of conditional variances of b ,

with total weight one, as in (3.8). This gives the convexity, and the upper bound in (5.8) follows from $b(\sigma)^2 \leq (|\widehat{x}|^2 + |\widehat{g}|^2)(|\sigma|^2 + 2t|v_\sigma|^2) = N(1 + 2t\xi(\lambda_N))$, by the Cauchy–Schwarz inequality and (2.1). $\square$

5.3. The averaged concentration estimate.

Proposition 5.4 (Variance of level averages). *Let q be regular. For every slice $\Omega_{N,k}$, every deterministic $a \in L^2([0, 1])$ and every species s , we have*

$$(5.9) \quad \text{Var} \left(\int_0^1 a(r) L_{r,s} dr \right) \leq \frac{4}{N^2} \int_0^1 a(r)^2 \mathbb{E} \text{tr}(K_{r,ss} C_{r,ss}) dr.$$

The main step of the proof is Lemma 5.6 below, which bounds the variance of the level average conditionally on the sampled configuration. Given σ , the law of the Gaussian data (g, X) has, by Lemma 5.2, a density with respect to γ which is the exponential of a linear function of (g, x) minus $\int_0^1 U(r, x_r; g) dr$, and U is convex in (x, g) by Lemma 5.3. The conditional law is therefore log-concave, and the Hessian of minus its log-density is the inverse covariance of γ plus the integral over r of the Hessian of $U(r, \cdot; \cdot)$. By the Brascamp–Lieb inequality, the conditional variance of a function of the Gaussian data is at most the expectation of its squared gradient, measured in the inverse of this Hessian. The level average is a quadratic function of the residuals $P_s(\sigma - \mathbf{m}_r)$, and $\mathbf{m}_r = \nabla U(r, X_r; g)$; its gradient in the Gaussian data is therefore an integral over r of the Hessian of $U(r, \cdot; \cdot)$ applied to the residual. The squared gradient thus contains two such Hessians, and the next lemma shows that the inverse Hessian cancels one of them, leaving the quadratic forms $P_s(\sigma - \mathbf{m}_r) \cdot K_r P_s(\sigma - \mathbf{m}_r)$. The symmetry of the slice then makes the level average independent of σ , so that its conditional variance is its variance, and the expectation of the quadratic forms is computed by conditioning on $\mathcal{G}_r$, where the residual is centered with covariance $C_{r,ss}$.

Lemma 5.5 (Cancellation of one Hessian). *Let $(A_r)_{0 \leq r \leq 1}$ be a bounded measurable family of positive semidefinite matrices of the same size, let $a \in L^2([0, 1])$, let $(v_r)_{0 \leq r \leq 1}$ be a bounded measurable family of vectors, and let Q be a positive definite matrix with $Q \geq \int_0^1 A_r dr$. The vector $z := \int_0^1 a(r) A_r v_r dr$ satisfies*

$$(5.10) \quad z \cdot Q^{-1} z \leq \int_0^1 a(r)^2 v_r \cdot A_r v_r dr.$$

Proof. Using the assumption on Q , we can write, for every vector u ,

$$2u \cdot z - u \cdot Qu \leq \int_0^1 (2a(r) u \cdot A_r v_r - u \cdot A_r u) dr \leq \int_0^1 a(r)^2 v_r \cdot A_r v_r dr,$$

where we used that $(u - a(r)v_r) \cdot A_r (u - a(r)v_r) \geq 0$ for the second inequality. Taking $u = Q^{-1}z$, the left-hand side equals $z \cdot Q^{-1}z$, and we thus obtain the announced result. $\square$

Lemma 5.6 (Concentration given the sampled configuration). *Let q be regular, let $\Omega = \Omega_{N,k}$ be a slice, let $a \in L^2([0, 1])$ be deterministic, and let s be a species. For every $\tau \in \Omega$,*

$$(5.11) \quad \text{Var} \left(\int_0^1 a(r) L_{r,s} dr \mid \sigma = \tau \right) \leq \frac{4}{N^2} \mathbb{E} \left[\int_0^1 a(r)^2 P_s(\sigma - \mathbf{m}_r) \cdot K_r P_s(\sigma - \mathbf{m}_r) dr \mid \sigma = \tau \right].$$

Proof. Step 1: the conditional law. Fix $\tau \in \Omega$, and let $f_\tau(g, x)$ denote the density (5.7) at (g, x, τ) . Then $c_\tau := \int f_\tau d\gamma = \mathbb{P}(\sigma = \tau) > 0$, and $\mu_\tau := c_\tau^{-1} f_\tau \gamma$ is the conditional law of

(g, X) given $\sigma = \tau$. For a realization g of the disorder and a continuous path x , set

$$(5.12) \quad \begin{aligned} \theta_r(g, x) &:= P_s(\tau - \nabla U(r, x_r; g)), & F(g, x) &:= \frac{1}{N} \int_0^1 a(r) \left(|I_{N,s}| - |\theta_r(g, x)|^2 \right) dr, \\ \mathcal{E}(g, x) &:= \int_0^1 a(r)^2 \theta_r(g, x) \cdot \nabla^2 U(r, x_r; g) \theta_r(g, x) dr. \end{aligned}$$

On the event $\{\sigma = \tau\}$, by (5.3), $F(g, X)$ is the level average $\int_0^1 a(r) L_{r,s} dr$, and $\mathcal{E}(g, X)$ is the integral on the right-hand side of (5.11). Thus (5.11) reads

$$(5.13) \quad \text{Var}_{\mu_\tau}(F) \leq \frac{4}{N^2} \mathbb{E}_{\mu_\tau}[\mathcal{E}].$$

Since $\nabla U \in [-1, 1]^N$, $|\theta_r| \leq 2\sqrt{|I_{N,s}|}$ and $0 \leq \nabla^2 U \leq N \text{Id}$ by (3.13), we have

$$(5.14) \quad |F| \leq 3\lambda_{N,s} \|a\|_{L^1}, \quad 0 \leq \mathcal{E} \leq 4N |I_{N,s}| \|a\|_{L^2}^2.$$

Moreover, ∇U and $\nabla^2 U$ are Lipschitz continuous in x , uniformly in r , by Lemma 2.1, so that $F(g, \cdot)$ and $\mathcal{E}(g, \cdot)$ are continuous for the uniform norm on paths.

Step 2: finite-dimensional approximations of the conditional law. For $n \geq 1$ and $r_j := j/n$, let $x^{[n]}$ denote the polygonal interpolation of the path x on the mesh $(r_j)_{0 \leq j \leq n}$, that is, the continuous path which agrees with x at the mesh points and is affine on each $[r_j, r_{j+1}]$. It is a linear function of $(x_{r_0}, \dots, x_{r_n})$, it satisfies $\|x^{[n]}\|_\infty \leq \|x\|_\infty$, and $x^{[n]} \rightarrow x$ uniformly as $n \rightarrow \infty$ for every continuous path x . Let

$$(5.15) \quad f_{n,\tau}(g, x) := 2^{-N} \exp \left(\sqrt{2t} g \cdot v_\tau + x_1 \cdot \tau - \int_0^1 U(r, x_r^{[n]}; g) dr \right),$$

which is f_τ with the path x replaced by $x^{[n]}$ inside the integral, the function U being unchanged. We first show that these functions are integrable. By Lemma 5.3, the first derivatives of U in (x, g) are bounded, and $r \mapsto U(r, 0; 0)$ is continuous on $[0, 1]$, so that $|U(r, x; g)| \leq C(1 + |x| + |g|)$ for some constant C . Since $\|x^{[n]}\|_\infty \leq \|x\|_\infty$, we obtain

$$(5.16) \quad f_{n,\tau}(g, x) \leq \exp(C'(1 + |g| + \|x\|_\infty))$$

for a constant C' independent of n , and the same bound holds for f_τ . Under γ , the vector g is Gaussian, and each coordinate $i \in I_{N,s}$ of the path is a Gaussian initial value plus a Brownian motion run with the clock $2(q_s(r) - q_s(0))$, so $|g|$ and $\|x\|_\infty$ have Gaussian tails, and the right-hand side of (5.16) is γ -integrable. Hence $c_{n,\tau} := \int f_{n,\tau} d\gamma$ is finite and positive, and $\mu_{n,\tau} := c_{n,\tau}^{-1} f_{n,\tau} \gamma$ is a probability measure. Since $|U(r, x; g) - U(r, x'; g)| \leq \sqrt{N}|x - x'|$, we also have $f_{n,\tau} \rightarrow f_\tau$ pointwise, and dominated convergence gives $c_{n,\tau} \rightarrow c_\tau$. Consequently, if (G_n) is a uniformly bounded sequence of measurable functions of (g, x) converging pointwise to G , then

$$(5.17) \quad \mathbb{E}_{\mu_{n,\tau}}[G_n] = \frac{\int G_n f_{n,\tau} d\gamma}{c_{n,\tau}} \longrightarrow \frac{\int G f_\tau d\gamma}{c_\tau} = \mathbb{E}_{\mu_\tau}[G],$$

again by dominated convergence. We write $F_n(g, x) := F(g, x^{[n]})$ and $\mathcal{E}_n(g, x) := \mathcal{E}(g, x^{[n]})$ for the functions in (5.12) evaluated along the interpolated path.

Step 3: Gaussian coordinates and the log-density. Let $Y := (g, x_{r_0}, \dots, x_{r_n}) \in \mathbb{R}^d$, where d is the dimension of the disorder space plus $(n+1)N$. Under γ , this is a centered Gaussian vector whose covariance matrix Σ is positive definite: g is a standard Gaussian vector, and, independently of g , the initial value x_{r_0} and the increments $x_{r_j} - x_{r_{j-1}}$ are independent and centered, with covariances $2 \text{diag}(q(0))$ and $2 \text{diag}(q(r_j) - q(r_{j-1}))$, which are positive definite since q is regular. The interpolated path $x^{[n]}$ is a linear function of Y , so the density $f_{n,\tau}$ and the functions F_n and $\mathcal{E}_n$ depend on (g, x) only through Y . The law of Y

under $\mu_{n,\tau}$ therefore has, relative to Lebesgue measure, a density proportional to $e^{-\Phi_{n,\tau}}$, where

$$\Phi_{n,\tau}(Y) := \frac{1}{2} Y \cdot \Sigma^{-1} Y - \sqrt{2t} g \cdot v_\tau - x_1 \cdot \tau + \int_0^1 U(r, x_r^{[n]}; g) dr.$$

By Lemma 5.3, $U(r, \cdot; \cdot)$ is smooth and convex, with bounded derivatives, and $Y \mapsto (g, x_r^{[n]})$ is linear. Hence $A_r := \nabla_Y^2 [U(r, x_r^{[n]}; g)]$ is positive semidefinite, bounded and continuous in (r, Y) , the function $\Phi_{n,\tau}$ is smooth, and, differentiating under the integral sign,

$$(5.18) \quad \nabla^2 \Phi_{n,\tau} = Q := \Sigma^{-1} + \int_0^1 A_r dr \geq \Sigma^{-1}.$$

The Hessian in (5.18) exceeds the inverse covariance Σ^{-1} of the Gaussian coordinates by the positive semidefinite term $\int_0^1 A_r dr$, which involves differentiation in the disorder g as well as in the field coordinates.

Step 4: the gradient of the level average. Let $J : \mathbb{R}^N \rightarrow \mathbb{R}^d$ be the linear map $Ju := (0, u, \dots, u)$, which leaves g unchanged and shifts every x_{r_j} , hence the whole path $x^{[n]}$, by u . Thus $U(r, x_r^{[n]}; g)$ evaluated at $Y + Ju$ equals $U(r, x_r^{[n]} + u; g)$, and differentiating once and twice in u at $u = 0$ gives, with $\mathbf{m}_r^{[n]} := \nabla U(r, x_r^{[n]}; g)$ and $K_r^{[n]} := \nabla^2 U(r, x_r^{[n]}; g)$,

$$(5.19) \quad J^\top \nabla_Y [U(r, x_r^{[n]}; g)] = \mathbf{m}_r^{[n]}, \quad J^\top A_r J = K_r^{[n]}.$$

Let $\theta_r^{[n]} := P_s(\tau - \mathbf{m}_r^{[n]}) = \theta_r(g, x^{[n]})$ be the residual along the interpolated path. By the first identity in (5.19), the Jacobian matrix of $\theta_r^{[n]}$ as a function of Y is $-P_s J^\top A_r$. Since A_r is symmetric and $P_s \theta_r^{[n]} = \theta_r^{[n]}$, differentiating F_n with τ fixed gives, with $v_r := J \theta_r^{[n]}$,

$$(5.20) \quad N \nabla_Y F_n = 2 \int_0^1 a(r) A_r v_r dr, \quad v_r \cdot A_r v_r = \theta_r^{[n]} \cdot K_r^{[n]} \theta_r^{[n]},$$

the second identity being the second one in (5.19). The positive initial variance $q(0) > 0$ was used to make Σ invertible, but Σ will disappear from the estimate once Lemma 5.5 is applied. Since A_r and $\theta_r^{[n]}$ are bounded, F_n is a smooth function of Y with a bounded gradient.

Step 5: the Brascamp–Lieb inequality. We use the Brascamp–Lieb inequality [4], in the following form: if a probability measure on $\mathbb{R}^d$ has a density proportional to $e^{-\Phi}$, with Φ smooth and $\nabla^2 \Phi \geq \lambda \text{Id}$ for some $\lambda > 0$, then $\text{Var}(G) \leq \mathbb{E}[\nabla G \cdot (\nabla^2 \Phi)^{-1} \nabla G]$ for every smooth G with bounded gradient. Applied to $\Phi_{n,\tau}$ and F_n , together with (5.18), (5.20) and Lemma 5.5 applied to $z := \frac{N}{2} \nabla_Y F_n$, it gives

$$(5.21) \quad \begin{aligned} \text{Var}_{\mu_{n,\tau}}(F_n) &\leq \mathbb{E}_{\mu_{n,\tau}} [\nabla_Y F_n \cdot Q^{-1} \nabla_Y F_n] \\ &\leq \frac{4}{N^2} \mathbb{E}_{\mu_{n,\tau}} \int_0^1 a(r)^2 \theta_r^{[n]} \cdot K_r^{[n]} \theta_r^{[n]} dr = \frac{4}{N^2} \mathbb{E}_{\mu_{n,\tau}} [\mathcal{E}_n]. \end{aligned}$$

Step 6: passage to the limit. As $n \rightarrow \infty$, $F_n \rightarrow F$ and $\mathcal{E}_n \rightarrow \mathcal{E}$ pointwise, by the uniform convergence of $x^{[n]}$ to x and the continuity noted in Step 1, and these functions are bounded uniformly in n , by (5.14) applied along the interpolated paths. By (5.17), applied to F_n , F_n^2 and $\mathcal{E}_n$, both sides of (5.21) converge to the corresponding sides of (5.13). $\square$

Proof of Proposition 5.4. Let $F := \int_0^1 a(r) L_{r,s} dr$. We first show that F is independent of σ . Let S be a permutation of $\{1, \dots, N\}$ that preserves each $I_{N,s}$, and let O_S be as in Lemma 2.7. Consider the transformation

$$(5.22) \quad (g, x, \sigma) \mapsto (O_S g, Sx, S\sigma)$$

of the disorder, the path and the configuration. It preserves γ , because O_S is orthogonal and S commutes with $\text{diag}(q(0))$ and with $D(r)$; it preserves the counting measure on

the slice; and it preserves the density (5.7), because $O_S g \cdot v_{S\sigma} = g \cdot v_\sigma$, $Sx_1 \cdot S\sigma = x_1 \cdot \sigma$, and $U(r, Sx; O_S g) = U(r, x; g)$ by Lemma 2.7(3), the weights 2^{-N} being invariant. Hence the law of (g, X, σ) is invariant under (5.22). Since $\nabla U(r, Sx; O_S g) = S \nabla U(r, x; g)$ and P_s commutes with S , each $L_{r,s}$, and therefore F , is invariant under (5.22). It follows that $\mathbb{E}[\varphi(F) \mathbb{1}_{\{\sigma=\tau\}}] = \mathbb{E}[\varphi(F) \mathbb{1}_{\{\sigma=S\tau\}}]$ for every bounded measurable function φ , and that $\mathbb{P}(\sigma = \tau) = \mathbb{P}(\sigma = S\tau)$, for every $\tau \in \Omega_{N,k}$. These permutations act transitively on $\Omega_{N,k}$, so the conditional law of F given $\sigma = \tau$ does not depend on τ , that is, F is independent of σ . In particular,

$$(5.23) \quad \text{Var}(F) = \text{Var}(F \mid \sigma = \tau) \quad (\tau \in \Omega_{N,k}).$$

Averaging (5.11) over the law of σ and using (5.23), we obtain

$$\text{Var}(F) \leq \frac{4}{N^2} \int_0^1 a(r)^2 \mathbb{E}[P_s(\sigma - \mathbf{m}_r) \cdot K_r P_s(\sigma - \mathbf{m}_r)] dr.$$

Conditionally on $\mathcal{G}_r$, the matrix K_r is fixed, and the residual $P_s(\sigma - \mathbf{m}_r)$ is centered with covariance $P_s C_r P_s$, whose only nonzero block is $C_{r,ss}$, by Lemma 3.1. Hence $\mathbb{E}[P_s(\sigma - \mathbf{m}_r) \cdot K_r P_s(\sigma - \mathbf{m}_r) \mid \mathcal{G}_r] = \text{tr}(K_r P_s C_r P_s) = \text{tr}(K_{r,ss} C_{r,ss})$, which proves (5.9). $\square$

The symmetry step (5.23) is the only place in this section where the fact that we work with a slice $\Omega = \Omega_{N,k}$ is used. What it requires is a group of transformations of (g, x, σ) that preserves the law of (g, X, σ) and the function F , and acts transitively on Ω . The permutations (5.22) preserve the law whenever Ω and w are invariant under them, which includes the full cube with the weights $2^{-N} e^{h^{(N)} \cdot \sigma}$, but their orbits on the full cube are the slices. On the full cube with an external field, the argument would therefore only show that $\mathbb{E}[F \mid \sigma]$ is a function of the magnetization of σ , and the term $\text{Var}(\mathbb{E}[F \mid \sigma])$ in the decomposition $\text{Var}(F) = \mathbb{E}[\text{Var}(F \mid \sigma)] + \text{Var}(\mathbb{E}[F \mid \sigma])$ is not controlled by the Brascamp–Lieb inequality, which bounds only the first term. Without external field, the spin flips $\sigma \mapsto \tau\sigma$, the coordinatewise products with $\tau \in \Sigma_N$, supply the missing transformations, as in [1]. A flip preserves every species overlap, so there is an orthogonal map O_τ of the disorder space with $O_\tau v_\sigma = v_{\tau\sigma}$, as for O_S ; the matrix $T_\tau := \text{diag}(\tau)$ commutes with $\text{diag}(q(0))$ and $D(r)$, so $(g, x, \sigma) \mapsto (O_\tau g, T_\tau x, \tau\sigma)$ preserves γ ; and the proof of Lemma 2.7(3) applies verbatim, because the uniform weights are invariant under $\sigma \mapsto \tau\sigma$, and gives $U(r, T_\tau x; O_\tau g) = U(r, x; g)$ and $\nabla U(r, T_\tau x; O_\tau g) = T_\tau \nabla U(r, x; g)$. Hence the density (5.7), each $L_{r,s}$ and F are invariant, and since the flips act transitively on Σ_N , the proof of Proposition 5.4 goes through on the full cube without field, with no other change; Lemma 5.6 and its proof do not use the slice. The weights $2^{-N} e^{h^{(N)} \cdot \sigma}$ are not invariant under the flips when $h \neq 0$: by Lemma 2.7(1), a flip maps the model with the field h to the model with the site-dependent field $\tau_i h_s$ at $i \in I_{N,s}$. On a slice, the field is constant and plays no role in this section.

5.4. Proof of Theorem 5.1. We bound the two terms of (5.6) in turn. Two consequences of the response bound $K_r \geq r C_r$ will be used for $r \geq \delta$. Principal blocks preserve the order of symmetric matrices, so $K_{r,ss} \geq \delta C_{r,ss}$; since the trace of a product of two positive semidefinite matrices is nonnegative, this gives $\text{tr}(C_{r,ss}^2) \leq \delta^{-1} \text{tr}(K_{r,ss} C_{r,ss}) \leq \delta^{-2} \text{tr}(K_{r,ss}^2)$, and $\sum_s \text{tr}(K_{r,ss}^2) \leq \text{tr} K_r^2$. Hence

$$(5.24) \quad \sum_s \text{tr}(C_{r,ss}^2) \leq \delta^{-2} \text{tr} K_r^2, \quad \sum_s \text{tr}(K_{r,ss} C_{r,ss}) \leq \delta^{-1} \text{tr} K_r^2 \quad (r \geq \delta).$$

Gibbs fluctuations of the overlap. For $r < \delta$, we use $\sum_s \text{tr}(C_{r,ss}^2) \leq \sum_s (\text{tr } C_{r,ss})^2 \leq N^2 \sum_s \lambda_{N,s}^2 \leq N^2$. For $r \geq \delta$ we use (5.24) and (3.17). Thus

$$(5.25) \quad \frac{1}{N^2} \int_0^1 \mathbb{E} \sum_s \text{tr}(C_{r,ss}^2) dr \leq \delta + \frac{1}{2\kappa N \delta^2}.$$

From averages to individual levels. For $u \leq r$, the variation of the realized gain across levels is controlled by the increments of the expected gain (3.15): writing $L_{r,s} - L_{u,s} = N^{-1} P_s(\mathbf{m}_r - \mathbf{m}_u) \cdot P_s(2\sigma - \mathbf{m}_r - \mathbf{m}_u)$, the second factor has norm at most $4\sqrt{|I_{N,s}|}$, and (3.15) with $\lambda_{N,s} \leq 1$ gives

$$(5.26) \quad \sum_s \mathbb{E} |L_{r,s} - L_{u,s}|^2 \leq 16 \sum_s \lambda_{N,s} (p_{N,s}(r) - p_{N,s}(u)) \leq 16(\Gamma(r) - \Gamma(u)).$$

We partition $[\delta, 1]$ into intervals Δ of length δ , and write $\bar{L}_{\Delta,s} := \delta^{-1} \int_{\Delta} L_{r,s} dr$. The exact variance identity

$$(5.27) \quad \int_{\Delta} \text{Var}(L_{r,s}) dr = \delta \text{Var}(\bar{L}_{\Delta,s}) + \frac{1}{2\delta} \int_{\Delta} \int_{\Delta} \text{Var}(L_{r,s} - L_{u,s}) dr du,$$

which splits the variance over Δ into that of the average and that of the fluctuations around the average, together with (5.26), bounds the sum over Δ and s of the second terms by

$$8\delta \sum_{\Delta} (\Gamma(\sup \Delta) - \Gamma(\inf \Delta)) \leq 8\delta.$$

For the first terms, Proposition 5.4 with $a = \delta^{-1} \mathbf{1}_{\Delta}$, summed over Δ and s , followed by (5.24) and (3.17), gives a total of at most

$$\frac{4}{\delta N^2} \int_{\delta}^1 \mathbb{E} \sum_s \text{tr}(K_{r,ss} C_{r,ss}) dr \leq \frac{2}{\kappa N \delta^2}.$$

Finally, $L_{r,s}$ takes values in an interval of length $4\lambda_{N,s}$, so $\sum_s \text{Var}(L_{r,s}) \leq 4$ and the discarded levels contribute at most 4δ . Altogether,

$$(5.28) \quad \int_0^1 \sum_s \text{Var}(L_{r,s}) dr \leq 12\delta + \frac{2}{\kappa N \delta^2}.$$

Inserting (5.25) and (5.28) into (5.6) gives (5.1).

For (5.2), recall from (4.7) and (4.9) that the left-hand side is $|\mathcal{E}_N(t, q)| \leq \frac{M_{\xi}}{2} \mathscr{W}_N(t, q)$. If $\kappa N \geq 1$, choose $\delta = 1/\lceil (\kappa N)^{1/3} \rceil$, so that $\delta \leq 2(\kappa N)^{-1/3}$ and $\delta^{-2} \leq (\kappa N)^{2/3}$; the right-hand side of (5.1) is then at most $(26 + \frac{5}{2})(\kappa N)^{-1/3}$. If $\kappa N < 1$, we use the trivial bound $\mathscr{W}_N \leq \sum_s (2\lambda_{N,s})^2 \leq 4$, which gives $|\mathcal{E}_N| \leq 2M_{\xi} \leq 2M_{\xi}(\kappa N)^{-1/3}$. $\square$

6. THE SCALAR INITIAL CONDITION AT FIXED MAGNETIZATION

In this section, we study the scalar function g_m appearing in the definition (1.9) of the slice initial condition ψ_m . We will see in Section 7 that g_m is the limit, as $n \rightarrow \infty$, of the free energy at $t = 0$ of a one-species system with n spins, when the configurations are constrained to have magnetization m . Three properties of g_m are needed later, each for a different interpolation of its argument.

The first property is that the map $m \mapsto g_m(\nu) - I(m)$ is convex and Lipschitz. It is used in Section 7 to identify the limit of the slice initial conditions. Since $I'' \geq 1$, the convexity makes g_m strongly convex in m , which forces the minimizing slice to have a magnetization close to the prescribed one. The Lipschitz bound also gives the continuity of f_m in m , which is used in Section 10 to pass from the minimum over slices to the infimum over m .

The second property is that, for every $T > 0$ and $\nu_0, \nu_1 \in \mathcal{P}([0, T])$, the map $[0, 1] \ni \theta \mapsto g_m((1-\theta)\nu_0 + \theta\nu_1)$ is concave. It comes from a representation of $g_m(\nu)$ as an infimum, over martingales started at m , of functionals that are affine in ν ; we deduce this representation from the variational formula of [24], which itself is based on [2]. In Section 9, we use this representation to pass from the Hopf formula to the martingale formula (1.13).

Recall that, for a path q , we write $g_m(q)$ to denote the function g_m evaluated at the law of $q(U)$, where U is a uniform random variable on $[0, 1]$. The last property we show is that, for every $q_0, q_1 \in \mathcal{Q}_\infty$, the map $[0, 1] \ni \theta \mapsto g_m((1-\theta)q_0 + \theta q_1)$ is convex. Since the path is the quantile function of the law it represents, the pointwise interpolation of paths is the displacement interpolation of laws, and not their mixture. The fact that the function $\psi_o(\cdot; 0) = g_0$ satisfies this convexity property in the path variable has already been shown in [7, 16], and here we extend this result to every m . It makes the slice initial condition ψ_m convex, which allows us to represent the solution of the Hamilton–Jacobi equation (1.11) by the Hopf formula (1.12) in Section 8. As recalled in the introduction, the unconstrained initial condition $\psi_o(\cdot; h)$ is not convex in the path when the field h is large; it is the constraint on the magnetization that restores this property.

6.1. Paths, laws, and the Parisi equation. We first fix the correspondence between paths and probability measures. Recall that $\mathcal{Q}$ is the set of right-continuous increasing maps from $[0, 1)$ to $\mathbb{R}_+$. For $q \in \mathcal{Q}_1$, we denote by ν_q the law of $q(U)$, where U is uniform on $[0, 1]$, so that $g_m(q) = g_m(\nu_q)$ with the convention recalled above. Conversely, for a probability measure ν on $\mathbb{R}_+$ with finite first moment, its quantile function $r \mapsto \inf\{v \geq 0 : \nu([0, v]) > r\}$ is the unique element $q \in \mathcal{Q}_1$ with $\nu_q = \nu$. Since the coupling $(q(U), q'(U))$ of ν_q and $\nu_{q'}$ is monotone, in the sense that both coordinates are increasing functions of the same variable U (see [23, Section 2]), the L^1 distance between two elements of $\mathcal{Q}_1$ is the 1-Wasserstein distance between their laws: $W_1(\nu_q, \nu_{q'}) = \|q - q'\|_{L^1}$ for all $q, q' \in \mathcal{Q}_1$. In particular, a functional of laws that is 1-Lipschitz for W_1 defines a functional of paths that is 1-Lipschitz for the L^1 norm, and conversely. We write $\mathcal{P}([0, T])$ for the set of probability measures on $[0, T]$, and $\mathcal{P}_1(\mathbb{R}_+)$ for the set of probability measures on $\mathbb{R}_+$ with finite first moment. For $\nu \in \mathcal{P}_1(\mathbb{R}_+)$, we write

$$\zeta_\nu(v) = \nu([0, v]), \quad v \geq 0,$$

for its cumulative distribution function. When $\nu = \nu_q$ for a path q , we have $\zeta_\nu(v) = |\{r \in [0, 1] : q(r) \leq v\}|$, and for $\nu, \nu' \in \mathcal{P}([0, T])$ we have $W_1(\nu, \nu') = \int_0^T |\zeta_\nu - \zeta_{\nu'}|$. We say that a law $\nu \in \mathcal{P}([0, T])$ has finite support if it is a finite combination of Dirac masses, and that a path is a finite-step path if it takes finitely many values. Finite-step paths correspond to laws with finite support.

We now define Φ_ν for every $\nu \in \mathcal{P}([0, T])$, using the well-posedness theory of [19]. A weak solution of (1.7) is a continuous function Φ on $[0, T] \times \mathbb{R}$ with bounded weak derivative $\partial_x \Phi$, which satisfies the equation in the sense of distributions on $(0, T) \times \mathbb{R}$ and the terminal condition $\Phi(T, \cdot) = \log \cosh$. Let μ be the image of ν under $v \mapsto v/T$, so that $\mu([0, t]) = \zeta_\nu(Tt)$, and let $\xi_T(t) := Tt^2$. Then Φ is a weak solution of (1.7) if and only if $u(t, x) := \Phi(Tt, x)$ is a weak solution of

$$(6.1) \quad -\partial_t u = \frac{\xi_T''(t)}{2} (\partial_x^2 u + \mu([0, t])(\partial_x u)^2) \quad \text{on } [0, 1] \times \mathbb{R}, \quad u(1, x) = \log \cosh x,$$

which is the Parisi equation of [19, 24] for the covariance function ξ_T and the measure $\mu \in \mathcal{P}([0, 1])$. Since the results of these papers are valid for every measure and every covariance function with nonnegative coefficients, they apply to (1.7) after this change of time. In particular, we have the following.

Lemma 6.1 (Well-posedness and regularity). *Let $\nu \in \mathcal{P}([0, T])$. The equation (1.7) has a unique weak solution, which we denote by Φ_ν . For every $j \geq 1$, the derivative $\partial_x^j \Phi_\nu$ is continuous and bounded on $[0, T] \times \mathbb{R}$, and Lipschitz in v uniformly in x . The equation obtained by differentiating (1.7) $j \geq 0$ times in x holds at every point of $(0, T) \times \mathbb{R}$, with one-sided derivatives in v and the corresponding one-sided limits of ζ_ν ; in particular, it holds classically at every v at which ζ_ν is continuous. Moreover,*

$$(6.2) \quad |\partial_x \Phi_\nu| < 1, \quad 0 < \partial_x^2 \Phi_\nu \leq 1 - (\partial_x \Phi_\nu)^2 \leq 1.$$

Finally, for every $j \geq 0$, Itô's formula applies to $\partial_x^j \Phi_\nu(v, Y_v)$ along every Itô process Y with bounded progressively measurable drift and diffusion coefficients, the time derivative being understood in the one-sided sense above.

Proof. The existence, uniqueness and regularity statements, including the pointwise form of the differentiated equations, are [19, Theorem 4 and Corollary 11], and Itô's formula at this level of regularity is [19, Proposition 22], see also the remark following it. The bounds $|\partial_x \Phi_\nu| < 1$ and $0 < \partial_x^2 \Phi_\nu \leq 1$ are [19, Lemma 16]; the sharper upper bound in (6.2) is proved in Lemma 6.2 below, by the same argument. $\square$

For a finitely supported law, the solution is explicit. For $r \in [0, 1]$ and $a \geq 0$, we define the operators

$$(6.3) \quad T_{r,a}f(x) := \frac{1}{r} \log \mathbb{E} \exp(rf(x + \sqrt{2a}Z)) \quad (r > 0), \quad T_{0,a}f(x) := \mathbb{E}f(x + \sqrt{2a}Z),$$

where Z is a standard Gaussian random variable. These operators act on functions with at most linear growth. By Hölder's inequality, each $T_{r,a}$ is monotone, commutes with the addition of constants, and is a contraction for the uniform norm; moreover $T_{r,a}T_{r,a'} = T_{r,a+a'}$. If f is smooth with bounded derivatives of all orders, then so is $T_{r,a}f$, and $(r, a, x) \mapsto T_{r,a}f(x)$ is smooth on $[0, 1] \times [0, \infty) \times \mathbb{R}$: for $r > 0$, $T_{r,a}f(x) = K(r)/r$, where $K(r) := \log \mathbb{E} \exp(rf(x + \sqrt{2a}Z))$ is smooth in (r, a, x) with $K(0) = 0$, so that $K(r)/r = \int_0^1 K'(\theta r) d\theta$ is smooth up to $r = 0$.

Let $\nu \in \mathcal{P}([0, T])$ have finite support, and let $0 \leq v_0 < v_1 < \dots < v_\ell \leq T$ be the atoms of ν . The function ζ_ν is then constant on each interval $[v_i, v_{i+1})$, with value $\zeta_\nu(v_i)$, and it equals 0 on $[0, v_0)$ and 1 on $[v_\ell, T]$. Define $\Phi : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ by backward recursion: set $\Phi(T, x) := \log \cosh x$, and, on each interval on which ζ_ν is constant with value r ,

$$(6.4) \quad \Phi(v, \cdot) := T_{r, v'-v} \Phi(v', \cdot), \quad v \leq v' \text{ in this interval.}$$

The semigroup property of the operators shows that this is consistent. By construction, Φ is continuous, with $|\partial_x \Phi| \leq 1$, it is smooth on each strip $[v_i, v_{i+1}) \times \mathbb{R}$ with bounded derivatives of all orders in x , and it solves there the Parisi equation (1.7), which we recall:

$$(6.5) \quad -\partial_v \Phi_\nu = \partial_x^2 \Phi_\nu + \zeta_\nu(v)(\partial_x \Phi_\nu)^2.$$

Indeed, if $r > 0$, the function $e^{r\Phi}$ solves the backward heat equation $-\partial_v h = \partial_x^2 h$ on the strip, and if $r = 0$ this is true of Φ itself. Integrating by parts on each strip, the boundary terms cancel by the continuity of Φ , so Φ is a weak solution, and $\Phi = \Phi_\nu$ by Lemma 6.1. On $[v_\ell, T]$, where $\zeta_\nu = 1$, we have $T_{1,a} \log \cosh = \log \cosh + a$, so that

$$(6.6) \quad \Phi_\nu(v, x) = \log \cosh x + T - v, \quad v \in [v_\ell, T].$$

This identity holds for every $\nu \in \mathcal{P}([0, T])$ on $[T_0, T] \times \mathbb{R}$, where $T_0 := \sup \text{supp } \nu$: the function equal to $\Phi_\nu^{T_0}(v, x) + T - T_0$ for $v \leq T_0$, where $\Phi_\nu^{T_0}$ is the solution with T_0 in place of T , and to $\log \cosh x + T - v$ for $v \geq T_0$ is a weak solution of (1.7), because $\zeta_\nu = 1$ on $[T_0, T]$ and $(\log \cosh)'' + ((\log \cosh)')^2 = 1$, so it is Φ_ν by uniqueness. In particular, the quantity $T - \Phi_\nu(0, x)$ does not depend on the choice of $T \geq T_0$, and we may always enlarge T .

The regularity of Lemma 6.1 allows us to use the Parisi equation through Itô's formula, along the diffusion whose drift is the magnetization $\partial_x \Phi_\nu$. Let $\nu \in \mathcal{P}([0, T])$, and write $\Phi := \Phi_\nu$ and $\zeta := \zeta_\nu$. The drift $2\zeta(v)\partial_x \Phi(v, x)$ is bounded, measurable in v , and 2-Lipschitz in x by (6.2), so that for every $s \in [0, T]$ and $x \in \mathbb{R}$, the stochastic differential equation

$$(6.7) \quad dX_v = 2\zeta(v)\partial_x \Phi(v, X_v) dv + \sqrt{2} dB_v, \quad v \in [s, T],$$

has a unique strong solution started from x at time s , which we denote by $X^{s,x}$; here B is a standard Brownian motion, with its completed natural filtration. We write $X^y := X^{0,y}$. This is the diffusion of [2], in the present normalization.

Lemma 6.2 (Itô identities along the field diffusion). *Let $\nu \in \mathcal{P}([0, T])$, $s \in [0, T]$, $x \in \mathbb{R}$, and $X := X^{s,x}$. For every $v \in [s, T]$,*

$$(6.8) \quad \partial_x \Phi(v, X_v) = \partial_x \Phi(s, x) + \sqrt{2} \int_s^v \partial_x^2 \Phi(w, X_w) dB_w,$$

so that $M_v := \partial_x \Phi(v, X_v)$ is a martingale with $M_T = \tanh X_T$. Moreover,

$$(6.9) \quad \partial_x^2 \Phi(s, x) = 1 - \partial_x \Phi(s, x)^2 - 2 \int_s^T (1 - \zeta(w)) \mathbb{E}[\partial_x^2 \Phi(w, X_w)^2] dw,$$

$$(6.10) \quad \Phi(s, x) = \mathbb{E}[\log \cosh X_T] - \int_s^T \zeta(w) \mathbb{E}[M_w^2] dw.$$

In particular, (6.2) holds, $\Phi(v, \cdot)$ is strictly convex, $\partial_x \Phi$ is 1-Lipschitz in x , and M takes values in $(-1, 1)$.

Proof. By Lemma 6.1, Itô's formula applies to $\partial_x^j \Phi(w, X_w)$ for $j \leq 2$, and the equations obtained by differentiating (6.5) j times in x hold at every w at which ζ is continuous, hence for almost every w . The drift of $\Phi(w, X_w)$ is $\partial_w \Phi + \partial_x^2 \Phi + 2\zeta(\partial_x \Phi)^2 = \zeta(\partial_x \Phi)^2$, that of $\partial_x \Phi(w, X_w)$ is $\partial_w \partial_x \Phi + \partial_x^3 \Phi + 2\zeta \partial_x \Phi \partial_x^2 \Phi = 0$, and that of $\partial_x^2 \Phi(w, X_w)$ is $\partial_w \partial_x^2 \Phi + \partial_x^4 \Phi + 2\zeta \partial_x \Phi \partial_x^3 \Phi = -2\zeta(\partial_x^2 \Phi)^2$, all evaluated at (w, X_w) . The stochastic integrals have bounded integrands, hence zero mean. The second identity gives (6.8), and the first gives (6.10), since $\Phi(T, \cdot) = \log \cosh$. Taking expectations in (6.8) at $v = T$ gives $\partial_x \Phi(s, x) = \mathbb{E}[\tanh X_T]$, so $|\partial_x \Phi| < 1$, and Itô's isometry gives

$$\mathbb{E}[\tanh^2 X_T] = \partial_x \Phi(s, x)^2 + 2 \int_s^T \mathbb{E}[\partial_x^2 \Phi(w, X_w)^2] dw.$$

The third identity gives $\partial_x^2 \Phi(s, x) = \mathbb{E}[1 - \tanh^2 X_T] + 2 \int_s^T \zeta(w) \mathbb{E}[\partial_x^2 \Phi(w, X_w)^2] dw$, which is positive, and substituting the previous display yields (6.9), hence $0 < \partial_x^2 \Phi \leq 1 - (\partial_x \Phi)^2$. The remaining assertions follow. $\square$

Lemma 6.3 (Continuity in the law). *For all $\nu, \nu' \in \mathcal{P}([0, T])$ and $v \in [0, T]$,*

$$(6.11) \quad \sup_{x \in \mathbb{R}} |\Phi_\nu(v, x) - \Phi_{\nu'}(v, x)| \leq \int_v^T |\zeta_\nu - \zeta_{\nu'}|,$$

$$\sup_{x \in \mathbb{R}} |\partial_x \Phi_\nu(v, x) - \partial_x \Phi_{\nu'}(v, x)| \leq 2 \left(\int_0^T |\zeta_\nu - \zeta_{\nu'}| \right)^{1/2}.$$

In particular, if $\zeta_{\nu_n} \rightarrow \zeta_\nu$ in $L^1([0, T])$, then $\Phi_{\nu_n} \rightarrow \Phi_\nu$ and $\partial_x \Phi_{\nu_n} \rightarrow \partial_x \Phi_\nu$ uniformly on $[0, T] \times \mathbb{R}$.

Proof. The difference $w := \Phi_\nu - \Phi_{\nu'}$ vanishes at $v = T$ and solves, for almost every v ,

$$-\partial_v w = \partial_x^2 w + \zeta_\nu(\partial_x \Phi_\nu + \partial_x \Phi_{\nu'}) \partial_x w + (\zeta_\nu - \zeta_{\nu'}) (\partial_x \Phi_{\nu'})^2.$$

Fix (v, x) , and let Y solve $dY_u = \zeta_\nu(u)(\partial_x \Phi_\nu + \partial_x \Phi_{\nu'})(u, Y_u) du + \sqrt{2} dB_u$ on $[v, T]$ with $Y_v = x$; the drift is bounded, measurable in u and Lipschitz in x , by (6.2), so this equation

has a unique strong solution. By Itô's formula (Lemma 6.1), the drift of $w(u, Y_u)$ is $-(\zeta_\nu - \zeta_{\nu'})^2(u)(\partial_x \Phi_{\nu'})^2(u, Y_u)$, and since $w(T, \cdot) = 0$,

$$w(v, x) = \mathbb{E} \int_v^T (\zeta_\nu(u) - \zeta_{\nu'}(u)) (\partial_x \Phi_{\nu'})^2(u, Y_u) du.$$

Since $|\partial_x \Phi_{\nu'}| \leq 1$, this gives the first bound. For the second, we use the elementary interpolation inequality $\|f'\|_\infty^2 \leq 4\|f\|_\infty \|f''\|_\infty$, valid for every bounded $f \in C^2(\mathbb{R})$ with bounded second derivative; it follows from the Taylor bound $h|f'(x)| \leq 2\|f\|_\infty + \frac{h^2}{2}\|f''\|_\infty$, optimized in $h > 0$. We apply it to $f = w(v, \cdot)$, whose second derivative is bounded by 1 in absolute value, by (6.2). The last assertion is immediate. $\square$

The derivative $\partial_x \Phi_\nu$ is the magnetization of a single spin in the field x at the covariance v . The following two-sided bound compares it with the magnetization of a free spin, in shifted fields. It will give the behavior of $\partial_x \Phi_\nu$ at infinity, and the Lipschitz dependence of g_m on m .

Lemma 6.4 (Barriers for the magnetization). *For every $\nu \in \mathcal{P}([0, T])$, $v \in [0, T]$ and $x \in \mathbb{R}$,*

$$(6.12) \quad \tanh(x - 2b_\nu(v)) \leq \partial_x \Phi_\nu(v, x) \leq \tanh(x + 2b_\nu(v)), \quad b_\nu(v) := \int_v^T (1 - \zeta_\nu(u)) du.$$

In particular, $\partial_x \Phi_\nu(v, x) \rightarrow \pm 1$ as $x \rightarrow \pm\infty$, for every $v \in [0, T]$, so that $\partial_x \Phi_\nu(v, \cdot)$ is an increasing bijection from $\mathbb{R}$ onto $(-1, 1)$. Moreover, $b_\nu(0) = \int u \nu(du)$.

Proof. Let $u := \partial_x \Phi_\nu$, which is continuous, with $\partial_x u$ and $\partial_x^2 u$ continuous and bounded, and which solves, for almost every v ,

$$(6.13) \quad -\partial_v u = \partial_x^2 u + 2\zeta_\nu u \partial_x u,$$

with $u(T, x) = \tanh x$, by Lemma 6.1. Let $u_\pm(v, x) := \tanh(x \pm 2b_\nu(v))$, which is smooth in x and Lipschitz in v , with $b'_\nu = -(1 - \zeta_\nu)$ almost everywhere. Writing $\theta := 1 - u_\pm^2$, we have $\partial_x u_\pm = \theta$, $\partial_x^2 u_\pm = -2u_\pm \theta$ and $\partial_v u_\pm = \mp 2(1 - \zeta_\nu)\theta$, so that

$$-\partial_v u_\pm - \partial_x^2 u_\pm - 2\zeta_\nu u_\pm \partial_x u_\pm = \pm 2(1 - \zeta_\nu)\theta + 2(1 - \zeta_\nu)u_\pm \theta = \pm 2(1 - \zeta_\nu)(1 \pm u_\pm)\theta,$$

which is nonnegative for u_+ and nonpositive for u_- . Hence the difference $d := u_+ - u_-$ satisfies, for almost every v ,

$$-\partial_v d = \partial_x^2 d + \zeta_\nu(u_+ + u_-) \partial_x d + \zeta_\nu \partial_x(u_+ + u_-) d + S, \quad S \geq 0,$$

with $d(T, \cdot) = 0$ and $|d| \leq 2$, where we used $2(u_+ \partial_x u_+ - u_- \partial_x u_-) = \partial_x[(u_+ - u_-)(u_+ + u_-)]$. Fix (v, x) , let Y solve $dY_w = \zeta_\nu(w)(u_+ + u_-)(w, Y_w) dw + \sqrt{2} dB_w$ on $[v, T]$ with $Y_v = x$, whose drift is bounded and Lipschitz in x , and let $c := \zeta_\nu \partial_x(u_+ + u_-)$, which is bounded by 2. By Itô's formula, which applies to u by Lemma 6.1 and to $u_\pm$ by its regularity, the drift of $e^{\int_v^w c(r, Y_r) dr} d(w, Y_w)$ is $-e^{\int_v^w c(r, Y_r) dr} S(w, Y_w) \leq 0$. Hence

$$d(v, x) \geq \mathbb{E}[e^{\int_v^T c(r, Y_r) dr} d(T, Y_T)] = 0,$$

which is the upper bound in (6.12); the lower bound is proved in the same way with u_- . The bijection statement follows from (6.2), and the last identity is Fubini's theorem, since $1 - \zeta_\nu(u) = \nu((u, T])$. $\square$

We now define the constrained initial condition. For $\nu \in \mathcal{P}([0, T])$, $y \in \mathbb{R}$ and $m \in [-1, 1]$, recall that

$$(6.14) \quad \psi_\circ(\nu; y) = T - \Phi_\nu(0, y), \quad g_m(\nu) = \sup_{y \in \mathbb{R}} \{\psi_\circ(\nu; y) + ym\} = T + \sup_{y \in \mathbb{R}} \{ym - \Phi_\nu(0, y)\}.$$

Neither quantity depends on T , by (6.6). In words, with the sign convention of (1.3), $\psi_\circ(\nu; y)$ is the free energy of a single spin in the field y and in the random hierarchical field encoded by ν , and $g_m(\nu)$ is the convex conjugate of $y \mapsto -\psi_\circ(\nu; y)$. The function $y \mapsto ym - \Phi_\nu(0, y)$ is strictly concave, by Lemma 6.2, and for $|m| < 1$ it tends to $-\infty$ at $\pm\infty$, by Lemma 6.4. For $|m| < 1$, we therefore denote by

$$(6.15) \quad y(\nu, m) \in \mathbb{R} \quad \text{the unique point with } \partial_x \Phi_\nu(0, y(\nu, m)) = m,$$

which is the unique maximizer in (6.14). We will see in Proposition 6.6 that $g_{\pm 1}(\nu) = \log 2 + \int v \nu(dv)$.

6.2. The one-spin branch in covariance time. We now relate the Parisi equation (6.5), which is parametrized by the covariance v , to the branch equation (2.5) for a single spin at $t = 0$, which is parametrized by the level r . This dictionary connects the scalar initial condition $\psi_\circ$ to the free energies of Sections 2 to 5, and it is used in Section 7.2. The two parametrizations discussed here are two instances of a more general freedom in the choice of the time variable; see for instance the decompositions of paths in [10].

Let $q : [0, 1] \rightarrow \mathbb{R}_+$ be a regular path with a single species, so that q is continuous, piecewise affine and strictly increasing with $q(0) > 0$, and let $\nu := \nu_q$ be the law of $q(U)$, which is supported in $[q(0), q(1)]$. Its cumulative function is

$$\zeta_\nu(v) = |\{r \in [0, 1] : q(r) \leq v\}| = \begin{cases} 0 & \text{if } 0 \leq v < q(0), \\ q^{-1}(v) & \text{if } q(0) \leq v \leq q(1), \end{cases}$$

which is continuous and piecewise affine on $[0, q(1)]$. Let u be the solution of the one-spin branch equation

$$(6.16) \quad \partial_r u + \dot{q}(r) \partial_x^2 u + r \dot{q}(r) (\partial_x u)^2 = 0 \quad \text{on } [0, 1) \times \mathbb{R}, \quad u(1, x) = \log \cosh x,$$

given by Lemma 2.1 with $N = 1$, $t = 0$, $\Omega = \{-1, 1\}$ and $w \equiv \frac{1}{2}$, for which the terminal function (2.2) is $\log \frac{1}{2}(e^x + e^{-x}) = \log \cosh x$.

Lemma 6.5 (Dictionary between the two parametrizations). *Let q , ν and u be as above, let $T := q(1)$, and let Z be a standard Gaussian random variable. We have*

$$(6.17) \quad \Phi_\nu(v, x) = \begin{cases} \mathbb{E} u(0, x + \sqrt{2(q(0) - v)} Z) & \text{if } 0 \leq v \leq q(0), \\ u(q^{-1}(v), x) & \text{if } q(0) \leq v \leq q(1). \end{cases}$$

In particular, for every $h \in \mathbb{R}$,

$$(6.18) \quad \psi_\circ(\nu_q; h) = q(1) - \mathbb{E} u(0, h + \sqrt{2q(0)} Z).$$

Proof. Let Φ denote the right-hand side of (6.17). The two expressions agree at $v = q(0)$, so Φ is continuous on $[0, T] \times \mathbb{R}$, with $\Phi(T, \cdot) = u(1, \cdot) = \log \cosh$, and $|\partial_x \Phi| \leq 1$ since $|\partial_x u| \leq 1$. On $[q(0), q(1)]$, the chain rule gives $\partial_v \Phi(v, x) = \partial_r u(q^{-1}(v), x) / \dot{q}(q^{-1}(v))$, so that (6.16) reads $-\partial_v \Phi = \partial_x^2 \Phi + q^{-1}(v) (\partial_x \Phi)^2 = \partial_x^2 \Phi + \zeta_\nu(v) (\partial_x \Phi)^2$. On $[0, q(0)]$, the function Φ solves the backward heat equation $-\partial_v \Phi = \partial_x^2 \Phi$, which is (6.5) with $\zeta_\nu = 0$. Thus Φ solves (6.5) classically on each of the finitely many strips delimited by 0, $q(0)$, the images under q of its breakpoints, and $q(1)$, with bounded derivatives in x on each strip by Lemma 2.1. Integrating by parts on each strip, the boundary terms cancel by the continuity of Φ , so Φ is a weak solution of (1.7), and $\Phi = \Phi_\nu$ by the uniqueness in Lemma 6.1. Finally, (6.18) is the first case of (6.17) at $v = 0$, combined with the definition (6.14) of $\psi_\circ$ with $T = q(1)$. $\square$

6.3. A martingale representation, and Legendre duality. We now represent g_m as the value of a stochastic control problem, in which the control is a martingale with prescribed initial value m . Fix $T > 0$, let B be the standard Brownian motion of (6.7), with its completed natural filtration, and let $\mathcal{M}_T$ denote the set of martingales $\alpha = (\alpha(v))_{v \in [0, T]}$ in this filtration with values in $[-1, 1]$. As in the introduction, where the subscript is reserved for the species, we write the time variable of the elements of $\mathcal{M}_T$ in parentheses, and accordingly $B(v) := B_v$. Since the initial σ -algebra is trivial, $\alpha(0)$ is deterministic. For $m \in [-1, 1]$, let $\mathcal{M}_T^m$ be the set of those $\alpha \in \mathcal{M}_T$ with $\alpha(0) = m$. For $\alpha \in \mathcal{M}_T$, we set

$$(6.19) \quad \begin{aligned} \chi_\alpha(x) &:= \int_0^x \mathbb{E}[\alpha(v \wedge T)^2] dv \quad (x \geq 0), \\ c_T(\alpha) &:= \mathbb{E}[I(\alpha(T)) - \sqrt{2}\alpha(T)B(T)] + \int_0^T (1 - \mathbb{E}[\alpha(v)^2]) dv. \end{aligned}$$

We often call functions of the form χ_α potentials. Since $v \mapsto \mathbb{E}[\alpha(v)^2]$ is increasing and takes values in $[\alpha(0)^2, 1]$, the function χ_α is convex, increasing, 1-Lipschitz, vanishes at zero, and has slope $\mathbb{E}[\alpha(T)^2]$ on $[T, \infty)$. For a vector of martingales $\alpha = (\alpha_s)_{s \in \mathcal{S}}$ as in the introduction, the cost there is $\mathcal{C}_T(\alpha) = \sum_s \lambda_s c_{T_s}(\alpha_s)$; the potentials χ_{α_s} are combined in the same way in Section 9.5. For $\nu \in \mathcal{P}([0, T])$, we write $\Phi := \Phi_\nu$ and $\zeta := \zeta_\nu$ until the end of the proof of Proposition 6.6. For $y \in \mathbb{R}$, recall that X^y solves (6.7) from $X_0^y = y$, and set $M_v^y := \partial_x \Phi(v, X_v^y)$, so that $M_0^y = \partial_x \Phi(0, y)$ and $M_T^y = \tanh X_T^y$. By Lemma 6.2, M^y is a martingale with values in $(-1, 1)$, so $M^y \in \mathcal{M}_T$.

Proposition 6.6 (Scalar martingale formula). *Let $\nu \in \mathcal{P}([0, T])$.*

(1) *For every $y \in \mathbb{R}$,*

$$(6.20) \quad \Phi(0, y) = \sup_{\alpha \in \mathcal{M}_T} \left\{ y \alpha(0) + \mathbb{E}[\sqrt{2}\alpha(T)B(T) - I(\alpha(T))] + \int_0^T \zeta(v) \mathbb{E}[\alpha(v)^2] dv \right\},$$

and the supremum is attained exactly at $\alpha = M^y$.

(2) *For every $m \in [-1, 1]$,*

$$(6.21) \quad g_m(\nu) = \inf_{\alpha \in \mathcal{M}_T^m} \left\{ c_T(\alpha) + \int \chi_\alpha d\nu \right\}.$$

For $|m| < 1$, the infimum is attained at $M^{y(\nu, m)}$. For $m = \pm 1$, it is attained at the constant martingale $\alpha \equiv m$, and

$$(6.22) \quad g_{\pm 1}(\nu) = \log 2 + \int v \nu(dv).$$

Proof. Step 1: the unconstrained formula.

In the normalization (6.1), with the Brownian motion $W_t := B_{Tt}/\sqrt{T}$ on $[0, 1]$ and the external field $h = y$, [24, Lemma 2.2] states that

$$\Phi(0, y) = \sup_{\tilde{\alpha}} \mathbb{E} \left[\frac{1}{2} \int_0^1 \xi_T''(t) \mu([0, t]) \tilde{\alpha}_t^2 dt + \tilde{\alpha}_1 \int_0^1 \sqrt{\xi_T''(t)} dW_t - \phi^*(\tilde{\alpha}_1) \right],$$

where the supremum is over the bounded martingales $\tilde{\alpha}$ on $[0, 1]$ in the completed filtration of W , and where $\phi^*(z) = I(z) - yz$ for $z \in [-1, 1]$ and $\phi^*(z) = +\infty$ otherwise is the convex conjugate of $x \mapsto \log \cosh(x + y)$; moreover, the supremum is attained exactly at $\tilde{\alpha}_t = \partial_x \Phi(Tt, \tilde{X}_t)$, where $d\tilde{X}_t = \xi_T''(t) \mu([0, t]) \partial_x \Phi(Tt, \tilde{X}_t) dt + \sqrt{\xi_T''(t)} dW_t$ and $\tilde{X}_0 = y$. The map $\alpha \mapsto \tilde{\alpha}$, $\tilde{\alpha}_t := \alpha(Tt)$, is a bijection from $\mathcal{M}_T$ onto the set of bounded martingales $\tilde{\alpha}$ with $|\tilde{\alpha}_1| \leq 1$ almost surely, and the other bounded martingales have objective $-\infty$. Substituting $t = v/T$, the objective becomes $\int_0^T \zeta(v) \alpha(v)^2 dv + \sqrt{2}\alpha(T)B(T) - I(\alpha(T)) + y\alpha(T)$, and $\mathbb{E}[\alpha(T)] = \alpha(0)$. This gives (6.20). Finally, $\tilde{X}_t = X_{Tt}^y$, so the maximizer is $\tilde{\alpha}_t = M_{Tt}^y$, that is, $\alpha = M^y$.

Step 2: the lower bound.

Let $\alpha \in \mathcal{M}_T$. By Fubini's theorem, since $\zeta(v) = \nu([0, v])$,

$$(6.23) \quad \int_0^T \zeta(v) \mathbb{E}[\alpha(v)^2] dv = \int \nu(dw) \int_w^T \mathbb{E}[\alpha(v)^2] dv = \chi_\alpha(T) - \int \chi_\alpha d\nu.$$

Since $T - \chi_\alpha(T) = \int_0^T (1 - \mathbb{E}[\alpha(v)^2]) dv$, the objective in (6.20) is $y\alpha(0) + T - c_T(\alpha) - \int \chi_\alpha d\nu$, and (6.20) reads

$$(6.24) \quad \Phi(0, y) = T + \sup_{\alpha \in \mathcal{M}_T} \left\{ y\alpha(0) - c_T(\alpha) - \int \chi_\alpha d\nu \right\}.$$

For $\alpha \in \mathcal{M}_T^m$, this gives $c_T(\alpha) + \int \chi_\alpha d\nu \geq T - \Phi(0, y) + ym = \psi_\circ(\nu; y) + ym$ for every $y \in \mathbb{R}$, and taking the supremum over y in (6.14), we obtain $c_T(\alpha) + \int \chi_\alpha d\nu \geq g_m(\nu)$.

Step 3: attainment for $|m| < 1$.

Let $y := y(\nu, m)$ and $\alpha := M^y$, which belongs to $\mathcal{M}_T^m$ by (6.15). Since α attains the supremum in (6.24), we have $c_T(\alpha) + \int \chi_\alpha d\nu = T - \Phi(0, y) + ym = g_m(\nu)$, by the choice of y in (6.14).

Step 4: the endpoints.

Let $m = 1$. For the constant martingale $\alpha \equiv 1$, we have $c_T(\alpha) = I(1) = \log 2$ and $\chi_\alpha(x) = x$, so Step 2 gives $g_1(\nu) \leq \log 2 + \int v \nu(dv)$. For the reverse inequality, let $y \in \mathbb{R}$, $X := X^y$, $M := M^y$ and $m_y := M_0$. By (6.10) and the identity $\log \cosh x = x - \log 2 + \log(1 + e^{-2x})$,

$$\Phi(0, y) \leq \mathbb{E}[X_T] - \log 2 + \mathbb{E}[\log(1 + e^{-2X_T})] - \int_0^T \zeta(v) \mathbb{E}[M_v^2] dv.$$

By (6.7) and the martingale property, $\mathbb{E}[X_T] = y + 2m_y \int_0^T \zeta$, and $2m_y - \mathbb{E}[M_v^2] \leq 2m_y - m_y^2 \leq 1$ by Jensen's inequality, so that

$$\Phi(0, y) \leq y - \log 2 + \int_0^T \zeta + \mathbb{E}[\log(1 + e^{-2X_T})].$$

Since the drift in (6.7) is at least -2 , we have $X_T \geq y - 2T + \sqrt{2}B_T$, so for $y \geq 2T$ the last integrand is bounded by $\log(1 + e^{-2\sqrt{2}B_T})$, which is integrable, and it tends to zero as $y \rightarrow \infty$. By dominated convergence, $\sup_y \{y - \Phi(0, y)\} \geq \log 2 - \int_0^T \zeta$, that is, $g_1(\nu) \geq T + \log 2 - \int_0^T \zeta = \log 2 + \int v \nu(dv)$, since $\int_0^T (1 - \zeta) = \int v \nu(dv)$. This proves (6.22) for $m = 1$, and shows that the infimum is attained at $\alpha \equiv 1$. The case $m = -1$ follows by the symmetry $x \mapsto -x$. $\square$

We next collect the elementary properties of g_m that follow from Legendre duality.

Lemma 6.7 (Legendre duality in the magnetization). *Let $\nu \in \mathcal{P}([0, T])$. The function $m \mapsto g_m(\nu)$ is finite and continuous on $[-1, 1]$, and the function*

$$(6.25) \quad m \mapsto g_m(\nu) - I(m)$$

is even and convex on $[-1, 1]$, with Lipschitz constant $2 \int v \nu(dv)$:

$$(6.26) \quad |g_m(\nu) - g_{m'}(\nu) - I(m) + I(m')| \leq 2|m - m'| \int v \nu(dv), \quad m, m' \in [-1, 1].$$

In particular, for $|m_0| < 1$ and $m \in [-1, 1]$,

$$(6.27) \quad g_m(\nu) \geq g_{m_0}(\nu) + y(\nu, m_0)(m - m_0) + \frac{1}{2}(m - m_0)^2.$$

For every $h \in \mathbb{R}$,

$$(6.28) \quad \psi_\circ(\nu; h) = \inf_{m \in [-1, 1]} \{g_m(\nu) - hm\},$$

where the infimum is attained exactly at $m = \partial_x \Phi_\nu(0, h) \in (-1, 1)$. Finally, $g_m(\delta_0) = I(m)$, and $|g_m(\nu) - g_m(\nu')| \leq W_1(\nu, \nu')$ for all $\nu, \nu' \in \mathcal{P}([0, T])$ and $m \in [-1, 1]$. Consequently, g_m extends uniquely to a function on $\mathcal{P}_1(\mathbb{R}_+)$ that is 1-Lipschitz for W_1 .

Proof. Step 1: elementary properties.

For $\nu = \delta_0$, we have $\Phi_\nu(0, x) = \log \cosh x + T$, so $m \mapsto g_m(\delta_0)$ is the convex dual of $\log \cosh$, which is I . The Lipschitz bound in ν stated at the very end follows from (6.11), since $\int_0^T |\zeta_\nu - \zeta_{\nu'}| = W_1(\nu, \nu')$, and this bound passes to the supremum over y . Since T is arbitrary and g_m does not depend on it, the bound holds for all compactly supported laws, and these are dense in $\mathcal{P}_1(\mathbb{R}_+)$ for W_1 : if ν^T is the image of $\nu \in \mathcal{P}_1(\mathbb{R}_+)$ under $x \mapsto x \wedge T$, then $W_1(\nu, \nu^T) = \int (x - T)_+ d\nu \rightarrow 0$ as $T \rightarrow \infty$. The extension follows. The function $m \mapsto g_m(\nu)$ is finite on $[-1, 1]$, by (6.22) and (6.15), and convex and lower semicontinuous, as a supremum of affine functions of m . It is also even, since $\Phi_\nu(0, \cdot)$ is even.

Step 2: convexity of (6.25).

The function $m \mapsto g_m(\nu) - T$ is the Legendre transform of the strictly convex function $\Phi_\nu(0, \cdot)$, whose derivative is a continuous increasing bijection onto $(-1, 1)$ with inverse $y(\nu, \cdot)$, by Lemma 6.4. Hence $m \mapsto g_m(\nu)$ is differentiable on $(-1, 1)$, with

$$(6.29) \quad \partial_m g_m(\nu) = y(\nu, m), \quad |m| < 1.$$

Since $\Phi_\nu(0, \cdot)$ is twice continuously differentiable with $\partial_x^2 \Phi_\nu > 0$, by Lemmas 6.1 and 6.2, the function $y(\nu, \cdot)$ is continuously differentiable on $(-1, 1)$ by the inverse function theorem, and differentiating the identity $\partial_x \Phi_\nu(0, y(\nu, m)) = m$ and using (6.2), we get

$$\partial_m^2 g_m(\nu) = \partial_m y(\nu, m) = \frac{1}{\partial_x^2 \Phi_\nu(0, y(\nu, m))} \geq \frac{1}{1 - m^2} = I''(m), \quad |m| < 1.$$

Thus (6.25) is convex on $(-1, 1)$. At the endpoints, lower semicontinuity and convexity give

$$\liminf_{m \rightarrow 1} g_m(\nu) \geq g_1(\nu), \quad g_m(\nu) \leq \frac{1-m}{2} g_{-1}(\nu) + \frac{1+m}{2} g_1(\nu),$$

so that $g_m(\nu) \rightarrow g_1(\nu)$ as $m \rightarrow 1$; the endpoint -1 is similar. Hence $m \mapsto g_m(\nu)$ is continuous on $[-1, 1]$, and the convexity of (6.25) extends to $[-1, 1]$. Since $I''(m) = (1 - m^2)^{-1} \geq 1$, the function $m \mapsto g_m(\nu) - m^2/2$ is convex as well, and (6.27) is its supporting inequality at m_0 , where its derivative is $y(\nu, m_0) - m_0$ by (6.29).

Step 3: the Lipschitz bound.

Let $\varphi(m) := g_m(\nu) - I(m)$, so that $\varphi'(m) = y(\nu, m) - \operatorname{arctanh} m$ for $|m| < 1$, by (6.29). Evaluating the barriers (6.12) at $v = 0$ and $x = y(\nu, m)$, where $\partial_x \Phi_\nu(0, x) = m$, and inverting the increasing function $\tanh$, we obtain

$$y(\nu, m) - 2b_\nu(0) \leq \operatorname{arctanh} m \leq y(\nu, m) + 2b_\nu(0), \quad b_\nu(0) = \int v \nu(dv),$$

that is, $|\varphi'| \leq 2 \int v \nu(dv)$ on $(-1, 1)$. Since φ is continuous on $[-1, 1]$ by Step 2, (6.26) follows.

Step 4: duality.

Let $m_0 := \partial_x \Phi_\nu(0, h)$, so that $y(\nu, m_0) = h$ and $g_{m_0}(\nu) - h m_0 = T - \Phi_\nu(0, h)$. For every other $m \in [-1, 1]$, taking $y = h$ in (6.14) gives $g_m(\nu) - h m \geq T - \Phi_\nu(0, h)$, and the minimizer is unique by (6.27). $\square$

6.4. Concavity in the law, and supporting functions. The right-hand side of (6.21) is an infimum of affine functions of ν , so we can readily deduce the concavity of the map $\nu \mapsto g_m(\nu)$. To describe the supporting functions, for $|m| < 1$, we denote by $M^{\nu, m} := M^{y(\nu, m)}$ (defined with $\Phi = \Phi_\nu$) the minimizing martingale of Proposition 6.6,

and for $m = \pm 1$ we let $M^{\nu,m} \equiv m$. We write $\chi_{\nu,m} := \chi_{M^{\nu,m}}$ for the associated potential, which is convex, increasing, 1-Lipschitz, and vanishes at zero.

Corollary 6.8 (Concavity in the law and supporting functions). *Let $m \in [-1, 1]$. The function $\nu \mapsto g_m(\nu)$ is concave on $\mathcal{P}([0, T])$ for the affine interpolation of measures, and for all $\nu, \nu' \in \mathcal{P}([0, T])$,*

$$(6.30) \quad g_m(\nu') \leq g_m(\nu) + \int \chi_{\nu,m} d(\nu' - \nu).$$

Each $\chi_{\nu,m}$ is continuously differentiable on $\mathbb{R}_+$, with $\chi'_{\nu,m}(v) = \mathbb{E}[(M_{v \wedge T}^{\nu,m})^2]$. If $\zeta_{\nu_n} \rightarrow \zeta_\nu$ in $L^1([0, T])$, then $\chi'_{\nu_n,m} \rightarrow \chi'_{\nu,m}$ uniformly on $\mathbb{R}_+$; in particular, $\chi_{\nu_n,m} \rightarrow \chi_{\nu,m}$ uniformly on $[0, T]$. Consequently, for all $\nu, \nu' \in \mathcal{P}([0, T])$, with $\nu_\varepsilon := (1 - \varepsilon)\nu + \varepsilon\nu'$,

$$(6.31) \quad \lim_{\varepsilon \downarrow 0} \frac{g_m(\nu_\varepsilon) - g_m(\nu)}{\varepsilon} = \int \chi_{\nu,m} d(\nu' - \nu).$$

Proof. Concavity is clear from (6.21). The inequality (6.30) follows from applying the inequality in (6.21) at ν' with the martingale $\alpha = M^{\nu,m}$, and the equality at ν with the same martingale. The function $v \mapsto \mathbb{E}[(M_v^{\nu,m})^2]$ is continuous on $[0, T]$, since $M^{\nu,m}$ is a bounded continuous process, which gives the differentiability of $\chi_{\nu,m}$. For the convergence statement, the case $m = \pm 1$ is trivial. Let $|m| < 1$, and let $y_n := y(\nu_n, m)$ and $y := y(\nu, m)$. Since $\partial_x \Phi_{\nu_n}(0, \cdot) \rightarrow \partial_x \Phi_\nu(0, \cdot)$ uniformly by Lemma 6.3, we have $\partial_x \Phi_\nu(0, y_n) \rightarrow m = \partial_x \Phi_\nu(0, y)$, and $y_n \rightarrow y$ because $\partial_x \Phi_\nu(0, \cdot)$ is a continuous increasing bijection. Let X^n solve (6.7) with $\zeta_n := \zeta_{\nu_n}$ and $\Phi_n := \Phi_{\nu_n}$ in place of ζ and Φ , driven by the same Brownian motion as $X := X^y$ and started at y_n , and let $M_v^n := \partial_x \Phi_n(v, X_v^n)$. Since the noise is additive, and $\partial_x \Phi_\nu$ is 1-Lipschitz in x with $|\partial_x \Phi_n| \leq 1$,

$$|X_v^n - X_v| \leq |y_n - y| + 2 \int_0^T |\zeta_n - \zeta_\nu| + 2T \sup |\partial_x \Phi_n - \partial_x \Phi_\nu| + 2 \int_0^v |X_w^n - X_w| dw,$$

and Gronwall's inequality shows that $\sup_v |X_v^n - X_v| \rightarrow 0$ for every realization of B , by Lemma 6.3. Since $\partial_x \Phi_n \rightarrow \partial_x \Phi_\nu$ uniformly, we also get $\sup_v |M_v^n - M_v^{\nu,m}| \rightarrow 0$ for every realization. Since these processes are bounded by 1, bounded convergence gives $\sup_{v \leq T} |\mathbb{E}[(M_v^n)^2] - \mathbb{E}[(M_v^{\nu,m})^2]| \rightarrow 0$, which is the uniform convergence of the derivatives of the potentials; the potentials themselves converge uniformly on $[0, T]$ since they vanish at 0. Finally, the supporting inequalities at ν and at ν_ε give

$$\int \chi_{\nu_\varepsilon,m} d(\nu' - \nu) \leq \frac{g_m(\nu_\varepsilon) - g_m(\nu)}{\varepsilon} \leq \int \chi_{\nu,m} d(\nu' - \nu),$$

and $\zeta_{\nu_\varepsilon} \rightarrow \zeta_\nu$ in L^1 as $\varepsilon \rightarrow 0$, so (6.31) follows from the continuity statement. $\square$

6.5. The susceptibility in magnetization coordinates. We now prepare the proof of the path convexity of g_m . For illustration, we start with the simplest example, which shows the effect of fixing the magnetization rather than the field. For the constant path $q \equiv v$, that is $\nu = \delta_v$, the recursion (6.4) gives $\psi_\circ(\delta_v; h) = v - \mathbb{E} \log \cosh(h + \sqrt{2v}Z)$, and the heat equation gives

$$\frac{d}{dv} \psi_\circ(\delta_v; h) = \mathbb{E} \tanh^2(h + \sqrt{2v}Z), \quad \left. \frac{d^2}{dv^2} \psi_\circ(\delta_v; h) \right|_{v=0} = 2(1 - \tanh^2 h)(1 - 3 \tanh^2 h),$$

which is negative when $\tanh^2 h > 1/3$: the function $\psi_\circ(\cdot; h)$ is not convex in the path when $|h|$ is large. Let now $|m| < 1$. The optimal field $y(\delta_v, m)$ in (6.14) solves $\mathbb{E} \tanh(y + \sqrt{2v}Z) = m$, so it equals $\operatorname{arctanh} m$ at $v = 0$ and, by implicit differentiation, moves with speed $2m$

there. By the envelope theorem, $\frac{d}{dv}g_m(\delta_v) = \mathbb{E} \tanh^2(y(\delta_v, m) + \sqrt{2v}Z)$, and differentiating once more, the motion of the optimal field contributes a second, positive term:

$$\left. \frac{d^2}{dv^2}g_m(\delta_v) \right|_{v=0} = 2(1-m^2)(1-3m^2) + 2m(1-m^2) \cdot 2m = 2(1-m^2)^2 > 0.$$

For a finite-step path with distinct heights $q_0 < \dots < q_\ell$, carried by intervals of lengths $w_0, \dots, w_\ell$, we will show in Proposition 6.10 that $\partial_{q_i}g_m(q) = w_i\mathbb{E}[M_{q_i}^2]$, where M is the magnetization martingale of Lemma 6.2 started at m , and convexity will follow from the following two observations. First, raising another height q_j , $j \neq i$, lowers the cumulative function ζ on the interval swept by the atom, which should make the magnetization at the height q_i less spread, so the off-diagonal entries of the Hessian of g_m should be nonpositive. Second, raising all the heights together amounts to starting the magnetization process from a more spread initial point, so the row sums of the Hessian of g_m should be nonnegative. These two properties imply that the Hessian of g_m is positive semidefinite. Lemma 6.9 below makes these monotonicity properties precise, in coordinates that we now introduce.

Throughout this subsection and the next, $\nu \in \mathcal{P}([0, T])$ has finite support and $T > \max \text{supp } \nu$; we write $\Phi := \Phi_\nu$ and $\zeta := \zeta_\nu$, so that Φ is given by the recursion (6.4), and is smooth on each strip $[v_i, v_{i+1}] \times \mathbb{R}$ delimited by the atoms of ν . By Lemma 6.4, for every $v \in [0, T]$ the map $x \mapsto \partial_x \Phi(v, x)$ is an increasing bijection from $\mathbb{R}$ onto $(-1, 1)$; we denote by $x(v, z)$ its inverse, which is smooth on each strip $[v_i, v_{i+1}] \times (-1, 1)$ by the inverse function theorem, since $\partial_x^2 \Phi > 0$. We define

$$(6.32) \quad A(v, z) := \partial_x^2 \Phi(v, x(v, z)), \quad W(v, z) := \sup_{x \in \mathbb{R}} \{xz - \Phi(v, x)\}, \quad z \in (-1, 1).$$

Thus A is the susceptibility $\partial_x^2 \Phi$, viewed as a function of the magnetization $\partial_x \Phi$, and $W(v, \cdot)$ is the Legendre transform of $\Phi(v, \cdot)$, so that $g_m(\nu) = T + W(0, m)$. Both functions are smooth on each strip $[v_i, v_{i+1}] \times (-1, 1)$. The envelope formula gives $\partial_z W(v, z) = x(v, z)$, hence $\partial_z^2 W = 1/A$, and $\partial_v W = -\partial_v \Phi(v, x(v, z))$, hence by (6.5)

$$(6.33) \quad \partial_v W = A + \zeta(v)z^2, \quad \partial_z^2 W = \frac{1}{A}, \quad W(T, z) = I(z),$$

on each strip. Moreover, by (6.2),

$$(6.34) \quad 0 < A(v, z) \leq 1 - z^2, \quad z \in (-1, 1),$$

so that A extends continuously to $[0, T] \times [-1, 1]$ by $A(v, \pm 1) := 0$.

The martingale M of Lemma 6.2 is the magnetization process of a single spin, and it is natural to describe its dynamics in its own coordinates. For $s \in [0, T]$ and $z \in (-1, 1)$, we define the magnetization process started at time s from z by

$$(6.35) \quad M_r^{s,z} := \partial_x \Phi(r, X_r^{s,x(s,z)}), \quad r \in [s, T],$$

and for $z = \pm 1$ we let $M^{s,\pm 1} \equiv \pm 1$. We write $\mathbb{E}_z$ for the expectation of functionals of $M^{s,z}$, the starting time s being understood; by the Markov property of (6.7), the process $M^{s,z}$ is a Markov process, and its law depends only on (s, z) . For $s = 0$, we have $x(0, z) = y(\nu, z)$ by (6.15), so $M^{0,z}$ is the martingale $M^{\nu,z}$ of Proposition 6.6. By Lemma 6.2, $M^{s,z}$ is a martingale started at z with values in $(-1, 1)$, and since $\partial_x^2 \Phi(r, X_r) = A(r, \partial_x \Phi(r, X_r))$, the identity (6.8) reads

$$(6.36) \quad dM_r = \sqrt{2} A(r, M_r) dB_r.$$

In these coordinates, the law ν enters only through the diffusion coefficient A , and the initial condition is the prescribed magnetization. The two comparison statements below say that a larger cumulative function ζ gives a larger diffusion coefficient, and

that a larger diffusion coefficient spreads the magnetization more. We first record the monotone dependence on the starting point. Let $z \leq z'$ in $(-1, 1)$, so that $x := x(s, z) \leq x' := x(s, z')$, and let X and X' solve (6.7) from (s, x) and (s, x') , driven by the same Brownian motion. Their difference $D := X' - X$ is absolutely continuous, with $D'_v = 2\zeta(v)[\partial_x \Phi(v, X'_v) - \partial_x \Phi(v, X_v)]$; since $\partial_x \Phi$ is increasing in x , the function D is increasing on every interval on which it is positive, so $D_v \geq x' - x \geq 0$ for every v . Hence $M_r^{s,z} \leq M_r^{s,z'}$ for every r , and since both are martingales, for every Lipschitz function φ on $[-1, 1]$ and every $v \in [s, T]$,

$$(6.37) \quad |\mathbb{E}_{z'}\varphi(M_v) - \mathbb{E}_z\varphi(M_v)| \leq \|\varphi'\|_\infty \mathbb{E}[M_v^{s,z'} - M_v^{s,z}] = \|\varphi'\|_\infty (z' - z).$$

This remains true when $z' = 1$ or $z = -1$, since $\mathbb{E}|1 - M_v^{s,z}| = 1 - z$ by the martingale property. In particular, $z \mapsto \mathbb{E}_z\varphi(M_v)$ is continuous on $[-1, 1]$, uniformly in s .

Lemma 6.9 (Diffusion comparisons). *Let $\nu_1, \nu_2 \in \mathcal{P}([0, T])$ have finite support, with susceptibilities A_1, A_2 and magnetization processes M^1, M^2 , and let $0 \leq s \leq v \leq T$. We recall that $\mathbb{E}_z$ denotes the expectation when the process starts from $z \in [-1, 1]$, the starting time s being understood.*

- (1) *For every continuous convex function φ on $[-1, 1]$, the map $z \mapsto \mathbb{E}_z[\varphi(M_v^1)]$ is convex on $[-1, 1]$.*
- (2) *If $\zeta_{\nu_1} \leq \zeta_{\nu_2}$ on $[0, T]$, then $A_1 \leq A_2$ on $[0, T] \times [-1, 1]$.*
- (3) *If $A_1 \leq A_2$, then for every continuous convex function φ on $[-1, 1]$,*

$$\mathbb{E}_z[\varphi(M_v^1)] \leq \mathbb{E}_z[\varphi(M_v^2)], \quad z \in [-1, 1].$$

Proof. Step 0: regularity.

We drop the index 1 in this step and the next. In items (1) and (3), we may assume that φ is smooth on $[-1, 1]$, by extending φ outside $[-1, 1]$ (e.g. via a dilation) and mollifying, which preserves convexity and gives uniform approximations; and the case $s = v$ is trivial, so let $s < v$. For $s' \in [s, v]$ and $x \in \mathbb{R}$, let $F(s', x) := \mathbb{E}[\varphi(\partial_x \Phi(v, X_v^{s', x}))]$. On each strip, the drift of (6.7) is smooth in x with bounded derivatives, so by the parabolic theory used in Lemma 2.1 and Itô's formula, F is smooth in x with bounded derivatives on each strip, continuous on $[s, v] \times \mathbb{R}$, and solves the backward Kolmogorov equation $-\partial_{s'} F = \partial_x^2 F + 2\zeta \partial_x \Phi \partial_x F$ classically on each strip. For $|z| < 1$, we have $u(s', z) := \mathbb{E}_z[\varphi(M_v)] = F(s', x(s', z))$, which is therefore smooth on each strip $[v_i, v_{i+1}] \times (-1, 1)$ intersected with $[s, v] \times (-1, 1)$, with derivatives bounded on $[-1 + \eta, 1 - \eta]$ for every $\eta > 0$; the flow of (6.7) and the inverse $x(\cdot, \cdot)$ are jointly continuous, so u is continuous on $[s, v] \times (-1, 1)$, and on $[s, v] \times [-1, 1]$ by (6.37). Writing $F(s', x) = u(s', \partial_x \Phi(s', x))$, the chain rule and (6.13) transform the equation for F into

$$(6.38) \quad -\partial_{s'} u = a \partial_z^2 u, \quad a := A^2,$$

on each strip, in the interior.

Step 1: convexity in the starting point.

Let $z \in (-1, 1)$, $x := x(s, z)$, $X := X^{s,x}$ and $M_r := \partial_x \Phi(r, X_r)$. Since the noise in (6.7) is additive and the drift is smooth in x on each strip, for every Brownian path the map $x' \mapsto X_r^{s,x'}$ is continuously differentiable, with derivative $J_r := \exp(2 \int_s^r \zeta(w) \partial_x^2 \Phi(w, X_w) dw) \in [1, e^{2T}]$ at $x' = x$. Differentiating under the expectation and using $\partial_z u = \partial_x F / \partial_x^2 \Phi$, we obtain

$$\partial_z u(s, z) = \mathbb{E}[\varphi'(M_v) N_v], \quad N_r := \frac{\partial_x^2 \Phi(r, X_r) J_r}{\partial_x^2 \Phi(s, x)}, \quad r \in [s, v].$$

By Itô's formula, the drift of $\partial_x^2 \Phi(r, X_r)$ is $-2\zeta(\partial_x^2 \Phi)^2$, as in the proof of Lemma 6.2, while $dJ_r = 2\zeta \partial_x^2 \Phi(r, X_r) J_r dr$, so that

$$dN_r = \sqrt{2} N_r \frac{\partial_x^3 \Phi}{\partial_x^2 \Phi}(r, X_r) dB_r.$$

Thus N is a positive continuous local martingale, started from 1 at time s and bounded by $e^{2T}/\partial_x^2 \Phi(s, x)$, hence a martingale with $\mathbb{E}[N_v] = 1$. Let $\tilde{\mathbb{P}} := N_v \cdot \mathbb{P}$ on the σ -algebra at time v . By Girsanov's theorem, $\tilde{B}_r := B_r - \sqrt{2} \int_s^r (\partial_x^3 \Phi / \partial_x^2 \Phi)(w, X_w) dw$ is a Brownian motion under $\tilde{\mathbb{P}}$ on $[s, v]$, and

$$dX_r = b(r, X_r) dr + \sqrt{2} d\tilde{B}_r, \quad b := 2\zeta \partial_x \Phi + 2 \partial_x \log \partial_x^2 \Phi,$$

while $\partial_z u(s, z) = \tilde{\mathbb{E}}[\varphi'(\partial_x \Phi(v, X_v))]$. The function b is smooth in x on each strip, since $\partial_x^2 \Phi > 0$. For a Brownian motion B' and $x' \in \mathbb{R}$, the equation $dY_r = b(r, Y_r) dr + \sqrt{2} dB'_r$ on $[s, v]$, started from x' at time s , is, for every Brownian path, an ordinary differential equation for $Y - \sqrt{2}B'$ with a locally Lipschitz vector field: it has a unique solution up to its explosion time, which is a measurable functional of the Brownian path, and two solutions driven by the same Brownian motion from $x' < x''$ remain ordered as long as both exist. Under $\tilde{\mathbb{P}}$, the process X is such a solution, driven by $\tilde{B}$ and started at x , and it does not explode before v , since it is defined on $[s, v]$ almost surely under $\mathbb{P}$, hence under $\tilde{\mathbb{P}}$. By uniqueness in law, the solution Y^x started at x and driven by any Brownian motion does not explode before v either, and $\partial_z u(s, z) = \mathbb{E}[\varphi'(\partial_x \Phi(v, Y_v^x))]$. Now let $z < z'$ in $(-1, 1)$, so that $x = x(s, z) < x' := x(s, z')$. Driving Y^x and $Y^{x'}$ by the same Brownian motion, we get $Y_v^x < Y_v^{x'}$, and since φ' and $\partial_x \Phi(v, \cdot)$ are increasing, $\partial_z u(s, z) \leq \partial_z u(s, z')$. Hence $u(s, \cdot)$ is convex on $(-1, 1)$, and on $[-1, 1]$ by continuity. This proves item (1).

Step 2: comparison of the coefficients.

Adding atoms of mass zero, we may assume that ν_1 and ν_2 have the same atoms $v_0 < \dots < v_\ell$, so that $\zeta_1 := \zeta_{\nu_1}$ and $\zeta_2 := \zeta_{\nu_2}$ are constant on the same strips. For $\theta \in [0, 1]$, let $\nu_\theta := (1 - \theta)\nu_1 + \theta\nu_2$, whose cumulative function is $\zeta_\theta = \zeta_1 + \theta\eta$ with $\eta := \zeta_2 - \zeta_1 \geq 0$, and let Φ_θ , x_θ , A_θ , W_θ and M^θ be the corresponding objects. On each strip, by the recursion (6.4), $\Phi_\theta(v, x)$ is a finite composition of operators $T_{r,a}$, whose parameters $r = \zeta_\theta(v_i)$ are affine in θ , applied to $\log \cosh$. By the smoothness of $(r, a, x) \mapsto T_{r,a}f(x)$ noted after (6.3), and by differentiation under the expectation, $\Phi_\theta(v, x)$ is smooth in (θ, v, x) on $[0, 1] \times [v_i, v_{i+1}] \times \mathbb{R}$, with derivatives in x of all orders bounded. By the inverse function theorem, $x_\theta(v, z)$, $A_\theta(v, z)$ and $W_\theta(v, z)$ are then smooth in (θ, v, z) on $[0, 1] \times [v_i, v_{i+1}] \times (-1, 1)$. Let $H := \partial_\theta W_\theta$, which is continuous on $[0, T] \times (-1, 1)$ for each θ . Differentiating (6.33) in θ , and using $\partial_\theta(1/\partial_z^2 W_\theta) = -A_\theta^2 \partial_z^2 H$, we find that on each strip,

$$(6.39) \quad -\partial_v H = A_\theta^2 \partial_z^2 H - \eta(v) z^2, \quad H(T, z) = 0,$$

the terminal condition coming from $W_\theta(T, \cdot) = I$. Moreover, by Lemma 6.3, $\sup_x |\Phi_\theta(v, x) - \Phi_{\theta'}(v, x)| \leq |\theta - \theta'| \int_v^T \eta$, and the Legendre transform is 1-Lipschitz for the uniform norm, so

$$(6.40) \quad |H(v, z)| \leq \int_v^T \eta, \quad (v, z) \in [0, T] \times (-1, 1).$$

Fix θ , and for this step let $s \in [0, T]$ be arbitrary. Let $z \in (-1, 1)$, let $M := M^\theta$ be started at (s, z) , and for small $\varepsilon > 0$ let τ_ε be its exit time from $[-1 + \varepsilon, 1 - \varepsilon]$. On each strip, H is smooth with bounded derivatives on this interval, so Itô's formula on $[s, T \wedge \tau_\varepsilon]$, together

with (6.36) and (6.39), gives

$$H(s, z) = \mathbb{E}[H(T \wedge \tau_\varepsilon, M_{T \wedge \tau_\varepsilon})] - \mathbb{E} \int_s^{T \wedge \tau_\varepsilon} \eta(r) M_r^2 dr.$$

Since M takes values in $(-1, 1)$ and is continuous, $T \wedge \tau_\varepsilon = T$ for every sufficiently small ε , for every realization. Letting $\varepsilon \rightarrow 0$, by the continuity of H , the bound (6.40) and bounded convergence,

$$(6.41) \quad H(s, z) = - \int_s^T \eta(r) \mathbb{E}_z[(M_r^\theta)^2] dr.$$

By item (1) applied to ν_θ and $\varphi(z) = z^2$, each function $z \mapsto \mathbb{E}_z[(M_r^\theta)^2]$ is convex, and $\eta \geq 0$, so $H(s, \cdot)$ is concave on $(-1, 1)$, that is, $\partial_z^2 H \leq 0$. Hence $\partial_\theta A_\theta = \partial_\theta(1/\partial_z^2 W_\theta) = -A_\theta^2 \partial_z^2 H \geq 0$: the function $\theta \mapsto A_\theta(s, z)$ is increasing on $[0, 1]$, which gives $A_1 \leq A_2$ on $[0, T] \times (-1, 1)$. Both vanish at $z = \pm 1$. This proves item (2).

Step 3: comparison of the processes.

Let $u_i(s', z) := \mathbb{E}_z[\varphi(M_v^i)]$ for the process started at time s' from z , and $a_i := A_i^2$. By Step 0 and item (1), u_1 is continuous on $[s, v] \times [-1, 1]$, smooth on each strip in the interior, convex in z , and solves $-\partial_{s'} u_1 = a_1 \partial_z^2 u_1$. Let $|z| < 1$ and $\varepsilon > 0$, and let τ_ε be the exit time of M^2 , started at time s from z , from $[-1 + \varepsilon, 1 - \varepsilon]$. By Itô's formula on $[s, v \wedge \tau_\varepsilon]$ and (6.36), the drift of $u_1(r, M_r^2)$ is $(a_2 - a_1) \partial_z^2 u_1 \geq 0$, so that $u_1(s, z) \leq \mathbb{E}[u_1(v \wedge \tau_\varepsilon, M_{v \wedge \tau_\varepsilon}^2)]$. Letting $\varepsilon \rightarrow 0$ and using the continuity of u_1 , we obtain $u_1(s, z) \leq \mathbb{E}[u_1(v, M_v^2)] = u_2(s, z)$ by bounded convergence. The case $|z| = 1$ is trivial. $\square$

6.6. Convexity in the path. We now regard g_m as a function of the path. For a finite-step path

$$(6.42) \quad q = \sum_{i=0}^{\ell} q_i \mathbb{1}_{[r_i, r_{i+1})}, \quad 0 = r_0 < r_1 < \dots < r_{\ell+1} = 1, \quad 0 \leq q_0 \leq q_1 \leq \dots \leq q_\ell,$$

the law ν_q has atoms q_i with masses $w_i := r_{i+1} - r_i$, and we write $g_m(q) = g_m(\nu_q)$. In this subsection, the partition (r_i) is fixed and we regard $g_m(q)$ as a function of the heights $(q_0, \dots, q_\ell)$; the derivative $\partial_{q_i} g_m(q)$ refers to this function. When the heights are distinct, ζ_{ν_q} takes the value r_{i+1} on $[q_i, q_{i+1})$. We write M for the magnetization process (6.35) associated with ν_q and started at time 0, and $\mathbb{E}_z$ for the expectation when it starts from z ; in particular, $\mathbb{E}_m[\varphi(M_v)] = \mathbb{E}[\varphi(M_v^{\nu_q, m})]$.

Proposition 6.10 (Convexity at fixed magnetization). *For every $m \in [-1, 1]$, the function g_m is convex on $\mathcal{Q}_1$ and satisfies*

$$(6.43) \quad |g_m(q) - g_m(q')| \leq \|q - q'\|_{L^1}, \quad q, q' \in \mathcal{Q}_1.$$

For a finite-step path (6.42) with $0 < q_0 < \dots < q_\ell$ and $|m| < 1$, its derivative is

$$(6.44) \quad \partial_{q_i} g_m(q) = (r_{i+1} - r_i) \mathbb{E}_m[M_{q_i}^2], \quad i \in \{0, \dots, \ell\}.$$

More generally, for every $a \in \mathcal{Q}_\infty$, the path

$$(6.45) \quad z_a(r) := \mathbb{E}[(M_{a(r)}^{\nu_a, m})^2] = \chi'_{\nu_a, m}(a(r)), \quad r \in [0, 1],$$

is increasing with values in $[m^2, 1]$, and it supports g_m at a : for every $b \in \mathcal{Q}_\infty$,

$$(6.46) \quad g_m(b) \geq g_m(a) + \int_0^1 z_a(r) (b(r) - a(r)) dr,$$

and the integral in (6.46) is the right derivative of $\theta \mapsto g_m(a + \theta(b - a))$ at $\theta = 0$.

As explained at the beginning of Section 6.5, the proof of convexity consists in showing that the Hessian of g_m in the heights has nonpositive off-diagonal entries and nonnegative row sums. The sign of the off-diagonal entries comes from items (2) and (3) of Lemma 6.9, and the sign of the row sums from item (1). For $m = \pm 1$, we have $M \equiv m$ and $z_a \equiv 1$, and (6.46) is an equality by (6.22).

Proof of Proposition 6.10. Step 1: smoothness, and directional derivatives.

Let $|m| < 1$ and $0 < q_0 < \dots < q_\ell < T$. On the open cone of such heights, $\Phi_{\nu_q}(0, y)$ is a smooth function of $(q_0, \dots, q_\ell, y)$, as a finite composition of the operators (6.3), which depend smoothly on their parameter a ; and $y(\nu_q, m)$ is a smooth function of the heights by the implicit function theorem, since $\partial_x^2 \Phi_{\nu_q} > 0$. Hence $g_m(q) = T - \Phi_{\nu_q}(0, y(\nu_q, m)) + my(\nu_q, m)$ is smooth on this cone.

We now compute directional derivatives. Let $a \in \mathcal{Q}_\infty$, and let γ be a bounded measurable function on $[0, 1]$ such that $a_\theta := a + \theta\gamma$ belongs to $\mathcal{Q}_\infty$ for all small $\theta > 0$, with uniform norms bounded by some T' , and let $\nu_\theta := \nu_{a_\theta} \in \mathcal{P}([0, T'])$. We claim that

$$(6.47) \quad \lim_{\theta \downarrow 0} \frac{g_m(a_\theta) - g_m(a)}{\theta} = \int_0^1 \chi'_{\nu_0, m}(a(r)) \gamma(r) dr.$$

For every potential χ , we have $\int \chi d(\nu_\theta - \nu_0) = \int_0^1 [\chi(a_\theta(r)) - \chi(a(r))] dr$, so the supporting inequality (6.30), applied at ν_θ and at ν_0 , gives

$$\int_0^1 [\chi_{\nu_\theta, m}(a_\theta(r)) - \chi_{\nu_\theta, m}(a(r))] dr \leq g_m(a_\theta) - g_m(a) \leq \int_0^1 [\chi_{\nu_0, m}(a_\theta(r)) - \chi_{\nu_0, m}(a(r))] dr.$$

Each integrand is $\theta\gamma(r)$ times the average of the derivative of the corresponding potential over the segment between $a(r)$ and $a_\theta(r)$. Since $\int_0^{T'} |\zeta_{\nu_\theta} - \zeta_{\nu_0}| = W_1(\nu_\theta, \nu_0) \leq \theta \|\gamma\|_{L^1}$, Corollary 6.8 gives $\chi'_{\nu_\theta, m} \rightarrow \chi'_{\nu_0, m}$ uniformly as $\theta \downarrow 0$, and $\chi'_{\nu_0, m}$ is continuous. Dividing by θ and letting $\theta \downarrow 0$, both bounds converge to the right-hand side of (6.47), by bounded convergence.

For the finite-step path q with distinct heights and $\gamma = \mathbb{1}_{[r_i, r_{i+1}]}$, the paths a_θ are increasing for small θ , and $\chi'_{\nu_q, m}(q_i) = \mathbb{E}_m[M_{q_i}^2]$ by Corollary 6.8, so (6.47) proves (6.44). For general $a \in \mathcal{Q}_\infty$, the function $v \mapsto \mathbb{E}[(M_v^{\nu_a, m})^2] = \chi'_{\nu_a, m}(v)$ is increasing with values in $[m^2, 1]$, by the martingale property and Jensen's inequality, and composing it with the increasing path a gives the properties of z_a stated in (6.45). Taking $\gamma = b - a$ with $b \in \mathcal{Q}_\infty$, for which a_θ is a convex combination of a and b , (6.47) says that $\int_0^1 z_a(b - a)$ is the right derivative of $\theta \mapsto g_m(a + \theta(b - a))$ at $\theta = 0$.

Step 2: off-diagonal entries of the Hessian.

Let $K_{ij} := \partial_{q_i} \partial_{q_j} g_m(q)$ and $U_i(q, z) := \mathbb{E}_z[M_{q_i}^2]$, so that $K_{ij} = w_i \partial_{q_j} U_i(q, m)$ by Step 1. Let $j \neq i$. Increasing q_j lowers ζ_{ν_q} on the interval swept by the atom, and leaves the observation time q_i unchanged. By Lemma 6.9(2), the diffusion coefficient A decreases, and then, by Lemma 6.9(3), $\mathbb{E}_m[\varphi(M_{q_i})]$ decreases for the convex function $\varphi(z) = z^2$. Hence $\partial_{q_j} U_i(q, m) \leq 0$, that is, $K_{ij} \leq 0$ for all $j \neq i$.

Step 3: row sums of the Hessian.

Let $0 < \varepsilon < T - q_\ell$, and consider the path $q + \varepsilon \mathbf{1}$, with all heights shifted by ε . We write $\tilde{\Phi}$ and $\tilde{A}$ for the associated objects, and $\tilde{M}$ for its magnetization process started at time 0 from m . The cumulative function of $\nu_{q+\varepsilon \mathbf{1}}$ is $\zeta_{\nu_q}(\cdot - \varepsilon)$ on $[\varepsilon, T]$, and $\Phi(T - \varepsilon, \cdot) = \log \cosh + \varepsilon$ by (6.6), since $T - \varepsilon > q_\ell$. Since the operators $T_{r, a}$ commute with the addition of constants, the recursion (6.4) gives $\tilde{\Phi}(v, x) = \Phi(v - \varepsilon, x) - \varepsilon$ for $v \in [\varepsilon, T]$, so that $\tilde{A}(v, z) = A(v - \varepsilon, z)$. By (6.7) and (6.35), it follows that for every $z \in (-1, 1)$, the processes $(\tilde{M}_{\varepsilon+r}^{\varepsilon, z})_{r \in [0, T-\varepsilon]}$ and $(M_r^{0, z})_{r \in [0, T-\varepsilon]}$ have the same law, where $M^{0, z}$ is the magnetization process for q . The

i -th height of $q + \varepsilon \mathbf{1}$ is $q_i + \varepsilon$, so by the Markov property, the convexity in the starting point of Lemma 6.9(1), and Jensen's inequality for the martingale $\widetilde{M}$,

$$U_i(q + \varepsilon \mathbf{1}, m) = \mathbb{E}[\widetilde{M}_{q_i + \varepsilon}^2] = \mathbb{E}[U_i(q, \widetilde{M}_\varepsilon)] \geq U_i(q, \mathbb{E}\widetilde{M}_\varepsilon) = U_i(q, m).$$

Differentiating at $\varepsilon = 0$ gives $\sum_j \partial_{q_j} U_i(q, m) \geq 0$, that is, $\sum_j K_{ij} \geq 0$.

Step 4: conclusion.

For every real vector $\theta = (\theta_0, \dots, \theta_\ell)$,

$$\theta^\top K \theta = \sum_i \left(\sum_j K_{ij} \right) \theta_i^2 + \sum_{i < j} (-K_{ij}) (\theta_i - \theta_j)^2 \geq 0,$$

by Steps 2 and 3. This proves that g_m is convex on the open cone of distinct positive heights. The Lipschitz bound (6.43) is the W_1 bound of Lemma 6.7, since $W_1(\nu_q, \nu_{q'}) = \|q - q'\|_{L^1}$. Convexity extends by continuity to coincident or zero heights, hence to all finite-step paths, any two of which share a common partition, and then by density and (6.43) to $\mathcal{Q}_1$. For $m = \pm 1$, the function g_m is affine by (6.22).

Step 5: the supporting inequality.

Let $|m| < 1$ and $a, b \in \mathcal{Q}_\infty$. Since g_m is convex on $\mathcal{Q}_1$, the function $\theta \mapsto g_m(a + \theta(b - a))$ is convex on $[0, 1]$, so it lies above its tangent line at $\theta = 0$, which by Step 1 is $\theta \mapsto g_m(a) + \theta \int_0^1 z_a(b - a)$. Taking $\theta = 1$ gives (6.46). $\square$

We now return to the multi-species setting. Recall from (1.9) that, for $m \in [-1, 1]^\mathcal{S}$ and $q \in \mathcal{Q}_1^\mathcal{S}$, we have $\psi_m(q) = \sum_s \lambda_s g_{m_s}(q_s)$. We write L^2 for $L^2([0, 1]; \mathbb{R}^\mathcal{S})$, and $\langle \cdot, \cdot \rangle_{L^2}$ for its inner product. Let $(\mathcal{Q}_2^\mathcal{S})^*$ denote the dual cone of $\mathcal{Q}_2^\mathcal{S}$ in L^2 , that is, the set of $\kappa \in L^2$ with $\langle \kappa, q \rangle_{L^2} \geq 0$ for every $q \in \mathcal{Q}_2^\mathcal{S}$. It has the concrete description [11, Lemma 3.4]: for $\kappa \in L^2$,

$$(6.48) \quad \kappa \in (\mathcal{Q}_2^\mathcal{S})^* \iff \int_r^1 \kappa_s(u) du \geq 0 \quad \text{for every } s \in \mathcal{S} \text{ and } r \in [0, 1].$$

Indeed, the condition is necessary by testing against the paths $\mathbf{1}_{[r, 1]} e_s \in \mathcal{Q}_2^\mathcal{S}$, and it is sufficient because every finite-step path in $\mathcal{Q}_2^\mathcal{S}$ is a combination of such paths with nonnegative coefficients, and these are dense. In particular, $\mathcal{Q}_2^\mathcal{S} \subseteq (\mathcal{Q}_2^\mathcal{S})^*$. We say that a function u on $\mathcal{Q}_2^\mathcal{S}$ is $(\mathcal{Q}_2^\mathcal{S})^*$ -increasing if $u(q') \geq u(q)$ whenever $q' - q \in (\mathcal{Q}_2^\mathcal{S})^*$. For every $a \geq 0$, we write $\mathcal{Q}_{\infty, \leq a}$ for the set of paths in $\mathcal{Q}_\infty$ that remain bounded by a , and we set

$$(6.49) \quad \mathcal{Q}_{\infty, \leq \lambda}^\mathcal{S} := \prod_{s \in \mathcal{S}} \mathcal{Q}_{\infty, \leq \lambda_s},$$

that is, the set of $p \in \mathcal{Q}_\infty^\mathcal{S}$ such that $0 \leq p_s \leq \lambda_s$ for every $s \in \mathcal{S}$.

Corollary 6.11 (The slice initial condition). *For every $m \in [-1, 1]^\mathcal{S}$, the function ψ_m is convex and $(\mathcal{Q}_2^\mathcal{S})^*$ -increasing on $\mathcal{Q}_2^\mathcal{S}$, and*

$$(6.50) \quad |\psi_m(q) - \psi_m(q')| \leq \sum_s \lambda_s \|q_s - q'_s\|_{L^1} \leq \|q - q'\|_{L^1}, \quad q, q' \in \mathcal{Q}_1^\mathcal{S}.$$

For every $q \in \mathcal{Q}_\infty^\mathcal{S}$, the path $p := (\lambda_s z_{q_s})_{s \in \mathcal{S}}$, where z_{q_s} is the path (6.45) for the magnetization m_s , belongs to $\mathcal{Q}_{\infty, \leq \lambda}^\mathcal{S}$ and supports ψ_m at q :

$$(6.51) \quad \psi_m(q') \geq \psi_m(q) + \langle q' - q, p \rangle_{L^2}, \quad q' \in \mathcal{Q}_\infty^\mathcal{S}.$$

Finally, let $\mathcal{I}_\lambda(m) := \sum_s \lambda_s I(m_s)$. For every $q \in \mathcal{Q}_1^\mathcal{S}$, the function $m \mapsto \psi_m(q) - \mathcal{I}_\lambda(m)$ is convex on $[-1, 1]^\mathcal{S}$, and

$$(6.52) \quad |\psi_m(q) - \psi_{m'}(q) - \mathcal{I}_\lambda(m) + \mathcal{I}_\lambda(m')| \leq 2 \sum_s \lambda_s \|q_s\|_{L^1} |m_s - m'_s|, \quad m, m' \in [-1, 1]^\mathcal{S};$$

in particular, $m \mapsto \psi_m(q)$ is continuous on $[-1, 1]^{\mathcal{S}}$.

Proof. Convexity and the Lipschitz bound follow from Proposition 6.10, and so does (6.51), by summing the supporting inequalities (6.46) with the weights λ_s . For the monotonicity, let $q, q' \in \mathcal{Q}_\infty^{\mathcal{S}}$ with $q' - q \in (\mathcal{Q}_2^{\mathcal{S}})^*$. Since p belongs to $\mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}} \subseteq \mathcal{Q}_2^{\mathcal{S}}$, we have $\langle q' - q, p \rangle_{L^2} \geq 0$, and (6.51) gives $\psi_m(q') \geq \psi_m(q)$. For general $q, q' \in \mathcal{Q}_2^{\mathcal{S}}$ with $q' - q \in (\mathcal{Q}_2^{\mathcal{S}})^*$, we approximate q and q' in L^1 by their averages over the dyadic intervals of length 2^{-n} , which are bounded paths; this preserves the relation $q' - q \in (\mathcal{Q}_2^{\mathcal{S}})^*$, by (6.48), and we conclude by (6.50). The statements in m follow from Lemma 6.7 for bounded paths, since $\int v \nu_{q_s}(dv) = \|q_s\|_{L^1}$, and they extend to $\mathcal{Q}_1^{\mathcal{S}}$ by (6.50), since the right-hand side of (6.52) is continuous in q for the L^1 norm. $\square$

7. THE INITIAL CONDITION ON THE SLICES

In this section, we identify the initial condition of the constrained models: we show that $\bar{F}_{N,k}(0, \cdot)$ is close to $\psi_{m_{N,k}}$ as N tends to infinity, uniformly over the slices. We first show that, for every $t \geq 0$, the free energy of the full model is, up to a vanishing error, the minimum of the slice free energies, each corrected by the contribution of the external field. This relation will be used again in Section 10 to identify the limit free energy of the full model; here, we only need it at $t = 0$. At $t = 0$, the free energies of the full model and of the slices factorize over the species, which reduces the problem to a single species, and the free energy of the full model with external field $h \in \mathbb{R}$ is then exactly $\psi_\circ(\cdot; h)$. The relation above thus states that $h \mapsto -\psi_\circ(\cdot; h)$ is, up to a small error, the convex conjugate of the slice free energies, viewed as functions of the magnetization; by Lemma 6.7, it is also the convex conjugate of $m \mapsto g_m$. A priori, a convex conjugate only determines the convex envelope of a function, but here the strong convexity (and continuity) of g_m in m allows us to conclude.

7.1. Gluing the slices. Recall that $\bar{F}_N$ denotes the free energy of the full cube with the weights $w(\sigma) = 2^{-N} e^{h^{(N)} \cdot \sigma}$, where $h_i^{(N)} = h_s$ for $i \in I_{N,s}$, and that, by Lemma 4.3, its value at $q = 0$ is exactly (1.3). Let U^{Σ_N} and $U^{\Omega_{N,k}}$ denote the solutions of (2.5) for the full cube with these weights, and for the slices with the weights 2^{-N} . Since the field $h^{(N)} \cdot \sigma$ is constant on each slice, with value $N\lambda_N \cdot (hm_{N,k})$, where $hm := (h_s m_s)_s$, the terminal function of the full cube is a logarithmic sum of the slice terminal functions:

$$(7.1) \quad U^{\Sigma_N}(1, x; g) = \log \sum_k \exp(N\lambda_N \cdot (hm_{N,k}) + U^{\Omega_{N,k}}(1, x; g)).$$

Throughout this section, we write $n_{N,s} := |I_{N,s}|$ for the number of spins of species s . The number of slices is $M_N := \prod_s (n_{N,s} + 1) \leq (N+1)^{|\mathcal{S}|}$, and we set

$$(7.2) \quad a_N := \frac{|\mathcal{S}| \log(N+1)}{N} \geq \frac{\log M_N}{N}.$$

Lemma 7.1 (Gluing the slices). *For every N , $t \geq 0$, $h \in \mathbb{R}^{\mathcal{S}}$, and every $q \in \mathcal{Q}_1^{\mathcal{S}}$,*

$$(7.3) \quad 0 \leq \min_k \{ \bar{F}_{N,k}(t, q) - \lambda_N \cdot (hm_{N,k}) \} - \bar{F}_N(t, q) \leq a_N + 2\sqrt{a_N(t\xi(\mathbf{1}) + \|q\|_{L^1})}.$$

This lemma is perhaps most easily understood through the cascade representation of Appendix A, where the quenched partition function of the full model is clearly the sum of the partition functions over the slices. This would easily yield a statement in the spirit of (7.3) at the level of the quenched free energies, but then a concentration argument, uniform over the slices, would need to be supplied in order to conclude. This is delicate, since the randomness of the cascade includes its weights, which are not Gaussian, and

since the total variance of the field at a leaf is $2N\lambda_N \cdot q(1)$, which is not controlled by the L^1 norm of the path.

The proof given here avoids any concentration argument over the cascade, as we work directly with the branch equation. The terminal datum of the full cube is the logarithmic sum of those of the slices, and the lower bound is the comparison principle. For the upper bound, we would like the logarithmic sum of the slice solutions to stay above the solution of the full cube for every r . But a logarithmic sum of solutions is not a solution, and the computation (7.5) shows that it is a subsolution, which lies below the solution of the full cube and only reproduces the lower bound. We therefore use instead the tempered logarithmic sum $\rho^{-1} \log \sum_k e^{\rho V_k}$ with $\rho < 1$, which is a supersolution on $[\rho, 1]$, and on $[0, \rho)$ we add a correction which involves only the values of the path on $[0, \rho]$. The concentration inequality is then used only at the root, for g and X_0 , where it is Gaussian, to compare the tempered logarithmic sum with the maximum; the value at the root of the correction cancels the contribution of the initial field to this estimate, leaving an error controlled by the L^1 norm of the path, and the optimization over ρ produces the term of order $\sqrt{a_N}$ in (7.3).

Proof of Lemma 7.1. We first assume that q is regular, and we denote by $\mathcal{G}_N(q)$ the middle term in (7.3). Let

$$V_k(r, x) := N\lambda_N \cdot (hm_{N,k}) + U^{\Omega_{N,k}}(r, x; g), \quad V(r, x) := U^{\Sigma_N}(r, x; g).$$

By Lemma 2.7(1), the functions V_k solve the branch equation (2.5) for the path q on $[0, 1] \times \mathbb{R}^N$, and so does V ; by (7.1), their terminal data satisfy $V(1, \cdot) = \log \sum_k e^{V_k(1, \cdot)}$. With $X_0 = \sqrt{2 \operatorname{diag}(q(0))} Z$, the identity (2.13) gives

$$\begin{aligned} \bar{F}_{N,k}(t, q) - \lambda_N \cdot (hm_{N,k}) &= t\xi(\lambda_N) + \lambda_N \cdot q(1) - \frac{1}{N} \mathbb{E} V_k(0, X_0), \\ \bar{F}_N(t, q) &= t\xi(\lambda_N) + \lambda_N \cdot q(1) - \frac{1}{N} \mathbb{E} V(0, X_0), \end{aligned}$$

so that

$$(7.4) \quad \mathcal{G}_N(q) = \frac{1}{N} \left(\mathbb{E} V(0, X_0) - \max_k \mathbb{E} V_k(0, X_0) \right).$$

Step 1: a logarithmic sum of solutions.

Fix $\rho \in (0, 1]$, and let $W := \rho^{-1} \log \sum_k e^{\rho V_k}$ and $\pi_k := e^{\rho V_k} / \sum_j e^{\rho V_j}$, so that $\sum_k \pi_k = 1$. Then $\partial_r W = \sum_k \pi_k \partial_r V_k$, $\nabla W = \sum_k \pi_k \nabla V_k$, and

$$\nabla^2 W = \sum_k \pi_k \nabla^2 V_k + \rho \sum_k \pi_k (\nabla V_k - \nabla W)(\nabla V_k - \nabla W)^\top.$$

Since each V_k solves the branch equation, we have $\mathcal{L}[V_k] = 0$, where $\mathcal{L}$ is the operator defined in (2.6). Using this and the identity $\sum_k \pi_k \nabla V_k \cdot D \nabla V_k - \nabla W \cdot D \nabla W = \sum_k \pi_k (\nabla V_k - \nabla W) \cdot D (\nabla V_k - \nabla W)$, we find

$$(7.5) \quad \mathcal{L}[W] = (\rho - r) \mathcal{V}, \quad \mathcal{V} := \sum_k \pi_k (\nabla V_k - \nabla W) \cdot D (\nabla V_k - \nabla W).$$

The matrix $D(r)$ is diagonal with entries $\dot{q}_s(r)$ on $I_{N,s}$, and $|\partial_{x_i} V_k| \leq 1$ for every i and k , so that the variance of $\partial_{x_i} V_k$ under the weights π_k is at most 1, and

$$0 \leq \mathcal{V} \leq \operatorname{tr} D(r) = N\lambda_N \cdot \dot{q}(r).$$

The sign of $\mathcal{L}[W]$ is favorable only for $r \geq \rho$. For $r < \rho$, we compensate it by the function $c(r) := N \int_r^1 (\rho - u)_+ \lambda_N \cdot \dot{q}(u) du$, which satisfies $\mathcal{L}[W + c] = (\rho - r) \mathcal{V} - (\rho - r)_+ N\lambda_N \cdot \dot{q}(r) \leq 0$ on $[0, 1] \times \mathbb{R}^N$. The function $W + c$ is continuous with bounded gradient and classical on each piece of q . At $r = 1$, we have $W(1, \cdot) = \rho^{-1} \log \sum_k e^{\rho V_k(1, \cdot)} \geq \log \sum_k e^{V_k(1, \cdot)} = V(1, \cdot)$,

because $p \mapsto (\sum_k a_k^p)^{1/p}$ is decreasing for positive a_k and $\rho \leq 1$. Lemma 2.3, applied to $W_1 = W + c$ and $W_2 = V$, gives $V \leq W + c$; and since $V_k(1, \cdot) \leq V(1, \cdot)$, it also gives $V_k \leq V$ for every k . Hence, for every $x \in \mathbb{R}^N$,

$$(7.6) \quad \max_k V_k(0, x) \leq V(0, x) \leq \frac{1}{\rho} \log \sum_k e^{\rho V_k(0, x)} + c(0),$$

where an integration by parts gives $c(0) = N \left(\int_0^\rho \lambda_N \cdot q(u) du - \rho \lambda_N \cdot q(0) \right)$.

Step 2: Gaussian concentration at the root.

We evaluate (7.6) at $x = X_0$ and take expectations. The lower bound and (7.4) give $\mathcal{G}_N(q) \geq 0$. For the upper bound, each $Y_k := V_k(0, X_0)$ is a function of the standard Gaussian vector (g, Z) . By (3.7), applied to the branch of the slice $\Omega_{N,k}$ started from $(0, x)$, the gradient of $U^{\Omega_{N,k}}(0, x; g)$ in g is $\sqrt{2t}$ times a conditional mean of v_σ , hence of norm at most $\sqrt{2tN\xi(\lambda_N)}$ by (2.1); and the gradient of $V_k(0, X_0)$ in Z is $\sqrt{2 \operatorname{diag}(q(0))} \nabla_x V_k$, of squared norm at most $2 \sum_s n_{N,s} q_s(0) = 2N \lambda_N \cdot q(0)$, since $|\partial_{x_i} V_k| \leq 1$. Thus Y_k is a Lipschitz function of (g, Z) with constant $\Lambda := \sqrt{2N(t\xi(\lambda_N) + \lambda_N \cdot q(0))}$, and the Gaussian concentration inequality, in its exponential form $\mathbb{E} e^{\rho(Y_k - \mathbb{E} Y_k)} \leq e^{\rho^2 \Lambda^2/2}$, together with Jensen's inequality, gives

$$\mathbb{E} \frac{1}{\rho} \log \sum_k e^{\rho Y_k} \leq \frac{1}{\rho} \log \sum_k \mathbb{E} e^{\rho Y_k} \leq \max_k \mathbb{E} Y_k + \frac{\log M_N}{\rho} + \rho N (t\xi(\lambda_N) + \lambda_N \cdot q(0)).$$

We insert this into (7.6) and divide by N . The term $\rho \lambda_N \cdot q(0)$ cancels against $c(0)/N$, and (7.4) and (7.2) yield

$$(7.7) \quad \mathcal{G}_N(q) \leq \frac{a_N}{\rho} + \rho t\xi(\lambda_N) + \int_0^\rho \lambda_N \cdot q(u) du \leq \frac{a_N}{\rho} + \rho (t\xi(\mathbf{1}) + \|q\|_{L^1}),$$

where we used that $\xi(\lambda_N) \leq \xi(\mathbf{1})$, and that the average of the increasing function $\lambda_N \cdot q$ over $[0, \rho]$ is at most its average over $[0, 1]$, which is $\sum_s \lambda_{N,s} \|q_s\|_{L^1} \leq \|q\|_{L^1}$.

Step 3: choice of ρ , and general paths.

Let $C := t\xi(\mathbf{1}) + \|q\|_{L^1}$. If $C \leq a_N$, we take $\rho = 1$ in (7.7); otherwise, we take $\rho = \sqrt{a_N/C}$. In both cases, $\mathcal{G}_N(q) \leq a_N + 2\sqrt{a_N C}$, which is (7.3) for regular paths. For $q \in \mathcal{Q}_1^\mathcal{S}$, let q^n be regular paths with $\|q^n - q\|_{L^1} \rightarrow 0$, which exist by the density noted in the proof of Proposition 4.2. All the free energies in (7.3) satisfy the Lipschitz bound (4.12), and so does their minimum over k , so that $|\mathcal{G}_N(q) - \mathcal{G}_N(q^n)| \leq 2\|q^n - q\|_{L^1}$; since $\|q^n\|_{L^1} \rightarrow \|q\|_{L^1}$, (7.3) for q^n gives (7.3) for q . $\square$

7.2. Factorization at $t = 0$. At $t = 0$, the branch equation separates over blocks of coordinates. We make this precise, and we use it to express the initial conditions of the full cube and of the slices through one-species quantities. For $n \geq 1$, $0 \leq k \leq n$ and $a \in \mathcal{Q}_1$, let

$$(7.8) \quad G_{n,k}(a) := \bar{F}_{n,k}(0, a) \quad \text{for the system with a single species and } n \text{ sites,}$$

that is, the quantity (2.10) at $t = 0$, extended to $\mathcal{Q}_1$ by Proposition 4.2, with $\mathcal{S}$ a singleton, $N = n$, $\lambda_N = 1$, and $\Omega = \Omega_{n,k} = \{\sigma \in \{-1, 1\}^n : |\{i : \sigma_i = 1\}| = k\}$ with the weights 2^{-n} . We write $m_k := 2k/n - 1$ for the magnetization of this slice. Recall that $\psi_\circ$ is defined in (6.14).

Lemma 7.2 (Factorization at $t = 0$). *For every N , $q \in \mathcal{Q}_1^\mathcal{S}$, $h \in \mathbb{R}^\mathcal{S}$ and k , we have*

$$(7.9) \quad \bar{F}_N(0, q) = \sum_s \lambda_{N,s} \psi_\circ(q_s; h_s),$$

$$(7.10) \quad \bar{F}_{N,k}(0, q) = \sum_s \lambda_{N,s} G_{n_{N,s}, k_s}(q_s).$$

Moreover, for every $n \geq 1$ and $a \in \mathcal{Q}_1$,

$$(7.11) \quad G_{n,0}(a) = G_{n,n}(a) = \log 2 + \int_0^1 a(r) dr.$$

Proof. By the Lipschitz bound (4.12), by Lemma 6.3 and by the density of regular paths, it suffices to consider a regular path q . At $t = 0$, the slice $\Omega_{N,k}$ is the product over s of the one-species slices $\Omega_{n_{N,s},k_s}$ in the coordinates $I_{N,s}$, and the terminal function (2.2) is

$$U(1, x) = \sum_s \log \sum_{\sigma_s \in \Omega_{n_{N,s},k_s}} 2^{-n_{N,s}} \exp(x_{I_{N,s}} \cdot \sigma_s),$$

the sum over the blocks $I_{N,s}$ of the terminal functions of the one-species slices. By Lemma 2.7(2), $U(0, x) = \sum_s U_s(0, x_{I_{N,s}})$, where U_s is the solution of the branch equation for the one-species slice $\Omega_{n_{N,s},k_s}$ and the path q_s ; moreover the blocks $X_{0,I_{N,s}} = \sqrt{2q_s(0)} Z_{I_{N,s}}$ of the initial field are independent, each with the law of the initial field of the corresponding one-species system. Hence $\mathbb{E}U(0, X_0)$ is the sum over s of the corresponding one-species quantities, and (7.10) follows from (2.13), since $\lambda_N \cdot q(1) = \sum_s \lambda_{N,s} q_s(1)$ and $N^{-1} = \lambda_{N,s} n_{N,s}^{-1}$.

For the full cube with the weights $2^{-N} e^{h^{(N)} \cdot \sigma}$, we take the sites as blocks. The terminal function is $U^{\Sigma_N}(1, x) = \sum_s \sum_{i \in I_{N,s}} \log \cosh(x_i + h_s)$. By Lemma 2.7(1) and (2), $U^{\Sigma_N}(0, x) = \sum_s \sum_{i \in I_{N,s}} u_s(0, x_i + h_s)$, where u_s is the solution of the one-spin branch equation (6.16) for the path q_s , and the coordinates of X_0 are independent with $X_{0,i} = \sqrt{2q_s(0)} Z_i$ for $i \in I_{N,s}$. By Lemma 6.5, $\mathbb{E}u_s(0, h_s + \sqrt{2q_s(0)} Z) = q_s(1) - \psi_\circ(q_s; h_s)$, hence

$$\mathbb{E}U^{\Sigma_N}(0, X_0) = \sum_s n_{N,s} [q_s(1) - \psi_\circ(q_s; h_s)],$$

and (7.9) follows from (2.13).

Finally, the slice $\Omega_{n,n}$ consists of the single configuration $\mathbf{1}$, so that the overlap $R_n(\sigma^1, \sigma^2) = \frac{1}{n} \sigma^1 \cdot \sigma^2$ of two replicas is identically 1. For a regular path a and $\theta > 0$, the path θa is regular, and $\theta \mapsto \theta a$ is an admissible variation with $h = a$, so Theorem 4.1 gives

$$\frac{d}{d\theta} G_{n,n}(\theta a) = \int_0^1 a(r) \mathbb{E}\langle R_n \rangle_r dr = \int_0^1 a(r) dr \quad (\theta > 0).$$

The function $\theta \mapsto G_{n,n}(\theta a)$ is continuous on $[0, \infty)$ by (4.12), and $G_{n,n}(0) = -\frac{1}{n} \log 2^{-n} = \log 2$ by Lemma 4.3. Integrating over $[0, 1]$ proves (7.11) for $k = n$ and regular a ; the slice $\Omega_{n,0} = \{-1\}$ is treated identically, and the general case follows by (4.12). $\square$

7.3. Continuity in the magnetization after subtracting the entropy. We next compare two slices of the one-species system. Near $m = \pm 1$, the slice free energies depend on the magnetization in a singular way, through the entropy of the slice, and the comparison is made after subtracting this entropy. We will use a coupling argument between the two slices that incorporates the appropriate entropy correction, and proceed by interpolation. The interpolation goes through a system with $2n$ coordinates in which, at the endpoints, the path of one of the two blocks vanishes; we first identify the free energy of such a system.

Lemma 7.3 (A block with zero path). *Let $t = 0$, let $\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_1$ be a partition of the species, and let $q = (q^{(0)}, 0)$, where $q^{(0)}$ is a regular path indexed by $\mathcal{S}_0$ and $q_s \equiv 0$ for $s \in \mathcal{S}_1$. Let N_0 be the number of coordinates of the species in $\mathcal{S}_0$, write $\sigma_{(0)}$ for the restriction of $\sigma \in \Omega$ to these coordinates, let $\Omega_{(0)} := \{\sigma_{(0)} : \sigma \in \Omega\}$, and let $w_{(0)}(\tau) := \sum_{\sigma \in \Omega: \sigma_{(0)} = \tau} w(\sigma)$. Then*

$$(7.12) \quad N \overline{F}_{N,\Omega}(0, q) = N_0 \overline{F}_{N_0, \Omega_{(0)}}(0, q^{(0)}),$$

where the right-hand side is the free energy of the system with the N_0 coordinates of the species in $\mathcal{S}_0$, the configuration set $\Omega_{(0)}$ and the weights $w_{(0)}$.

Proof. For $\varepsilon > 0$, let $q^\varepsilon := (q^{(0)}, \varepsilon(1+r)\mathbf{1})$, which is regular and converges to q in L^1 , so that $\bar{F}_{N,\Omega}(0, q^\varepsilon) \rightarrow \bar{F}_{N,\Omega}(0, q)$ by (4.12). Let $N_1 := N - N_0$, write $x = (x_{(0)}, x_{(1)})$ according to the two groups of species, let U^ε be the solution of (2.5) for q^ε in dimension N , and recall from (3.13) that $0 \leq \text{tr } \nabla^2 U^\varepsilon \leq N$ and $|\nabla U^\varepsilon|^2 \leq N$. Let V be the solution of the branch equation in dimension N_0 , for the path $q^{(0)}$, with the terminal datum $x_{(0)} \mapsto U^\varepsilon(1, x_{(0)}, 0) = \log \sum_{\tau \in \Omega_{(0)}} w_{(0)}(\tau) e^{x_{(0)} \cdot \tau}$, which is the terminal function of the reduced system. The function $W(r, x_{(0)}) := U^\varepsilon(r, x_{(0)}, 0)$ solves the equation of the reduced system up to the source $-\varepsilon(\text{tr } \nabla_{(1)}^2 U^\varepsilon + r|\nabla_{(1)} U^\varepsilon|^2)(r, x_{(0)}, 0)$, which is bounded in absolute value by $2\varepsilon N$, so Lemma 2.3 gives $|W - V| \leq 2\varepsilon N$. Moreover $X_0^\varepsilon = (X_{0,(0)}, \sqrt{2\varepsilon}Z_{(1)})$, with $X_{0,(0)} \sim \mathcal{N}(0, 2\text{diag}(q^{(0)}(0)))$ the initial field of the reduced system and $Z_{(1)}$ a standard Gaussian vector in $\mathbb{R}^{N_1}$, so that (4.13), applied in the second group of coordinates, gives $0 \leq \mathbb{E} U^\varepsilon(0, X_0^\varepsilon; g) - \mathbb{E} W(0, X_{0,(0)}) \leq N\varepsilon$. By (2.13), since $N\lambda_{N,s} = n_{N,s}$,

$$\begin{aligned} N\bar{F}_{N,\Omega}(0, q^\varepsilon) &= \sum_{s \in \mathcal{S}_0} n_{N,s} q_s(1) + 2\varepsilon N_1 - \mathbb{E} U^\varepsilon(0, X_0^\varepsilon; g), \\ N_0 \bar{F}_{N_0, \Omega_{(0)}}(0, q^{(0)}) &= \sum_{s \in \mathcal{S}_0} n_{N,s} q_s(1) - \mathbb{E} V(0, X_{0,(0)}), \end{aligned}$$

and letting $\varepsilon \rightarrow 0$ proves (7.12). $\square$

Lemma 7.4 (Entropy-corrected continuity in the magnetization). *For every $n \geq 1$, $0 \leq k, k' \leq n$, and $a \in \mathcal{Q}_1$, we have*

$$(7.13) \quad |(G_{n,k}(a) - I_{n,k}) - (G_{n,k'}(a) - I_{n,k'})| \leq 2\|a\|_{L^1} |m_k - m_{k'}|,$$

where $I_{n,k} := \log 2 - \frac{1}{n} \log \binom{n}{k}$ satisfies $|I_{n,k} - I(m_k)| \leq \frac{\log(n+1)}{n}$.

Proof. We first prove the entropy estimate. Let $h(p) := -p \log p - (1-p) \log(1-p)$ for $p \in [0, 1]$, with $0 \log 0 = 0$. Writing $p = \frac{1+z}{2}$ in the definition of I , we get $I(z) = p \log(2p) + (1-p) \log(2(1-p)) = \log 2 - h(\frac{1+z}{2})$, and since $\frac{1+m_k}{2} = \frac{k}{n}$, we have

$$I_{n,k} - I(m_k) = h\left(\frac{k}{n}\right) - \frac{1}{n} \log \binom{n}{k}.$$

Let $p := k/n$ and $b_j := \binom{n}{j} p^j (1-p)^{n-j}$ for $0 \leq j \leq n$, so that $\sum_j b_j = 1$ and $b_k = \binom{n}{k} e^{-nh(p)}$. In particular, $b_k \leq 1$. If $k \in \{0, n\}$, then $b_k = 1$. Otherwise, for $0 \leq j < n$, the ratio $b_{j+1}/b_j = \frac{(n-j)k}{(j+1)(n-k)}$ is at least 1 for $j < k$ and less than 1 for $j \geq k$, since $(n-j)k - (j+1)(n-k) = n(k-1-j) + k$. So b_k is the largest of the $n+1$ numbers b_j , and $b_k \geq (n+1)^{-1}$. Thus $(n+1)^{-1} e^{nh(k/n)} \leq \binom{n}{k} \leq e^{nh(k/n)}$, that is, $0 \leq I_{n,k} - I(m_k) \leq \frac{\log(n+1)}{n}$.

We now turn to the proof of (7.13). By (4.12), both sides of this inequality are continuous in a for the L^1 norm, so it suffices to consider a regular path a .

Step 1: uniform weights, and a coupling of the two slices.

Since $2^{-n} = e^{-nI_{n,k}} \binom{n}{k}^{-1}$, and since adding a constant to the terminal datum adds it to the solution of the branch equation, by Lemma 2.7(1), the quantity $G_{n,k}(a) - I_{n,k}$ is the free energy at $t = 0$ of the slice $\Omega_{n,k}$ with the uniform weights $\binom{n}{k}^{-1}$. Let $U_1, \dots, U_n$ be independent random variables, uniformly distributed on $[0, 1]$. Let $\sigma \in \Omega_{n,k}$ be the configuration whose plus spins are the k sites with the smallest values of U_i , and define $\sigma' \in \Omega_{n,k'}$ in the same way with k' . Then σ and σ' are each uniformly distributed on their slices, and they differ at exactly $d := |k - k'|$ sites. Let $\tilde{\Omega} \subseteq \{-1, 1\}^{2n}$ be the set of

pairs (σ, σ') obtained in this way, with the weights $w(\sigma, \sigma')$ given by their probabilities. Since $\sum_{\sigma'} w(\sigma, \sigma') = \binom{n}{k}^{-1}$ for every $\sigma \in \Omega_{n,k}$, the slice $\Omega_{n,k}$ with the uniform weights is the projection of $(\tilde{\Omega}, w)$ on the first block of n coordinates, in the sense of Lemma 7.3, and $\Omega_{n,k'}$ with the uniform weights is its projection on the second block.

Step 2: a doubled system.

For $u \in [0, 1]$, consider the system with $2n$ coordinates, partitioned into two species, the first and the second block of n coordinates, with the configuration set $\tilde{\Omega}$ and the weights w , at $t = 0$, with the path $q^u := ((1-u)a, ua)$. Its free energy, normalized by $2n$ as in Definition 2.4, is $\bar{F}_{2n, \tilde{\Omega}}(0, q^u)$, and we set $G(u) := 2\bar{F}_{2n, \tilde{\Omega}}(0, q^u)$. The auxiliary field of this system has two blocks, driven by independent Brownian motions, with covariance paths $(1-u)a$ and ua ; the terminal weights couple the two blocks. Since $\lambda = (\frac{1}{2}, \frac{1}{2})$ for this system, the self-overlap correction $2\lambda \cdot q^u(1) = a(1)$ does not depend on u . The function G is continuous on $[0, 1]$ by (4.12). At $u = 0$, the second path vanishes, and Lemma 7.3 together with Step 1 gives $2n\bar{F}_{2n, \tilde{\Omega}}(0, (a, 0)) = n[G_{n,k}(a) - I_{n,k}]$, that is, $G(0) = G_{n,k}(a) - I_{n,k}$; similarly, $G(1) = G_{n,k'}(a) - I_{n,k'}$. For $u \in (0, 1)$, the path q^u is regular, and $u \mapsto q^u$ is an admissible variation with $h = (-a, a)$, so by Theorem 4.1, with the overlaps $\frac{1}{2n}\sigma^1 \cdot \sigma^2$ and $\frac{1}{2n}\sigma'^1 \cdot \sigma'^2$ of the two blocks, the function G is differentiable on $(0, 1)$, with

$$G'(u) = \frac{1}{n} \int_0^1 a(r) \mathbb{E} \langle \sigma'^1 \cdot \sigma'^2 - \sigma^1 \cdot \sigma^2 \rangle_r dr,$$

where (σ^1, σ'^1) and (σ^2, σ'^2) are two replicas at level r . Since σ^b and σ'^b differ at d sites, at most $2d$ terms of the sum defining $\sigma'^1 \cdot \sigma'^2 - \sigma^1 \cdot \sigma^2$ are nonzero, each bounded by 2. Hence $|G'(u)| \leq \frac{4d}{n} \|a\|_{L^1} = 2|m_k - m_{k'}| \|a\|_{L^1}$ on $(0, 1)$, and (7.13) follows by the continuity of G on $[0, 1]$. $\square$

7.4. Identification of the initial condition. We now identify the limit of the slice initial conditions $\bar{F}_{N,k}(0, \cdot)$ as N tends to infinity. By the factorization (7.10), this is a statement about a single species, which we prove first.

Proposition 7.5 (Initial data on one-species slices). *For every $R < \infty$, we have*

$$(7.14) \quad \lim_{n \rightarrow \infty} \sup_{0 \leq k \leq n} \sup_{\substack{a \in \mathcal{Q}_1 \\ \|a\|_{L^1} \leq R}} |G_{n,k}(a) - g_{m_k}(a)| = 0.$$

More precisely, with $\varepsilon_n(R) := \log(n+1)/n + 2\sqrt{R \log(n+1)/n}$, for every $0 \leq k \leq n$ and every $a \in \mathcal{Q}_1$ with $\|a\|_{L^1} \leq R$,

$$(7.15) \quad 0 \leq G_{n,k}(a) - g_{m_k}(a) \leq E_n(R) := \varepsilon_n(R) + 4R\sqrt{2\varepsilon_n(R)} + \frac{2\log(n+1)}{n}.$$

Proof. We first assume that a is bounded, so that $\nu_a \in \mathcal{P}([0, T])$ for some $T > 0$ and the results of Section 6 apply to it; note that $\int v \nu_a(dv) = \|a\|_{L^1} \leq R$. We write $d_k := G_{n,k}(a) - g_{m_k}(a)$, and we begin with two preliminary bounds.

Lemma 7.1 at $t = 0$, applied to the system with a single species and n sites, for which a_N becomes $\log(n+1)/n$, together with (7.9) for this system, which reads $\bar{F}_n(0, a) = \psi_\circ(a; h)$, gives

$$(7.16) \quad 0 \leq \min_{0 \leq j \leq n} \{G_{n,j}(a) - hm_j\} - \psi_\circ(a; h) \leq \varepsilon_n(R) \quad \text{for every } h \in \mathbb{R}.$$

The lower bound gives $G_{n,k}(a) \geq \psi_\circ(a; h) + hm_k$ for every k and every $h \in \mathbb{R}$; taking the supremum over h and recalling the definition (6.14) of g_m , we obtain $d_k \geq 0$ for every k .

Next, by Lemma 7.4, including its entropy estimate, and by (6.26), for all j, k ,

$$(7.17) \quad |d_k - d_j| \leq |(G_{n,k}(a) - I_{n,k}) - (G_{n,j}(a) - I_{n,j})| + \frac{2\log(n+1)}{n} \\ + |(g_{m_k}(a) - I(m_k)) - (g_{m_j}(a) - I(m_j))| \leq 4R|m_k - m_j| + \frac{2\log(n+1)}{n}.$$

By (6.28), the upper bound in (7.16) says that the Legendre transforms in the magnetization of $k \mapsto G_{n,k}(a)$ and of $m \mapsto g_m(a)$ nearly coincide. A Legendre transform identifies in general only a convex envelope; here, the strong convexity of g_m in m localizes the minimizing slice. Let $0 < k < n$, so that $|m_k| < 1$, let $h := y(\nu_a, m_k)$, so that $\psi_\circ(a; h) = g_{m_k}(a) - hm_k$ by (6.28), and let j attain the minimum in (7.16). The upper bound there reads

$$d_j + [g_{m_j}(a) - g_{m_k}(a) - h(m_j - m_k)] \leq \varepsilon_n(R),$$

and by the strong convexity (6.27) at $m_0 = m_k$, where $y(\nu_a, m_k) = h$, the bracket is at least $\frac{1}{2}(m_j - m_k)^2$. Since $d_j \geq 0$, we obtain

$$(7.18) \quad d_j + \frac{1}{2}(m_j - m_k)^2 \leq \varepsilon_n(R) :$$

the minimizing slice has both a small error and a magnetization close to m_k . Transporting its error to the slice k by (7.17),

$$d_k \leq d_j + 4R|m_j - m_k| + \frac{2\log(n+1)}{n} \leq \varepsilon_n(R) + 4R\sqrt{2\varepsilon_n(R)} + \frac{2\log(n+1)}{n} = E_n(R).$$

For $k \in \{0, n\}$, we have $m_k = \pm 1$, and (7.11) and (6.22) give $d_0 = d_n = 0$. This proves (7.15) for bounded a .

Let finally $a \in \mathcal{Q}_1$ with $\|a\|_{L^1} \leq R$, and let $a^M := a \wedge M$ for $M > 0$. Then a^M is bounded, $\|a^M\|_{L^1} \leq R$, and $\|a^M - a\|_{L^1} \rightarrow 0$ as $M \rightarrow \infty$. Both $G_{n,k}$ and g_{m_k} are 1-Lipschitz for the L^1 norm, by (4.12) and (6.43), so d_k at a is the limit of d_k at a^M , which lies in $[0, E_n(R)]$. Since $E_n(R) \rightarrow 0$ as $n \rightarrow \infty$, this also proves (7.14). $\square$

We now return to the system with several species, and compare $\bar{F}_{N,k}(0, \cdot)$ with the function $\psi_m(q) = \sum_s \lambda_s g_{m_s}(q_s)$ of (1.9), evaluated at the magnetization $m = m_{N,k}$ of the slice.

Corollary 7.6 (Initial data on slices). *For every $R < \infty$, we have*

$$(7.19) \quad \lim_{N \rightarrow \infty} \sup_k \sup_{\substack{q \in \mathcal{Q}_1^{\mathcal{S}} \\ \|q\|_{L^1} \leq R}} |\bar{F}_{N,k}(0, q) - \psi_{m_{N,k}}(q)| = 0.$$

Proof. By (7.10) and the definition (1.9) of ψ_m ,

$$\bar{F}_{N,k}(0, q) - \psi_{m_{N,k}}(q) = \sum_s \lambda_{N,s} [G_{n_{N,s},k_s}(q_s) - g_{m_{N,k,s}}(q_s)] + \sum_s (\lambda_{N,s} - \lambda_s) g_{m_{N,k,s}}(q_s).$$

Since $\|q_s\|_{L^1} \leq \|q\|_{L^1} \leq R$ for every s , the first sum is bounded in absolute value by $\sum_s \lambda_{N,s} E_{n_{N,s}}(R)$, by Proposition 7.5, and this tends to zero because $n_{N,s} = N\lambda_{N,s} \rightarrow \infty$ for every s . For the second, the Lipschitz bounds of Lemma 6.7 and Proposition 6.10 give $|g_m(q_s)| \leq g_m(0) + \|q_s\|_{L^1} \leq \log 2 + R$, so the second sum is at most $(\log 2 + R) \sum_s |\lambda_{N,s} - \lambda_s|$, which tends to zero because $\lambda_N \rightarrow \lambda$. $\square$

8. CONVERGENCE OF THE SLICE FREE ENERGIES

In this section, we prove assertions (1.10)–(1.12) of Theorem 1.2: the slice free energies converge, uniformly over the slices, to the viscosity solution f_m of the Hamilton–Jacobi equation (1.11) with initial condition ψ_m , and f_m is given by the Hopf formula. By the remainder estimate (5.2), the enriched free energy is an approximate solution of the equation, and by Corollary 7.6, its initial condition is close to $\psi_{m_{N,k}}$ on bounded sets of paths. A comparison principle with error terms, which also expresses the finite speed of propagation of the equation, turns these two facts into an explicit convergence estimate. We start by recalling the analytic framework.

8.1. The equation on the closed cone. We view $\mathcal{Q}_2^\mathcal{S}$ as a closed convex cone in the Hilbert space $L^2 = L^2([0, 1]; \mathbb{R}^\mathcal{S})$, with dual cone $(\mathcal{Q}_2^\mathcal{S})^*$ described in (6.48). This cone has empty interior in L^2 , which is the source of the technical difficulties in the definition of viscosity solutions; we only recall the outcome of [11], in the form used in [7, Section 7]. The solution on $\mathcal{Q}_2^\mathcal{S}$ is constructed there as a limit of solutions on finite-dimensional cones, which we also use. Let L be a power of two, let $J_1, \dots, J_L$ be the dyadic intervals of length $1/L$ partitioning $[0, 1)$, and let

$$(8.1) \quad \mathbf{C}_L := \{x \in (\mathbb{R}_+^\mathcal{S})^L : x_1 \leq \dots \leq x_L\}, \quad \langle x, y \rangle_L := \frac{1}{L} \sum_{j=1}^L x_j \cdot y_j, \quad |x|_L^2 := \langle x, x \rangle_L,$$

which is a closed convex cone in $(\mathbb{R}^\mathcal{S})^L$, with dual cone $\mathbf{C}_L^*$. For $x \in (\mathbb{R}^\mathcal{S})^L$ and $p \in L^2$, let

$$(8.2) \quad \iota_L x := \sum_j x_j \mathbf{1}_{J_j}, \quad \pi_L p := \left(L \int_{J_j} p(r) \, dr \right)_{j \leq L}$$

be the step path with values x_j , and the vector of cell averages. These maps are adjoint to each other, $\langle \iota_L x, p \rangle_{L^2} = \langle x, \pi_L p \rangle_L$; the map ι_L is an isometry, and it sends $\mathbf{C}_L$ into $\mathcal{Q}_2^\mathcal{S}$; the map π_L sends $\mathcal{Q}_2^\mathcal{S}$ into $\mathbf{C}_L$, with $|\pi_L p|_L \leq \|p\|_{L^2}$ by Jensen’s inequality; and ι_L sends $\mathbf{C}_L^*$ into $(\mathcal{Q}_2^\mathcal{S})^*$, since $\langle \iota_L z, p \rangle_{L^2} = \langle z, \pi_L p \rangle_L \geq 0$ for $z \in \mathbf{C}_L^*$ and $p \in \mathcal{Q}_2^\mathcal{S}$.

The function ξ is proper on $\mathbb{R}_+^\mathcal{S}$, in the sense that it is coordinatewise increasing and so is $x \mapsto \xi(x + y) - \xi(x)$ for every $y \in \mathbb{R}_+^\mathcal{S}$; this follows from the nonnegativity of its power-series coefficients. It is locally but not globally Lipschitz. By [11, Lemma 4.4], there exists a Lipschitz, proper function $\bar{\xi} : \mathbb{R}_+^\mathcal{S} \rightarrow \mathbb{R}$ that agrees with ξ on the positive part of the closed unit ball; we fix such a regularization and define, for every $\kappa \in L^2$,

$$(8.3) \quad \mathbf{H}(\kappa) := \inf \left\{ \int_0^1 \bar{\xi}(p(r)) \, dr : p \in \mathcal{Q}_2^\mathcal{S} \cap \left(\kappa + (\mathcal{Q}_2^\mathcal{S})^* \right) \right\}.$$

By [11, Lemma 4.6], $\mathbf{H}$ is Lipschitz and $(\mathcal{Q}_2^\mathcal{S})^*$ -increasing on L^2 , and $\mathbf{H}(p) = \int_0^1 \bar{\xi}(p(r)) \, dr$ for $p \in \mathcal{Q}_2^\mathcal{S}$; in particular, $\mathbf{H}(p) = \mathcal{H}(p)$ for $p \in \mathcal{Q}_2^\mathcal{S}$ with $|p(r)| \leq 1$ for every r , where we recall that $\mathcal{H}(p) = \int_0^1 \xi(p(r)) \, dr$ is the nonlinearity in (4.7). We denote by $L_{\mathbf{H}}$ the Lipschitz constant of $\mathbf{H}$ for the norm of L^2 , and we set $\mathbf{H}_L(x) := \mathbf{H}(\iota_L x)$ for $x \in (\mathbb{R}^\mathcal{S})^L$, which is Lipschitz with the constant $L_{\mathbf{H}}$ for the norm $|\cdot|_L$, since ι_L is an isometry, and increasing for the dual cone $\mathbf{C}_L^*$, since ι_L maps $\mathbf{C}_L^*$ into $(\mathcal{Q}_2^\mathcal{S})^*$.

The following definitions apply to both settings: the cone $C = \mathcal{Q}_2^\mathcal{S}$ in $E = L^2$, with the nonlinearity $H = \mathbf{H}$; and the cone $C = \mathbf{C}_L$ in $E = (\mathbb{R}^\mathcal{S})^L$, with the inner product $\langle \cdot, \cdot \rangle_L$ and the nonlinearity $H = \mathbf{H}_L$. We write $\|\cdot\|$ for the norm of E . We say that a function $\phi : (0, \infty) \times C \rightarrow \mathbb{R}$ is smooth if there is a unique $(\partial_t \phi, \nabla \phi)(t, x) \in \mathbb{R} \times E$ such that $\phi(t', x') - \phi(t, x) = \partial_t \phi(t, x)(t' - t) + \langle \nabla \phi(t, x), x' - x \rangle + O(|t' - t|^2 + \|x' - x\|^2)$ as

$(t', x') \rightarrow (t, x)$ in $(0, \infty) \times C$, and if this differential is continuous; this is [11, Definition 1.2]. A continuous function u on $\mathbb{R}_+ \times C$ is a viscosity subsolution of

$$(8.4) \quad \partial_t u - H(\nabla u) = 0 \quad \text{on } \mathbb{R}_+ \times C$$

if, whenever a smooth ϕ is such that $u - \phi$ has a local maximum at some $(t, x) \in (0, \infty) \times C$, we have $\partial_t \phi(t, x) - H(\nabla \phi(t, x)) \leq 0$; it is a viscosity supersolution if the reverse inequality holds at local minima; and it is a viscosity solution if it is both. These are [11, Definitions 1.4 and 4.3]. No boundary condition is imposed at the boundary of the cone: the contacts with test functions are required at all points of C , and the monotonicity of H for the dual-cone order is what makes this consistent. A function on $\mathbb{R}_+ \times C$ is Lipschitz if it is so for the norm of $\mathbb{R} \times E$. In Theorem 1.2, a viscosity solution of (1.11) means a viscosity solution of (8.4) on $\mathcal{Q}_2^\mathcal{J}$; by [11, Remark 4.10], the solutions below do not depend on the choice of the regularization $\bar{\xi}$. We use the following results from [11].

Theorem 8.1 (Well-posedness, comparison, and Hopf formula [11]). *Let $\psi : \mathcal{Q}_2^\mathcal{J} \rightarrow \mathbb{R}$ be $(\mathcal{Q}_2^\mathcal{J})^*$ -increasing and 1-Lipschitz for the L^1 norm.*

- (1) *There exists a unique Lipschitz viscosity solution f of (8.4) on $\mathcal{Q}_2^\mathcal{J}$ with $f(0, \cdot) = \psi$.*
- (2) *If u and v are respectively a Lipschitz viscosity subsolution and a Lipschitz viscosity supersolution of (8.4) on $\mathcal{Q}_2^\mathcal{J}$, and $u(0, \cdot) \leq v(0, \cdot)$ on $\mathcal{Q}_2^\mathcal{J}$, then $u \leq v$ on $\mathbb{R}_+ \times \mathcal{Q}_2^\mathcal{J}$.*
- (3) *If ψ is convex, then for every $(t, q) \in \mathbb{R}_+ \times \mathcal{Q}_2^\mathcal{J}$,*

$$(8.5) \quad f(t, q) = \sup_{p \in \mathcal{Q}_2^\mathcal{J}} \inf_{q' \in \mathcal{Q}_2^\mathcal{J}} \left\{ \psi(q') + \langle q - q', p \rangle_{L^2} + t \int_0^1 \xi(p(r)) dr \right\}.$$

- (4) *For every L that is a power of 2, the equation (8.4) on $\mathbb{C}_L$, with initial condition $x \mapsto \psi(\iota_L x)$, has a unique Lipschitz viscosity solution f_L , the comparison principle of item (2) holds on $\mathbb{C}_L$, and $f_L(t, \pi_L q) \rightarrow f(t, q)$ as $L \rightarrow \infty$, for every $(t, q) \in \mathbb{R}_+ \times \mathcal{Q}_2^\mathcal{J}$.*

Items (1) to (3) are [11, Theorem 4.7], together with [11, Proposition 3.8 and Remark 4.8] for the comparison principle, specialized to the present $\mathbb{R}^\mathcal{J}$ -valued paths as explained in [7, Section 7]. In item (4), the existence of a unique Lipschitz viscosity solution on the closed cone $\mathbb{C}_L$ and its comparison principle are [11, Theorem 2.2 and Propositions 2.3 and 2.5], applied to the Lipschitz and monotone nonlinearity H_L . The dyadic partitions form a good collection of partitions in the sense of [11, Section 3.1], and the convergence of $f_L(t, \pi_L q)$ to $f(t, q)$ is [11, Proposition 3.10]; see also [11, Theorem 4.7(1) and Remark 4.11], where the solution on the cone is constructed as the limit of these projections.

By Corollary 6.11, the initial condition ψ_m is $(\mathcal{Q}_2^\mathcal{J})^*$ -increasing, 1-Lipschitz for the L^1 norm, and convex, for every $m \in [-1, 1]^\mathcal{J}$. Theorem 8.1 therefore yields a unique Lipschitz viscosity solution of (1.11), which we take as the definition of f_m , together with its Hopf representation (1.12). The second assertion of Theorem 1.2 amounts to the existence and uniqueness of this solution, and the third to the Hopf formula; both are now established, and it remains to prove the convergence (1.10), the continuity in m , and the representation (1.13).

We will compare the free energy with the finite-dimensional solutions $f_{m,L}$ of item (4) of Theorem 8.1, although the former solves the equation only approximately, and although the two initial conditions are close only on a bounded set. The comparison principle handles both errors at once in the following form, which also expresses the finite speed of propagation of the equation.

Lemma 8.2 (Comparison with error terms). *In either of the two settings above, let v be a Lipschitz viscosity solution of (8.4), let $\varepsilon \geq 0$, and let u be a Lipschitz function on $\mathbb{R}_+ \times C$ such that $u - \varepsilon t$ is a viscosity subsolution and $u + \varepsilon t$ a viscosity supersolution of (8.4). Let $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be increasing and Lipschitz, and suppose that $|u(0, x) - v(0, x)| \leq \omega(\|x\|)$ for every $x \in C$. Then*

$$(8.6) \quad |u(t, x) - v(t, x)| \leq \varepsilon t + \omega(\|x\| + L_H t) \quad \text{for every } (t, x) \in \mathbb{R}_+ \times C.$$

Proof. We first assume that ω is smooth with bounded derivative. For $a > 0$, let $n_a(x) := (\|x\|^2 + a^2)^{1/2}$, which is smooth with $\|\nabla n_a\| \leq 1$ and $n_a \geq \|x\|$, and let

$$\Gamma(t, x) := \omega(n_a(x) + L_H t).$$

The function Γ is smooth and Lipschitz, and $\partial_t \Gamma = L_H \omega' \geq L_H \|\nabla \Gamma\|$, the derivative of ω being evaluated at $n_a(x) + L_H t$. We claim that $v + \Gamma$ is a Lipschitz viscosity supersolution. Let ϕ be smooth and such that $(v + \Gamma) - \phi$ has a local minimum at (t, x) . Then $\phi - \Gamma$ is smooth and $v - (\phi - \Gamma)$ has a local minimum at (t, x) , so $\partial_t \phi - \partial_t \Gamma - H(\nabla \phi - \nabla \Gamma) \geq 0$, and the Lipschitz continuity of H gives

$$\partial_t \phi - H(\nabla \phi) \geq \partial_t \Gamma - |H(\nabla \phi) - H(\nabla \phi - \nabla \Gamma)| \geq \partial_t \Gamma - L_H \|\nabla \Gamma\| \geq 0.$$

In the same way, $v - \Gamma$ is a Lipschitz viscosity subsolution. At $t = 0$, since ω is increasing and $n_a \geq \|x\|$, the assumption gives $u(0, \cdot) \leq v(0, \cdot) + \Gamma(0, \cdot)$ and $v(0, \cdot) - \Gamma(0, \cdot) \leq u(0, \cdot)$. The comparison principle of Theorem 8.1, applied to the pairs $(u - \varepsilon t, v + \Gamma)$ and $(v - \Gamma, u + \varepsilon t)$, gives $|u - v| \leq \varepsilon t + \Gamma$, and letting $a \rightarrow 0$ yields (8.6). For a general ω , we extend it by $\omega(0)$ to the negative half-line, and we apply the previous case to $\omega_b(s) := \int \omega(s + b - y) \eta_b(y) dy$, where $b > 0$ and η_b is a smooth probability density supported in $[-b, b]$. Each ω_b is smooth and increasing, with derivative bounded by the Lipschitz constant of ω , it satisfies $\omega_b \geq \omega$ since ω is increasing, and $\omega_b \rightarrow \omega$ pointwise as $b \rightarrow 0$; the conclusion for ω_b passes to the limit. $\square$

8.2. The free energy as an approximate solution in finite dimensions. The remainder estimate (5.2) applies to regular paths with slopes bounded below. We now show that the enriched free energy, evaluated at regular interpolations of the vectors of C_L , is an approximate viscosity solution of the finite-dimensional equation, including at the boundary of the cone. Besides the maps (8.2), we use a continuous interpolation. Let $c_j := (j - \frac{1}{2})/L$, for $j \in \{1, \dots, L\}$, be the midpoints of the dyadic intervals. For $x \in (\mathbb{R}^{\mathcal{S}})^L$, let $\mathcal{T}_L x$ be the continuous path which is piecewise affine with the values x_j at the midpoints c_j , and constant on $[0, c_1]$ and on $[c_L, 1]$. Thus

$$(8.7) \quad \mathcal{T}_L x(r) = \sum_j w_j(r) x_j, \quad w_j \geq 0, \quad \sum_j w_j = 1, \quad \int_0^1 w_j(r) dr = \frac{1}{L},$$

where w_j is supported in the union of J_j and its neighbors, and the probability densities $L w_j$ are stochastically ordered from left to right. For $p \in L^2$, let

$$(8.8) \quad A_L p := \left(L \int_0^1 w_j(r) p(r) dr \right)_{j \leq L}$$

be the vector of averages of p against these densities. The interpolation $\mathcal{T}_L$ produces, after the addition of a positive slope, the regular paths to which the derivative identities and the remainder estimate apply, and A_L , which is its adjoint, gives the gradient of the interpolated free energy. For $x \in C_L$, the path $\mathcal{T}_L x$ belongs to $\mathcal{Q}_\infty^{\mathcal{S}}$, and

$$(8.9) \quad \|\mathcal{T}_L x - \iota_L x\|_{L^1} \leq \frac{1}{4L} \sum_{j=1}^{L-1} |x_{j+1} - x_j| \leq \frac{\sqrt{|\mathcal{S}|}}{4\sqrt{L}} |x|_L \quad \text{and} \quad \|\mathcal{T}_L x\|_{L^2} \leq |x|_L,$$

the last bound being valid for every $x \in (\mathbb{R}^{\mathcal{S}})^L$. Indeed, on $[c_j, c_{j+1}]$, the two paths differ by two triangles of total area $|x_{j+1} - x_j|/(4L)$, and $\sum_j |x_{j+1} - x_j| \leq \sum_s (x_{L,s} - x_{1,s}) \leq \sqrt{|\mathcal{S}|} |x_L| \leq \sqrt{|\mathcal{S}|L} |x_L|$; and $\mathcal{T}_L x(r)$ is an average of the x_j with the weights $w_j(r)$, so Jensen's inequality gives $\|\mathcal{T}_L x\|_{L^2}^2 \leq \int_0^1 \sum_j w_j(r) |x_j|^2 dr = |x|_L^2$. Moreover, $\mathcal{T}_L \pi_L q \rightarrow q$ in L^1 as $L \rightarrow \infty$, for every $q \in \mathcal{Q}_2^{\mathcal{S}}$, since the dyadic averages $\iota_L \pi_L q$ of an integrable function converge to it in L^1 , and $|\pi_L q|_L \leq \|q\|_{L^2}$ is bounded.

Lemma 8.3 (The interpolation and its adjoint). *The map A_L sends $\mathcal{Q}_2^{\mathcal{S}}$ into $\mathbb{C}_L$, and*

$$(8.10) \quad \langle \mathcal{T}_L z, p \rangle_{L^2} = \langle z, A_L p \rangle_L \quad \text{for all } z \in (\mathbb{R}^{\mathcal{S}})^L \text{ and } p \in L^2.$$

Proof. For $p \in \mathcal{Q}_2^{\mathcal{S}}$, we have $(A_L p)_j = \int p d\mu_j$, where the probability measures $\mu_j := L w_j(r) dr$ are stochastically ordered from left to right, so $(A_L p)_j$ is nonnegative and increasing in j . The identity (8.10) is immediate from (8.7). $\square$

Proposition 8.4 (Approximate solution in finite dimensions). *Fix N , a slice $\Omega_{N,k}$, $0 < \kappa \leq 1$, and a power of two L . For $x \in \mathbb{C}_L$, let*

$$(8.11) \quad q_x^\kappa(r) := \mathcal{T}_L x(r) + \kappa(1+r)\mathbf{1}, \quad u(t, x) := \bar{F}_{N,k}(t, q_x^\kappa).$$

The function u is Lipschitz on $\mathbb{R}_+ \times \mathbb{C}_L$, and the functions $u - \varepsilon t$ and $u + \varepsilon t$ are respectively a viscosity subsolution and a viscosity supersolution of (8.4) on $\mathbb{C}_L$, where

$$(8.12) \quad \varepsilon := 15M_\xi(\kappa N)^{-1/3} + \frac{3L\xi}{L}, \quad \text{with } L_\xi := \sup_{[0,1]^{\mathcal{S}}} |\nabla \xi|,$$

and where we recall that $M_\xi = \sup_{[-1,1]^{\mathcal{S}}} \|\nabla^2 \xi\|$.

Proof. Step 1: the residual.

The shift by $\kappa(1+r)\mathbf{1}$ in (8.11) has two purposes: it supplies the positive path slope required by Theorem 5.1, and it makes u differentiable in a neighborhood of the closed cone. Indeed, for $x \in \mathbb{C}_L$, every coordinate of q_x^κ starts at $x_{1,s} + \kappa \geq \kappa$ and has slope at least κ . Moreover, q_x^κ remains a regular path, with positive initial value and positive slopes, on the open neighborhood of $\mathbb{C}_L$ where

$$x_{1,s} > -\kappa, \quad x_{j+1,s} - x_{j,s} > -\kappa/L.$$

On this neighborhood, the variations $x \mapsto q_x^\kappa$ are admissible in the sense of Section 4, and the derivatives of Theorem 4.1 depend continuously on (t, x) by the parameter dependence noted after Lemma 2.3. Consequently u is continuously differentiable on this neighborhood for $t > 0$, with a right derivative in t at $t = 0$. We use the scalar product $\langle \cdot, \cdot \rangle_L$ to define gradients. By Theorem 4.1 and (8.10), the projected gradient is still an average of expected Gibbs overlaps,

$$(8.13) \quad \nabla_L u = A_L p_N, \quad p_N(r) = \mathbb{E}\langle R_N \rangle_r, \quad \partial_t u = \int_0^1 \mathbb{E}\langle \xi(R_N) \rangle_r dr,$$

where the Gibbs averages are those of the path q_x^κ . By Lemmas 3.2 and 8.3, $A_L p_N \in \mathbb{C}_L$; also $0 \leq (A_L p_N)_{j,s} \leq \lambda_{N,s}$, so that each $(A_L p_N)_j$ lies in the positive part of the closed unit ball, and $\mathbf{H}_L(A_L p_N) = L^{-1} \sum_j \xi((A_L p_N)_j)$. Since w_j is supported in the cell J_j and its neighbors, the L^1 distance between p_N and $\iota_L A_L p_N$ is at most $3/L$ times the sum over s of the total variations of $p_{N,s}$, which is at most $\sum_s \lambda_{N,s} = 1$. This gives

$$(8.14) \quad |\mathcal{H}(p_N) - \mathbf{H}_L(\nabla_L u)| = \left| \int_0^1 \xi(p_N(r)) dr - \frac{1}{L} \sum_j \xi((A_L p_N)_j) \right| \leq \frac{3L\xi}{L}.$$

Combining (4.8), Theorem 5.1 and (8.14), we obtain, for every $x \in \mathbb{C}_L$ and $t > 0$,

$$(8.15) \quad |\partial_t u - \mathbf{H}_L(\nabla_L u)| \leq \varepsilon.$$

The function u is Lipschitz in x by (4.12) and (8.9), since $\|\mathcal{T}_L(x-y)\|_{L^1} \leq \|\mathcal{T}_L(x-y)\|_{L^2} \leq |x-y|_L$, and in t since $|\partial_t \bar{F}_{N,k}| \leq \xi(1)$.

Step 2: contacts at the boundary.

Suppose that a smooth function ϕ is such that $u - \phi$ has a local maximum at $(t, x) \in (0, \infty) \times \mathbb{C}_L$, relative to the closed cone. The extension of u to an open neighborhood gives $\partial_t \phi = \partial_t u$ and, for every $y \in \mathbb{C}_L$, the point $x + \theta y$ belongs to $\mathbb{C}_L$ for $\theta > 0$, so $\langle \nabla_L \phi - \nabla_L u, y \rangle_L \geq 0$; that is, $\nabla_L \phi - \nabla_L u \in \mathbb{C}_L^*$. Since H_L is increasing for the dual-cone order, (8.15) gives $\partial_t \phi - H_L(\nabla_L \phi) \leq \partial_t u - H_L(\nabla_L u) \leq \varepsilon$. At a local minimum, the dual-cone relation is reversed and gives $\partial_t \phi - H_L(\nabla_L \phi) \geq -\varepsilon$. Hence $u - \varepsilon t$ is a viscosity subsolution and $u + \varepsilon t$ a viscosity supersolution of the equation on $\mathbb{C}_L$. $\square$

8.3. Convergence, and continuity in the magnetization. For $R > 0$, let

$$(8.16) \quad \delta_N(R) := \sup_k \sup_{\substack{q \in \mathcal{Q}_2^{\mathcal{S}} \\ \|q\|_{L^2} \leq R}} |\bar{F}_{N,k}(0, q) - \psi_{m_{N,k}}(q)|,$$

which tends to zero as $N \rightarrow \infty$ by Corollary 7.6, since $\|q\|_{L^1} \leq \|q\|_{L^2}$.

Proposition 8.5 (Convergence estimate). *For every N , every slice $\Omega_{N,k}$, every $t \geq 0$, every $q \in \mathcal{Q}_2^{\mathcal{S}}$, every $0 < \kappa \leq 1$, and every $R > \|q\|_{L^2} + L_H t$,*

$$(8.17) \quad |\bar{F}_{N,k}(t, q) - f_{m_{N,k}}(t, q)| \leq \delta_N(R) + 3\kappa + 15M_\xi t (\kappa N)^{-1/3}.$$

Proof. Write $m := m_{N,k}$, fix a power of two L , let u be the function (8.11), and let $f_{m,L}$ be the solution of Theorem 8.1(4) with initial condition $x \mapsto \psi_m(\iota_L x)$. By (4.12), for every $q \in \mathcal{Q}_1^{\mathcal{S}}$ and $t \geq 0$,

$$(8.18) \quad |\bar{F}_{N,k}(t, q + \kappa(1+r)\mathbf{1}) - \bar{F}_{N,k}(t, q)| \leq \sum_s \lambda_{N,s} \int_0^1 \kappa(1+r) dr = \frac{3}{2}\kappa,$$

which is the cost of the shift in (8.11).

Step 1: the initial data.

Let $x \in \mathbb{C}_L$. By (4.12), (8.18) and (8.9),

$$|u(0, x) - \bar{F}_{N,k}(0, \iota_L x)| \leq \frac{3}{2}\kappa + \|\mathcal{T}_L x - \iota_L x\|_{L^1} \leq \frac{3}{2}\kappa + \frac{\sqrt{|\mathcal{S}|}}{4\sqrt{L}} |x|_L.$$

If $|x|_L \leq R$, then $\|\iota_L x\|_{L^2} = |x|_L \leq R$, and $|\bar{F}_{N,k}(0, \iota_L x) - \psi_m(\iota_L x)| \leq \delta_N(R)$ by definition. Otherwise, let $x' := (R/|x|_L)x$, which belongs to $\mathbb{C}_L$ and satisfies $\|\iota_L(x - x')\|_{L^1} \leq \|\iota_L(x - x')\|_{L^2} = |x|_L - R$; since $\bar{F}_{N,k}(0, \cdot)$ and ψ_m are 1-Lipschitz for the L^1 norm, by Proposition 4.2 and (6.50), we get $|\bar{F}_{N,k}(0, \iota_L x) - \psi_m(\iota_L x)| \leq \delta_N(R) + 2(|x|_L - R)$. Altogether,

$$(8.19) \quad |u(0, x) - \psi_m(\iota_L x)| \leq \omega(|x|_L), \quad \text{where } \omega(s) := \delta_N(R) + \frac{3}{2}\kappa + \frac{\sqrt{|\mathcal{S}|}}{4\sqrt{L}} s + 2(s - R)_+.$$

Step 2: comparison.

The function ω is increasing and Lipschitz. By Proposition 8.4 and Lemma 8.2, applied on $\mathbb{C}_L$ with $v = f_{m,L}$, we have

$$(8.20) \quad |u(t, x) - f_{m,L}(t, x)| \leq \varepsilon t + \omega(|x|_L + L_H t) \quad \text{for every } (t, x) \in \mathbb{R}_+ \times \mathbb{C}_L,$$

with ε as in (8.12).

Step 3: passage to the limit.

Let $x := \pi_L q$. Since $|x|_L \leq \|q\|_{L^2}$ and $R > \|q\|_{L^2} + L_H t$, the last term of ω vanishes in (8.20), which gives

$$|u(t, \pi_L q) - f_{m,L}(t, \pi_L q)| \leq \varepsilon t + \delta_N(R) + \frac{3}{2}\kappa + \frac{\sqrt{|\mathcal{S}|}}{4\sqrt{L}}(\|q\|_{L^2} + L_H t).$$

By (4.12) and (8.18), $|\bar{F}_{N,k}(t, q) - u(t, \pi_L q)| \leq \|q - \mathcal{T}_L \pi_L q\|_{L^1} + \frac{3}{2}\kappa$. We now let $L \rightarrow \infty$ along the powers of two: $\|q - \mathcal{T}_L \pi_L q\|_{L^1} \rightarrow 0$, the term $3L_\xi t/L$ in εt and the last term of the previous display vanish, and $f_{m,L}(t, \pi_L q) \rightarrow f_m(t, q)$ by Theorem 8.1(4). This proves (8.17). $\square$

Proof of (1.10) in Theorem 1.2, and continuity in m . Let $t \geq 0$ and $q \in \mathcal{Q}_2^\mathcal{S}$, and choose $R > \|q\|_{L^2} + L_H t$. Proposition 8.5 with $\kappa = N^{-1/4}$ gives

$$\sup_k |\bar{F}_{N,k}(t, q) - f_{m_{N,k}}(t, q)| \leq \delta_N(R) + (3 + 15M_\xi t)N^{-1/4} \xrightarrow{N \rightarrow \infty} 0,$$

which is (1.10).

It remains to show that $m \mapsto f_m(t, q)$ is continuous. Let $m, m' \in [-1, 1]^\mathcal{S}$, and let $c := \mathcal{I}_\lambda(m) - \mathcal{I}_\lambda(m')$, with $\mathcal{I}_\lambda$ as in Corollary 6.11. The function $\psi_{m'} + c$ satisfies the assumptions of Theorem 8.1, and its solution is $f_{m'} + c$. By (6.52), for every $q' \in \mathcal{Q}_2^\mathcal{S}$,

$$|\psi_m(q') - (\psi_{m'}(q') + c)| \leq 2 \sum_s \lambda_s \|q'_s\|_{L^1} |m_s - m'_s| \leq 2 \|q'\|_{L^2} |m - m'|_\infty,$$

where $|m - m'|_\infty := \max_s |m_s - m'_s|$. Lemma 8.2 on $\mathcal{Q}_2^\mathcal{S}$, with $u = f_m$, $v = f_{m'} + c$, $\varepsilon = 0$ and $\omega(s) = 2s|m - m'|_\infty$, gives

$$(8.21) \quad |f_m(t, q) - f_{m'}(t, q) - \mathcal{I}_\lambda(m) + \mathcal{I}_\lambda(m')| \leq 2(\|q\|_{L^2} + L_H t) |m - m'|_\infty.$$

Since $\mathcal{I}_\lambda$ is continuous on $[-1, 1]^\mathcal{S}$, so is $m \mapsto f_m(t, q)$. $\square$

The estimate (8.21) is more precise than the continuity asserted in Theorem 1.2: the entropy $\mathcal{I}_\lambda(m)$ carries all the singular dependence of the slice free energy on the magnetization, and the remainder is Lipschitz in m , uniformly over bounded sets of (t, q) .

9. FROM THE HOPF FORMULA TO MARTINGALES

In this section, we prove the martingale representation (1.13) of $f_m(t, q)$, starting from the Hopf formula (1.12). The two convexity properties of g_m established in Section 6 play complementary roles. The convexity in the path gives the Hopf formula, and we first rewrite this formula in conjugate form, in terms of the path conjugate of g_m (Section 9.1). The concavity in the law, in the form of the martingale representation of Proposition 6.6, expresses g_m as an infimum of affine functionals of the law, parametrized by the potentials χ_α . We then show that the Hopf formula can be written as an infimum over such potentials (Sections 9.2 to 9.4), and we replace the potentials by martingales (Section 9.5). The passage from the Hopf formula to the potentials requires exchanging a supremum and an infimum; this is done with Sion's minimax theorem, after enlarging the set of paths to a compact set of couplings. The enlargement does not change the value, by a rearrangement inequality that uses the nonnegativity of the coefficients of ξ .

Three different conjugacies appear in this section, and we fix the notation for them now. For increasing paths, we use the path conjugates ψ_m^* and J_m of Section 9.1, which pair a path with an increasing slope. We give the path conjugate J_m of g_m its own letter, rather than an upper star, because it is conjugated once more in Section 9.3, for a different pairing. For laws, we pair a probability measure ν with a potential χ through $\int \chi d\nu$, and we represent the concave functional g_m by its supporting potentials; the corresponding concave conjugate is denoted by g_* below, with a lower star. Finally, for

functions of a finite-dimensional variable, such as a potential χ on $\mathbb{R}_+^{\mathcal{S}}$ or a function v on $[0, 1]$, we write χ^* and v^* for the usual Legendre transforms. Upper stars always denote convex conjugates. The two scalar identities of Section 9.3 relate the first two conjugacies through the third one, and they also show that the path conjugate J_m is convex for the affine interpolation of the laws of slopes.

9.1. The conjugate form of the Hopf formula. It is convenient to rewrite the Hopf formula (1.12) in conjugate form, and to record the duality in the path variable that comes with the supporting paths of Proposition 6.10. For $m \in [-1, 1]$ and $z \in \mathcal{Q}_\infty$ with $0 \leq z \leq 1$, we define the path conjugate

$$(9.1) \quad J_m(z) := \sup_{a \in \mathcal{Q}_\infty} \left\{ \int_0^1 z(r) a(r) dr - g_m(a) \right\} \in [-g_m(0), +\infty].$$

Lemma 9.1 (Path biconjugacy). *For every $m \in [-1, 1]$ and $a \in \mathcal{Q}_\infty$, the path z_a of (6.45) satisfies $J_m(z_a) = \int_0^1 z_a a - g_m(a)$, and*

$$(9.2) \quad g_m(a) = \sup_{\substack{z \in \mathcal{Q}_\infty \\ 0 \leq z \leq 1}} \left\{ \int_0^1 z(r) a(r) dr - J_m(z) \right\},$$

the supremum being attained at $z = z_a$.

Proof. By (6.46), $\int_0^1 z_a b - g_m(b) \leq \int_0^1 z_a a - g_m(a)$ for every $b \in \mathcal{Q}_\infty$, with equality for $b = a$; this is the first claim. The inequality $\geq$ in (9.2) holds for every admissible z by the definition of J_m , and the first claim gives equality at $z = z_a$. $\square$

In words, g_m is the conjugate of its own conjugate, and the dual variable may be restricted to increasing paths with values in $[0, 1]$. For $p \in \mathcal{Q}_\infty^{\mathcal{S}}$, let

$$(9.3) \quad \psi_m^*(p) := \sup_{q' \in \mathcal{Q}_\infty^{\mathcal{S}}} \{ \langle p, q' \rangle_{L^2} - \psi_m(q') \} \in [-\psi_m(0), +\infty].$$

Recall the set $\mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}}$ of paths with $0 \leq p_s \leq \lambda_s$ from (6.49).

Lemma 9.2 (Conjugate form of the Hopf formula). *For every $m \in [-1, 1]^{\mathcal{S}}$, $t \geq 0$ and $q \in \mathcal{Q}_2^{\mathcal{S}}$, we have*

$$(9.4) \quad f_m(t, q) = \sup_{p \in \mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}}} \left\{ \langle q, p \rangle_{L^2} + t \int_0^1 \xi(p(r)) dr - \psi_m^*(p) \right\}.$$

Moreover, for every $p \in \mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}}$, we have

$$(9.5) \quad \psi_m^*(p) = \sum_s \lambda_s J_{m_s} \left(\frac{p_s}{\lambda_s} \right),$$

where J_{m_s} is the path conjugate (9.1).

Proof. The Hopf formula (1.12) reads $f_m(t, q) = \sup_{p \in \mathcal{Q}_\infty^{\mathcal{S}}} \{ \langle q, p \rangle_{L^2} + t \int_0^1 \xi(p) - \psi_m^*(p) \}$. If $p \in \mathcal{Q}_\infty^{\mathcal{S}} \setminus \mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}}$, there are $s \in \mathcal{S}$ and $r_0 \in [0, 1]$ with $p_s(r_0) > \lambda_s$, by right-continuity, so $\int_{r_0}^1 (p_s - \lambda_s) > 0$ by monotonicity. Testing (9.3) with $q' = a \mathbb{1}_{[r_0, 1]} e_s$ and using (6.50), we get $\psi_m^*(p) \geq a \int_{r_0}^1 (p_s - \lambda_s) - \psi_m(0) \rightarrow +\infty$ as $a \rightarrow \infty$. Such p do not contribute to the supremum. For $p \in \mathcal{Q}_{\infty, \leq \lambda}^{\mathcal{S}}$, the separable form (9.5) follows from (1.9), since the components of q' vary independently; the lower bound on ψ_m^* is obtained by taking $q' = 0$. $\square$

At $t = 0$, where $f_m(0, \cdot) = \psi_m$, formula (9.4) reduces to the biconjugacy of Lemma 9.1, summed over the species.

9.2. Potentials, and localization to the horizon. Throughout the rest of this section, we fix $m \in [-1, 1]^{\mathcal{S}}$, $t \geq 0$ and $q \in \mathcal{Q}_{\infty}^{\mathcal{S}}$, and horizons $T = (T_s)_{s \in \mathcal{S}}$ such that

$$(9.6) \quad T_s > 0, \quad T_s \geq \|q_s\|_{\infty} + t \partial_s \xi(\lambda) \quad \text{for every } s \in \mathcal{S}.$$

We write $[0, T] := \prod_s [0, T_s]$ and $\Lambda_{\lambda} := \prod_s [0, \lambda_s]$.

Let $\mathfrak{X}_1$ be the set of convex, increasing, 1-Lipschitz functions $\chi : \mathbb{R}_+ \rightarrow \mathbb{R}$ with $\chi(0) = 0$, and let

$$(9.7) \quad \mathfrak{X} := \left\{ x \mapsto \sum_s \lambda_s \chi_s(x_s) : \chi_s \in \mathfrak{X}_1 \right\}$$

be the corresponding separable functions on $\mathbb{R}_+^{\mathcal{S}}$. We often call elements of $\mathfrak{X}$ potentials. For $\chi \in \mathfrak{X}$, we define its conjugate and its evolution at time t by

$$(9.8) \quad \chi^*(p) := \sup_{x \in \mathbb{R}_+^{\mathcal{S}}} \{p \cdot x - \chi(x)\}, \quad \mathcal{S}_t \chi(x) := \sup_{p \in \mathbb{R}_+^{\mathcal{S}}} \{p \cdot x + t \xi(p) - \chi^*(p)\}.$$

If $p_s > \lambda_s$ for some s , then $\chi^*(p) = +\infty$, since χ has slope at most λ_s in the direction e_s ; so the supremum defining $\mathcal{S}_t \chi$ is over Λ_{λ} . For a functional $g(\nu) = \sum_s \lambda_s g_s(\nu_s)$ of the marginals of $\nu \in \mathcal{P}_1(\mathbb{R}_+^{\mathcal{S}})$, we define its conjugate for the pairing with potentials, and the value associated with the pair (g, q) , by

$$(9.9) \quad g_*(\chi) := \inf_{\nu \in \mathcal{P}_1(\mathbb{R}_+^{\mathcal{S}})} \left\{ \int \chi d\nu - g(\nu) \right\},$$

$$(9.10) \quad \mathcal{U}_t[g](q) := \inf_{\chi \in \mathfrak{X}} \left\{ \int_0^1 \mathcal{S}_t \chi(q(r)) dr - g_*(\chi) \right\}.$$

Since χ and g are separable, and since every family of laws in $\mathcal{P}_1(\mathbb{R}_+)$ is the family of marginals of some $\nu \in \mathcal{P}_1(\mathbb{R}_+^{\mathcal{S}})$, the infimum in (9.9) decouples over the species:

$$(9.11) \quad g_*(\chi) = \sum_s \lambda_s (g_s)_*(\chi_s), \quad (g_s)_*(\chi_s) := \inf_{\nu_s \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ \int \chi_s d\nu_s - g_s(\nu_s) \right\}.$$

Our goal is to prove that $f_m(t, q) = \mathcal{U}_t[\psi_m](q)$, where ψ_m is regarded as the functional $\nu \mapsto \sum_s \lambda_s g_{m_s}(\nu_s)$ of laws. The next lemma shows that, thanks to the horizon condition (9.6), the problem (9.10) can be localized to the box $[0, T]$.

Lemma 9.3 (Localization). *Let $g(\nu) = \sum_s \lambda_s g_s(\nu_s)$ be a functional of the marginals, as above.*

(1) *For every $\chi \in \mathfrak{X}$ and every $x \in \mathbb{R}_+^{\mathcal{S}}$ with $x_s + t \partial_s \xi(\lambda) \leq T_s$ for all s , we have*

$$(9.12) \quad \mathcal{S}_t \chi(x) = \sup_{p \in \Lambda_{\lambda}} \left\{ p \cdot x + t \xi(p) - \sup_{y \in [0, T]} (p \cdot y - \chi(y)) \right\}.$$

In particular, $\mathcal{S}_t \chi(x)$ depends only on the restriction of χ to $[0, T]$; and if $\chi \leq \chi' + c$ on $[0, T]$ for a constant c , then $\mathcal{S}_t \chi(x) \leq \mathcal{S}_t \chi'(x) + c$.

(2) *Suppose that each g_s is 1-Lipschitz for the 1-Wasserstein distance. Then the infimum in (9.10) may be restricted to the potentials $\chi = \sum_s \lambda_s \chi_s$ such that χ_s has slope 1 on $[T_s, \infty)$ for every s , and for these potentials, we have*

$$(9.13) \quad g_*(\chi) = \inf_{\nu \in \mathcal{P}([0, T])} \left\{ \int \chi d\nu - g(\nu) \right\}.$$

Proof. Step 1: the evolution.

Write $\chi_T^*(p) := \sup_{y \in [0, T]} (p \cdot y - \chi(y))$, which is a finite convex function of p , separable in the coordinates. Since $\chi_T^* \leq \chi^*$, the right-hand side of (9.12) is at least $\mathcal{S}_t \chi(x)$. For the converse, let $c_s := \lambda_s \chi_s'(T_s-)$ be the left derivative of $\lambda_s \chi_s$ at T_s . If $p_s \geq c_s$, the concave function $y_s \mapsto p_s y_s - \lambda_s \chi_s(y_s)$ is increasing on $[0, T_s]$, so its supremum over $[0, T_s]$

is attained at $y_s = T_s$, and χ_T^* is affine in p_s with slope T_s on this range. There, the derivative in p_s of the objective in (9.12) is $x_s + t\partial_s \xi(p) - T_s \leq x_s + t\partial_s \xi(\lambda) - T_s \leq 0$, because $\partial_s \xi$ is coordinatewise increasing on $\mathbb{R}_+^{\mathcal{J}}$. Hence decreasing p_s to c_s does not decrease the objective. After doing this in every coordinate, we obtain a point p with $p_s \leq c_s$ for all s ; for such p , the function $y_s \mapsto p_s y_s - \lambda_s \chi_s(y_s)$ is decreasing on $[T_s, \infty)$, so $\chi^*(p) = \chi_T^*(p)$. The right-hand side of (9.12) is therefore at most $S_t \chi(x)$. The supremum is attained by continuity and compactness. The last two assertions follow, since χ_T^* only involves χ on $[0, T]$, and $\chi \leq \chi' + c$ on $[0, T]$ implies $\chi_T^* \geq \chi_T'^* - c$.

Step 2: the measure conjugate.

Let $\chi \in \mathfrak{X}$, and let $\widehat{\chi}_s$ coincide with χ_s on $[0, T_s]$ and have slope 1 on $[T_s, \infty)$. Since χ_s is 1-Lipschitz, $\widehat{\chi}_s \geq \chi_s$, so $g_*(\widehat{\chi}) \geq g_*(\chi)$, while $S_t \widehat{\chi}(q(r)) = S_t \chi(q(r))$ for every r by Step 1 and (9.6). Thus replacing χ by $\widehat{\chi}$ does not increase the objective in (9.10). Now let χ have slope 1 beyond the horizons, let $\nu \in \mathcal{P}_1(\mathbb{R}_+^{\mathcal{J}})$, and let ν^T be its image under the coordinatewise truncation $x \mapsto x \wedge T$. For each s , $\int \chi_s d\nu_s - \int \chi_s d\nu_s^T = \int (x_s - T_s)_+ d\nu_s = W_1(\nu_s, \nu_s^T)$, and $g_s(\nu_s) \leq g_s(\nu_s^T) + W_1(\nu_s, \nu_s^T)$. Hence $\int \chi d\nu^T - g(\nu^T) \leq \int \chi d\nu - g(\nu)$. Finally, $\mathcal{P}([0, T])$ is compact for the weak topology, and $\nu \mapsto \int \chi d\nu - g(\nu)$ is continuous on it, since χ is continuous on $[0, T]$ and each g_s is continuous for the weak topology on $\mathcal{P}([0, T_s])$. This proves (9.13). $\square$

By Lemma 6.7, the functions g_m are 1-Lipschitz for W_1 , so the lemma applies to ψ_m . The supporting inequality (6.30) also yields a representation of g_m by potentials. Here $(g_m)_*(\chi) = \inf_{\nu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ \int \chi d\nu - g_m(\nu) \right\}$ is the scalar conjugate appearing in (9.11).

Lemma 9.4 (Dual representation of g_m). *For every $m \in [-1, 1]$, $T > 0$ and $\nu \in \mathcal{P}([0, T])$,*

$$(9.14) \quad g_m(\nu) = \inf_{\chi \in \mathfrak{X}_1} \left\{ \int \chi d\nu - (g_m)_*(\chi) \right\},$$

and the infimum is attained at the potential $\widehat{\chi}_{\nu, m} \in \mathfrak{X}_1$ that coincides with the supporting potential $\chi_{\nu, m}$ of Corollary 6.8 on $[0, T]$ and has slope 1 on $[T, \infty)$.

Proof. The inequality $\leq$ in (9.14) holds for every $\chi \in \mathfrak{X}_1$, by the definition of $(g_m)_*$. Conversely, let $\nu' \in \mathcal{P}_1(\mathbb{R}_+)$, and let ν'^T be its image under $x \mapsto x \wedge T$. By Lemma 6.7, $g_m(\nu') \leq g_m(\nu'^T) + W_1(\nu', \nu'^T) = g_m(\nu'^T) + \int (x - T)_+ d\nu'$, and by (6.30) at $\nu'^T \in \mathcal{P}([0, T])$, $g_m(\nu'^T) \leq g_m(\nu) + \int \chi_{\nu, m} d(\nu'^T - \nu)$. Since $\widehat{\chi}_{\nu, m}(x) = \chi_{\nu, m}(x \wedge T) + (x - T)_+$ and ν is supported in $[0, T]$, the two bounds combine into $g_m(\nu') \leq g_m(\nu) + \int \widehat{\chi}_{\nu, m} d\nu' - \int \widehat{\chi}_{\nu, m} d\nu$. Taking the infimum over ν' , we get $(g_m)_*(\widehat{\chi}_{\nu, m}) \geq \int \widehat{\chi}_{\nu, m} d\nu - g_m(\nu)$, which is the reverse inequality at $\chi = \widehat{\chi}_{\nu, m}$. $\square$

9.3. Two scalar conjugacies. We fix a scalar magnetization $m_0 \in [-1, 1]$ and abbreviate $g := g_{m_0}$ and $J := J_{m_0}$, where J is the path conjugate (9.1). For $\rho \in \mathcal{P}([0, 1])$, we write $J(\rho) := J(p_\rho)$, where p_ρ is the quantile function of ρ ; recall that the quantile functions are exactly the right-continuous increasing functions on $[0, 1)$ with values in $[0, 1]$, so that $J(\rho)$ is defined for every ρ and every such z is p_ρ for the law ρ of $z(U)$. The function g is convex in the path, with conjugate J ; and g is concave in the law, with concave conjugate g_* for the pairing with potentials. The following lemma relates the two conjugates. Let $\mathcal{V}$ be the convex set of convex, increasing, Lipschitz functions $v : [0, 1] \rightarrow \mathbb{R}$ with $v(0) = 0$. For $v \in \mathcal{V}$, its conjugate $v^*(x) := \sup_{0 \leq z \leq 1} (xz - v(z))$, $x \geq 0$, is convex, increasing, 1-Lipschitz, vanishes at 0, and has slope 1 beyond the Lipschitz constant of v ; conversely, if $\chi \in \mathfrak{X}_1$ has slope 1 beyond some point, then $\chi^*|_{[0, 1]} \in \mathcal{V}$ and $(\chi^*)^* = \chi$. Thus $v \mapsto v^*$ is a bijection from $\mathcal{V}$ onto the potentials in $\mathfrak{X}_1$ with eventual slope 1, and $v^{**} = v$ on $[0, 1]$. We also regard J as a function on $\mathcal{P}([0, 1])$, and we define its convex conjugate for the

pairing $\int v d\rho$ between functions and measures on $[0, 1]$:

$$(9.15) \quad J^*(v) := \sup_{\rho \in \mathcal{P}([0,1])} \left\{ \int v d\rho - J(\rho) \right\}, \quad v \in \mathcal{V}.$$

The following lemma identifies J^* with the concave conjugate g_* , and shows that J is its own biconjugate.

Lemma 9.5 (Scalar dual identities). *For every $v \in \mathcal{V}$ and every $\rho \in \mathcal{P}([0, 1])$,*

$$(9.16) \quad J^*(v) = -g_*(v^*),$$

$$(9.17) \quad J(\rho) = \sup_{v \in \mathcal{V}} \left\{ \int v d\rho - J^*(v) \right\}.$$

In particular, J^* is finite, convex and 1-Lipschitz for the uniform norm on $\mathcal{V}$.

Proof. Step 1: a pointwise conjugacy.

Let $\chi \in \mathfrak{X}_1$ have slope 1 beyond $x_1 \geq 0$, and let $z \in \mathcal{Q}_\infty$ with $0 \leq z \leq 1$. We claim that

$$(9.18) \quad \sup_{a \in \mathcal{Q}_\infty} \int_0^1 [z(r)a(r) - \chi(a(r))] dr = \int_0^1 \chi^*(z(r)) dr.$$

The inequality $\leq$ is clear. For the converse, let $a(r)$ be the smallest maximizer of $x \mapsto z(r)x - \chi(x)$ on $\mathbb{R}_+$; it exists and belongs to $[0, x_1]$, since this function is concave and decreasing beyond x_1 , and it is increasing in r because the function has increasing differences in (x, z) and z is increasing. Replacing a by its right-continuous version, which differs from it on a countable set only, we obtain an element of $\mathcal{Q}_\infty$ attaining the supremum.

Step 2: a first formula for J .

Let $z := p_\rho$ be the quantile function of ρ . Inserting (9.14) into the definition of J , where by Lemma 9.4 the infimum may be restricted to potentials with eventual slope 1, and exchanging the two suprema,

$$J(\rho) = \sup_{a \in \mathcal{Q}_\infty} \sup_{\chi} \left\{ \int_0^1 [z(r)a(r) - \chi(a(r))] dr + g_*(\chi) \right\} = \sup_{\chi} \left\{ \int_0^1 \chi^*(z(r)) dr + g_*(\chi) \right\},$$

by (9.18). Writing $v := \chi^*|_{[0,1]} \in \mathcal{V}$, we have $\chi = v^*$ and $\int_0^1 v(z(r)) dr = \int v d\rho$, so that

$$(9.19) \quad J(\rho) = \sup_{v \in \mathcal{V}} \left\{ \int v d\rho + g_*(v^*) \right\}.$$

Step 3: proof of (9.16) and (9.17).

Since v^* has slope 1 beyond the Lipschitz constant of v , the truncation argument of Lemma 9.3 shows that the infimum defining $g_*(v^*)$ may be taken over compactly supported laws, that is, $-g_*(v^*) = \sup_{a \in \mathcal{Q}_\infty} \{g(a) - \int_0^1 v^*(a(r)) dr\}$. Inserting the biconjugacy (9.2) of Lemma 9.1 and exchanging the suprema,

$$\begin{aligned} -g_*(v^*) &= \sup_{\substack{z \in \mathcal{Q}_\infty \\ 0 \leq z \leq 1}} \left\{ \sup_{a \in \mathcal{Q}_\infty} \int_0^1 [z(r)a(r) - v^*(a(r))] dr - J(z) \right\} \\ &= \sup_z \left\{ \int_0^1 v(z(r)) dr - J(z) \right\}, \end{aligned}$$

by (9.18) applied to $\chi = v^*$, whose conjugate is v . Since the paths z in the supremum are exactly the quantile functions of the laws $\rho \in \mathcal{P}([0, 1])$, the right-hand side is $J^*(v)$, which proves (9.16). Inserting (9.16) into (9.19) gives (9.17). Finally, J^* is finite by (9.16), since $g(\nu) \leq g(\delta_0) + \int x d\nu$ by the Lipschitz bound of Lemma 6.7 and $x - v(1) \leq v^*(x) \leq x$,

so that $-v(1) - g(\delta_0) \leq g_*(v^*) \leq -g(\delta_0)$; and (9.15) represents J^* as a supremum of affine functions of v , each 1-Lipschitz for the uniform norm. $\square$

The identity (9.17) will be used to express the Hopf value through potentials, and the convexity of J^* will be used when appealing to the minimax theorem. The lemma also shows that J , which was defined as a conjugate in the path variable, is convex for the affine interpolation of the laws ρ , as a supremum of affine functions of ρ .

9.4. Minimax and a common increasing rearrangement. We now express the Hopf value through potentials.

Proposition 9.6 (The Hopf value through potentials). *For every $m \in [-1, 1]^\mathcal{S}$, $t \geq 0$ and $q \in \mathcal{Q}_\infty^\mathcal{S}$, we have*

$$(9.20) \quad f_m(t, q) = \mathcal{U}_t[\psi_m](q).$$

Proof. Step 1: from increasing paths to couplings.

For each $s \in \mathcal{S}$, let J_s and J_s^* be the functions of Section 9.3 with magnetization m_s , and for $z \in [0, 1]^\mathcal{S}$, let $\lambda z := (\lambda_s z_s)_s$. By Lemma 9.2, writing $p = \lambda z$ with z an increasing path with values in $[0, 1]^\mathcal{S}$,

$$f_m(t, q) = \sup_z \left\{ \int_0^1 [q(r) \cdot \lambda z(r) + t\xi(\lambda z(r))] dr - \sum_s \lambda_s J_s(z_s) \right\}.$$

Let μ_q be the law of $q(U)$ on $\mathbb{R}_+^\mathcal{S}$, and let Π_q be the set of probability measures π on $\mathbb{R}_+^\mathcal{S} \times [0, 1]^\mathcal{S}$ whose first marginal is μ_q . Since q is bounded, every $\pi \in \Pi_q$ is supported in the compact set $K \times [0, 1]^\mathcal{S}$, where $K := \Pi_s[0, \|q_s\|_\infty]$; hence Π_q is a convex, weakly closed subset of $\mathcal{P}(K \times [0, 1]^\mathcal{S})$, and thus weakly compact, and the integrands below are bounded and continuous on $K \times [0, 1]^\mathcal{S}$. We write (x, z) for a generic point of $\mathbb{R}_+^\mathcal{S} \times [0, 1]^\mathcal{S}$, with $x = (x_s)_{s \in \mathcal{S}}$ and $z = (z_s)_{s \in \mathcal{S}}$, and for $\pi \in \Pi_q$ and $s \in \mathcal{S}$, we write $\pi_s \in \mathcal{P}([0, 1])$ for the image of π under the map $(x, z) \mapsto z_s$. For an increasing path z as in the last display, let $\pi^z \in \Pi_q$ be the law of $(q(U), z(U))$; we call these couplings monotone. The marginal π_s^z is the law of $z_s(U)$, so that $J_s(\pi_s^z) = J_s(z_s)$, and the integral above is $\int [x \cdot \lambda z + t\xi(\lambda z)] d\pi^z$. We claim that

$$(9.21) \quad f_m(t, q) = \sup_{\pi \in \Pi_q} \left\{ \int [x \cdot \lambda z + t\xi(\lambda z)] d\pi - \sum_s \lambda_s J_s(\pi_s) \right\},$$

that is, that the supremum over all couplings is already attained on the monotone ones. Let $\pi \in \Pi_q$, with marginals μ_q and $(\rho_s)_s$ for the coordinates z_s , and let $\pi^\dagger := \pi^z$ for $z := (p_{\rho_s})_s$, where p_{ρ_s} is the quantile function of ρ_s . This monotone coupling has the same marginals as π , so the costs J_s are unchanged, and it suffices to show that the first integral in (9.21) does not decrease when π is replaced by $\pi^\dagger$. For nonnegative random variables $Y_1, \dots, Y_n$ with prescribed marginal laws, the layer-cake formula gives

$$\mathbb{E}\left[\prod_i Y_i\right] = \int_{\mathbb{R}_+^n} \mathbb{P}(Y_1 > u_1, \dots, Y_n > u_n) du \leq \int_{\mathbb{R}_+^n} \min_i \mathbb{P}(Y_i > u_i) du,$$

and the monotone coupling $Y_i = p_{\text{Law}(Y_i)}(U)$ achieves the bound, since for it the joint tail probability equals the smallest marginal tail probability. We apply this to each product $x_s z_s$ and to each monomial $\prod_s z_s^{a_s}$ in the power series of $\xi(\lambda z)$. Since the coefficients of this series are nonnegative, and since it converges absolutely, summing the inequalities proves the claim.

Step 2: exchange of the supremum and the infimum.

By (9.17), $J_s(\pi_s) = \sup_{v_s \in \mathcal{V}} \{ \int v_s d\pi_s - J_s^*(v_s) \}$, so that (9.21) becomes

$$(9.22) \quad f_m(t, q) = \sup_{\pi \in \Pi_q} \inf_{v \in \mathcal{V}^{\mathcal{S}}} \left\{ \int [x \cdot \lambda z + t\xi(\lambda z)] d\pi - \sum_s \lambda_s \int v_s d\pi_s + \sum_s \lambda_s J_s^*(v_s) \right\}.$$

The expression in braces is affine in π , and weakly continuous since its integrands are bounded and continuous on $K \times [0, 1]^{\mathcal{S}}$; it is convex and continuous for the uniform norm in v , by Lemma 9.5. Since Π_q is convex and compact, Sion's minimax theorem [32] allows us to exchange the supremum and the infimum.

Step 3: the supremum over couplings at a fixed potential.

Let $v = (v_s)_{s \in \mathcal{S}} \in \mathcal{V}^{\mathcal{S}}$, and let $\chi_v := \sum_s \lambda_s v_s^* \in \mathfrak{X}$. Then, for $z \in [0, 1]^{\mathcal{S}}$,

$$(9.23) \quad \chi_v^*(\lambda z) = \sum_s \lambda_s v_s(z_s), \quad -(\psi_m)_*(\chi_v) = \sum_s \lambda_s J_s^*(v_s).$$

The first identity follows from $v_s^{**} = v_s$, and the second from (9.11) and (9.16). The supremum over $\pi \in \Pi_q$ of the integral of a continuous function of (x, z) is the integral against μ_q of its pointwise supremum over $z \in [0, 1]^{\mathcal{S}}$. By the first identity in (9.23), and since $\chi_v^* = +\infty$ outside Λ_λ , the definition (9.8) gives

$$\sup_{\pi \in \Pi_q} \int \left[x \cdot \lambda z + t\xi(\lambda z) - \sum_s \lambda_s v_s(z_s) \right] d\pi(x, z) = \int_0^1 S_{t\chi_v}(q(r)) dr.$$

Together with the second identity in (9.23), the exchanged form of (9.22) therefore reads

$$f_m(t, q) = \inf_{v \in \mathcal{V}^{\mathcal{S}}} \left\{ \int_0^1 S_{t\chi_v}(q(r)) dr - (\psi_m)_*(\chi_v) \right\}.$$

By Lemma 9.3, the infimum defining $\mathcal{U}_t[\psi_m](q)$ may be restricted to the potentials with slope 1 beyond the horizons, and these are of the form χ_v with $v \in \mathcal{V}^{\mathcal{S}}$. The infimum over all $v \in \mathcal{V}^{\mathcal{S}}$ lies between this restricted infimum and the infimum over all of $\mathfrak{X}$, so it equals $\mathcal{U}_t[\psi_m](q)$. This proves (9.20). $\square$

9.5. Replacing potentials by martingales. Recall from the introduction that $\mathcal{M}_T$ is the set of families $\alpha = (\alpha_s)_{s \in \mathcal{S}}$ of $[-1, 1]$ -valued martingales on $[0, T_s]$, α_s being a martingale for the filtration $\mathcal{F}^s$ of the Brownian motion B_s , and that $\mathcal{M}_T^m$ is the subset of those with $\alpha_s(0) = m_s$ for every s . For $\alpha \in \mathcal{M}_T$, we define the potential

$$(9.24) \quad \chi_\alpha(x) := \sum_s \lambda_s \chi_{\alpha_s}(x_s) = \sum_s \lambda_s \int_0^{x_s} \mathbb{E}[\alpha_s(v \wedge T_s)^2] dv, \quad x \in \mathbb{R}_+^{\mathcal{S}},$$

with χ_{α_s} as in (6.19). It belongs to $\mathfrak{X}$. The cost $\mathcal{C}_T(\alpha)$ of the introduction is $\sum_s \lambda_s \mathcal{C}_{T_s}(\alpha_s)$, with $\mathcal{C}_T$ as in (6.19).

Proposition 9.7 (Transfer and attainment). *For every m, t, q and T as in (9.6),*

$$(9.25) \quad \mathcal{U}_t[\psi_m](q) = \inf_{\alpha \in \mathcal{M}_T^m} \left\{ \mathcal{C}_T(\alpha) + \int_0^1 S_{t\chi_\alpha}(q(r)) dr \right\},$$

and the infimum is attained.

Proof. The compactness used for the attainment is that of the potentials, not of the martingales: we first find an optimal potential, and we then construct a martingale from a law attaining the concave conjugate at this potential.

Step 1: an optimal potential.

By Lemma 9.3(2), the infimum defining $\mathcal{U}_t[\psi_m](q)$ may be restricted to the potentials $\chi = \sum_s \lambda_s \chi_s$ with $\chi_s \in \mathfrak{X}_1$ of slope 1 beyond T_s . Such a potential is determined by its restriction to $[0, T]$, and these restrictions form a compact set for the uniform norm, by the Arzelà–Ascoli theorem, since they are 1-Lipschitz, vanish at 0, and form a closed class.

Both terms of the objective in (9.10) are 1-Lipschitz for the uniform norm on $[0, T]$, by (9.12) and (9.13). An optimal potential χ therefore exists; we fix it for the rest of the proof.

Step 2: from the optimal potential to a martingale.

By (9.11) and Lemma 9.3(2), $(\psi_m)_*(\chi) = \sum_s \lambda_s (g_{m_s})_*(\chi_s)$, and each $(g_{m_s})_*(\chi_s)$ is an infimum over $\mathcal{P}([0, T_s])$, attained at some ν_s . Let $\alpha_s := M^{\nu_s, m_s}$ be the minimizing martingale of Proposition 6.6 on $[0, T_s]$, so that $\alpha \in \mathcal{M}_T^m$ and $g_{m_s}(\nu_s) = c_{T_s}(\alpha_s) + \int \chi_{\alpha_s} d\nu_s$. We claim that

$$(9.26) \quad \chi_s(x) - \chi_{\alpha_s}(x) \geq \vartheta_s \quad \text{for every } x \in [0, T_s], \quad \vartheta_s := \int (\chi_s - \chi_{\alpha_s}) d\nu_s.$$

This is the first-order optimality condition for ν_s . Indeed, let $x \in [0, T_s]$, $\varepsilon \in (0, 1)$, and $\nu' := (1 - \varepsilon)\nu_s + \varepsilon\delta_x \in \mathcal{P}([0, T_s])$. By the minimality of ν_s ,

$$\frac{g_{m_s}(\nu') - g_{m_s}(\nu_s)}{\varepsilon} \leq \frac{1}{\varepsilon} \int \chi_s d(\nu' - \nu_s) = \chi_s(x) - \int \chi_s d\nu_s,$$

and letting $\varepsilon \rightarrow 0$ with (6.31) turns the left-hand side into $\chi_{\alpha_s}(x) - \int \chi_{\alpha_s} d\nu_s$. In words, χ_s touches $\chi_{\alpha_s} + \vartheta_s$ from above on $[0, T_s]$, and the two integrate to the same value against ν_s , so ν_s is supported on the contact set. Consequently,

$$(\psi_m)_*(\chi) = \sum_s \lambda_s \left[\int \chi_s d\nu_s - g_{m_s}(\nu_s) \right] = \sum_s \lambda_s \vartheta_s - \mathcal{C}_T(\alpha),$$

and, since $\chi \geq \chi_\alpha + \sum_s \lambda_s \vartheta_s$ on $[0, T]$, the ordering property of Lemma 9.3(1) gives $S_t \chi(q(r)) \geq S_t \chi_\alpha(q(r)) + \sum_s \lambda_s \vartheta_s$ for every r . Therefore

$$(9.27) \quad \mathcal{U}_t[\psi_m](q) = \int_0^1 S_t \chi(q(r)) dr - (\psi_m)_*(\chi) \geq \mathcal{C}_T(\alpha) + \int_0^1 S_t \chi_\alpha(q(r)) dr.$$

Step 3: every martingale gives a potential.

Let $\alpha' \in \mathcal{M}_T^m$ be arbitrary, and let $\widehat{\chi}_{\alpha'} := \sum_s \lambda_s \widehat{\chi}_{\alpha'_s}$, where $\widehat{\chi}_{\alpha'_s}(x) := \chi_{\alpha'_s}(x \wedge T_s) + (x - T_s)_+$ coincides with $\chi_{\alpha'_s}$ on $[0, T_s]$ and has slope 1 beyond T_s . Let $\nu \in \mathcal{P}_1(\mathbb{R}^{\mathcal{I}})$, and let ν^T be its image under the coordinatewise truncation $x \mapsto x \wedge T$. By Lemma 6.7, $g_{m_s}(\nu_s) \leq g_{m_s}(\nu_s^T) + \int (x_s - T_s)_+ d\nu_s$, and by the inequality in (6.21) at $\nu_s^T \in \mathcal{P}([0, T_s])$, $g_{m_s}(\nu_s^T) \leq c_{T_s}(\alpha'_s) + \int \chi_{\alpha'_s} d\nu_s^T$. Since $\int \chi_{\alpha'_s} d\nu_s^T + \int (x_s - T_s)_+ d\nu_s = \int \widehat{\chi}_{\alpha'_s} d\nu_s$, this gives $g_{m_s}(\nu_s) \leq c_{T_s}(\alpha'_s) + \int \widehat{\chi}_{\alpha'_s} d\nu_s$, hence $\psi_m(\nu) \leq \mathcal{C}_T(\alpha') + \int \widehat{\chi}_{\alpha'} d\nu$ and $(\psi_m)_*(\widehat{\chi}_{\alpha'}) \geq -\mathcal{C}_T(\alpha')$. Moreover, $S_t \widehat{\chi}_{\alpha'}(q(r)) = S_t \chi_{\alpha'}(q(r))$ by Lemma 9.3(1). Thus

$$\mathcal{U}_t[\psi_m](q) \leq \int_0^1 S_t \widehat{\chi}_{\alpha'}(q(r)) dr - (\psi_m)_*(\widehat{\chi}_{\alpha'}) \leq \mathcal{C}_T(\alpha') + \int_0^1 S_t \chi_{\alpha'}(q(r)) dr.$$

Together with (9.27), this proves (9.25), with the infimum attained at the martingale α of Step 2. $\square$

The minimizer constructed in Step 2 has a special form: each α_s is the martingale M^{ν_s, m_s} of Proposition 6.6, for some law $\nu_s \in \mathcal{P}([0, T_s])$. The infimum in (9.25), and hence in (1.13), may therefore always be sought among these martingales.

Proof of (1.13) in Theorem 1.2. Let $q \in \mathcal{Q}_\infty^{\mathcal{I}}$ and let T satisfy (9.6). We claim that $S_t \chi_\alpha(q(r)) = \mathcal{K}_{t, \alpha}^T(q(r))$ for every $\alpha \in \mathcal{M}_T$ and $r \in [0, 1]$. Since $q_s(r) \leq \|q_s\|_\infty$, the point $x = q(r)$ satisfies the condition of Lemma 9.3(1), so $S_t \chi_\alpha(q(r))$ is given by (9.12), which only involves the conjugate of χ_α truncated to $[0, T]$. This truncated conjugate is explicit: for an increasing function $\varrho: [0, T_s] \rightarrow [0, 1]$ and $z \in [0, 1]$,

$$(9.28) \quad \sup_{0 \leq y \leq T_s} \left\{ zy - \int_0^y \varrho(v) dv \right\} = \int_0^{T_s} (z - \varrho(v))_+ dv,$$

since the function of y on the left increases while $\varrho < z$ and decreases afterwards. Applying this in each coordinate of (9.12), with $p = \lambda z$ and $\varrho(v) = \mathbb{E}[\alpha_s(v)^2]$, which is increasing since α_s is a martingale, and comparing with (1.4), proves the claim. Combining it with Propositions 9.6 and 9.7 gives

$$f_m(t, q) = \inf_{\alpha \in \mathcal{M}_T^m} \left\{ \mathcal{C}_T(\alpha) + \int_0^1 \mathcal{K}_{t, \alpha}^T(q(r)) dr \right\},$$

with the infimum attained. This completes the proof of Theorem 1.2. $\square$

10. RESTORING THE EXTERNAL FIELDS

We now sum over the slices and prove Theorem 1.1. Recall that $\bar{F}_N(t, q)$ denotes the enriched free energy of the full cube with the weights $2^{-N} e^{h^{(N)} \cdot \sigma}$, which coincides with (1.3) at $q = 0$ by Lemma 4.3. The gluing lemma expresses it through the slice free energies, and Theorem 1.2 identifies their limits.

Theorem 10.1 (Enriched free energy with external fields). *Let $t \geq 0$ and $h \in \mathbb{R}^{\mathcal{S}}$. For every $q \in \mathcal{Q}_2^{\mathcal{S}}$, the identity (1.14) holds, and the infimum there is attained. For every $q \in \mathcal{Q}_\infty^{\mathcal{S}}$ and every T satisfying (9.6), we have*

$$(10.1) \quad \lim_{N \rightarrow \infty} \bar{F}_N(t, q) = \inf_{\alpha \in \mathcal{M}_T} \left\{ \mathcal{C}_T(\alpha) - \sum_s \lambda_s h_s \alpha_s(0) + \int_0^1 \mathcal{K}_{t, \alpha}^T(q(r)) dr \right\},$$

and the infimum is attained.

In the case $q = 0$, the condition (9.6) reads $T_s > 0$ and $T_s \geq t \partial_s \xi(\lambda)$, and (10.1) reduces to (1.5). Theorem 1.1 is thus the special case $q = 0$ of Theorem 10.1.

Proof of Theorem 10.1. We start with the proof of (1.14). Let $t \geq 0$, $h \in \mathbb{R}^{\mathcal{S}}$ and $q \in \mathcal{Q}_2^{\mathcal{S}}$. By Lemma 7.1,

$$\bar{F}_N(t, q) - \min_k \{ \bar{F}_{N, k}(t, q) - \lambda_N \cdot (h m_{N, k}) \} \rightarrow 0.$$

By Theorem 1.2, $\sup_k |\bar{F}_{N, k}(t, q) - f_m(t, q)| \rightarrow 0$, the function $m \mapsto f_m(t, q)$ is continuous on $[-1, 1]^{\mathcal{S}}$, the set of compositions $\{m_{N, k}\}_k$ becomes dense in $[-1, 1]^{\mathcal{S}}$ as $N \rightarrow \infty$, and $\lambda_N \rightarrow \lambda$. Hence

$$\lim_{N \rightarrow \infty} \min_k \{ \bar{F}_{N, k}(t, q) - \lambda_N \cdot (h m_{N, k}) \} = \inf_{m \in [-1, 1]^{\mathcal{S}}} \{ f_m(t, q) - \lambda \cdot (h m) \},$$

where the infimum on the right is attained by continuity and compactness. This shows (1.14).

Turning now to the proof of (10.1), we let $t \geq 0$, $h \in \mathbb{R}^{\mathcal{S}}$ and $q \in \mathcal{Q}_\infty^{\mathcal{S}}$, and let T satisfy (9.6). By (1.13), we have for every $m \in [-1, 1]^{\mathcal{S}}$ that

$$f_m(t, q) - \lambda \cdot (h m) = \inf_{\alpha \in \mathcal{M}_T^m} \left\{ \mathcal{C}_T(\alpha) - \sum_s \lambda_s h_s \alpha_s(0) + \int_0^1 \mathcal{K}_{t, \alpha}^T(q(r)) dr \right\},$$

since $\alpha_s(0) = m_s$ on $\mathcal{M}_T^m$. The sets $\mathcal{M}_T^m$, $m \in [-1, 1]^{\mathcal{S}}$, form a partition of $\mathcal{M}_T$, because the initial value of a martingale in $\mathcal{M}_T$ is deterministic. Minimizing over m in (1.14) therefore gives (10.1), and the infimum is attained: choose a minimizing m in (1.14), and then a minimizing martingale in $\mathcal{M}_T^m$. $\square$

We conclude with a few comments.

At $h = 0$, the functions $m \mapsto g_m(\nu)$ are even and convex, so $\psi_0 \leq \psi_m$ for every m , and the comparison principle gives $f_0 \leq f_m$. Thus $m = 0$ is a minimizing slice in (1.14), and the formula reduces to the result of [7] for centered spins: the free energy of the

centered model is the slice free energy at zero magnetization, which is the solution of the Hamilton–Jacobi equation with the centered initial condition ψ_0 , and this is consistent with $g_0 = \psi_0(\cdot; 0)$.

For general h , the limit free energy of the unconstrained system also admits a representation by potentials, as conjectured in [18] and shown in [6, Theorem 1.4] for convex models. To see this, we set

$$\psi^h(q) := \sum_s \lambda_s \psi_0(q_s; h_s), \quad q \in \mathcal{Q}_1^{\mathcal{J}},$$

which is the limit of $\bar{F}_N(0, q)$ by (7.9). We regard ψ^h as a functional of laws, as in Section 9.2, and we recall the definition of the lower-star operator from (9.9), of $\mathfrak{X}$ from (9.7), and of S_t from (9.8).

Theorem 10.2 (Representation by potentials). *Let $t \geq 0$ and $h \in \mathbb{R}^{\mathcal{J}}$. For every $q \in \mathcal{Q}_{\infty}^{\mathcal{J}}$,*

$$(10.2) \quad \lim_{N \rightarrow \infty} \bar{F}_N(t, q) = \mathcal{U}_t[\psi^h](q) = \inf_{\chi \in \mathfrak{X}} \left\{ \int_0^1 S_t \chi(q(r)) \, dr - (\psi^h)_*(\chi) \right\}.$$

Proof. Let $t \geq 0$, $h \in \mathbb{R}^{\mathcal{J}}$ and $q \in \mathcal{Q}_{\infty}^{\mathcal{J}}$. By Lemma 6.7 applied in each coordinate,

$$\psi^h(\nu) = \inf_{m \in [-1, 1]^{\mathcal{J}}} \{ \psi_m(\nu) - \lambda \cdot (hm) \}$$

for every law ν . Each $\psi_0(\cdot; h_s)$ is 1-Lipschitz for W_1 by Lemma 6.3, so Lemma 9.3 applies to ψ^h : the infimum in (10.2) may be restricted to potentials χ with slope 1 beyond the horizons T of (9.6), and for these potentials,

$$-(\psi^h)_*(\chi) = \sup_{\nu \in \mathcal{P}([0, T])} \inf_{m \in [-1, 1]^{\mathcal{J}}} \left\{ \psi_m(\nu) - \lambda \cdot (hm) - \int \chi \, d\nu \right\}.$$

The function in braces is continuous on the compact convex set $\mathcal{P}([0, T]) \times [-1, 1]^{\mathcal{J}}$, by the Lipschitz bounds in ν and in m of Lemma 6.7; it is concave in ν by Corollary 6.8, and convex in m by Lemma 6.7. By Sion’s minimax theorem, the supremum and the infimum may be exchanged, and Lemma 9.3 applied to ψ_m then gives

$$-(\psi^h)_*(\chi) = \inf_{m \in [-1, 1]^{\mathcal{J}}} \{ -(\psi_m)_*(\chi) - \lambda \cdot (hm) \}.$$

Taking the infimum over χ and exchanging the two infima, we obtain, by Proposition 9.6,

$$\mathcal{U}_t[\psi^h](q) = \inf_{m \in [-1, 1]^{\mathcal{J}}} \{ \mathcal{U}_t[\psi_m](q) - \lambda \cdot (hm) \} = \inf_{m \in [-1, 1]^{\mathcal{J}}} \{ f_m(t, q) - \lambda \cdot (hm) \},$$

and (10.2) follows from (1.14). $\square$

APPENDIX A. THE CASCADE INTERPRETATION OF THE BRANCH

This appendix relates the enriched free energy of Definition 2.4 to its usual description through a Poisson–Dirichlet cascade; its results are not used elsewhere in the paper. We prove three statements. First, $\bar{F}_{N, \Omega}$ agrees with the cascade free energy on all of $\mathcal{Q}_1^{\mathcal{J}}$ (Proposition A.8). Second, in a cascade with finitely many levels, suppose that we sample a leaf and a configuration from the Gibbs measure; as the number of levels tends to infinity, the joint law of the fields along this leaf and of this configuration converges in total variation to that of the values of the branch at the levels and of its sampled configuration (Proposition A.5). Third, concerning the Gibbs measures at level r of Section 3, the law of two configurations under $\mathbb{E}\langle \cdot \rangle_r$ is the limit of the law of two replicas of the cascade Gibbs measure, conditioned on their leaves separating at level r (Lemma A.4 and Proposition A.6). The agreement of the free energies does not by itself imply the other claims, and it is the second statement that justifies calling X a branch of the

cascade. The third statement makes explicit the sense in which the derivative identities of Theorem 4.1 are the familiar cascade computation, disintegrated according to the level at which two replicas separate.

Throughout, (Ω, w) is as in Section 2, $t \geq 0$, and q is a regular path. We write γ_q for the law of the Gaussian path $Z_r = X_0 + \int_0^r \sqrt{2D(u)} dB_u$ of Section 5.2, so that the measure γ of Lemma 5.2 is the product of the law of g and of γ_q .

A.1. The frozen-coefficient problem. Let $n \geq 1$, $r_j := j/n$, and let $\rho_n(r) := r_j$ for $r \in [r_j, r_{j+1})$, with $\rho_n(1) := 1$. Let $U^{(n)}$ be the solution of (2.5) with $\rho_n(r)$ in place of the coefficient r , and with the terminal condition (2.2). Lemmas 2.1 and 2.3 apply to this equation, with the same proofs, the pieces being now delimited by the breakpoints of q and the mesh points; in particular, $|\partial_{x_i} U^{(n)}| \leq 1$ for every i . For $r \in [0, 1]$ and $a \in \mathbb{R}_+^{\mathcal{S}}$, we define the Gaussian operator $T_{r,a}$ on functions of $x \in \mathbb{R}^N$ with at most linear growth by

$$(A.1) \quad \begin{aligned} T_{r,a}f(x) &:= \frac{1}{r} \log \mathbb{E}_z \exp(r f(x + \sqrt{2 \operatorname{diag}(a)} z)) \quad (r > 0), \\ T_{0,a}f(x) &:= \mathbb{E}_z f(x + \sqrt{2 \operatorname{diag}(a)} z), \end{aligned}$$

where z is a standard Gaussian vector in $\mathbb{R}^N$. This is the N -dimensional version of the scalar operator (6.3), with the same semigroup property $T_{r,a}T_{r,a'} = T_{r,a+a'}$.

Lemma A.1 (Cole–Hopf recursion). *For every $x \in \mathbb{R}^N$ and $0 \leq j \leq n-1$,*

$$(A.2) \quad U^{(n)}(r_j, x; g) = T_{r_j, q(r_{j+1}) - q(r_j)} U^{(n)}(r_{j+1}, \cdot; g)(x),$$

with $U^{(n)}(1, \cdot; g)$ the terminal function (2.2). Moreover, let $X^{(n)}$ solve (2.8) with $\rho_n(r)$ in place of the coefficient r and $U^{(n)}$ in place of U , driven by X_0 and B . For every realization of g , the law of $X^{(n)}$ has the density

$$(A.3) \quad \begin{aligned} \Lambda_n(x; g) &:= \prod_{j=1}^{n-1} \exp\left(r_j [U^{(n)}(r_{j+1}, x_{r_{j+1}}; g) - U^{(n)}(r_j, x_{r_j}; g)]\right) \\ &= \exp\left(\left(1 - \frac{1}{n}\right) U(1, x_1; g) - \frac{1}{n} \sum_{j=1}^{n-1} U^{(n)}(r_j, x_{r_j}; g)\right) \end{aligned}$$

at the path x with respect to γ_q . In particular, the values $(X_{r_1}^{(n)}, \dots, X_{r_n}^{(n)})$ form a Markov chain, whose initial value $X_{r_1}^{(n)}$ is a centered Gaussian vector with covariance $2 \operatorname{diag}(q(r_1))$, and in which, for $1 \leq j \leq n-1$ and given $X_{r_j}^{(n)} = x$, the increment $X_{r_{j+1}}^{(n)} - x$ has density

$$(A.4) \quad \exp\left(r_j [U^{(n)}(r_{j+1}, x + z; g) - U^{(n)}(r_j, x; g)]\right)$$

at z with respect to the centered Gaussian law with covariance $2 \operatorname{diag}(q(r_{j+1}) - q(r_j))$. Conditionally on these values, $X^{(n)}$ has the conditional law of the Gaussian path Z given the same values.

Proof. On $[r_j, r_{j+1}]$ the coefficient is the constant r_j , and for $j \geq 1$ the function $\exp(r_j U^{(n)})$ solves the backward heat equation $\partial_r w + \operatorname{tr}(D \nabla^2 w) = 0$, whose solution is the Gaussian convolution with covariance $2 \int_{r_j}^{r_{j+1}} D = 2 \operatorname{diag}(q(r_{j+1}) - q(r_j))$. For $j = 0$ the equation for $U^{(n)}$ itself is this heat equation. This proves (A.2).

For the density, we repeat the Girsanov computation of the proof of Lemma 5.2 with the coefficient ρ_n . Given g , the frozen branch has the bounded drift $2\rho_n D \nabla U^{(n)}$ relative to the Gaussian path, so its law has density

$$\exp\left(\int_0^1 \rho_n(r) \nabla U^{(n)} \cdot \sqrt{2D} d\tilde{B}_r - \int_0^1 \rho_n(r)^2 \nabla U^{(n)} \cdot D \nabla U^{(n)} dr\right)$$

with respect to γ_q , where $dx_r = \sqrt{2D(r)} d\tilde{B}_r$ under γ_q , and where $U^{(n)}$ is evaluated at $(r, x_r; g)$. Under γ_q , Itô's formula and the frozen equation give $dU^{(n)}(r, x_r; g) = -\rho_n(r) \nabla U^{(n)} \cdot D \nabla U^{(n)} dr + \nabla U^{(n)} \cdot \sqrt{2D} d\tilde{B}_r$ on each interval between consecutive mesh points and breakpoints of q , and $U^{(n)}$ is continuous across these points. Since $\rho_n = r_j$ on $[r_j, r_{j+1})$ and $r_0 = 0$, the exponent equals

$$\sum_{j=1}^{n-1} r_j [U^{(n)}(r_{j+1}, x_{r_{j+1}}; g) - U^{(n)}(r_j, x_{r_j}; g)],$$

which is the first expression in (A.3); the second follows by summation by parts, using $r_j - r_{j-1} = 1/n$ and $U^{(n)}(1, \cdot; g) = U(1, \cdot; g)$. Under γ_q , the value $x_{r_1} = X_0 + (Z_{r_1} - Z_0)$ is a centered Gaussian vector with covariance $2 \text{diag}(q(r_1))$, and the increments $x_{r_{j+1}} - x_{r_j}$, $1 \leq j \leq n-1$, are independent of it and of each other, centered Gaussian with covariances $2 \text{diag}(q(r_{j+1}) - q(r_j))$. The density Λ_n depends on the path only through its values at the mesh points, and it is a product of functions of two consecutive values, each of which integrates to one in the later value by (A.2). The two remaining assertions follow. $\square$

The recursion (A.2) is the one that computes the free energy of a cascade with $n-1$ levels, and the transition (A.4) is the law of the increment of the field along a leaf of this cascade selected according to the Gibbs weights; this is the content of Lemma A.3 below. We first record the analytic estimate that controls the effect of freezing the coefficient r .

Lemma A.2 (Convergence of the frozen problem). *Let $\|\dot{q}\|_\infty$ denote the largest slope of q over all pieces and species. For every $n \geq 1$ and every realization of g ,*

$$(A.5) \quad \sup_{[0,1] \times \mathbb{R}^N} |U^{(n)} - U| \leq \frac{N \|\dot{q}\|_\infty}{n}.$$

Proof. Since $U^{(n)}$ solves the frozen equation, the operator (2.6) gives

$$\mathcal{L}[U^{(n)}] = (r - \rho_n(r)) \nabla U^{(n)} \cdot D(r) \nabla U^{(n)} \in \left[0, \frac{N \|\dot{q}\|_\infty}{n}\right],$$

because $0 \leq r - \rho_n(r) \leq 1/n$ and $|\nabla U^{(n)}|^2 \leq N$, while $\mathcal{L}[U] = 0$ and the two functions have the same terminal datum. The estimate (2.7) of Lemma 2.3, applied on the pieces delimited by the breakpoints of q and the mesh points, gives (A.5). $\square$

A.2. Finite-level cascades and the selected leaf. Consider a step path $\bar{q}$ taking the value $q_j \in \mathbb{R}_+^\mathcal{S}$ on $[r_{j-1}, r_j)$, $1 \leq j \leq n$, where $0 = r_0 < r_1 < \dots < r_n = 1$ is an arbitrary partition and $q_1 \leq \dots \leq q_n$ coordinatewise. Let $(v_\alpha)_{\alpha \in \mathbb{N}^{n-1}}$ be a Poisson–Dirichlet cascade with parameters $r_1 < \dots < r_{n-1}$, independent of g : for $n = 1$, it has a single leaf $\alpha = \emptyset$ of weight one, and in general the weights are the normalized products, along the branches of the tree, of the points of independent Poisson–Dirichlet point processes, with parameter r_j at depth j ; see [27, Section 2.3] or [13, Section 5.6]. For a leaf $\alpha = (\alpha_1, \dots, \alpha_{n-1})$ and $0 \leq j \leq n-1$, let $\alpha_{|j} := (\alpha_1, \dots, \alpha_j)$ denote its ancestor at depth j , with $\alpha_{|0} = \emptyset$ the root. We attach to every vertex β of the tree $\bigcup_{j=0}^{n-1} \mathbb{N}^j$ an independent standard Gaussian vector z_β of $\mathbb{R}^N$, independent of the weights and of g and which we call a mark, and we define the field at a vertex β of depth j by

$$(A.6) \quad x_\beta := \sqrt{2 \text{diag}(q_1)} z_\emptyset + \sum_{\ell=1}^j \sqrt{2 \text{diag}(q_{\ell+1} - q_\ell)} z_{\beta_{|\ell}}.$$

Thus the fields $(x_{\alpha_{|0}}, \dots, x_{\alpha_{|n-1}})$ along the branch leading to the leaf α have independent centered Gaussian increments, $x_{\alpha_{|j}}$ has covariance $2 \text{diag}(q_{j+1})$, and the fields $x_\alpha, x_{\alpha'}$ at

two leaves whose most recent common ancestor has depth j have covariance $2 \operatorname{diag}(q_{j+1}) = 2 \operatorname{diag}(\bar{q}(r_j))$. The cascade free energy at $\bar{q}$ is

$$\begin{aligned} \bar{F}_{N,\Omega}^{\text{cas}}(t, \bar{q}) &:= t\xi(\lambda_N) + \lambda_N \cdot q_n - \frac{1}{N} \mathbb{E} \log \sum_{\alpha} \mathbf{v}_{\alpha} \sum_{\sigma \in \Omega} w(\sigma) e^{\sqrt{2t} g \cdot v_{\sigma} + x_{\alpha} \cdot \sigma} \\ (A.7) \quad &= t\xi(\lambda_N) + \lambda_N \cdot q_n - \frac{1}{N} \mathbb{E} \log \sum_{\alpha} \mathbf{v}_{\alpha} e^{U(1, x_{\alpha}; g)}. \end{aligned}$$

For the full cube without external field, this is the enriched free energy of [7] at the step path $\bar{q}$. There, the Poisson–Dirichlet cascade has a continuum of levels, with uniform overlap distribution, and the field is $\sqrt{2}$ times a Gaussian field indexed by the cascade with covariance $N\bar{q}(\alpha \wedge \alpha') \cdot R_N(\sigma, \sigma')$; for a step path, this continuous cascade may be replaced by the cascade with $n-1$ levels just described without changing the free energy, by [8, Proposition 4.6]. For two step paths $\bar{q}, \bar{q}'$, we have

$$(A.8) \quad |\bar{F}_{N,\Omega}^{\text{cas}}(t, \bar{q}) - \bar{F}_{N,\Omega}^{\text{cas}}(t, \bar{q}')| \leq \sum_{s \in \mathcal{J}} \lambda_{N,s} \|\bar{q}_s - \bar{q}'_s\|_{L^1},$$

by Gaussian integration by parts in the cascade, exactly as in [8, Proposition 5.1]: the derivative of the free energy in the direction of a path is the expected Gibbs average of the overlap of two replicas, evaluated at the level where their leaves separate, and each species overlap lies in $[-\lambda_{N,s}, \lambda_{N,s}]$; the argument does not depend on the set Ω or on the weights w . Hence $\bar{F}_{N,\Omega}^{\text{cas}}(t, \cdot)$ extends by continuity to $\mathcal{Q}_1^{\mathcal{J}}$, and, for the full cube without external field, the extension is the enriched free energy of [7], which is defined on $\mathcal{Q}_1^{\mathcal{J}}$ by the same continuity.

Given g and the cascade, the Gibbs measure of the enriched Hamiltonian on the leaves and the configurations has weights proportional to $\mathbf{v}_{\alpha} w(\sigma) e^{\sqrt{2t} g \cdot v_{\sigma} + x_{\alpha} \cdot \sigma}$; the self-overlap corrections in (A.7) do not affect it. Its marginal on the leaves is

$$(A.9) \quad \tilde{\mathbf{v}}_{\alpha} := \frac{\mathbf{v}_{\alpha} e^{U(1, x_{\alpha}; g)}}{\sum_{\beta} \mathbf{v}_{\beta} e^{U(1, x_{\beta}; g)}},$$

and, given the leaf α , the configuration has the law

$$(A.10) \quad \mathbb{P}(\sigma = \tau \mid g, \text{weights, marks, } \alpha) = w(\tau) \exp(\sqrt{2t} g \cdot v_{\tau} + x_{\alpha} \cdot \tau - U(1, x_{\alpha}; g)), \quad \tau \in \Omega,$$

which is (3.1) with X_1 replaced by x_{α} . We say that the leaf α is selected according to the Gibbs weights when it is sampled from (A.9), and we write $\mathbb{P}^{\text{cas}}$ and $\mathbb{E}^{\text{cas}}$ for the probability and the expectation over the weights, the marks, the selected leaf and the configuration, the realization of g being fixed; this law depends on g and on the step path, and we keep this dependence implicit. The law of the fields along the selected leaf under $\mathbb{P}^{\text{cas}}$ is thus averaged over the environment (by “environment” we mean the weights and the marks); for a fixed realization of the environment, the selected leaf has a discrete law.

Lemma A.3 (The selected leaf of a finite-level cascade). *Let q be regular, let $n \geq 1$, $r_j := j/n$, and let $\bar{q}^{(n)}$ be the step path with the values $q_j := q(r_j)$ on $[r_{j-1}, r_j]$, $1 \leq j \leq n$. Consider the cascade associated with $\bar{q}^{(n)}$, a leaf α selected according to the Gibbs weights, and a configuration σ drawn from (A.10). For every realization of g :*

- (1) $\mathbb{E}^{\text{cas}} \log \sum_{\alpha} \mathbf{v}_{\alpha} e^{U(1, x_{\alpha}; g)} = \mathbb{E}[U^{(n)}(0, X_0; g) \mid g]$, where $X_0 \sim \mathcal{N}(0, 2 \operatorname{diag}(q(0)))$ is the initial value of the branch, independent of g ;
- (2) under $\mathbb{E}^{\text{cas}}$, the fields $(x_{\alpha_0}, \dots, x_{\alpha_{|n-1}})$ along the selected leaf and the configuration σ have the law of $(X_{r_1}^{(n)}, \dots, X_{r_n}^{(n)}, \sigma^{(n)})$, where $X^{(n)}$ is the frozen branch of Lemma A.1, and where $\sigma^{(n)}$ is drawn, given $X^{(n)}$, from the Gibbs measure (3.1) in the terminal field $X_1^{(n)}$.

Proof. We use the invariance of Poisson–Dirichlet cascades under a tilt of the weights by a function of the marks, [31, Lemma 3], through its two consequences [31, (3.13) and Theorem 6]; the first one is the recursive formula of [13, Theorem 5.25], and the formulation with a mark at the root that we use is that of [9, Appendix A]. Fix g and the root mark $z_\emptyset$, and regard $\Theta_\alpha := U(1, x_\alpha; g)$ as a function of the marks $(z_{\alpha_1}, \dots, z_{\alpha_{|n-1}})$ along the leaf α . This function is Lipschitz, since $|\nabla U(1, \cdot; g)| \leq \sqrt{N}$, so that e^{Θ_α} is integrable, as required in [31]. With the parameters $m_j = r_j$, the recursion (3.4) of [31] sets $\Theta_{\alpha_{|n-1}} := \Theta_\alpha$ and, for j from $n-1$ down to 1,

$$\Theta_{\alpha_{|j-1}} := \frac{1}{r_j} \log \mathbb{E}_j \exp(r_j \Theta_{\alpha_{|j}}),$$

where $\mathbb{E}_j$ integrates the mark $z_{\alpha_{|j}}$ only. Since $x_{\alpha_{|j}} = x_{\alpha_{|j-1}} + \sqrt{2 \operatorname{diag}(q(r_{j+1}) - q(r_j))} z_{\alpha_{|j}}$ by (A.6), a backward induction based on (A.2) gives

$$(A.11) \quad \Theta_{\alpha_{|j}} = U^{(n)}(r_{j+1}, x_{\alpha_{|j}}; g), \quad 0 \leq j \leq n-1,$$

and the weights $\mathbf{v}_\alpha e^{\Theta_\alpha} / \sum_\beta \mathbf{v}_\beta e^{\Theta_\beta}$ of the tilted cascade in [31] are the Gibbs weights (A.9). In particular, $\Theta_{\alpha_0} = U^{(n)}(r_1, x_\emptyset; g)$ depends on the root mark only, and [31, (3.13)] gives

$$\mathbb{E}^{\text{cas}} \left[\log \sum_\alpha \mathbf{v}_\alpha e^{U(1, x_\alpha; g)} \mid z_\emptyset \right] = U^{(n)}(r_1, x_\emptyset; g).$$

Since $x_\emptyset = \sqrt{2 \operatorname{diag}(q(r_1))} z_\emptyset$ has the law of $X_0 + \sqrt{2 \operatorname{diag}(q(r_1) - q(0))} z$, where z is a standard Gaussian vector independent of X_0 and of g , averaging over $z_\emptyset$ and using the case $j=0$ of (A.2), in which $T_{r_0, a} = T_{0, a}$ is a plain Gaussian expectation, gives (1). Next, [31, Theorem 6] states that, conditionally on $z_\emptyset$ and under $\mathbb{E}^{\text{cas}}$, the marks $(z_{\alpha_1}, \dots, z_{\alpha_{|n-1}})$ along the selected leaf have the law under which, for $1 \leq j \leq n-1$ and given the marks of smaller depth, $z_{\alpha_{|j}}$ has density $\exp(r_j [\Theta_{\alpha_{|j}} - \Theta_{\alpha_{|j-1}}])$ with respect to the standard Gaussian law. By (A.11) and (A.6), this means that, given $x_{\alpha_{|j-1}} = x$, the increment $x_{\alpha_{|j}} - x$ has density (A.4) with respect to the centered Gaussian law with covariance $2 \operatorname{diag}(q(r_{j+1}) - q(r_j))$, while $x_{\alpha_0} = x_\emptyset$ is a centered Gaussian vector with covariance $2 \operatorname{diag}(q(r_1))$. Thus the fields along the selected leaf form the Markov chain of Lemma A.1, and $(x_{\alpha_0}, \dots, x_{\alpha_{|n-1}})$ has the law of $(X_{r_1}^{(n)}, \dots, X_{r_n}^{(n)})$. Finally, the conditional law (A.10) of the configuration depends on the environment and on the leaf only through $x_\alpha = x_{\alpha_{|n-1}}$, and it is the conditional law of $\sigma^{(n)}$ given $X_1^{(n)} = x_\alpha$. This proves (2). $\square$

By Lemmas A.1 and A.3, the frozen branch $X^{(n)}$ is, in law and after averaging over the environment of the cascade, the Gaussian field along a leaf of the cascade with $n-1$ levels selected according to the Gibbs weights, placed at the mesh points and interpolated between them according to the conditional law of the Gaussian path Z . Its sampled configuration $\sigma^{(n)}$ is the configuration drawn from the Gibbs measure of the selected leaf.

We now select two leaves. Given g and the environment, let (α^1, σ^1) and (α^2, σ^2) be two independent samples from the Gibbs measure on the leaves and the configurations, so that each α^a has the law (A.9), and each σ^a is drawn from (A.10) with $\alpha = \alpha^a$; we let $\mathbb{E}^{\text{cas}}$ and $\mathbb{P}^{\text{cas}}$ also integrate the second replica. Let

$$(A.12) \quad J := \max \{ j \in \{0, \dots, n-1\} : \alpha_{|j}^1 = \alpha_{|j}^2 \}$$

be the depth of the most recent common ancestor of the two selected leaves, so that $J = n-1$ exactly when $\alpha^1 = \alpha^2$, and the fields $x_{\alpha^1}, x_{\alpha^2}$ have covariance $2 \operatorname{diag}(\bar{q}^{(n)}(r_J))$. In the cascade, the two replicas thus separate at the level r_J . On the side of the branch, we define the Gibbs averages of the frozen problem exactly as in Section 3: $\sigma^{(n)}$ is drawn,

given $X^{(n)}$, from the Gibbs measure (3.1) in the terminal field $X_1^{(n)}$, as in Lemma A.3, and for $u \in [0, 1]$ and a function b of two configurations, $\langle b \rangle_u^{(n)}$ is the integral of b against the twofold product of the conditional law of $\sigma^{(n)}$ given $\mathcal{G}_u$, where $\mathcal{G}_u$ is generated by g , X_0 and $(B_v)_{v \leq u}$. Since the frozen branch is a Markov process, this conditional law is a function of $X_u^{(n)}$ and g , and $\mathbb{E}[\langle b \rangle_u^{(n)} \mid g] = \mathbb{E}[b(\sigma^{(n),1}, \sigma^{(n),2}) \mid g]$, where $\sigma^{(n),1}, \sigma^{(n),2}$ are drawn independently from the Gibbs measures in the terminal fields of two frozen branches that coincide up to the level u and are driven by independent Brownian motions afterwards. Precisely, $X^{(n),1} := X^{(n)}$, and $X^{(n),2}$ solves the frozen equation with the same initial value X_0 , driven by the Brownian motion that coincides with B on $[0, u]$ and whose increments after u are those of an independent Brownian motion; the two solutions coincide on $[0, u]$ by pathwise uniqueness. This is the description of two replicas at level u given in Section 3.1.

Lemma A.4 (Two leaves selected from a finite-level cascade). *Consider the setting of Lemma A.3, and let $0 \leq j \leq n-1$. For every realization of g :*

- (1) $\mathbb{P}^{\text{cas}}(J = j) = 1/n$;
- (2) *conditionally on $J = j$ and under $\mathbb{E}^{\text{cas}}$, the fields along the two selected leaves and the two configurations, that is, the two families $(x_{\alpha_0^a}, \dots, x_{\alpha_{n-1}^a}, \sigma^a)$, $a = 1, 2$, have the law of the two families $(X_{r_1}^{(n),a}, \dots, X_{r_n}^{(n),a}, \sigma^{(n),a})$, $a = 1, 2$, where the two frozen branches coincide up to the level r_{j+1} and are driven by independent Brownian motions afterwards, and where the configurations are drawn independently in their terminal fields, as above;*
- (3) *consequently, for every $b : \Omega^2 \rightarrow \mathbb{R}$,*

$$(A.13) \quad \mathbb{E}^{\text{cas}}[b(\sigma^1, \sigma^2) \mathbf{1}_{\{J=j\}}] = \frac{1}{n} \mathbb{E}[\langle b \rangle_{r_{j+1}}^{(n)} \mid g].$$

Proof. We keep the notation of the proof of Lemma A.3: g and the root mark $z_\emptyset$ are fixed, the recursion (3.4) of [31] with the parameters $m_\ell = r_\ell$ produces $\Theta_{\alpha_{|\ell}} = U^{(n)}(r_{\ell+1}, x_{\alpha_{|\ell}}; g)$, the weights of the tilted cascade are the Gibbs weights (A.9), and the factor $W_{\alpha_{|\ell}} := \exp(r_\ell[\Theta_{\alpha_{|\ell}} - \Theta_{\alpha_{|\ell-1}}])$ of [31, (3.5)] is, by (A.11) and (A.6), the density (A.4) of the increment $x_{\alpha_{|\ell}} - x_{\alpha_{|\ell-1}}$ given $x_{\alpha_{|\ell-1}}$, with respect to its Gaussian law. In [31], the overlap of two distinct leaves is the index of the first coordinate at which they differ, which is $J+1$ in our notation. Let Y be a bounded function of the marks along two leaves. For $0 \leq j \leq n-2$, [31, Theorem 7] gives

$$(A.14) \quad \mathbb{E}^{\text{cas}}[Y \mathbf{1}_{\{J=j\}}] = (r_{j+1} - r_j) \mathbb{E}\left[\prod_{\ell=1}^j W_{\alpha_{|\ell}} \prod_{\ell=j+1}^{n-1} W_{\alpha_{|\ell}}^1 W_{\alpha_{|\ell}}^2 Y\right],$$

where, on the right, α^1 and α^2 are two fixed leaves whose most recent common ancestor $\alpha_{|j} := \alpha_{|j}^1 = \alpha_{|j}^2$ has depth j , and the marks are independent standard Gaussian vectors, the two leaves sharing the marks of depth at most j . For $j = n-1$, the invariance [31, (3.14)] of the pair formed by the weights and the marks, which is the content of [31, Lemma 3], gives

$$(A.15) \quad \mathbb{E}^{\text{cas}}[Y \mathbf{1}_{\{\alpha^1 = \alpha^2\}}] = \mathbb{E}\left[\sum_{\alpha} v_{\alpha}^2\right] \mathbb{E}\left[\prod_{\ell=1}^{n-1} W_{\alpha_{|\ell}} Y\right],$$

with a single fixed leaf α on the right, since the weights and the marks of the second process in [31, (3.14)] are independent; and $\mathbb{E} \sum_{\alpha} v_{\alpha}^2 = 1 - r_{n-1} = 1/n$, by [31, (3.17)] summed over the overlap. Since $r_{j+1} - r_j = 1/n$, taking $Y \equiv 1$ in (A.14) and (A.15) proves (1).

The density on the right of (A.14), multiplied by the Gaussian law of the marks and read through (A.6), is the following law of the fields along the two leaves: the common fields $(x_{\alpha_{|0}}, \dots, x_{\alpha_{|j}})$ have the law of $(X_{r_1}^{(n)}, \dots, X_{r_{j+1}}^{(n)})$, since $x_{\alpha_{|0}} = x_{\emptyset}$ has the law of $X_{r_1}^{(n)}$ once $z_{\emptyset}$ is averaged, and since the factors $W_{\alpha_{|\ell}}$, $\ell \leq j$, are the transition densities (A.4) of the chain of Lemma A.1; and, given the common fields, the two families $(x_{\alpha_{|j+1}}^a, \dots, x_{\alpha_{|n-1}}^a)$, $a = 1, 2$, are independent, each with the conditional law of $(X_{r_{j+2}}^{(n)}, \dots, X_{r_n}^{(n)})$ given $X_{r_{j+1}}^{(n)} = x_{\alpha_{|j}}$, by the same transition densities, each factor integrating to one in the next mark by (A.2). On the side of the branch, $X^{(n),1}$ and $X^{(n),2}$ coincide up to the level r_{j+1} , and, given $\mathcal{G}_{r_{j+1}}$, their values at the later mesh points are two independent copies of the chain started from $X_{r_{j+1}}^{(n)}$, by the Markov property of the frozen equation and the independence of the two driving noises after r_{j+1} . The laws of the fields in (2) thus agree. The conditional law (A.10) of each configuration depends on the environment and on its leaf only through the terminal field of that leaf, and the two configurations are drawn independently; the same holds for $\sigma^{(n),1}, \sigma^{(n),2}$. This proves (2), and (3) follows from (1), (2) and the identity $\mathbb{E}[\langle b \rangle_u^{(n)} | g] = \mathbb{E}[b(\sigma^{(n),1}, \sigma^{(n),2}) | g]$ recalled before the statement. $\square$

A.3. Convergence to the branch. We now let the number of levels tend to infinity.

Proposition A.5 (The branch as the limit of the selected leaves). *Let q be regular, and for $n \geq 1$ let α be a leaf selected according to the Gibbs weights and σ the configuration drawn at this leaf, as in Lemma A.3. For every realization of g , the law under $\mathbb{P}^{\text{cas}}$ of the fields along the selected leaf and of its configuration converges in total variation, as $n \rightarrow \infty$, to the conditional law given g of the values of the branch X at the mesh points $r_1, \dots, r_n$ and of its sampled configuration (3.1), that is,*

$$(A.16) \quad \lim_{n \rightarrow \infty} \sup_f \left| \mathbb{E}^{\text{cas}}[f(x_{\alpha_{|0}}, \dots, x_{\alpha_{|n-1}}, \sigma)] - \mathbb{E}[f(X_{r_1}, \dots, X_{r_n}, \sigma) | g] \right| = 0,$$

where the supremum is over all measurable $f : (\mathbb{R}^N)^n \times \Omega \rightarrow [-1, 1]$. The statement remains true after averaging over g , with f then also allowed to depend on g .

Proof. Let $X^{(n)}$ and $\sigma^{(n)}$ be the frozen branch of Lemma A.1 and its sampled configuration, as in Lemma A.3(2), so that the law of $(x_{\alpha_{|0}}, \dots, x_{\alpha_{|n-1}}, \sigma)$ under $\mathbb{P}^{\text{cas}}$ is the conditional law of $(X_{r_1}^{(n)}, \dots, X_{r_n}^{(n)}, \sigma^{(n)})$ given g . The two laws compared in (A.16) are thus the images of the conditional laws of $(X^{(n)}, \sigma^{(n)})$ and of (X, σ) given g under the map $(x, \sigma) \mapsto (x_{r_1}, \dots, x_{r_n}, \sigma)$, and it suffices to show that the conditional law of $(X^{(n)}, \sigma^{(n)})$ given g converges in total variation to the conditional law of (X, σ) given g .

The proof of Lemma 5.2 uses neither the slice nor the value of the weights, and the Girsanov density there is computed at fixed g : for every (Ω, w) and every realization of g , the conditional law of (X, σ) given g has density

$$(A.17) \quad p(g, x, \sigma) := w(\sigma) \exp \left(\sqrt{2t} g \cdot v_{\sigma} + x_1 \cdot \sigma - \int_0^1 U(r, x_r; g) dr \right)$$

at (x, σ) with respect to the product of γ_q and the counting measure on Ω . Likewise, multiplying the density (A.3) of the frozen branch by the Gibbs probability (3.1) of $\sigma^{(n)}$ given $X_1^{(n)}$, which contains the factor $e^{-U(1, x_1; g)}$, we see that the conditional law of $(X^{(n)}, \sigma^{(n)})$ given g has density

$$(A.18) \quad p_n(g, x, \sigma) := w(\sigma) \exp \left(\sqrt{2t} g \cdot v_{\sigma} + x_1 \cdot \sigma - \frac{1}{n} \sum_{j=1}^n U^{(n)}(r_j, x_{r_j}; g) \right)$$

with respect to the same reference measure, the term $j = n$ of the sum being $U(1, x_1; g)$. For every g , every continuous path x and every σ , we have $p_n(g, x, \sigma) \rightarrow p(g, x, \sigma)$ as $n \rightarrow \infty$: by Lemma A.2, the sum in (A.18) differs by at most $N\|\dot{q}\|_\infty/n$ from $\frac{1}{n} \sum_{j=1}^n U(r_j, x_{r_j}; g)$, which is a Riemann sum of the function $r \mapsto U(r, x_r; g)$, continuous on $[0, 1]$ by Lemma 2.1. For every g , both $p_n(g, \cdot, \cdot)$ and $p(g, \cdot, \cdot)$ are probability densities with respect to the same measure, so Scheffé's lemma gives

$$(A.19) \quad \varepsilon_n(g) := \int \sum_{\sigma \in \Omega} |p_n(g, x, \sigma) - p(g, x, \sigma)| \gamma_q(dx) \rightarrow 0,$$

and $\varepsilon_n(g)$ is twice the total variation distance between the conditional laws of $(X^{(n)}, \sigma^{(n)})$ and (X, σ) given g . The supremum in (A.16) is therefore at most $\varepsilon_n(g)$, and tends to zero. Since γ is the product of the law of g and of γ_q , the total variation distance between the laws of $(g, X^{(n)}, \sigma^{(n)})$ and (g, X, σ) is $\frac{1}{2}\mathbb{E}\varepsilon_n(g)$, which tends to 0 by dominated convergence, since $\varepsilon_n \leq 2$; this gives the statement averaged over g . $\square$

We now turn to the two replicas of Lemma A.4, and let the number of levels tend to infinity. The conditional law of the two configurations given the depth of their common ancestor converges, uniformly in the depth, to the law of two replicas at the corresponding level of the branch.

Proposition A.6 (Two replicas at level r as conditioned cascade replicas). *Let q be regular, let $b : \Omega^2 \rightarrow \mathbb{R}$, and for $n \geq 1$ let (α^1, σ^1) , (α^2, σ^2) and J be as in Lemma A.4. For every realization of g ,*

$$(A.20) \quad \lim_{n \rightarrow \infty} \max_{0 \leq j \leq n-1} \sup_{u \in [r_j, r_{j+1}]} \left| \mathbb{E}^{\text{cas}}[b(\sigma^1, \sigma^2) \mid J = j] - \mathbb{E}[\langle b \rangle_u \mid g] \right| = 0,$$

and consequently, for every continuous $\varphi : [0, 1] \rightarrow \mathbb{R}$,

$$(A.21) \quad \lim_{n \rightarrow \infty} \mathbb{E}^{\text{cas}}[\varphi(r_J) b(\sigma^1, \sigma^2)] = \int_0^1 \varphi(r) \mathbb{E}[\langle b \rangle_r \mid g] dr.$$

Both statements remain true after averaging over g ; since $\mathbb{P}^{\text{cas}}(J = j) = 1/n$ does not depend on g , the averaged conditional expectation in (A.20) is the conditional expectation given $J = j$ under the law that also averages over g .

Proof. Fix a realization of g , and let $\delta_n(g)$ denote the supremum in (A.16), so that $\delta_n(g) \rightarrow 0$ by Proposition A.5. We proceed in three steps.

Step 1: stability of two replicas. For a signed measure ρ , let $\|\rho\|$ denote the total mass of $|\rho|$, so that $|\int b d\rho| \leq \|b\|_\infty \|\rho\|$. Let μ, μ' be probability measures on $\mathbb{R}^N$, and let $\mathbf{k}, \mathbf{k}'$ be probability kernels from $\mathbb{R}^N$ to Ω . We claim that

$$(A.22) \quad \|\mu(dx) \mathbf{k}(x, \cdot)^{\otimes 2} - \mu'(dx) \mathbf{k}'(x, \cdot)^{\otimes 2}\| \leq 5 \|\mu(dx) \mathbf{k}(x, \cdot) - \mu'(dx) \mathbf{k}'(x, \cdot)\|.$$

Indeed, for probability measures ρ, ρ' on Ω , we have $\|\rho^{\otimes 2} - \rho'^{\otimes 2}\| \leq \|(\rho - \rho') \otimes \rho\| + \|\rho' \otimes (\rho - \rho')\| = 2\|\rho - \rho'\|$. Hence the left side of (A.22) is at most

$$\|\mu - \mu'\| + \int \mu'(dx) \|\mathbf{k}(x, \cdot)^{\otimes 2} - \mathbf{k}'(x, \cdot)^{\otimes 2}\| \leq \|\mu - \mu'\| + 2 \|\mu'(dx) \mathbf{k}(x, \cdot) - \mu'(dx) \mathbf{k}'(x, \cdot)\|,$$

and the last norm is at most $\|\mu' - \mu\|$ plus the right side of (A.22) without the factor 5. Since $\|\mu - \mu'\|$ is the norm of the difference of the first marginals, it is also at most this right side, and (A.22) follows.

Step 2: uniform convergence of the two-replica averages. For $u \in [0, 1]$, let μ_u be the conditional law of X_u given g . By the Markov property of the branch, the conditional law of σ given $\mathcal{G}_u$ is a function of X_u : there is a probability kernel $\mathbf{k}_u$ from $\mathbb{R}^N$ to Ω , namely $\mathbf{k}_u(x, \tau)$ is the probability that the configuration sampled in the terminal field of the

branch started from (u, x) equals τ , such that $\mathbb{P}(\sigma = \tau \mid \mathcal{G}_u) = k_u(X_u, \tau)$ for every $\tau \in \Omega$. By (3.7) applied to $b = \mathbf{1}_{\{\sigma=\tau\}}$ and the proof of Lemma 3.1, $k_u(x, \tau)$ is the value at (u, x) of the solution of a linear parabolic equation, and $(u, x) \mapsto k_u(x, \tau)$ is continuous and bounded on $[0, 1] \times \mathbb{R}^N$. Hence the conditional law of (X_u, σ) given g is $\mu_u(dx) k_u(x, \cdot)$, and, by the definition of $\langle \cdot \rangle_u$,

$$\mathbb{E}[\langle b \rangle_u \mid g] = \int \mu_u(dx) \sum_{\tau, \tau' \in \Omega} b(\tau, \tau') k_u(x, \tau) k_u(x, \tau').$$

The same holds for the frozen branch, with the conditional law $\mu_u^{(n)}$ of $X_u^{(n)}$ given g and the kernel $k_u^{(n)}$ defined with the frozen equation. At a mesh point $u = r_{j+1}$, $0 \leq j \leq n-1$, the conditional law of $(X_u^{(n)}, \sigma^{(n)})$ given g is the law of (x_{α_j}, σ) under $\mathbb{P}^{\text{cas}}$, by Lemma A.3(2), so the conditional laws of $(X_u^{(n)}, \sigma^{(n)})$ and (X_u, σ) given g are the images, under the projection onto the $(j+1)$ -th field and the configuration, of the two laws compared in (A.16). The norm of their difference is thus at most $\delta_n(g)$, and (A.22) gives

$$(A.23) \quad \max_{1 \leq j \leq n} \left| \mathbb{E}[\langle b \rangle_{r_j}^{(n)} \mid g] - \mathbb{E}[\langle b \rangle_{r_j} \mid g] \right| \leq 5 \|b\|_\infty \delta_n(g).$$

Step 3: conclusion. Since $(u, x) \mapsto k_u(x, \tau)$ is continuous and bounded for every τ and $u \mapsto X_u$ is continuous, the function $u \mapsto \mathbb{E}[\langle b \rangle_u \mid g] = \mathbb{E}[\sum_{\tau, \tau'} b(\tau, \tau') k_u(X_u, \tau) k_u(X_u, \tau') \mid g]$ is continuous on $[0, 1]$, by dominated convergence. Let ω_g denote its modulus of continuity. By Lemma A.4, $\mathbb{E}^{\text{cas}}[b(\sigma^1, \sigma^2) \mid J = j] = \mathbb{E}[\langle b \rangle_{r_{j+1}}^{(n)} \mid g]$, so that, for $u \in [r_j, r_{j+1}]$, (A.23) at the mesh point r_{j+1} gives

$$\left| \mathbb{E}^{\text{cas}}[b(\sigma^1, \sigma^2) \mid J = j] - \mathbb{E}[\langle b \rangle_u \mid g] \right| \leq 5 \|b\|_\infty \delta_n(g) + \omega_g(1/n),$$

which proves (A.20). Next, since $\mathbb{P}^{\text{cas}}(J = j) = 1/n$, we have $\mathbb{E}^{\text{cas}}[\varphi(r_J) b(\sigma^1, \sigma^2)] = \frac{1}{n} \sum_{j=0}^{n-1} \varphi(r_j) \mathbb{E}^{\text{cas}}[b(\sigma^1, \sigma^2) \mid J = j]$, so the previous display gives

$$\left| \mathbb{E}^{\text{cas}}[\varphi(r_J) b(\sigma^1, \sigma^2)] - \frac{1}{n} \sum_{j=0}^{n-1} \varphi(r_j) \mathbb{E}[\langle b \rangle_{r_j} \mid g] \right| \leq \|\varphi\|_\infty (5 \|b\|_\infty \delta_n(g) + \omega_g(1/n)),$$

and the Riemann sum converges to the right side of (A.21), the integrand being continuous. Finally, all the quantities in (A.20) and (A.21) are bounded by $\|\varphi\|_\infty \|b\|_\infty$ (with $\varphi \equiv 1$ in the first case), so the statements averaged over g follow by dominated convergence. $\square$

Remark A.7 (Disintegration of the derivative formulas). For the cascade with $n-1$ levels at the step path $\bar{q}^{(n)}$, a Gaussian integration by parts gives, as for (A.8), $\partial_t \bar{F}_{N, \Omega}^{\text{cas}}(t, \bar{q}^{(n)}) = \mathbb{E} \mathbb{E}^{\text{cas}}[\xi(R_N(\sigma^1, \sigma^2))]$: the derivatives of the cascade free energy are expectations of two-replica observables under the Gibbs measure, with no conditioning on the separation level. Taking $b = \xi(R_N)$ and $\varphi \equiv 1$ in (A.21), we see that the limit of this expression is $\int_0^1 \mathbb{E}(\xi(R_N))_r dr = \partial_t \bar{F}_{N, \Omega}(t, q)$. Similarly, taking $b = R_{N, s}$ in (A.20), we obtain, for every $r \in [0, 1]$ and every species s ,

$$(A.24) \quad p_{N, s}(r) = \mathbb{E}(R_{N, s})_r = \lim_{n \rightarrow \infty} \mathbb{E} \mathbb{E}^{\text{cas}}[R_{N, s}(\sigma^1, \sigma^2) \mid J = \min(\lfloor nr \rfloor, n-1)].$$

In words, the quantity $p_N(r)$ is the limit of the mean overlap of two cascade replicas conditioned on separating at level r . Theorem 4.1 thus disintegrates the derivative formulas of the cascade, in which the separation level is uniformly distributed, according to the value of this level.

A.4. Agreement of the free energies.

Proposition A.8 (Agreement with the cascade). *For every $t \geq 0$ and $q \in \mathcal{Q}_1^{\mathcal{J}}$, we have*

$$\overline{F}_{N,\Omega}(t, q) = \overline{F}_{N,\Omega}^{\text{cas}}(t, q).$$

Proof. Let q be regular, and let $\bar{q}^{(n)}$ be as in Lemma A.3. Averaging Lemma A.3(1) over g , and since $q_n = q(1)$, we obtain

$$\overline{F}_{N,\Omega}^{\text{cas}}(t, \bar{q}^{(n)}) = t\xi(\lambda_N) + \lambda_N \cdot q(1) - \frac{1}{N} \mathbb{E} U^{(n)}(0, X_0; g).$$

As $n \rightarrow \infty$, the right side converges to $\overline{F}_{N,\Omega}(t, q)$ by Lemma A.2 and (2.13). On the left side, $0 \leq \bar{q}^{(n)} - q \leq q(r_{j+1}) - q(r_j)$ on $[r_j, r_{j+1})$, so that $\|\bar{q}_s^{(n)} - q_s\|_{L^1} \leq n^{-1}(q_s(1) - q_s(0))$ for every species s , and the left side converges to $\overline{F}_{N,\Omega}^{\text{cas}}(t, q)$ by (A.8). This proves the proposition for regular paths. The general case follows from the density of regular paths in $\mathcal{Q}_1^{\mathcal{J}}$, by Proposition 4.2 and (A.8). The identity concerns the corrected free energies as a whole: the terminal value $q(1)$ alone is not continuous in L^1 , and it may be infinite for a general path in $\mathcal{Q}_1^{\mathcal{J}}$. $\square$

Remark A.9 (Cost and entropy). For the finite-level cascade of Lemma A.3, the relative entropy of the Gibbs weights (A.9) with respect to the cascade weights is

$$(A.25) \quad \text{Ent}(\tilde{\nu} | \nu) := \sum_{\alpha} \tilde{\nu}_{\alpha} \log \frac{\tilde{\nu}_{\alpha}}{\nu_{\alpha}} = \sum_{\alpha} \tilde{\nu}_{\alpha} U(1, x_{\alpha}; g) - \log \sum_{\alpha} \nu_{\alpha} e^{U(1, x_{\alpha}; g)}.$$

Under $\mathbb{E}^{\text{cas}}$, the first term on the right of (A.25) has mean $\mathbb{E}[U(1, X_1^{(n)}; g) | g]$ by Lemma A.3(2), and the second has mean $\mathbb{E}[U^{(n)}(0, X_0; g) | g]$ by Lemma A.3(1). Applying Itô's formula to the frozen problem as in the proof of Lemma 2.5, we obtain, for every realization of g ,

$$(A.26) \quad \begin{aligned} \mathbb{E}^{\text{cas}} \text{Ent}(\tilde{\nu} | \nu) &= \mathbb{E}[U(1, X_1^{(n)}; g) - U^{(n)}(0, X_0; g) | g] \\ &= \mathbb{E}\left[\int_0^1 \rho_n(r) \nabla U^{(n)} \cdot D(r) \nabla U^{(n)}(r, X_r^{(n)}; g) dr \mid g\right], \end{aligned}$$

all terms being integrable, since $U(1, \cdot; g)$ has linear growth and $X_1^{(n)}$ has Gaussian tails. The right side of (A.26) is the cost A_q of Definition 2.4 for the frozen branch. As $n \rightarrow \infty$, it converges to $\mathbb{E}[A_q | g] = \mathbb{E}[U(1, X_1; g) - U(0, X_0; g) | g]$, by (2.12): the second term in (A.26) converges by Lemma A.2, and the first by Lemma A.3(2) and Proposition A.5, since $|U(1, x; g)| \leq |U(1, 0; g)| + \sqrt{N}|x|$ and the second moments of $X_1^{(n)}$ are bounded uniformly in n , the drift of the frozen branch being bounded by $2\sqrt{N}\|\dot{q}\|_{\infty}$. Thus the cost is, in expectation, the limit of the relative entropy of the Gibbs weights of the leaves with respect to the cascade weights.

REFERENCES

- [1] Peter M. Aronow and Patrick Lopatto. Quantitative Parisi formulas and fluctuations in the Sherrington–Kirkpatrick model. *Preprint, arXiv:2609.09103*, 2026.
- [2] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.
- [3] Adriano Barra, Pierluigi Contucci, Emanuele Mingione, and Daniele Tantari. Multi-species mean field spin glasses. Rigorous results. *Ann. Henri Poincaré*, 16(3):691–708, 2015.
- [4] Herm Jan Brascamp and Elliott H. 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. *J. Funct. Anal.*, 22:366–389, 1976.
- [5] Hong-Bin Chen. On free energy of non-convex multi-species spin glasses. *ALEA Lat. Am. J. Probab. Math. Stat.*, 23(1):429–473, 2026.

- [6] Hong-Bin Chen, Victor Issa, and Jean-Christophe Mourrat. The convex structure of the Parisi formula for multi-species spin glasses. *Preprint, arXiv:2508.06397*, 2025.
- [7] Hong-Bin Chen, Victor Issa, and Jean-Christophe Mourrat. Free energy of non-convex multi-species spin glasses with centered Ising spins. *Preprint, arXiv:2606.16636*, 2026.
- [8] Hong-Bin Chen and Jean-Christophe Mourrat. On the free energy of vector spin glasses with nonconvex interactions. *Probab. Math. Phys.*, 6(1):1–80, 2025.
- [9] Hong-Bin Chen and Jean-Christophe Mourrat. Free energy of non-convex multi-species spherical spin glasses. *Preprint, arXiv:2609.32710*, 2026.
- [10] Hong-Bin Chen and Jean-Christophe Mourrat. Simultaneous replica-symmetry breaking for vector spin glasses. *Comm. Math. Phys.*, 407(6):Paper No. 109, 2026.
- [11] Hong-Bin Chen and Jiaming Xia. Hamilton–Jacobi equations from mean-field spin glasses. *Probab. Theory Related Fields*, 192(3):803–873, 2025.
- [12] Yuansi Chen and Ronen Eldan. Localization schemes: a framework for proving mixing bounds for Markov chains. *Duke Math. J.*, 174(8):1431–1510, 2025.
- [13] Tomas Dominguez and Jean-Christophe Mourrat. *Statistical mechanics of mean-field disordered systems: a Hamilton–Jacobi approach*. Zurich Lectures in Advanced Mathematics. EMS Press, 2024.
- [14] Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.*, 23(2):532–569, 2013.
- [15] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.
- [16] Fu-Hsuan Ho. A concavity theorem for the Parisi PDE. *Preprint, arXiv:2606.15432*, 2026.
- [17] Fu-Hsuan Ho. Convergence of the free energy of vector spin glasses with non-convex interactions. *Preprint, arXiv:2609.39307*, 2026.
- [18] Victor Issa. A Hopf-like formula for mean-field spin glass models. *Comm. Math. Phys.*, 407(5):Paper No. 96, 2026.
- [19] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.
- [20] Olga A. Ladyzhenskaya, Vsevolod A. Solonnikov, and Nina N. Ural’tseva. *Linear and quasilinear equations of parabolic type*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1968.
- [21] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.
- [22] Jean-Christophe Mourrat. The Parisi formula is a Hamilton–Jacobi equation in Wasserstein space. *Canad. J. Math.*, 74(3):607–629, 2022.
- [23] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(3):1143–1182, 2023.
- [24] Jean-Christophe Mourrat. Un-inverting the Parisi formula. *Ann. Inst. Henri Poincaré Probab. Stat.*, 61(4):2709–2720, 2025.
- [25] Jean-Christophe Mourrat and Dmitry Panchenko. Extending the Parisi formula along a Hamilton–Jacobi equation. *Electron. J. Probab.*, 25:Paper No. 23, 17, 2020.
- [26] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.
- [27] Dmitry Panchenko. *The Sherrington–Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.
- [28] Dmitry Panchenko. The free energy in a multi-species Sherrington–Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.
- [29] Dmitry Panchenko. Free energy in the Potts spin glass. *Ann. Probab.*, 46(2):829–864, 2018.
- [30] Dmitry Panchenko. Free energy in the mixed p -spin models with vector spins. *Ann. Probab.*, 46(2):865–896, 2018.
- [31] Dmitry Panchenko and Michel Talagrand. Guerra’s interpolation using Derrida–Ruelle cascades. *Preprint, arXiv:0708.3641*, 2007.
- [32] Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.
- [33] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

(Jeanne Boursier) DEPARTMENT OF MATHEMATICS, MIT, CAMBRIDGE MA, USA

(Jean-Christophe Mourrat) DEPARTMENT OF MATHEMATICS, ENS LYON AND CNRS, LYON, FRANCE