

A FORMULA FOR THE FREE ENERGY OF NON-CONVEX VECTOR SPIN GLASSES

HONG-BIN CHEN , FU-HSUAN HO , AND VICTOR ISSA 

ABSTRACT. We consider vector mean-field spin glass models with independent and identically distributed spins without assuming the covariance function is convex. We identify the limit free energy as the minimum of a functional that can be written as the sum of a relative entropy and the solution of a Hamilton–Jacobi equation. The proof of this result relies on the observation that when the spin configurations of the system are constrained to have a given empirical measure, the corresponding free energy approximately satisfies a Hamilton–Jacobi equation with a uniformly bounded error.

1. INTRODUCTION

The model. Write $\mathbb{R}_+ := [0, \infty)$. Fix $D \in \mathbb{N}$, put $\mathbf{B} := \{a \in \mathbb{R}^D : |a| \leq 1\}$, and let μ be a probability measure supported on $\mathbf{B}$. For $N \in \mathbb{N}$, we view a spin configuration $\sigma \in (\text{supp } \mu)^N$ as a D -by- N matrix, with columns $\sigma_{\cdot, i} \in \text{supp } \mu$ for $i \in \{1, \dots, N\}$. For matrices or vectors a and b of the same dimension, let $a \cdot b := \text{tr}(ab^\top)$ and $|a| := (a \cdot a)^{1/2}$. The overlap between two configurations σ and τ is defined by

$$R_N(\sigma, \tau) := \frac{\sigma \tau^\top}{N}.$$

Let $\xi : \mathbb{R}^{D \times D} \rightarrow \mathbb{R}$ be smooth, admit a power-series expansion converging absolutely at every matrix in $\mathbb{R}^{D \times D}$, and satisfy $\xi(0) = 0$. For every $N \in \mathbb{N}$, we consider the Hamiltonian $H_N : \mathbb{R}^{D \times N} \rightarrow \mathbb{R}$ which is a centered Gaussian process with covariance

$$(1.1) \quad \mathbb{E} H_N(\sigma) H_N(\tau) = N \xi(R_N(\sigma, \tau)), \quad \sigma, \tau \in \mathbb{R}^{D \times N}.$$

These are the covariance assumptions in [19, Sections 1.1 and 1.5]. We do not assume convexity of ξ .

For $N \in \mathbb{N}$ and $t \in \mathbb{R}_+$, define the free energy

$$(1.2) \quad \bar{F}_{N, \mu^{\otimes N}}(t, 0) := -\frac{1}{N} \mathbb{E} \log \int \exp\left(\sqrt{2t} H_N(\sigma) - tN \xi(R_N(\sigma, \sigma))\right) \mu^{\otimes N}(d\sigma).$$

The self-overlap correction term $-tN \xi(R_N(\sigma, \sigma))$ is introduced as a convenience to simplify the expression of the free energy when differentiating with respect to t . Theorem 1.1 identifies the limiting free energy; the limit without this correction can be recovered by the procedure in [16].

The main result. Let S^D be the space of real symmetric D -by- D matrices and $S_+^D \subseteq S^D$ the cone of positive semi-definite matrices. Given $a, b \in S^D$, we write $a \leq b$ when $b - a \in S_+^D$. We denote by $\mathcal{Q}$ the set of right-continuous increasing paths $q : [0, 1) \rightarrow S_+^D$, where increasing means that $q(r) \leq q(r')$ for every $0 \leq r \leq r' < 1$. For $r \in [1, \infty]$, let

$$\mathcal{Q}_r := \mathcal{Q} \cap L^r([0, 1]; S^D).$$

In Section 2, we define an enriched free energy $\bar{F}_{N, \lambda_N}(t, q)$ for every nonzero finite measure λ_N on $\mathbf{B}^N$ and every $q \in \mathcal{Q}_1$. In particular, $\bar{F}_{N, \mu^{\otimes N}}(t, q)$ denotes the enriched free energy associated with the original product reference measure $\mu^{\otimes N}$. At $q = 0$, this definition agrees with (1.2).

Write $\mathcal{P}(\text{supp } \mu)$ for the probability measures on $\text{supp } \mu$ and $\mathcal{B}_b(\text{supp } \mu)$ for the bounded Borel functions from $\text{supp } \mu$ to $\mathbb{R}$. For $\nu \in \mathcal{P}(\text{supp } \mu)$ and $h \in \mathcal{B}_b(\text{supp } \mu)$, let $e^h \nu$ denote the finite measure with density e^h relative to ν . Define $\psi_{\nu,h} : \mathcal{Q}_1 \rightarrow \mathbb{R}$ and $\varphi^\nu : \mathcal{Q}_1 \rightarrow \mathbb{R}$ by

$$(1.3) \quad \psi_{\nu,h}(q) := \bar{F}_{1,e^h \nu}(0, q), \quad \text{and} \quad \varphi^\nu(q) := \sup_{h \in \mathcal{B}_b(\text{supp } \mu)} \left\{ \psi_{\nu,h}(q) + \int_{\text{supp } \mu} h(a) \nu(da) \right\}.$$

For $\nu \in \mathcal{P}(\text{supp } \mu)$, let $f^\nu : \mathbb{R}_+ \times \mathcal{Q}_2 \rightarrow \mathbb{R}$ denote the solution of the following Hamilton–Jacobi equation

$$(1.4) \quad \begin{cases} \partial_t f^\nu(t, q) - \int_0^1 \xi(\partial_q f^\nu(t, q)(s)) ds = 0 \\ f^\nu(0, q) = \varphi^\nu(q). \end{cases}$$

The precise meaning of this equation is explained in Section 5. In particular, the notion of solution that we use is the notion of viscosity solution; see Definition 5.1 and Definition 5.2.

Denote by $\text{Ent}(\cdot | \mu) : \mathcal{P}(\text{supp } \mu) \rightarrow [0, \infty]$ the relative entropy

$$(1.5) \quad \text{Ent}(\nu | \mu) := \begin{cases} \int_{\text{supp } \mu} \log\left(\frac{d\nu}{d\mu}(a)\right) \nu(da), & \nu \ll \mu, \\ +\infty, & \nu \not\ll \mu. \end{cases}$$

Theorem 1.1. *For every $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_2$, we have*

$$(1.6) \quad \lim_{N \rightarrow \infty} \bar{F}_{N, \mu^{\otimes N}}(t, q) = \min_{\nu \in \mathcal{P}(\text{supp } \mu)} \{ \text{Ent}(\nu | \mu) + f^\nu(t, q) \}.$$

The minimum is attained. For every $T, Q \in \mathbb{R}_+$, the convergence in (1.6) is uniform over $t \in [0, T]$ and $q \in \mathcal{Q}_\infty$ with $|q|_{L^\infty} \leq Q$.

Background and related works. The Sherrington–Kirkpatrick model corresponds to choosing $D = 1$, taking μ to be the uniform measure on ± 1 and setting $\xi(x) = x^2$. The Parisi formula was derived in the physics literature in the context of the Sherrington–Kirkpatrick model using the non-rigorous replica method [41, 42]. This formula expresses the free energy of the model as the maximum value of a functional defined on $\mathcal{P}(\mathbb{R}_+)$. A rigorous proof of the lower bound in the Parisi formula is given in [28] using a technique now known as Guerra’s interpolation. A rigorous proof of the upper bound was then obtained in [44], completing the proof of the Parisi formula.

In [37] it is proven that the limiting Gibbs measure admits some structural ultrametric property. This made it possible to simplify the previous proof of the Parisi formula and to generalize it to the models we study in the present paper with the additional assumption that ξ is convex [38, 40]. The $D > 1$ case also brought to light the synchronization mechanism, first discovered in multi-species mean-field spin glasses [39] as introduced in [6].

Related results have also been obtained in the context of models with spherical spins both in the one-species case [43] and the multi-species case [9].

Finding expressions for the limit free energy beyond the cases where ξ is convex is a difficult task. Some progress has been made in the context of balanced multi-species spin glasses [10, 11] as in some cases the free energy of the full system reduces to the free energy of a system with $D = 1$ and ξ convex. However, no Parisi type formula is known in general for models with nonconvex ξ and some of those formulas are even provably false.

The Hamilton–Jacobi approach to spin glasses was introduced as an alternative method to compute and characterize the limit free energy. In essence, this approach consists in connecting the free energy with the solution of a Hamilton–Jacobi PDE. It was first developed in the high-temperature regime in [27, 8, 1, 7], in which case the corresponding partial differential equation is finite-dimensional. Later this method was revisited in [36, 33, 35, 23, 21] to include the whole temperature range with the aim of finding a suitable limiting object to describe the free energy beyond the convex setting. What emerged from this line of work is the following conjecture.

Conjecture 1.2 ([34]). *Let $\psi(q) = \bar{F}_{1,\mu}(0, q)$. For every $t \geq 0$ we have*

$$(1.7) \quad \lim_{N \rightarrow \infty} \bar{F}_{N, \mu^{\otimes N}}(t, 0) = f(t, 0),$$

where $f : \mathbb{R}_+ \times \mathcal{Q}_2 \rightarrow \mathbb{R}$ is the solution of

$$(1.8) \quad \begin{cases} \partial_t f(t, q) - \int_0^1 \xi(\partial_q f(t, q)(s)) ds = 0 \\ f(0, q) = \psi(q). \end{cases}$$

An important result toward proving this conjecture is obtained in [19]. The main result of [19] is that if the free energy converges, then its limiting value is given by a certain functional $\mathcal{J}_{t,q}$ evaluated at a critical point. The functional $\mathcal{J}_{t,q}$ is strongly related to the Hamilton–Jacobi equation of the conjecture. This result has two important caveats that we will explain in the following.

The first caveat is that the functional $\mathcal{J}_{t,q}$ in general has several critical points, so the main result of [19] does not fully characterize the limiting value of the free energy.

The second caveat is that [19] did not prove the convergence of the free energy. However, this missing convergence was provided in [29]. Several key ingredients in the proof of this convergence are also used in the present paper.

The first key ingredient is to represent the free energy in terms of the solution of the so-called Parisi PDE. The Parisi PDE and the associated diffusion have a substantial history in scalar spin glasses ($D = 1$). In particular, this PDE yields a stochastic-control representation which was used to establish the strict convexity of the Parisi functional [3, 4]; related dynamic-programming and stochastic-control approaches appear in [30, 22]. A treatment of the Parisi PDE for multidimensional spins is given in [12, Section 6]. For vector spins, as considered here, the corresponding Parisi PDE and its diffusion and Hessian identities are developed in [15, Lemma 3.3] and [18, Lemma 5.3].

The second key ingredient is a log-concavity that was recently revealed in [2]. This allows us to control fluctuations under a conditional log-concave law of Gaussian disorder and diffusion coordinates rather than directly under a Gibbs measure of the spins.

Other ingredients include: the realization of the centered Gaussian Hamiltonian as $H_N(\sigma; \mathbf{g}) = h(\sigma) \cdot \mathbf{g}$ where $\mathbf{g}$ is a Gaussian vector with independent coordinates and $h(\sigma)$ is deterministic. This is an instance of covariance factorization; a related Hilbert-space representation for mixed p -spin models appears in [25, Section 1]. The Brascamp–Lieb variance inequality [13], used previously for long-range spin glasses [26], concentration of multioverlaps under log-concave Gibbs measures [5, Theorem 1], and variational bounds for perceptron models [14, Section 3.2 and Appendix B] are also used.

Recently, a special case of Conjecture 1.2 was established in [17] for multi-species models when the reference measure μ is the uniform measure on ± 1 . Those assumptions guarantee that the function ψ is convex, which is the key ingredient to verifying Conjecture 1.2 in that case. The corresponding result has also been established for multi-species spherical spin glasses [20] using the same technique. Note that in general, the function ψ is not convex so this technique cannot be used to solve the full conjecture.

Theorem 1.1 gives an unconditional variational formula for arbitrary μ through the constrained solutions f^ν . Conjecture 1.2 asks whether this variational value at $q = 0$ equals the single solution f with initial condition ψ . That identification, though plausible, remains open in general.

Proof strategy and organization. In Section 2, we give a precise definition of the enriched free energy $\overline{F}_{N,\lambda_N}(t, q)$ and verify its Lipschitz continuity with respect to t, q , and λ_N (see Lemma 2.1 and Lemma 2.3).

For most of the manuscript, we assume that μ has finite support. Given a probability measure ν that is absolutely continuous with respect to μ , we introduce $\lambda_N(\nu)$, which is the uniform measure on $\Sigma_N(\nu)$, the set of spin configurations σ whose empirical measure is equal to ν .

In Section 3 we prove the key result of our proof strategy: Proposition 3.1. In the traditional approach of [32, 33] one wishes to verify that $\overline{F}_{N,\mu^{\otimes N}}$ approximately satisfies the partial differential equation (1.8). In this case, a major difficulty is that the corresponding error term is small but

not *uniformly* small in (t, q) . We verify in Lemma 3.6 that as $N \rightarrow +\infty$ we have

$$(1.9) \quad \bar{F}_{N, \mu^{\otimes N}}(t, q) = \min_{\nu} \left\{ \bar{F}_{N, \lambda_N(\nu)}(t, q) - \frac{1}{N} \log \mu^{\otimes N}(\Sigma_N(\nu)) \right\} + o(1).$$

Proposition 3.1 then guarantees that, on the set of paths $q \in \mathcal{C}^1([0, 1], S_+^D)$ satisfying $\dot{q}(s) \geq \varepsilon I_D$, we have that $\bar{F}_{N, \lambda_N(\nu)}$ satisfies, up to an error term, the partial differential equation

$$(1.10) \quad \partial_t f^\nu(t, q) - \int_0^1 \xi(\partial_q f^\nu(t, q)(s)) ds = 0.$$

What is remarkable is that this error term goes to 0 *uniformly* in (t, q) as $N \rightarrow +\infty$. This uniformity is proven using a new concentration result for a conditional version of the overlap (see Lemma 3.4 and Lemma 3.5). The main result, Theorem 1.1, then directly follows (up to a few technicalities) from (1.9), Proposition 3.1, and the fact that

$$(1.11) \quad \lim_{N \rightarrow +\infty} \bar{F}_{N, \lambda_N(\nu)}(0, q) = \varphi^\nu(q),$$

proven in Section 4 (see Proposition 4.1). The few technicalities in question are addressed in Section 5 and involve understanding how solutions of (1.10) can be approximated by solutions of finite-dimensional partial differential equations.

Finally, in Section 6 we remove the finiteness assumption on the support of μ by approximation. This uses the continuity of the free energy with respect to λ_N proven in Section 2.

Acknowledgments. All three authors are deeply indebted to Jean-Christophe Mourrat for numerous insightful discussions on the Hamilton–Jacobi program and its ramifications. Moreover, FHH would like to thank Mourrat for his openness in sharing his ideas on multiple occasions, including the CRM–PIMS Summer School 2021 and the St. Flour Summer School 2024, from which FHH has benefited.

HBC acknowledges funding from the NYU Shanghai Start-Up Fund and support from the NYU–ECNU Institute of Mathematical Sciences at NYU Shanghai. FHH gratefully acknowledges support from Eliran Subag’s ISF grant (No. 2055/21) and ERC grant (No. 101165541, Horizon Europe). VI acknowledges support from the FSMP postdoctoral program.

AI usage statement. During the preparation of this manuscript, the authors used large language models to support the development of certain ideas and improve certain sections of text. All AI-generated inputs were critically reviewed, revised, and validated by the authors to ensure their accuracy, relevance, and alignment with the objectives of the study.

2. ENRICHED MODELS AND BASIC PROPERTIES

2.1. Enriched free energies. We now define the enrichment used throughout the paper. For $q \in \mathcal{Q}_\infty$, let $q(1) := \lim_{s \uparrow 1} q(s)$. Let $\mathfrak{R}$ be the continuous **Ruelle probability cascade** of [19, Section 4], defined on an infinite-dimensional separable Hilbert space with uniform overlap distribution. Its support $\text{supp } \mathfrak{R}$ is contained in the unit sphere for almost every realization of $\mathfrak{R}$. If α and α' are independent samples from $\mathfrak{R}$, write $\alpha \wedge \alpha'$ for their Hilbert inner product. Under $\mathbb{E} \mathfrak{R}^{\otimes 2}$, the random variable $\alpha \wedge \alpha'$ is uniformly distributed on $[0, 1]$. Here and henceforth, we enlarge the probability space to include the independent random objects H_N and $\mathfrak{R}$. Conditionally on $\mathfrak{R}$, by further enlarging the probability space, let $W_N^q : \text{supp } \mathfrak{R} \rightarrow \mathbb{R}^{D \times N}$ be a centered Gaussian process whose columns are independent and satisfy

$$\mathbb{E} [W_{N, \cdot, i}^q(\alpha) W_{N, \cdot, j}^q(\alpha')^\top \mid \mathfrak{R}] = \mathbf{1}_{\{i=j\}} q(\alpha \wedge \alpha'), \quad i, j \in \{1, \dots, N\}.$$

The field H_N is independent of $\mathfrak{R}$, and conditionally on $\mathfrak{R}$, the Gaussian processes H_N and W_N^q are independent. For $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_\infty$, define $H_N^{t, q} : \mathbb{B}^N \times \text{supp } \mathfrak{R} \rightarrow \mathbb{R}$ by

$$H_N^{t, q}(\sigma, \alpha) := \sqrt{2t} H_N(\sigma) - t N \xi(\mathbf{R}_N(\sigma, \sigma)) + \sqrt{2} W_N^q(\alpha) \cdot \sigma - q(1) \cdot \sigma \sigma^\top.$$

For a nonzero finite measure λ_N on $\mathbb{B}^N$, define $\bar{F}_{N, \lambda_N} : \mathbb{R}_+ \times \mathcal{Q}_\infty \rightarrow \mathbb{R}$ by

$$(2.1) \quad \bar{F}_{N, \lambda_N}(t, q) := -\frac{1}{N} \mathbb{E} \log \iint \exp(H_N^{t, q}(\sigma, \alpha)) \lambda_N(d\sigma) \mathfrak{R}(d\alpha).$$

For fixed disorder and cascade, write $\langle \cdot \rangle$ for expectation under the probability measure proportional to $\exp(H_N^{t,q}(\sigma, \alpha)) \lambda_N(d\sigma) \mathfrak{R}(d\alpha)$. Primed or numbered replicas are conditionally independent samples from this measure. Lemma 2.1 shows that (2.1) extends uniquely from $\mathcal{Q}_\infty$ to $\mathcal{Q}_1$ by L^1 continuity, and we use the same notation for this extension. For a step path q , the same enrichment can equivalently be constructed from the finite-level Ruelle probability cascade associated with the jump points of q ; see [19, Section 4]. The continuous cascade formulation above is convenient because it applies directly to every bounded path.

We equip $L^2([0, 1]; S^D)$ with the inner product $\langle p, q \rangle_{L^2} := \int_0^1 p(s) \cdot q(s) ds$. The dual cone of $\mathcal{Q}_2$ is

$$(2.2) \quad \mathcal{Q}_2^* := \{r \in L^2([0, 1]; S^D) : \langle r, q \rangle_{L^2} \geq 0 \text{ for every } q \in \mathcal{Q}_2\}.$$

2.2. Gaussian comparisons and regularity. We first specify the notion of differentiability used below. Let E be a Hilbert space, let $G \subseteq E$, and let $g : G \rightarrow \mathbb{R}$. Following [19, Definition 2.1], we say that g is **Fréchet differentiable relative to G** at $x \in G$ if there exists a unique $y \in E$ such that

$$\lim_{\substack{x' \in G, \ x' \rightarrow x \\ x' \neq x}} \frac{|g(x') - g(x) - \langle y, x' - x \rangle_E|}{|x' - x|_E} = 0.$$

We call y the derivative of g at x . For $G = \mathbb{R}_+ \times \mathcal{Q}_2$, the ambient Hilbert space is $E := \mathbb{R} \times L^2([0, 1]; S^D)$. We write the two components of the derivative as $\partial_t g$ and $\partial_q g$. For a subset $G \subseteq L^2([0, 1]; S^D)$, we say that a function $g : G \rightarrow \mathbb{R}$ is **$\mathcal{Q}_2^*$ -increasing** if

$$q - q' \in \mathcal{Q}_2^* \implies g(q) \geq g(q'), \quad q, q' \in G;$$

cf. [19, Equation (3.19)]. The following lemma can be extracted from [19, Propositions 3.1, 3.6, and 5.1]. We adopt the following abbreviations for the derivatives of ξ :

$$(2.3) \quad M_0 := \sup_{a \in \mathbb{R}^{D \times D} : |a| \leq 1} |\xi(a)|, \quad M_1 := \sup_{a \in \mathbb{R}^{D \times D} : |a| \leq 1} |\nabla \xi(a)|, \quad \text{and} \quad M_2 := \sup_{a \in \mathbb{R}^{D \times D} : |a| \leq 1} \|\nabla^2 \xi(a)\|_{\text{op}}.$$

Lemma 2.1. *For every nonzero finite measure λ_N supported on $\mathbb{B}^N$ and all $(t, q), (t', q') \in \mathbb{R}_+ \times \mathcal{Q}_1$, we have*

$$(2.4) \quad |\bar{F}_{N, \lambda_N}(t, q) - \bar{F}_{N, \lambda_N}(t', q')| \leq M_0 |t - t'| + |q - q'|_{L^1}.$$

The restriction of $\bar{F}_{N, \lambda_N}$ to $\mathbb{R}_+ \times \mathcal{Q}_2$ is Fréchet differentiable relative to its domain. The derivative $\partial_q \bar{F}_{N, \lambda_N}(t, q)$ belongs to $\mathcal{Q}_\infty$ viewed as a path

$$(2.5) \quad [0, 1] \ni s \mapsto \partial_q \bar{F}_{N, \lambda_N}(t, q, s) \in S_+^D$$

and satisfies $|\partial_q \bar{F}_{N, \lambda_N}(t, q)|_{L^\infty} \leq 1$. For every $t \in \mathbb{R}_+$, $q \in \mathcal{Q}_\infty$, and $\kappa \in L^2([0, 1]; S^D)$, we have

$$(2.6) \quad \partial_t \bar{F}_{N, \lambda_N}(t, q) = \mathbb{E}(\xi(R_N(\sigma, \sigma'))) \quad \text{and} \quad \langle \kappa, \partial_q \bar{F}_{N, \lambda_N}(t, q) \rangle_{L^2} = \mathbb{E}(\langle \kappa, \alpha \wedge \alpha' \rangle \cdot R_N(\sigma, \sigma')).$$

For fixed $N \in \mathbb{N}$ and $t \in \mathbb{R}_+$, we have

$$(2.7) \quad |\partial_q \bar{F}_{N, \lambda_N}(t, q) - \partial_q \bar{F}_{N, \lambda_N}(t, q')|_{L^1} \leq 16N |q - q'|_{L^1}, \quad q, q' \in \mathcal{Q}_2.$$

For fixed $N \in \mathbb{N}$ and $t \in \mathbb{R}_+$, the function $q \mapsto \bar{F}_{N, \lambda_N}(t, q)$ is $\mathcal{Q}_2^$ -increasing on $\mathcal{Q}_2$.*

Lemma 2.2. *For every $\nu \in \mathcal{P}(\text{supp } \mu)$, we have $\varphi^\nu(0) = 0$ and*

$$(2.8) \quad 0 \leq \varphi^\nu(q) \leq |q|_{L^1}, \quad |\varphi^\nu(q) - \varphi^\nu(q')| \leq |q - q'|_{L^1}, \quad q, q' \in \mathcal{Q}_1.$$

Its restriction to $\mathcal{Q}_2$ is $\mathcal{Q}_2^$ -increasing.*

Proof. At $q = 0$, Jensen's inequality gives

$$\psi_{\nu, h}(0) + \int h d\nu = -\log \int e^h d\nu + \int h d\nu \leq 0.$$

Equality holds for $h = 0$, so $\varphi^\nu(0) = 0$. By Lemma 2.1, each function $q \mapsto \psi_{\nu, h}(q) + \int h d\nu$ is 1-Lipschitz in L^1 and $\mathcal{Q}_2^*$ -increasing on $\mathcal{Q}_2$. Taking the supremum in (1.3) preserves these properties and gives $\varphi^\nu(q) \leq |q|_{L^1}$. Since $q \in \mathcal{Q}_2$ implies $q \in \mathcal{Q}_2^*$, monotonicity gives $\varphi^\nu(q) \geq \varphi^\nu(0) = 0$ there. Approximation by bounded paths extends this lower bound to $\mathcal{Q}_1$. $\square$

Lemma 2.3. *Let λ_N and λ'_N be probability measures on $\mathbf{B}^N$. Suppose that they admit a coupling supported on pairs (σ, τ) satisfying $N^{-1} \sum_{i=1}^N |\sigma_{\cdot, i} - \tau_{\cdot, i}| \leq a$. Then, for every $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_\infty$, we have*

$$(2.9) \quad |\bar{F}_{N, \lambda_N}(t, q) - \bar{F}_{N, \lambda'_N}(t, q)| \leq 2a(tM_1 + |q|_{L^\infty}).$$

Proof. Let Π_N be a coupling of λ_N and λ'_N satisfying the assumption of the lemma. We use $x = (\sigma, \tau, \alpha)$ and $x' = (\sigma', \tau', \alpha')$ for two points in $\mathbf{B}^N \times \mathbf{B}^N \times \text{supp } \mathfrak{R}$. Conditionally on $\mathfrak{R}$, let X_0 and X_1 be independent centered Gaussian processes indexed by this space with covariance kernels

$$(2.10) \quad \begin{aligned} C_0(x, x') &:= 2tN\xi(\mathbf{R}_N(\sigma, \sigma')) + 2Nq(\alpha \wedge \alpha') \cdot \mathbf{R}_N(\sigma, \sigma'), \\ C_1(x, x') &:= 2tN\xi(\mathbf{R}_N(\tau, \tau')) + 2Nq(\alpha \wedge \alpha') \cdot \mathbf{R}_N(\tau, \tau'). \end{aligned}$$

These are the covariance kernels of the centered Gaussian parts of the enriched Hamiltonians evaluated respectively at σ and τ . In particular, we have

$$\frac{1}{2}C_0(x, x) = tN\xi(\mathbf{R}_N(\sigma, \sigma)) + q(1) \cdot \sigma\sigma^\top \quad \text{and} \quad \frac{1}{2}C_1(x, x) = tN\xi(\mathbf{R}_N(\tau, \tau)) + q(1) \cdot \tau\tau^\top.$$

Thus the diagonal terms in the enriched Hamiltonian are exactly $-C_0(x, x)/2$ and $-C_1(x, x)/2$ at the two endpoints of the interpolation.

For $\theta \in [0, 1]$, define $\tilde{X}_\theta := \sqrt{\theta}X_0 + \sqrt{1-\theta}X_1$, $\tilde{C}_\theta := \theta C_0 + (1-\theta)C_1$, and define $\phi : [0, 1] \rightarrow \mathbb{R}$ by

$$\phi(\theta) := -\frac{1}{N} \mathbb{E} \log \iint \exp\left(\tilde{X}_\theta(x) - \frac{1}{2}\tilde{C}_\theta(x, x)\right) \Pi_N(d\sigma, d\tau) \mathfrak{R}(d\alpha).$$

Since the first and second marginals of Π_N are λ_N and λ'_N , respectively, the endpoint values are

$$\phi(1) = \bar{F}_{N, \lambda_N}(t, q) \quad \text{and} \quad \phi(0) = \bar{F}_{N, \lambda'_N}(t, q).$$

Let $\langle \cdot \rangle_\theta$ denote the Gibbs expectation associated with $\phi(\theta)$, and let x^1, x^2 be two replicas. Gaussian integration by parts, together with the derivative of the compensation $-\tilde{C}_\theta(x, x)/2$, gives

$$(2.11) \quad \phi'(\theta) = \frac{1}{2N} \mathbb{E} \langle (C_0 - C_1)(x^1, x^2) \rangle_\theta.$$

We next bound the covariance difference in (2.11). If (σ, τ) and (σ', τ') belong to the support of Π_N , then

$$|\mathbf{R}_N(\sigma, \sigma') - \mathbf{R}_N(\tau, \tau')| \leq \frac{1}{N} \sum_{i=1}^N |\sigma_{\cdot, i}(\sigma'_{\cdot, i})^\top - \tau_{\cdot, i}(\tau'_{\cdot, i})^\top| \leq \frac{1}{N} \sum_{i=1}^N (|\sigma_{\cdot, i} - \tau_{\cdot, i}| + |\sigma'_{\cdot, i} - \tau'_{\cdot, i}|) \leq 2a.$$

Every overlap in this display belongs to the unit ball of $\mathbb{R}^{D \times D}$. Hence, the definition of M_1 gives

$$|\xi(\mathbf{R}_N(\sigma, \sigma')) - \xi(\mathbf{R}_N(\tau, \tau'))| \leq 2M_1a \quad \text{and} \quad |q(u) \cdot (\mathbf{R}_N(\sigma, \sigma') - \mathbf{R}_N(\tau, \tau'))| \leq 2a|q|_{L^\infty}.$$

Using (2.10), we therefore obtain $|(C_0 - C_1)(x, x')| \leq 4Na(tM_1 + |q|_{L^\infty})$. Substituting this estimate into (2.11) yields $|\phi'(\theta)| \leq 2a(tM_1 + |q|_{L^\infty})$ for every $\theta \in [0, 1]$. Integrating over $\theta \in [0, 1]$ proves (2.9). $\square$

Assume that μ has finite support and write $\text{supp } \mu = \{a_1, \dots, a_m\}$, with $a_1, \dots, a_m$ distinct. For $\nu \in \mathcal{P}(\text{supp } \mu)$, write $\nu_j := \nu(\{a_j\})$ for $j \in \{1, \dots, m\}$. For $N \in \mathbb{N}$, define $\Delta_{m, N} \subseteq \mathcal{P}(\text{supp } \mu)$ and $\Sigma_N : \Delta_{m, N} \rightarrow 2^{(\text{supp } \mu)^N}$ by

$$(2.12) \quad \begin{aligned} \Delta_{m, N} &:= \{\nu \in \mathcal{P}(\text{supp } \mu) : N\nu_j \in \{0, \dots, N\}, \forall j \in \{1, \dots, m\}\}, \\ \Sigma_N(\nu) &:= \left\{ \sigma \in (\text{supp } \mu)^N : \frac{1}{N} \#\{i \in \{1, \dots, N\} : \sigma_{\cdot, i} = a_j\} = \nu_j, \forall j \in \{1, \dots, m\} \right\}. \end{aligned}$$

For each $\nu \in \Delta_{m, N}$, let $\lambda_N(\nu)$ be the uniform probability measure on $\Sigma_N(\nu)$. For these admissible empirical measures, write

$$(2.13) \quad \bar{F}_N^\nu := \bar{F}_{N, \lambda_N(\nu)}.$$

Define $\mu_j := \mu(\{a_j\}) > 0$ for $j \in \{1, \dots, m\}$. Thus $\mu = \sum_{j=1}^m \mu_j \delta_{a_j}$. For $\nu \in \Delta_{m,N}$, define $\bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)} : \mathbb{R}_+ \times \mathcal{Q}_\infty \rightarrow \mathbb{R}$ by

$$(2.14) \quad \bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}(t, q) := -\frac{1}{N} \mathbb{E} \log \int_{\Sigma_N(\nu)} \int_{\text{supp } \mathfrak{R}} \exp(H_N^{t,q}(\sigma, \alpha)) \mathfrak{R}(d\alpha) \mu^{\otimes N}(d\sigma).$$

By Lemma 2.1, $\bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}$ has a unique L^1 -Lipschitz extension to $\mathbb{R}_+ \times \mathcal{Q}_1$; we use the same notation for this extension.

For every $\nu \in \Delta_{m,N}$, $\mu^{\otimes N}$ is constant on $\Sigma_N(\nu)$ and we have

$$(2.15) \quad |\Sigma_N(\nu)| = \frac{N!}{\prod_{j=1}^m (N\nu_j)!} \quad \text{and} \quad \mu^{\otimes N}(\Sigma_N(\nu)) = \frac{N!}{\prod_{j=1}^m (N\nu_j)!} \prod_{j=1}^m \mu_j^{N\nu_j}.$$

In particular, we have

$$(2.16) \quad \mathbf{1}_{\Sigma_N(\nu)} \mu^{\otimes N} = \mu^{\otimes N}(\Sigma_N(\nu)) \lambda_N(\nu).$$

Equations (2.14) and (2.16) give

$$(2.17) \quad \bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}(t, q) = \bar{F}_N^\nu(t, q) - \frac{1}{N} \log \mu^{\otimes N}(\Sigma_N(\nu)).$$

Lemma 2.4. *For every $N \in \mathbb{N}$ and $\nu \in \Delta_{m,N}$,*

$$(2.18) \quad 0 \leq -\frac{1}{N} \log \mu^{\otimes N}(\Sigma_N(\nu)) - \text{Ent}(\nu | \mu) \leq \frac{m \log(N+1)}{N}.$$

Proof. We have the exact identity

$$(2.19) \quad \mu^{\otimes N}(\Sigma_N(\nu)) = e^{-N \text{Ent}(\nu | \mu)} \nu^{\otimes N}(\Sigma_N(\nu)).$$

To conclude, it suffices to show that $\nu^{\otimes N}(\Sigma_N(\nu)) \geq (N+1)^{-m}$. First observe that by multinomial factorization, we have

$$(2.20) \quad \sum_{\ell} \frac{N!}{\prod_{j=1}^m \ell_j!} \prod_{j=1}^m \nu_j^{\ell_j} = (\nu_1 + \dots + \nu_m)^N = 1,$$

where the sum is taken over $\ell \in \{0, \dots, N\}^m$ such that $\sum_j \ell_j = N$. Also observe that the map $\ell \mapsto (N\nu_j)^\ell / \ell!$ is maximized over $\mathbb{N}$ at $\ell = N\nu_j$. Hence, each term in the sum of the previous display is $\leq \nu^{\otimes N}(\Sigma_N(\nu))$. Since there is less than $(N+1)^m$ terms in the sum, we obtain

$$(2.21) \quad (N+1)^m \nu^{\otimes N}(\Sigma_N(\nu)) \geq \sum_{\ell} \frac{N!}{\prod_{j=1}^m \ell_j!} \prod_{j=1}^m \nu_j^{\ell_j} = (\nu_1 + \dots + \nu_m)^N = 1,$$

thus the result is proven. $\square$

Let $\nu \in \Delta_{m,N}$, assume that there exists $x_1, \dots, x_N \in \mathcal{B}$ such that $\nu = N^{-1} \sum_{i=1}^N \delta_{x_i}$. Then, we have

$$(2.22) \quad \Sigma_N(\nu) = \{(x_{\pi(i)})_{1 \leq i \leq N} : \pi \in \mathfrak{S}_N\}.$$

In particular, if Π is a uniform random variable on the symmetric group $\mathfrak{S}_N$ and $\sigma_{\cdot, i} = x_{\Pi(i)}$, then σ is a random variable with law $\lambda_N(\nu)$.

Lemma 2.5. *If $\nu = N^{-1} \sum_{i=1}^N \delta_{x_i}$ and $\nu' = N^{-1} \sum_{i=1}^N \delta_{y_i}$ for $x_i, y_i \in \mathcal{B}$, then for every $(t, q) \in \mathbb{R}_+ \times \mathcal{Q}_\infty$,*

$$(2.23) \quad |\bar{F}_N^\nu(t, q) - \bar{F}_N^{\nu'}(t, q)| \leq 2(tM_1 + |q|_{L^\infty}) W_1(\nu, \nu').$$

Here W_1 is the Wasserstein distance on $\mathcal{P}(\mathcal{B})$ with cost $|a - b|$.

Proof. After relabeling $(y_i)_{i=1}^N$ (by the Birkhoff–von Neumann theorem), we may assume that

$$W_1(\nu, \nu') = \frac{1}{N} \sum_{i=1}^N |x_i - y_i|.$$

Let Π be uniform on the symmetric group $\mathfrak{S}_N$, and define $\sigma, \tau \in \mathcal{B}^N$ by

$$\sigma_{\cdot, i} := x_{\Pi(i)} \quad \text{and} \quad \tau_{\cdot, i} := y_{\Pi(i)}, \quad i \in \{1, \dots, N\}.$$

The law of σ is $\lambda_N(\nu)$, and the law of τ is $\lambda_N(\nu')$. Moreover, every pair in the support of this coupling satisfies

$$\frac{1}{N} \sum_{i=1}^N |\sigma_{\cdot,i} - \tau_{\cdot,i}| = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| = W_1(\nu, \nu').$$

Lemma 2.3, with $a = W_1(\nu, \nu')$, gives (2.23). $\square$

3. AN APPROXIMATE HAMILTON–JACOBI EQUATION FOR CONSTRAINED FREE ENERGY

The goal of this section is to prove the following approximate Hamilton–Jacobi equation for the constrained free energy.

Proposition 3.1. *Assume $\text{supp } \mu = \{a_1, \dots, a_m\}$. There is a constant $C = C_{D, M_2} < \infty$, where M_2 was defined in (2.3), such that for every $N \in \mathbb{N}$, $\nu \in \Delta_{m, N}$, $t > 0$, $\varepsilon > 0$, and $q \in C^1([0, 1]; S_+^D)$ satisfying $\dot{q}(s) \geq \varepsilon I_D$ for all $s \in [0, 1]$,*

$$(3.1) \quad \left| \partial_t \bar{F}_N^\nu(t, q) - \int_0^1 \xi(\partial_q \bar{F}_N^\nu(t, q, s)) \, ds \right| \leq C(N\varepsilon)^{-1/3}.$$

As explained in the introduction, the proof uses ingredients similar to those of the proof of [29, Proposition 3.9].

Throughout the first two subsections, we fix $N, \nu, t, \varepsilon, q$ satisfying the assumptions of Proposition 3.1.

3.1. Parisi representation and derivative identities. Let Cov_N be the covariance matrix on $(\text{supp } \mu)^N$, with entries $\text{Cov}_N(\sigma, \tau) := N\xi(R_N(\sigma, \tau))$. By (1.1), $\text{Cov}_N \geq 0$. For the standard basis $(\mathbf{e}_\sigma)_{\sigma \in (\text{supp } \mu)^N}$ of $\mathbb{R}^{(\text{supp } \mu)^N}$, put $\mathbf{h}_N(\sigma) := \text{Cov}_N^{1/2} \mathbf{e}_\sigma$. If $\mathbf{g}$ is a standard Gaussian vector in $\mathbb{R}^{(\text{supp } \mu)^N}$, then

$$(3.2) \quad H_N(\sigma; \mathbf{g}) := \mathbf{h}_N(\sigma) \cdot \mathbf{g}, \quad \mathbf{h}_N(\sigma) \cdot \mathbf{h}_N(\tau) = N\xi(R_N(\sigma, \tau)).$$

We use this realization to differentiate in the disorder and to exploit permutation symmetry.

For $\pi \in \mathfrak{S}_N$, write $\sigma_{\cdot, i}^\pi := \sigma_{\cdot, \pi(i)}$ and let $Q_\pi \mathbf{e}_\sigma := \mathbf{e}_{\sigma^\pi}$. Since Cov_N and $\text{Cov}_N^{1/2}$ commute with Q_π ,

$$(3.3) \quad H_N(\sigma^\pi; Q_\pi \mathbf{g}) = H_N(\sigma; \mathbf{g}).$$

For $\mathbf{g} \in \mathbb{R}^{(\text{supp } \mu)^N}$, define a positive finite measure $\Lambda_{\mathbf{g}}$ on $\Sigma_N(\nu)$ by

$$\Lambda_{\mathbf{g}}(d\sigma) := \exp\left(\sqrt{2t} H_N(\sigma; \mathbf{g}) - tN\xi(R_N(\sigma, \sigma))\right) \lambda_N(\nu)(d\sigma).$$

Let $\Phi : [0, 1] \times \mathbb{R}^{D \times N} \times \mathbb{R}^{(\text{supp } \mu)^N} \rightarrow \mathbb{R}$ solve the following Parisi PDE

$$(3.4) \quad \begin{cases} -\partial_s \Phi(s, \mathbf{x}; \mathbf{g}) = \left(\sum_{i=1}^N \dot{q}(s) \cdot \nabla_{\mathbf{x}, i}^2 \Phi + s \dot{q}(s) \cdot \nabla_{\mathbf{x}} \Phi (\nabla_{\mathbf{x}} \Phi)^\top \right) (s, \mathbf{x}; \mathbf{g}), & (s, \mathbf{x}; \mathbf{g}) \in [0, 1] \times \mathbb{R}^{D \times N} \times \mathbb{R}^{(\text{supp } \mu)^N} \\ \Phi(1, \mathbf{x}; \mathbf{g}) = \log \int \exp(\mathbf{x} \cdot \sigma - q(1) \cdot \sigma \sigma^\top) \Lambda_{\mathbf{g}}(d\sigma), & (\mathbf{x}; \mathbf{g}) \in \mathbb{R}^{D \times N} \times \mathbb{R}^{(\text{supp } \mu)^N}. \end{cases}$$

Let $\mathbf{z}_0$ be standard Gaussian in $\mathbb{R}^{D \times N}$, and let $B : [0, 1] \rightarrow \mathbb{R}^{D \times N}$ be a standard Brownian motion, with $\mathbf{g}, \mathbf{z}_0, B$ mutually independent and independent of the cascade.

Lemma 3.2. *The equation (3.4) has a unique classical solution Φ with bounded spatial derivatives of positive order, and $|\nabla_{\mathbf{x}, i} \Phi| \leq 1$ for every $i \in \{1, \dots, N\}$. Moreover, we have*

$$(3.5) \quad \bar{F}_N^\nu(t, q) = -\frac{1}{N} \mathbb{E} \Phi(0, \sqrt{2q(0)} \mathbf{z}_0; \mathbf{g}).$$

The stochastic differential equation

$$(3.6) \quad d\mathbf{X}_s = 2s\dot{q}(s) \nabla_{\mathbf{x}} \Phi(s, \mathbf{X}_s; \mathbf{g}) \, ds + \sqrt{2\dot{q}(s)} \, dB_s, \quad \text{and} \quad \mathbf{X}_0 = \sqrt{2q(0)} \mathbf{z}_0$$

admits a unique strong solution on $[0, 1]$.

Proof. In column-stacked coordinates, (3.4) has covariance path $I_N \otimes q$, coefficient $\alpha(s) = s$, and terminal measure $\Lambda_{\mathbf{g}}$. Existence and regularity follow from [18, Proposition 4.7(1), (3), and (4)]. Its cascade representation [18, Lemma 4.5, (4.10)] gives (3.5) after averaging over $(\mathbf{g}, \mathbf{z}_0)$. Differentiating in $\mathbf{x}_{\cdot i}$ gives a linear backward equation whose terminal gradient lies in the unit ball. The maximum principle gives the column bound. The bounded Hessian makes the drift of (3.6) globally Lipschitz, giving a unique strong solution; see also [15, Proposition 3.2]. The finite terminal sum and the same regularity estimates justify the parameter differentiations below: derivatives in $(\mathbf{x}, \mathbf{g})$ and in bounded terminal perturbations are bounded, while the t derivative has at most linear growth in $\mathbf{g}$ locally on $t > 0$. $\square$

Define the filtration $\mathcal{G}_s := \sigma(\mathbf{g}, \mathbf{z}_0, (B_r)_{0 \leq r \leq s})$ for $s \in [0, 1]$. For all $s \in [0, 1]$, let σ_s have the conditional law

$$\mathbb{P}(\sigma_s = \sigma \mid \mathcal{G}_s) = \Lambda_{\mathbf{g}}(\{\sigma\}) \mathbb{E}[\exp(\mathbf{X}_1 \cdot \sigma - q(1) \cdot \sigma \sigma^\top - \Phi(1, \mathbf{X}_1; \mathbf{g})) \mid \mathcal{G}_s].$$

Here σ_1 denotes the terminal spin.

Let vec stack the columns of a D -by- N matrix. Spatial Hessians are taken in these column-stacked coordinates.

Lemma 3.3. *The following statements hold.*

- (1) *The function $\Phi(s, \cdot; \cdot)$ is jointly convex in $(\mathbf{x}, \mathbf{g})$.*
- (2) *The process $\mathbb{E}[\sigma_1 \mid \mathcal{G}_s]$ is a bounded martingale, and for every $s \in [0, 1]$,*

$$\mathbb{E}[\sigma_1 \mid \mathcal{G}_s] = \nabla_{\mathbf{x}} \Phi(s, \mathbf{X}_s; \mathbf{g}).$$

- (3) *We have*

$$\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g}) = s \text{Cov}(\text{vec}(\sigma_1) \mid \mathcal{G}_s) + \mathbb{E} \left[\int_s^1 \text{Cov}(\text{vec}(\sigma_1) \mid \mathcal{G}_r) \, dr \mid \mathcal{G}_s \right].$$

- (4) *We have*

$$\int_0^1 \mathbb{E} |\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})|^2 \, ds \leq \frac{N}{2\varepsilon}.$$

Proof. For a bounded $v : \Sigma_N(\nu) \rightarrow \mathbb{R}$, let Φ^y have terminal measure $e^{y\nu} \Lambda_{\mathbf{g}}$. All parameter derivatives below are evaluated at $y = 0$ along the unperturbed diffusion. Differentiating (3.4) once in y and applying Itô's formula show that the first derivative is a martingale with terminal value $\mathbb{E}[v(\sigma_1) \mid \mathcal{G}_1]$. Hence

$$\partial_y \Phi^y(s, \mathbf{X}_s; \mathbf{g})|_{y=0} = \mathbb{E}[v(\sigma_1) \mid \mathcal{G}_s].$$

Differentiating twice shows that the second derivative has drift $-s \, d[\mathbb{E}[v(\sigma_1) \mid \mathcal{G}_s]]_s$, whereas the conditional variance $\text{Var}(v(\sigma_1) \mid \mathcal{G}_s)$ has drift $-d[\mathbb{E}[v(\sigma_1) \mid \mathcal{G}_s]]_s$. Subtracting $s \, \text{Var}(v(\sigma_1) \mid \mathcal{G}_s)$ from the second derivative therefore gives a process with drift $-\text{Var}(v(\sigma_1) \mid \mathcal{G}_s) \, ds$ and zero terminal value. Integration and conditioning give

$$(3.7) \quad \partial_y^2 \Phi^y(s, \mathbf{X}_s; \mathbf{g})|_{y=0} = s \, \text{Var}(v(\sigma_1) \mid \mathcal{G}_s) + \mathbb{E} \left[\int_s^1 \text{Var}(v(\sigma_1) \mid \mathcal{G}_r) \, dr \mid \mathcal{G}_s \right].$$

We now prove the four items.

Proof of (1): For $a \in \mathbb{R}^{D \times N}$ and $b \in \mathbb{R}^{(\text{supp } \mu)^N}$, take

$$v(\sigma) = a \cdot \sigma + \sqrt{2t} b \cdot \mathbf{h}_N(\sigma).$$

Uniqueness gives $\Phi^y(s, \mathbf{x}; \mathbf{g}) = \Phi(s, \mathbf{x} + ya; \mathbf{g} + yb)$, so $\partial_y^2 \Phi^y(s, \mathbf{X}_s; \mathbf{g})|_{y=0}$ is the second directional derivative in direction (a, b) . By (3.7), it is nonnegative. Restarting the diffusion from any deterministic $(s, \mathbf{x})$, with $\mathbf{g}$ fixed, proves joint convexity at every $(s, \mathbf{x}, \mathbf{g})$.

Proof of (2): For $v(\sigma) = a \cdot \sigma$ with $a \in \mathbb{R}^{D \times N}$, uniqueness gives $\Phi^y(s, \mathbf{x}; \mathbf{g}) = \Phi(s, \mathbf{x} + ya; \mathbf{g})$. Thus

$$a \cdot \nabla_{\mathbf{x}} \Phi(s, \mathbf{X}_s; \mathbf{g}) = \partial_y \Phi^y(s, \mathbf{X}_s; \mathbf{g})|_{y=0} = \mathbb{E}[a \cdot \sigma_1 \mid \mathcal{G}_s].$$

Since a is arbitrary, this proves the gradient identity. The conditional-expectation process is a martingale and is bounded because $|\sigma_1|^2 \leq N$.

Proof of (3): For the same spatial perturbation, the second parameter derivative is the Hessian quadratic form in direction a . Applying (3.7) with $v(\sigma) = a \cdot \sigma$ gives

$$\text{vec}(a)^\top \nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g}) \text{vec}(a) = s \text{Var}(a \cdot \sigma_1 \mid \mathcal{G}_s) + \mathbb{E} \left[\int_s^1 \text{Var}(a \cdot \sigma_1 \mid \mathcal{G}_r) dr \mid \mathcal{G}_s \right].$$

Since a is arbitrary, this proves the conditional covariance representation of $\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})$.

Proof of (4): Differentiating (3.4) once in $\mathbf{x}$, applying Itô's formula, and using Lemma 3.3(2) give

$$(3.8) \quad d \text{vec}(\mathbb{E}[\sigma_1 \mid \mathcal{G}_s]) = \nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g}) \sqrt{2(I_N \otimes \dot{q}(s))} d \text{vec}(B_s);$$

see [15, Lemma 3.3] and [18, Lemma 5.3]. Since $(I_N \otimes \dot{q}(s)) \geq \varepsilon I_{DN}$, Itô's isometry gives

$$2\varepsilon \int_0^1 \mathbb{E} |\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})|^2 ds \leq \mathbb{E} |\mathbb{E}[\sigma_1 \mid \mathcal{G}_1]|^2 - \mathbb{E} |\mathbb{E}[\sigma_1 \mid \mathcal{G}_0]|^2 \leq \mathbb{E} |\sigma_1|^2 \leq N.$$

Dividing by 2ε proves the estimate. $\square$

We now provide the derivative identities.

Lemma 3.4. *Let σ_s^1, σ_s^2 be conditionally independent copies of σ_s conditioned on $\mathcal{G}_s$. Let $R_s := N^{-1} \sigma_s^1 (\sigma_s^2)^\top$. Then,*

$$(3.9) \quad \partial_q \bar{F}_N^\nu(t, q, s) = \mathbb{E} R_s \quad \text{and} \quad \partial_t \bar{F}_N^\nu(t, q) = \int_0^1 \mathbb{E} \xi(R_s) ds.$$

Proof. We first prove the time identity. Differentiating (3.4) in t , applying the diffusion representation, and using (3.2) and (3.5) give

$$\partial_t \bar{F}_N^\nu(t, q) = \frac{1}{N} \mathbb{E} \left[|\mathbf{h}_N(\sigma_1)|^2 - \frac{\mathbf{g} \cdot \mathbf{h}_N(\sigma_1)}{\sqrt{2t}} \right].$$

Since $\mathbf{X}_0$ is independent of $\mathbf{g}$, Gaussian integration by parts and (3.7) in the proof of Lemma 3.3 at $s = 0$, with $v = \sqrt{2t} \mathbf{h}_{N,\ell}$, yield

$$\begin{aligned} \frac{1}{\sqrt{2t}} \mathbb{E} [\mathbf{g} \cdot \mathbf{h}_N(\sigma_1)] &= \frac{1}{2t} \sum_{\ell} \mathbb{E} \partial_{\mathbf{g}_\ell}^2 \Phi(0, \mathbf{X}_0; \mathbf{g}) \\ &= \int_0^1 \mathbb{E} \left[|\mathbf{h}_N(\sigma_1)|^2 - |\mathbb{E}[\mathbf{h}_N(\sigma_1) \mid \mathcal{G}_s]|^2 \right] ds. \end{aligned}$$

The self-overlap terms cancel, and conditional independence and (3.2) give

$$\partial_t \bar{F}_N^\nu(t, q) = \frac{1}{N} \int_0^1 \mathbb{E} |\mathbb{E}[\mathbf{h}_N(\sigma_1) \mid \mathcal{G}_s]|^2 ds = \int_0^1 \mathbb{E} \xi(R_s) ds.$$

For the path derivative, we first take q smooth; the approximation argument of [18, Lemma 5.5, Steps 1–5], together with Lemma 2.1, gives the general case. Fix $\kappa \in C_c^\infty((0, 1); S^D)$. Since $\dot{q} \geq \varepsilon I_D$, the paths $q + h\kappa$ are admissible for small $|h|$. Summing the diagonal D -by- D blocks in Lemma 3.3(3) and using conditional independence give

$$\sum_{i=1}^N \mathbb{E} \nabla_{\mathbf{x}, i}^2 \Phi(s, \mathbf{X}_s; \mathbf{g}) + N s \mathbb{E} R_s = \mathbb{E} [\sigma_1 \sigma_1^\top] - N \int_s^1 \mathbb{E} R_r dr.$$

Differentiating (3.4) in h , applying Itô's formula along $\mathbf{X}$, and using Lemma 3.3(2) therefore yield

$$\begin{aligned} (3.10) \quad \partial_h \bar{F}_N^\nu(t, q + h\kappa) \Big|_{h=0} &= -\frac{1}{N} \int_0^1 \dot{\kappa}(s) \cdot \left(\mathbb{E} [\sigma_1 \sigma_1^\top] - N \int_s^1 \mathbb{E} R_r dr \right) ds \\ &= \int_0^1 \kappa(s) \cdot \mathbb{E} R_s ds. \end{aligned}$$

The last equality follows by integration by parts, since $\kappa(0) = \kappa(1) = 0$. Density of these directions in $L^2([0, 1]; S^D)$ and Lemma 2.1 identify the path derivative almost everywhere. $\square$

3.2. Concentration of the overlap and proof of Proposition 3.1. We prove the concentration of the conditional overlap introduced in Lemma 3.4.

Lemma 3.5. *There is a constant $C_D < \infty$, depending only on D , such that*

$$(3.11) \quad \int_0^1 \mathbb{E} |R_s - \mathbb{E} R_s|^2 ds \leq C_D (N\varepsilon)^{-1/3}.$$

Proof. Since $|R_s| \leq 1$, the left-hand side of (3.11) is at most 1. It therefore suffices to consider $N\varepsilon > 1$. Fix $\delta \in (0, 1)$.

Step 1: Thermal fluctuations. For a deterministic $v \in \mathbb{R}^{DN}$, Lemma 3.3(3) gives

$$(3.12) \quad \text{Var}(v \cdot \text{vec}(\sigma_1) \mid \mathcal{G}_s) = v^\top \text{Cov}(\text{vec}(\sigma_1) \mid \mathcal{G}_s) v \leq \frac{1}{s} v^\top \nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g}) v \leq \frac{|v|^2}{s} |\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})|.$$

With one replica fixed at τ , each entry of R_s is a linear form in the other replica. The sum of the squared coefficient norms is

$$(3.13) \quad \frac{1}{N^2} \sum_{\alpha, \beta=1}^D \sum_{i=1}^N \tau_{\alpha i}^2 = \frac{D}{N^2} |\tau|^2 \leq \frac{D}{N}.$$

Conditional variance tensorization over the two replicas, followed by summation over the entries, therefore gives

$$(3.14) \quad \mathbb{E}[|R_s - \mathbb{E}[R_s \mid \mathcal{G}_s]|^2 \mid \mathcal{G}_s] \leq \frac{2D}{Ns} |\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})|.$$

Taking expectations, integrating, and applying Cauchy–Schwarz and Lemma 3.3(4), we obtain

$$(3.15) \quad \int_\delta^1 \mathbb{E} |R_s - \mathbb{E}[R_s \mid \mathcal{G}_s]|^2 ds \leq \frac{2D}{N} \left(\int_\delta^1 s^{-2} ds \right)^{1/2} \left(\int_\delta^1 \mathbb{E} |\nabla_{\mathbf{x}}^2 \Phi(s, \mathbf{X}_s; \mathbf{g})|^2 ds \right)^{1/2} \leq \frac{C_D}{\sqrt{N\varepsilon\delta}}.$$

Step 2: Concentration of the conditional mean.

Fix $a \in S^D$ with $|a| = 1$ and let $T_a := I_N \otimes a$, a self-adjoint contraction in column-stacked coordinates.

For $s \in (\delta, 1]$, define

$$(3.16) \quad U_s := \frac{1}{N\delta} \int_{s-\delta}^s a \cdot \left(\sigma_1 \mathbb{E}[\sigma_1 \mid \mathcal{G}_r]^\top - \frac{1}{2} \mathbb{E}[\sigma_1 \mid \mathcal{G}_r] \mathbb{E}[\sigma_1 \mid \mathcal{G}_r]^\top \right) dr.$$

Since $\mathbb{E}[\sigma_1 \mid \mathcal{G}_r]$ is $\mathcal{G}_s$ -measurable for $r \leq s$, the tower property and symmetry of a give

$$(3.17) \quad a \cdot \mathbb{E}[R_s \mid \mathcal{G}_s] - 2\mathbb{E}[U_s \mid \mathcal{G}_s] = \frac{1}{N\delta} \int_{s-\delta}^s a \cdot ((\mathbb{E}[\sigma_1 \mid \mathcal{G}_s] - \mathbb{E}[\sigma_1 \mid \mathcal{G}_r])(\mathbb{E}[\sigma_1 \mid \mathcal{G}_s] - \mathbb{E}[\sigma_1 \mid \mathcal{G}_r])^\top) dr.$$

The martingale identity from Lemma 3.3(2),

$$\mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_s] - \mathbb{E}[\sigma_1 \mid \mathcal{G}_r]]^2 = \mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_s]]^2 - \mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_r]]^2,$$

the bounds $0 \leq \mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_r]]^2 \leq N$ and $|\mathbb{E}[\sigma_1 \mid \mathcal{G}_s] - \mathbb{E}[\sigma_1 \mid \mathcal{G}_r]|^2 \leq 4N$, and time translation imply

$$(3.18) \quad \begin{aligned} \int_\delta^1 \mathbb{E} |a \cdot \mathbb{E}[R_s \mid \mathcal{G}_s] - 2\mathbb{E}[U_s \mid \mathcal{G}_s]|^2 ds &\leq \frac{4}{N\delta} \int_0^\delta \int_\delta^1 (\mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_s]]^2 - \mathbb{E}[\mathbb{E}[\sigma_1 \mid \mathcal{G}_{s-v}]]^2) ds dv \\ &\leq \frac{4}{\delta} \int_0^\delta v dv = 2\delta. \end{aligned}$$

Equation (3.3) and invariance of $\lambda_N(\nu)$ give the terminal equivariance of Φ . Uniqueness for the PDE then gives

$$(3.19) \quad \Phi(r, \mathbf{x}^\pi; Q_\pi \mathbf{g}) = \Phi(r, \mathbf{x}; \mathbf{g}), \quad \pi \in \mathfrak{S}_N.$$

Consequently the joint law is invariant under

$$(3.20) \quad (\mathbf{g}, \mathbf{z}_0, B, \sigma_1) \mapsto (Q_\pi \mathbf{g}, \mathbf{z}_0^\pi, B^\pi, \sigma_1^\pi).$$

The Gaussian inputs keep their laws, the SDE solution is column-permuted, and the terminal Gibbs law is carried to its corresponding permutation. Thus $\mathbb{E}[\sigma_1 \mid \mathcal{G}_r]$ is also column-permuted and U_s is unchanged. Since $\mathfrak{S}_N$ acts transitively on $\Sigma_N(\nu)$, $\mathbb{E}[U_s \mid \sigma_1]$ is constant, so

$$(3.21) \quad \text{Var}(U_s) = \mathbb{E} \text{Var}(U_s \mid \sigma_1).$$

Centering (3.17), applying Jensen's inequality, and using $|\mathbb{E}[R_s | \mathcal{G}_s]| \leq 1$ on $[0, \delta]$ yield

$$(3.22) \quad \int_0^1 \text{Var}(a \cdot \mathbb{E}[R_s | \mathcal{G}_s]) ds \leq C\delta + C \int_\delta^1 \mathbb{E} \text{Var}(U_s | \sigma_1) ds.$$

We next bound the conditional variance in (3.22). Let Γ be the product law of the independent inputs $(\mathbf{g}, \mathbf{z}_0, B)$, and put

$$\mathbf{Y}_r := \sqrt{2q(0)}\mathbf{z}_0 + \int_0^r \sqrt{2\dot{q}(u)} dB_u.$$

Itô's formula for $r\Phi(r, \mathbf{Y}_r; \mathbf{g})$ identifies

$$(3.23) \quad \exp\left(\Phi(1, \mathbf{Y}_1; \mathbf{g}) - \int_0^1 \Phi(r, \mathbf{Y}_r; \mathbf{g}) dr\right)$$

as the stochastic exponential with integrand $r\sqrt{2\dot{q}(r)}\nabla_{\mathbf{x}}\Phi(r, \mathbf{Y}_r; \mathbf{g})$. The bounded gradient satisfies Novikov's condition [31, Chapter 3, Corollary 5.13], and Girsanov's theorem identifies $(\mathbf{g}, \mathbf{z}_0, \mathbf{Y})$ under this change of measure with $(\mathbf{g}, \mathbf{z}_0, \mathbf{X})$ in (3.6). Multiplying by the terminal Gibbs mass at $\sigma \in \Sigma_N(\nu)$ cancels $\exp(\Phi(1, \mathbf{Y}_1; \mathbf{g}))$. Hence the conditional law $\mathbb{P}_\sigma$ of the reference inputs given $\sigma_1 = \sigma$ has density

$$(3.24) \quad \frac{d\mathbb{P}_\sigma}{d\Gamma} \propto \exp\left(\sqrt{2t}\mathbf{h}_N(\sigma) \cdot \mathbf{g} + \mathbf{Y}_1 \cdot \sigma - \int_0^1 \Phi(r, \mathbf{Y}_r; \mathbf{g}) dr\right).$$

The normalizing constant is positive and finite; terms depending only on σ are absorbed into it. In these reference coordinates, evaluate U_s by representing $\mathbb{E}[\sigma_1 | \mathcal{G}_r]$ as $\nabla_{\mathbf{x}}\Phi(r, \mathbf{Y}_r; \mathbf{g})$ using Lemma 3.3(2); the change of measure identifies this with the original conditional mean.

Choose an orthonormal basis of $L^2([0, 1]; \mathbb{R}^{D \times N})$ and write $(\eta_k)_{k \in \mathbb{N}}$ for the corresponding independent standard Gaussian coordinates of the Brownian noise. Put $\mathcal{F}_n := \sigma(\eta_{n+1}, \eta_{n+2}, \dots)$. Let ζ collect $\mathbf{g}, \mathbf{z}_0, \eta_1, \dots, \eta_n$, and condition on $\mathcal{F}_n$. Under Γ , ζ is standard Gaussian and $(\mathbf{Y}_r, \mathbf{g})$ is affine in ζ . Under $\mathbb{P}_\sigma$, its conditional Lebesgue density is proportional to $\exp(-V_{\sigma,n})$, where

$$V_{\sigma,n}(\zeta) = \frac{1}{2}|\zeta|^2 - \sqrt{2t}\mathbf{h}_N(\sigma) \cdot \mathbf{g} - \mathbf{Y}_1 \cdot \sigma + \int_0^1 \Phi(r, \mathbf{Y}_r; \mathbf{g}) dr.$$

The middle terms are affine, and Lemma 3.3(1) gives

$$(3.25) \quad \nabla_\zeta^2 V_{\sigma,n} = I + \int_0^1 \nabla_\zeta^2 [\Phi(r, \mathbf{Y}_r; \mathbf{g})] dr \geq I.$$

For fixed σ , the derivative of $a \cdot (\sigma \mathbb{E}[\sigma_1 | \mathcal{G}_r]^\top - \frac{1}{2} \mathbb{E}[\sigma_1 | \mathcal{G}_r] \mathbb{E}[\sigma_1 | \mathcal{G}_r]^\top)$ with respect to its conditional-mean argument is $a(\sigma - \mathbb{E}[\sigma_1 | \mathcal{G}_r])$. Cauchy-Schwarz for the joint Hessian of Φ and then in time therefore gives, for every ζ -direction w ,

$$(3.26) \quad |w \cdot \nabla_\zeta U_s|^2 \leq b_s w^\top (\nabla_\zeta^2 V_{\sigma,n}) w,$$

$$b_s := \frac{1}{N^2 \delta^2} \int_{s-\delta}^s \text{vec}(a(\sigma - \mathbb{E}[\sigma_1 | \mathcal{G}_r]))^\top \nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{Y}_r; \mathbf{g}) \text{vec}(a(\sigma - \mathbb{E}[\sigma_1 | \mathcal{G}_r])) dr.$$

Indeed, the first quadratic factor is $w^\top \nabla_\zeta^2 [\Phi(r, \mathbf{Y}_r; \mathbf{g})] w$, whose time integral is bounded by $w^\top (\nabla_\zeta^2 V_{\sigma,n}) w$. Taking $w = (\nabla_\zeta^2 V_{\sigma,n})^{-1} \nabla_\zeta U_s$ and applying the Brascamp-Lieb variance inequality [13, Section 4] give

$$(3.27) \quad \text{Var}_{\mathbb{P}_\sigma}(U_s | \mathcal{F}_n) \leq \mathbb{E}_{\mathbb{P}_\sigma}[\nabla_\zeta U_s^\top (\nabla_\zeta^2 V_{\sigma,n})^{-1} \nabla_\zeta U_s | \mathcal{F}_n] \leq \mathbb{E}_{\mathbb{P}_\sigma}[b_s | \mathcal{F}_n].$$

Taking $\mathbb{P}_\sigma$ -expectations first and using total variance yields

$$\text{Var}_{\mathbb{P}_\sigma}(U_s) - \text{Var}_{\mathbb{P}_\sigma}(\mathbb{E}_{\mathbb{P}_\sigma}[U_s | \mathcal{F}_n]) \leq \mathbb{E}_{\mathbb{P}_\sigma} b_s.$$

The tail sigma-algebra is trivial under Γ and hence under the equivalent measure $\mathbb{P}_\sigma$. Reverse martingale convergence (see Theorem 4.7.3 in [24]) gives convergence in $L^1(\mathbb{P}_\sigma)$. Since U_s and its conditional expectations are uniformly bounded, squared differences are bounded by a constant times absolute differences, so convergence also holds in $L^2(\mathbb{P}_\sigma)$ and the second variance vanishes. The right-hand side is independent of n and integrable, since the spatial Hessian is bounded. Returning to the original diffusion $\mathbf{X}$, with $\mathbf{Y}$ replaced by $\mathbf{X}$ in b_s , gives

$$(3.28) \quad \text{Var}(U_s | \sigma_1 = \sigma) \leq \mathbb{E}[b_s | \sigma_1 = \sigma].$$

Conditioning the quadratic integrand on $\mathcal{G}_r$ gives

$$(3.29) \quad \begin{aligned} & \mathbb{E}[\text{vec}(a(\sigma_1 - \mathbb{E}[\sigma_1 | \mathcal{G}_r]))^\top \nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}) \text{vec}(a(\sigma_1 - \mathbb{E}[\sigma_1 | \mathcal{G}_r])) | \mathcal{G}_r] \\ &= \text{tr}(\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}) T_a \text{Cov}(\text{vec}(\sigma_1) | \mathcal{G}_r) T_a) \leq \frac{1}{r} |\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g})|^2. \end{aligned}$$

Here Lemma 3.3(3) gives

$$\text{Cov}(\text{vec}(\sigma_1) | \mathcal{G}_r) \leq r^{-1} \nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}).$$

Since T_a is a contraction,

$$\begin{aligned} \text{tr}(\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}) T_a \nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}) T_a) &\leq |\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g})| |T_a \nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g}) T_a| \\ &\leq |\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g})|^2. \end{aligned}$$

Averaging (3.28) over σ , applying Fubini and using Lemma 3.3(4) yield

$$(3.30) \quad \begin{aligned} \int_{\delta}^1 \mathbb{E} \text{Var}(U_s | \sigma_1) ds &\leq \frac{1}{N^2 \delta^2} \int_{\delta}^1 \int_{s-\delta}^s \frac{1}{r} \mathbb{E} |\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g})|^2 dr ds \\ &\leq \frac{1}{N^2 \delta^2} \int_0^1 \frac{\min\{r, \delta\}}{r} \mathbb{E} |\nabla_{\mathbf{x}}^2 \Phi(r, \mathbf{X}_r; \mathbf{g})|^2 dr \\ &\leq \frac{1}{2N\epsilon \delta^2}. \end{aligned}$$

Here we used $[\delta, 1] \cap [r, r + \delta] \leq \min\{r, \delta\}$ for $r > 0$. Substituting into (3.22) and summing over an orthonormal basis of S^D give

$$(3.31) \quad \int_0^1 \mathbb{E} |\mathbb{E}[R_s | \mathcal{G}_s] - \mathbb{E} R_s|^2 ds \leq C_D \left(\delta + \frac{1}{N\epsilon \delta^2} \right).$$

Step 3: Combination of the estimates. Conditional variance decomposition gives

$$(3.32) \quad \mathbb{E} |R_s - \mathbb{E} R_s|^2 = \mathbb{E} |R_s - \mathbb{E}[R_s | \mathcal{G}_s]|^2 + \mathbb{E} |\mathbb{E}[R_s | \mathcal{G}_s] - \mathbb{E} R_s|^2.$$

Since $|R_s|, |\mathbb{E}[R_s | \mathcal{G}_s]| \leq 1$, (3.15) in Step 1 and (3.31) in Step 2 imply

$$\int_0^1 \mathbb{E} |R_s - \mathbb{E} R_s|^2 ds \leq C_D \left(\delta + \frac{1}{N\epsilon \delta^2} + \frac{1}{\sqrt{N\epsilon \delta}} \right) \leq \frac{3C_D}{2} \left(\delta + \frac{1}{N\epsilon \delta^2} \right).$$

Taking $\delta = (N\epsilon)^{-1/3}$ proves (3.11). $\square$

Proof of Proposition 3.1. By Lemma 3.4 and Taylor's theorem,

$$(3.33) \quad \left| \partial_t \bar{F}_N^\nu(t, q) - \int_0^1 \xi(\partial_q \bar{F}_N^\nu(t, q, s)) ds \right| \leq \frac{M_2}{2} \int_0^1 \mathbb{E} |R_s - \mathbb{E} R_s|^2 ds.$$

Lemma 3.5 now gives (3.1). $\square$

3.3. A finite-union estimate. We now allow arbitrary $N \in \mathbb{N}$, $t \geq 0$, and $q \in \mathcal{Q}_\infty$.

For a nonzero finite reference measure ρ on $(\text{supp } \mu)^N$, define $\Lambda_{\mathbf{g}}$ with ρ in place of $\lambda_N(\nu)$. For smooth increasing q , the representation (3.5) holds with $\bar{F}_{N, \rho}$ in place of $\bar{F}_N^\nu$, and we have

$$(3.34) \quad |\nabla_{\mathbf{x}, i} \Phi| \leq 1, \quad |\nabla_{\mathbf{g}} \Phi|^2 \leq 2t \sup_{\sigma \in (\text{supp } \mu)^N} |\mathbf{h}_N(\sigma)|^2 \leq 2t N M_0.$$

Indeed, a directional derivative in $\mathbf{x}$ or $\mathbf{g}$ solves the linear backward equation obtained by differentiating (3.4). Its terminal value is the Gibbs average of the corresponding linear observable. The maximum principle preserves these terminal bounds. For degenerate q , apply the argument to $q(s) + \gamma s I_D$ and let $\gamma \downarrow 0$ using the cascade representation and Lemma 2.1.

The following estimate compares the full free energy with the minimum over empirical measures and is adapted from [29, Proof of Lemma 3.4, equations (3.17)–(3.18)].

Lemma 3.6. *Suppose that μ has finite support. For $T, Q \in \mathbb{R}_+$, put $C := T M_0 + (1 + \sqrt{D})Q$. For every $N \in \mathbb{N}$, $t \in [0, T]$, $q \in \mathcal{Q}_\infty$ with $|q|_{L^\infty} \leq Q$, and $s_0 \in (0, 1/2]$, we have*

$$(3.35) \quad 0 \leq \min_{\nu \in \Delta_{m, N}} \bar{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(t, q) - \bar{F}_{N, \mu^{\otimes N}}(t, q) \leq C s_0 + \frac{m \log(N+1)}{N s_0}.$$

Proof. Restriction decreases the partition function, which proves the first inequality. By smooth approximation and Lemma 2.1, it suffices to prove the second inequality for smooth q .

In this proof, Φ and Φ_ν denote the solutions of (3.4) with reference measures $\mu^{\otimes N}$ and $\mathbf{1}_{\Sigma_N(\nu)}\mu^{\otimes N}$, respectively, using the Gaussian process fixed in (3.2). The restrictions retain their original masses as in (2.16). Since the sets $\Sigma_N(\nu)$ partition $(\text{supp } \mu)^N$,

$$(3.36) \quad \Phi(1, \mathbf{x}; \mathbf{g}) = \log \sum_{\nu \in \Delta_{m,N}} \exp(\Phi_\nu(1, \mathbf{x}; \mathbf{g})).$$

The representation (3.5) gives

$$(3.37) \quad \begin{aligned} \bar{F}_{N, \mu^{\otimes N}}(t, q) &= -\frac{1}{N} \mathbb{E} \Phi(0, \sqrt{2q(0)} \mathbf{z}_0; \mathbf{g}), \\ \bar{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(t, q) &= -\frac{1}{N} \mathbb{E} \Phi_\nu(0, \sqrt{2q(0)} \mathbf{z}_0; \mathbf{g}), \end{aligned}$$

where $\mathbf{z}_0$ is standard Gaussian in $\mathbb{R}^{D \times N}$ and independent of $\mathbf{g}$.

Comparison on $[s_0, 1]$. For $s \in [s_0, 1]$, let

$$\widehat{\Phi}(s, \mathbf{x}; \mathbf{g}) := \frac{1}{s} \log \sum_{\nu \in \Delta_{m,N}} e^{s\Phi_\nu(s, \mathbf{x}; \mathbf{g})}, \quad p_\nu := \frac{e^{s\Phi_\nu}}{\sum_{\nu'} e^{s\Phi_{\nu'}}}.$$

We claim that for all $(s, \mathbf{x}) \in [s_0, 1] \times \mathbb{R}^{D \times N}$

$$(3.38) \quad \Phi(s, \mathbf{x}; \mathbf{g}) \leq \widehat{\Phi}(s, \mathbf{x}; \mathbf{g}).$$

This follows from the following standard maximum principle argument.

We have $\nabla_{\mathbf{x}} \widehat{\Phi} = \sum_{\nu} p_\nu \nabla_{\mathbf{x}} \Phi_\nu$, and differentiation gives

$$\begin{aligned} \partial_s \widehat{\Phi} &= \sum_{\nu} p_\nu \partial_s \Phi_\nu + \frac{1}{s^2} \sum_{\nu} p_\nu \log p_\nu, \\ \sum_{i=1}^N \dot{q}(s) \cdot \nabla_{\mathbf{x}, i}^2 \widehat{\Phi} &= \sum_{\nu} p_\nu \sum_{i=1}^N \dot{q}(s) \cdot \nabla_{\mathbf{x}, i}^2 \Phi_\nu + s \dot{q}(s) \cdot \left(\sum_{\nu} p_\nu \nabla_{\mathbf{x}} \Phi_\nu (\nabla_{\mathbf{x}} \Phi_\nu)^\top - \nabla_{\mathbf{x}} \widehat{\Phi} (\nabla_{\mathbf{x}} \widehat{\Phi})^\top \right). \end{aligned}$$

Substituting the equations for Φ_ν , the quadratic terms cancel:

$$(3.39) \quad \partial_s \widehat{\Phi} + \sum_{i=1}^N \dot{q}(s) \cdot \nabla_{\mathbf{x}, i}^2 \widehat{\Phi} + s \dot{q}(s) \cdot \nabla_{\mathbf{x}} \widehat{\Phi} (\nabla_{\mathbf{x}} \widehat{\Phi})^\top = \frac{1}{s^2} \sum_{\nu} p_\nu \log p_\nu \leq 0.$$

Let $\widetilde{\Phi} := \widehat{\Phi} - \Phi$. By (3.36), the terminal condition equals $\widetilde{\Phi}(1, \cdot; \mathbf{g}) = 0$. Subtracting the equation for Φ from (3.39) gives

$$\partial_s \widetilde{\Phi} + \sum_{i=1}^N \dot{q}(s) \cdot \nabla_{\mathbf{x}, i}^2 \widetilde{\Phi} + \sum_{i=1}^N b_i \cdot \nabla_{\mathbf{x}, i} \widetilde{\Phi} \leq 0, \quad b_i := s \dot{q}(s) (\nabla_{\mathbf{x}, i} \widehat{\Phi} + \nabla_{\mathbf{x}, i} \Phi).$$

After the change of variables $\tau = 1 - s$, the parabolic maximum principle, together with the zero terminal value of $\widetilde{\Phi}$, gives $\widetilde{\Phi} \geq 0$ on $[s_0, 1]$. This proves (3.38).

Control of $[0, s_0]$. Let B be a standard Brownian motion in $\mathbb{R}^{D \times N}$ independent of $(\mathbf{g}, \mathbf{z}_0)$, and let

$$\mathbf{Y}_s := \sqrt{2q(0)} \mathbf{z}_0 + \int_0^s \sqrt{2\dot{q}(r)} dB_r.$$

For $\Psi \in \{\Phi\} \cup \{\Phi_\nu : \nu \in \Delta_{m,N}\}$, Itô's formula, the Parisi PDE, and the column bound in (3.34) give

$$(3.40) \quad \begin{aligned} 0 &\leq \mathbb{E} \Psi(0, \mathbf{Y}_0; \mathbf{g}) - \mathbb{E} \Psi(s_0, \mathbf{Y}_{s_0}; \mathbf{g}) \\ &= \mathbb{E} \int_0^{s_0} r \dot{q}(r) \cdot \nabla_{\mathbf{x}} \Psi(r, \mathbf{Y}_r; \mathbf{g}) (\nabla_{\mathbf{x}} \Psi(r, \mathbf{Y}_r; \mathbf{g}))^\top dr \leq N s_0 \text{tr } q(1) \leq N s_0 \sqrt{D} Q. \end{aligned}$$

Moreover, $\mathbf{Y}_{s_0}$ has the same law as $\sqrt{2q(s_0)} \mathbf{z}$, with $\mathbf{z}$ standard Gaussian and independent of $\mathbf{g}$.

Gaussian concentration. By (3.34) and the chain rule,

$$\left| \nabla_{(\mathbf{g}, \mathbf{z})} [\Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g})] \right|^2 \leq 2t N M_0 + 2N \|q(s_0)\|_{\text{op}} \leq 2N (T M_0 + Q).$$

Gaussian concentration therefore gives

$$\log \mathbb{E} e^{s_0 \Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g})} \leq s_0 \mathbb{E} \Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g}) + N s_0^2 (TM_0 + Q).$$

By Jensen's inequality,

$$\begin{aligned} \mathbb{E} \widehat{\Phi}(s_0, \mathbf{Y}_{s_0}; \mathbf{g}) &\leq \frac{1}{s_0} \log \sum_{\nu \in \Delta_{m,N}} \mathbb{E} e^{s_0 \Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g})} \\ (3.41) \quad &\leq \max_{\nu \in \Delta_{m,N}} \mathbb{E} \Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g}) + \frac{\log |\Delta_{m,N}|}{s_0} + N s_0 (TM_0 + Q). \end{aligned}$$

Combining (3.38)–(3.41), and using $\mathbb{E} \Phi_\nu(s_0, \sqrt{2q(s_0)} \mathbf{z}; \mathbf{g}) \leq \mathbb{E} \Phi_\nu(0, \mathbf{Y}_0; \mathbf{g})$ from (3.40), we obtain

$$(3.42) \quad \mathbb{E} \Phi(0, \mathbf{Y}_0; \mathbf{g}) \leq \max_{\nu \in \Delta_{m,N}} \mathbb{E} \Phi_\nu(0, \mathbf{Y}_0; \mathbf{g}) + \frac{\log |\Delta_{m,N}|}{s_0} + N s_0 (TM_0 + (1 + \sqrt{D})Q).$$

Together with (3.37) and $|\Delta_{m,N}| \leq (N+1)^m$, this proves the second inequality in (3.35). $\square$

4. INITIAL CONDITIONS FOR FINITE SPIN SETS

In this section, assume $\text{supp } \mu = \{a_1, \dots, a_m\}$ and retain the notation $\mu_j := \mu(\{a_j\}) > 0$ for $j \in \{1, \dots, m\}$. We equip $\mathcal{P}(\text{supp } \mu)$ with the W_1 topology. Since $\text{supp } \mu$ is finite, this is equivalent to the Euclidean topology under the identification $\nu \mapsto (\nu_1, \dots, \nu_m)$, where $\nu_j := \nu(\{a_j\})$. The path q belongs to $\mathcal{Q}_1$; the Gibbs computations below are first made for $q \in \mathcal{Q}_\infty$ and then extended by approximation.

For $h \in \mathbb{R}^m$, interpret h as the function on $\text{supp } \mu$ satisfying $h(a_j) = h_j$ for every $j \in \{1, \dots, m\}$, and write $h \cdot \nu := \sum_{j=1}^m h_j \nu_j$ for $\nu \in \mathcal{P}(\text{supp } \mu)$. By (1.3), we have $\psi_{\mu,h}(q) = \bar{F}_{1,e^h \mu}(0, q)$. The same identity holds at every system size:

$$(4.1) \quad \bar{F}_{N,(e^h \mu)^{\otimes N}}(0, q) = \psi_{\mu,h}(q), \quad N \in \mathbb{N};$$

this is the tensorization property in [19, Proposition 3.2].

Proposition 4.1. *Assume $\text{supp } \mu = \{a_1, \dots, a_m\}$ for some $m \in \mathbb{N}$. For every $q \in \mathcal{Q}_1$, there is a continuous function $g_q : \mathcal{P}(\text{supp } \mu) \rightarrow \mathbb{R}$ such that*

$$(4.2) \quad g_q(\nu) = \sup_{h \in \mathbb{R}^m} \{\psi_{\mu,h}(q) + h \cdot \nu\} = \text{Ent}(\nu \mid \mu) + \varphi^\nu(q), \quad \nu \in \mathcal{P}(\text{supp } \mu),$$

$$(4.3) \quad \lim_{N \rightarrow \infty} \sup_{\nu \in \Delta_{m,N}} |\bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}(0, q) - g_q(\nu)| = 0.$$

For every $Q \in \mathbb{R}_+$, the convergence in (4.3) is uniform over $q \in \mathcal{Q}_\infty$ satisfying $|q|_{L^\infty} \leq Q$. Moreover, φ^ν is 1-Lipschitz in L^1 , increasing along $\mathcal{Q}_2^*$, and satisfies $0 \leq \varphi^\nu(q) \leq |q|_{L^1}$.

Proof. We first fix $q \in \mathcal{Q}_\infty$.

Step 1: Compactness on $\Delta_{m,N}$ and the dual identity. Define $G_{N,q} : \Delta_{m,N} \rightarrow \mathbb{R}$ by

$$G_{N,q}(\nu) := \bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}(0, q).$$

Combining (2.17), Lemma 2.5, and (2.18), we obtain, for $\nu, \nu' \in \Delta_{m,N}$,

$$(4.4) \quad |G_{N,q}(\nu) - G_{N,q}(\nu')| \leq 2|q|_{L^\infty} W_1(\nu, \nu') + |\text{Ent}(\nu \mid \mu) - \text{Ent}(\nu' \mid \mu)| + \frac{m \log(N+1)}{N}.$$

The map $\nu \mapsto \text{Ent}(\nu \mid \mu) = \sum_{j=1}^m \nu_j \log(\nu_j / \mu_j)$ is continuous on the compact set $\mathcal{P}(\text{supp } \mu)$, with the convention $0 \log 0 := 0$. Thus (4.4) gives a common asymptotic modulus of continuity. Moreover, Lemma 2.1, applied to the probability measure $\lambda_N(\nu)$ at $t = 0$, gives $|\bar{F}_N^\nu(0, q)| \leq |q|_{L^1}$. Together with (2.17) and (2.18), this shows that $(G_{N,q})_{N \in \mathbb{N}}$ is uniformly bounded on its varying domains.

The grids $\Delta_{m,N}$ become dense in $\mathcal{P}(\text{supp } \mu)$ for W_1 : rounding the masses of any $\nu \in \mathcal{P}(\text{supp } \mu)$ to multiples of N^{-1} gives $\nu_N \in \Delta_{m,N}$ with $\lim_{N \rightarrow \infty} W_1(\nu_N, \nu) = 0$. Consequently, a finite-net and diagonal argument using (4.4) shows that every subsequence of $(G_{N,q})$ has a further subsequence, still indexed by N , and a continuous function $g_q : \mathcal{P}(\text{supp } \mu) \rightarrow \mathbb{R}$ such that

$$(4.5) \quad \lim_{N \rightarrow \infty} \sup_{\nu \in \Delta_{m,N}} |G_{N,q}(\nu) - g_q(\nu)| = 0.$$

We next identify g_q . Fix $h \in \mathbb{R}^m$ and let

$$Z_h := \sum_{j=1}^m \mu_j e^{h_j} \quad \text{and} \quad \mu_h := Z_h^{-1} e^h \mu.$$

The measure μ_h is a probability measure with support $\text{supp } \mu$. For every $\nu \in \Delta_{m,N}$, by (2.12), every $\sigma \in \Sigma_N(\nu)$ has exactly $N\nu_j$ coordinates equal to a_j . Hence the density of $\mu_h^{\otimes N}$ relative to $\mu^{\otimes N}$ is constant on $\Sigma_N(\nu)$ and equals $Z_h^{-N} e^{N h \cdot \nu}$. Therefore,

$$(4.6) \quad \bar{F}_{N, \mu_h^{\otimes N}, \Sigma_N(\nu)}(0, q) = G_{N,q}(\nu) + \log Z_h - h \cdot \nu.$$

Equation (4.1) also gives

$$(4.7) \quad \bar{F}_{N, \mu_h^{\otimes N}}(0, q) = \psi_{\mu,h}(q) + \log Z_h.$$

Applying Lemma 3.6 to the reference measure μ_h , with $t = 0$ and $s_{0,N} := \min\{\frac{1}{2}, \sqrt{\log(N+1)/N}\}$, and using (4.6) and (4.7), we obtain

$$\lim_{N \rightarrow \infty} \left| \min_{\nu \in \Delta_{m,N}} \{G_{N,q}(\nu) - h \cdot \nu\} - \psi_{\mu,h}(q) \right| = 0.$$

Passing along the subsequence in (4.5), using the density of $\Delta_{m,N}$ in $\mathcal{P}(\text{supp } \mu)$ and the continuity of g_q , gives

$$(4.8) \quad \psi_{\mu,h}(q) = \min_{\nu \in \mathcal{P}(\text{supp } \mu)} \{g_q(\nu) - h \cdot \nu\}, \quad h \in \mathbb{R}^m.$$

Step 2: Identification of g_q by convex duality. For bounded q , differentiation with respect to h gives

$$(4.9) \quad \partial_{h_j} \psi_{\mu,h}(q) = -\mathbb{E}\langle \mathbf{1}_{\{\sigma=a_j\}} \rangle, \quad j \in \{1, \dots, m\}.$$

A second differentiation expresses the Hessian as the negative of an expected covariance matrix of the indicators $(\mathbf{1}_{\{\sigma=a_j\}})_{j=1}^m$. Since these indicators are bounded by one, the Hessian is uniformly bounded. Thus $h \mapsto \psi_{\mu,h}(q)$ is continuously differentiable.

Fix $h \in \mathbb{R}^m$, and let $\eta \in \mathcal{P}(\text{supp } \mu)$ be a minimizer in (4.8); such a minimizer exists because $\mathcal{P}(\text{supp } \mu)$ is compact and g_q is continuous. For every $v \in \mathbb{R}^m$ and $\varepsilon \in \mathbb{R}$, we have

$$\psi_{\mu, h+\varepsilon v}(q) \leq g_q(\eta) - (h + \varepsilon v) \cdot \eta = \psi_{\mu,h}(q) - \varepsilon v \cdot \eta.$$

Dividing by ε for $\varepsilon > 0$ and $\varepsilon < 0$ and then sending $\varepsilon \rightarrow 0$ gives

$$(4.10) \quad \eta = -\nabla_h \psi_{\mu,h}(q).$$

We now fix ν in the relative interior of $\mathcal{P}(\text{supp } \mu)$ and show that there is an h for which ν is the minimizer in (4.8). Equation (2.4) and the identity $\psi_{\mu,h}(0) = -\log \sum_{j=1}^m \mu_j e^{h_j}$ give

$$(4.11) \quad \left| \psi_{\mu,h}(q) + \log \sum_{j=1}^m \mu_j e^{h_j} \right| \leq |q|_{L^1}.$$

Define $J_\nu(h) := \psi_{\mu,h}(q) + h \cdot \nu$. Translation invariance, $J_\nu(h + c\mathbf{1}) = J_\nu(h)$, allows us to normalize $\max_j h_j = 0$. Put $\mu_* := \min_j \mu_j > 0$ and $\nu_* := \min_j \nu_j > 0$. On this slice, (4.11) gives

$$J_\nu(h) \leq -\log \mu_* + |q|_{L^1} - \nu_* |h|_\infty.$$

Thus J_ν attains a maximum on the slice, and translation invariance makes the maximizer h_* global on $\mathbb{R}^m$. This also applies when $m = 1$, for which the slice is the single point 0. Differentiability at the global maximum gives

$$\nu = -\nabla_h \psi_{\mu,h_*}(q).$$

This is precisely (4.10) with $\eta = \nu$. Every minimizer in (4.8) at $h = h_*$ must therefore equal ν . Consequently,

$$(4.12) \quad g_q(\nu) = \psi_{\mu,h_*}(q) + h_* \cdot \nu = \sup_{h \in \mathbb{R}^m} \{\psi_{\mu,h}(q) + h \cdot \nu\}$$

for every ν in the relative interior of $\mathcal{P}(\text{supp } \mu)$.

Equation (4.12) shows that g_q is convex on the relative interior. Since g_q is continuous on $\mathcal{P}(\text{supp } \mu)$, it is convex on all of $\mathcal{P}(\text{supp } \mu)$. Extend g_q to $\mathbb{R}^m$ by setting it equal to $+\infty$ outside $\mathcal{P}(\text{supp } \mu)$. The resulting function is proper, lower semicontinuous, and convex. Equation (4.8) says that its convex conjugate is $h \mapsto -\psi_{\mu,h}(q)$. The finite-dimensional Fenchel–Moreau theorem therefore gives

$$(4.13) \quad g_q(\nu) = \sup_{h \in \mathbb{R}^m} \{ \psi_{\mu,h}(q) + h \cdot \nu \}, \quad \nu \in \mathcal{P}(\text{supp } \mu).$$

The right-hand side of (4.13) is independent of the subsequence chosen in (4.5). Hence every subsequence has the same subsequential limit, and the whole sequence satisfies (4.3). Equation (4.13) proves the first equality in (4.2).

Step 3: Identification with φ^ν and uniformity in q . If $\nu_j > 0$ for every j , change the tilt by

$$\widehat{h}_j := h_j + \log(\mu_j/\nu_j).$$

Then $e^h \mu = e^{\widehat{h}} \nu$ and $h \cdot \nu = \widehat{h} \cdot \nu + \text{Ent}(\nu \mid \mu)$, so (4.13) and (1.3) give $g_q(\nu) = \text{Ent}(\nu \mid \mu) + \varphi^\nu(q)$ directly.

For a boundary measure ν , restrict first to $\text{supp } \nu$. For any $h \in \mathbb{R}^m$, removing the mass outside $\text{supp } \nu$ increases the free energy, and the same tilt change on $\text{supp } \nu$ gives

$$\psi_{\mu,h}(q) + h \cdot \nu \leq \text{Ent}(\nu \mid \mu) + \varphi^\nu(q).$$

Conversely, fix a tilt k on $\text{supp } \nu$ and let $h_j = k_j + \log(\nu_j/\mu_j)$ for $a_j \in \text{supp } \nu$, and $h_j = -M$ for $a_j \notin \text{supp } \nu$. As $M \rightarrow \infty$, the partition functions decrease to those with reference measure $e^k \nu$. Their logarithms lie between the integrable logarithms for $M = 0$ and the restricted measure, so dominated convergence gives $\psi_{\mu,h}(q) \rightarrow \psi_{\nu,k}(q)$. Taking the supremum over k proves

$$(4.14) \quad g_q(\nu) = \text{Ent}(\nu \mid \mu) + \varphi^\nu(q), \quad \nu \in \mathcal{P}(\text{supp } \mu).$$

This proves the second equality in (4.2). The normalization identity (2.17) and the type-probability bound (2.18) transfer the uniform convergence in (4.3) to the normalized free energies. Since g_q and entropy are continuous,

$$(4.15) \quad \lim_{N \rightarrow \infty} \overline{F}_N^{\nu_N}(0, q) = \varphi^\nu(q) \quad \text{whenever } \nu_N \in \Delta_{m,N}, \quad \nu_N \rightarrow \nu \text{ in } W_1.$$

By Lemma 2.2, φ^ν is 1-Lipschitz in L^1 . Together with Lemma 2.1, this extends (4.2) and (4.3) from bounded paths to every $q \in \mathcal{Q}_1$ by L^1 approximation.

Finally, fix $Q \in \mathbb{R}_+$ and define

$$\mathcal{Q}_{\infty, \leq Q} := \{q \in \mathcal{Q}_\infty : |q|_{L^\infty} \leq Q\}.$$

The set $\mathcal{Q}_{\infty, \leq Q}$ is relatively compact in L^1 . Indeed, for every partition $0 = s_0 < \dots < s_k < 1$, positivity of the increments gives

$$\sum_{\ell=1}^k |q(s_\ell) - q(s_{\ell-1})| \leq \sum_{\ell=1}^k \text{tr}(q(s_\ell) - q(s_{\ell-1})) \leq \sqrt{D} Q.$$

Helly's selection argument and dominated convergence therefore give L^1 compactness after taking right-continuous representatives. By (2.17) and Lemma 2.1, the map $q \mapsto \overline{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(0, q)$ is 1-Lipschitz in L^1 , uniformly in N and ν . By (4.14) and the L^1 -Lipschitz bound for φ^ν , the map $q \mapsto g_q(\nu)$ has the same bound, uniformly in ν .

Fix $\varepsilon > 0$ and choose $q^1, \dots, q^R \in \mathcal{Q}_{\infty, \leq Q}$ such that every $q \in \mathcal{Q}_{\infty, \leq Q}$ satisfies $\min_{r \in \{1, \dots, R\}} |q - q^r|_{L^1} < \varepsilon$. By (4.3), for all sufficiently large N we have

$$\max_{r \in \{1, \dots, R\}} \sup_{\nu \in \Delta_{m,N}} |\overline{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(0, q^r) - g_{q^r}(\nu)| < \varepsilon.$$

For $q \in \mathcal{Q}_{\infty, \leq Q}$, choose r with $|q - q^r|_{L^1} < \varepsilon$. The two Lipschitz estimates and the preceding display give

$$\sup_{\nu \in \Delta_{m,N}} |\overline{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(0, q) - g_q(\nu)| < 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we obtain

$$\lim_{N \rightarrow \infty} \sup_{\substack{q \in \mathcal{Q}_\infty \\ |q|_{L^\infty} \leq Q}} \sup_{\nu \in \Delta_{m,N}} |\bar{F}_{N,\mu^{\otimes N},\Sigma_N(\nu)}(0,q) - g_q(\nu)| = 0.$$

The proof is complete. $\square$

Lemma 4.2. *Suppose that μ has finite support. For every $\nu, \nu' \in \mathcal{P}(\text{supp } \mu)$, the initial data satisfy*

$$(4.16) \quad |\varphi^\nu(q) - \varphi^{\nu'}(q)| \leq 2|q|_{L^\infty} W_1(\nu, \nu'), \quad q \in \mathcal{Q}_\infty.$$

Proof. Choose $\nu_N, \nu'_N \in \Delta_{m,N}$ converging to ν, ν' in W_1 . Lemma 2.5 at $t = 0$ gives

$$|\bar{F}_N^{\nu_N}(0,q) - \bar{F}_N^{\nu'_N}(0,q)| \leq 2|q|_{L^\infty} W_1(\nu_N, \nu'_N).$$

Pass to the limit using (4.15). $\square$

5. IDENTIFICATION OF THE CONSTRAINED VISCOSITY SOLUTIONS

We now give a precise meaning to the Hamilton–Jacobi solution f^ν appearing in Theorem 1.1 and identify f^ν as the limit of the normalized constrained free energies when μ has finite support.

5.1. Viscosity solutions and finite-dimensional approximations. Let E be a finite-dimensional Hilbert space, let $U \subseteq (0, \infty) \times E$ be open, and let $H : E \rightarrow \mathbb{R}$ be continuous.

Definition 5.1 (Viscosity solutions). A continuous function $u : U \rightarrow \mathbb{R}$ is a **viscosity subsolution** of $\partial_t u - H(D_x u) = 0$ if, for every $\phi \in C^1(U)$ and every local maximum $(t, x) \in U$ of $u - \phi$, we have

$$\partial_t \phi(t, x) - H(D_x \phi(t, x)) \leq 0.$$

A **viscosity supersolution** satisfies the reverse inequality at every local minimum of $u - \phi$. A **viscosity solution** is both a viscosity subsolution and a viscosity supersolution.

We use Definition 5.1 only for finite-dimensional equations; the solution on $\mathcal{Q}_2$ is defined by finite-dimensional approximations below.

The spin-glass Hamilton–Jacobi equation that we wish to interpret is

$$(5.1) \quad \partial_t f - \int_0^1 \xi(\partial_q f(s)) ds = 0 \quad \text{on } \mathbb{R}_+ \times \mathcal{Q}_2.$$

We define its solution through finite-dimensional approximations. For $K \in \mathbb{N}$, equip $E_K := (S^D)^K$ with the inner product $\langle a, b \rangle_K := K^{-1} \sum_{k=1}^K a_k \cdot b_k$, and write $|a|_{\ell_K^1} := K^{-1} \sum_{k=1}^K |a_k|$ and $|a|_{\ell^\infty} := \max_{k \in \{1, \dots, K\}} |a_k|$. Define the closed convex cone

$$\mathcal{Q}^K := \{a \in E_K : 0 \leq a_1 \leq \dots \leq a_K\}.$$

As in (2.2), define the dual cone $(\mathcal{Q}^K)^* := \{b \in E_K : \langle b, a \rangle_K \geq 0 \text{ for every } a \in \mathcal{Q}^K\}$. As for the notion of $\mathcal{Q}_2^*$ -increasing introduced after (2.2), a function $u : \mathcal{Q}^K \rightarrow \mathbb{R}$ is **$(\mathcal{Q}^K)^*$ -increasing** if $u(a) \geq u(b)$ whenever $a, b \in \mathcal{Q}^K$ satisfy $a - b \in (\mathcal{Q}^K)^*$. For the uniform partition of $[0, 1]$, define the lift $L_K : E_K \rightarrow L^2([0, 1]; S^D)$ and projection $P_K : L^2([0, 1]; S^D) \rightarrow E_K$ by

$$L_K a := \sum_{k=1}^K a_k \mathbf{1}_{[(k-1)/K, k/K)} \quad \text{and} \quad (P_K p)_k := K \int_{(k-1)/K}^{k/K} p(s) ds.$$

These maps preserve the corresponding cones in the sense that $P_K(\mathcal{Q}_2) = \mathcal{Q}^K$ and $L_K(\mathcal{Q}^K) \subseteq \mathcal{Q}_2$; see [21, Lemmas 3.3 and 3.5]. Let ∇_K denote the gradient in the weighted inner product on E_K ; thus it is K times the ordinary Euclidean gradient. Define the finite-dimensional Hamiltonian $\xi_K : E_K \rightarrow \mathbb{R}$ by

$$(5.2) \quad \xi_K(p) := \frac{1}{K} \sum_{k=1}^K \xi(p_k).$$

For every $\nu \in \mathcal{P}(\text{supp } \mu)$, Lemma 2.2 supplies the regularity and monotonicity of φ^ν required below. The restriction of ξ_K to $\mathcal{Q}^K$ is locally Lipschitz and increasing along $(\mathcal{Q}^K)^*$, by the

covariance assumptions on ξ ; see [21, proof of Theorem 4.7(1)]. Applying [21, Theorem 2.2(2)], let $f^{\nu,K} : \mathbb{R}_+ \times \mathcal{Q}^K \rightarrow \mathbb{R}$ be the unique Lipschitz, $(\mathcal{Q}^K)^*$ -increasing viscosity solution, in the sense of Definition 5.1 on $(0, \infty) \times \text{int } \mathcal{Q}^K$, of

$$(5.3) \quad \begin{cases} \partial_t f^{\nu,K} - \xi_K(\nabla_K f^{\nu,K}) = 0 & \text{on } (0, \infty) \times \text{int } \mathcal{Q}^K, \\ f^{\nu,K}(0, \cdot) = \varphi^\nu(L_K \cdot) & \text{on } \mathcal{Q}^K. \end{cases}$$

The monotonicity of ξ_K along $(\mathcal{Q}^K)^*$ makes a separate boundary condition unnecessary: existence and uniqueness follow from [21, Theorem 2.2], and the continuous extension from the interior solves the equation on the closed cone in the sense of [21, Proposition 2.3].

Definition 5.2 (Solution of the spin-glass Hamilton–Jacobi equation). For $\nu \in \mathcal{P}(\text{supp } \mu)$, define $f^\nu : \mathbb{R}_+ \times \mathcal{Q}_2 \rightarrow \mathbb{R}$ by

$$(5.4) \quad f^\nu(t, q) := \lim_{K \rightarrow \infty} f^{\nu,K}(t, P_K q), \quad (t, q) \in \mathbb{R}_+ \times \mathcal{Q}_2.$$

The limit in (5.4) exists locally uniformly. Indeed, the uniform-refinement estimate in [21, Proof of Proposition 3.10, (3.42)] allows one to compare the approximations on grids of sizes K and K' through their common refinement of size $\text{lcm}(K, K')$. Its identification as the unique Lipschitz viscosity solution of (5.1) in the infinite-dimensional sense is proved in [21, Definition 4.3 and Theorem 4.7(1)]; see also [21, Remark 4.11 and Proposition 4.13]. We do not repeat the infinite-dimensional definition, since (5.3) and (5.4) suffice for our purposes.

By [21, Proposition 2.4 and Theorem 4.7(1)], the finite-dimensional solutions preserve the 1-Lipschitz constant in the weighted ℓ^1 norm. Their time Lipschitz constants are at most M_0 (see (2.3)), because (5.2) is bounded by M_0 on $\mathcal{Q}^K \cap \{p : |p|_{\ell^\infty} \leq 1\}$. These bounds pass to (5.4); in particular, we have

$$(5.5) \quad |f^\nu(t, q) - f^\nu(t', q')| \leq M_0 |t - t'| + |q - q'|_{L^1}, \quad (t, q), (t', q') \in \mathbb{R}_+ \times \mathcal{Q}_2.$$

5.2. Finite-dimensional comparison. For the rest of this section, assume that $\text{supp } \mu = \{a_1, \dots, a_m\}$ is finite. We retain the normalized constrained free energies $\bar{F}_N^\nu$ (see (2.13)), with initial limits φ^ν from Proposition 4.1. We pass from the finite-size residual to viscosity comparison using the maps L_K and P_K defined above.

Lemma 5.3. *For every $p \in \mathcal{Q}_\infty$ with $|p|_{L^\infty} \leq 1$, we have*

$$|p - L_K P_K p|_{L^1} \leq \frac{\sqrt{D}}{K} \quad \text{and} \quad \left| \int_0^1 \xi(p(s)) \, ds - \xi_K(P_K p) \right| \leq \frac{M_1 \sqrt{D}}{K}.$$

The Hamiltonian ξ_K satisfies the Lipschitz bound

$$(5.6) \quad |\xi_K(p) - \xi_K(p')| \leq M_1 |p - p'|_{\ell_K^1}, \quad p, p' \in \mathcal{Q}^K, \quad |p|_{\ell^\infty} \vee |p'|_{\ell^\infty} \leq 1.$$

Proof. Fix $p \in \mathcal{Q}_\infty$ and let $I_k := [(k-1)/K, k/K)$ for $k \in \{1, \dots, K\}$. Since p is increasing in the positive-semidefinite order, we have

$$|p(s) - p(r)| \leq \text{tr}(p(s) - p(r)), \quad 0 \leq r \leq s < 1.$$

By Jensen's inequality and the definition of P_K , we obtain

$$\int_{I_k} |p(s) - (P_K p)_k| \, ds \leq K \int_{I_k} \int_{I_k} |p(s) - p(r)| \, dr \, ds \leq \frac{1}{K} \left(\sup_{s \in I_k} \text{tr } p(s) - \inf_{s \in I_k} \text{tr } p(s) \right).$$

The increments in the last line sum to at most $\sup_{s \in [0,1]} \text{tr } p(s)$, because $\text{tr } p$ is nonnegative and increasing. Consequently, for every $p \in \mathcal{Q}_\infty$, we have

$$(5.7) \quad |p - L_K P_K p|_{L^1} \leq \frac{1}{K} \sup_{s \in [0,1]} \text{tr } p(s) \leq \frac{\sqrt{D}}{K} |p|_{L^\infty}.$$

This proves the first assertion of the lemma when $|p|_{L^\infty} \leq 1$.

In this case, both $p(s)$ and $(L_K P_K p)(s)$ belong to the unit ball of S^D . By (2.3), the mean-value theorem gives

$$\left| \int_0^1 \xi(p(s)) ds - \xi_K(P_K p) \right| \leq M_1 |p - L_K P_K p|_{L^1} \stackrel{(5.7)}{\leq} \frac{M_1 \sqrt{D}}{K}.$$

Finally, for $p, p' \in \mathcal{Q}^K$ satisfying $|p|_{\ell^\infty} \vee |p'|_{\ell^\infty} \leq 1$, the line segment between p_k and p'_k lies in the unit ball. The definition of M_1 gives

$$|\xi_K(p) - \xi_K(p')| \leq \frac{1}{K} \sum_{k=1}^K |\xi(p_k) - \xi(p'_k)| \leq \frac{M_1}{K} \sum_{k=1}^K |p_k - p'_k| = M_1 |p - p'|_{\ell_K^1}.$$

This proves (5.6). $\square$

Lemma 5.4. *Proposition 3.1 also holds at $q = L_K a + \varepsilon \ell$, where $a \in \mathcal{Q}^K$ and $\ell : [0, 1] \rightarrow S_+^D$ is given by $\ell(s) := s I_D$.*

Proof. Fix $a \in \mathcal{Q}^K$ and let $\widehat{q}_a : \mathbb{R} \rightarrow S_+^D$ be the extension of $L_K a$ that is constant on $(-\infty, 0)$ and $[1, \infty)$. Choose a nonnegative $\kappa \in C_c^\infty(\mathbb{R})$ with $\int_{\mathbb{R}} \kappa = 1$. For $\eta > 0$, define $\kappa_\eta(r) := \eta^{-1} \kappa(r/\eta)$ and $q_\eta^0 := \kappa_\eta * \widehat{q}_a$. Since $\widehat{q}_a$ is positive semidefinite and increasing, the same holds for q_η^0 . In addition, q_η^0 is smooth, and the standard mollifying argument gives $\lim_{\eta \downarrow 0} |q_\eta^0 - L_K a|_{L^1} = 0$. Consequently, $q_\eta := q_\eta^0 + \varepsilon \ell$ satisfies $\dot{q}_\eta(s) \geq \varepsilon I_D$ for every $s \in [0, 1]$ and $\lim_{\eta \downarrow 0} |q_\eta - (L_K a + \varepsilon \ell)|_{L^1} = 0$.

We apply Proposition 3.1 to q_η and let $\eta \downarrow 0$, with N, t, K, a, ε fixed. Equation (2.7) gives convergence of the path derivatives in L^1 . The time derivative is continuous as well: the bounded replica representation in (2.6), differentiated along a Gaussian interpolation, gives

$$|\partial_t \bar{F}_N^\nu(t, q) - \partial_t \bar{F}_N^\nu(t, q')| \leq C N M_0 |q - q'|_{L^1}.$$

Let $q_* := L_K a + \varepsilon \ell$. For fixed N , both derivatives therefore converge as $\eta \downarrow 0$. Their path components have L^∞ norm at most one by Lemma 2.1, so

$$\left| \int_0^1 \xi(\partial_q \bar{F}_N^\nu(t, q_\eta, s)) ds - \int_0^1 \xi(\partial_q \bar{F}_N^\nu(t, q_*, s)) ds \right| \leq M_1 |\partial_q \bar{F}_N^\nu(t, q_\eta) - \partial_q \bar{F}_N^\nu(t, q_*)|_{L^1} \rightarrow 0.$$

The residual bound at q_η is independent of η and passes to q_* . $\square$

We now prove a *quantitative* comparison principle, in the style of [21, Proposition 2.5]. What is important here is that in the constraint on the ℓ^∞ -norm of b in (5.8) below the growth in time (given by M_1) is uniform in K .

Lemma 5.5 (Comparison principle for viscosity solutions). *Fix $K \in \mathbb{N}$ and $T, Q \in \mathbb{R}_+$. Let $u, v : \mathbb{R}_+ \times \mathcal{Q}^K \rightarrow \mathbb{R}$ be respectively a viscosity subsolution and a viscosity supersolution of $\partial_t w - \xi_K(\nabla_K w) = 0$ on $(0, \infty) \times \text{int } \mathcal{Q}^K$. Suppose that $u(t, \cdot)$ and $v(t, \cdot)$ are 1-Lipschitz in $|\cdot|_{\ell_K^1}$ and $(\mathcal{Q}^K)^*$ -increasing for every $t \in \mathbb{R}_+$. Assume also that there exists $C \in \mathbb{R}_+$ such that*

$$|u(t, a) - u(0, a)| \vee |v(t, a) - v(0, a)| \leq C t, \quad (t, a) \in \mathbb{R}_+ \times \mathcal{Q}^K.$$

Then

$$(5.8) \quad \sup_{\substack{0 \leq t \leq T, \\ a \in \mathcal{Q}^K \\ |a|_{\ell^\infty} \leq Q}} (u(t, a) - v(t, a)) \leq \sup_{\substack{b \in \mathcal{Q}^K \\ |b|_{\ell^\infty} \leq Q + M_1 T + 1}} (u(0, b) - v(0, b)).$$

The proof of Lemma 5.5 is a technical adaptation of the standard doubling-of-variables argument. The main difficulty is to establish comparison without imposing a boundary condition, which requires an auxiliary penalization to ensure that the contact points lie in the interior of the cone. We defer the details to Appendix A.

Assume that μ has finite support, fix $T, Q \in \mathbb{R}_+$, and put $Q' := Q + M_1 T + 1$.

Proposition 5.6. *We have*

$$(5.9) \quad \lim_{N \rightarrow \infty} \sup_{\substack{\nu' \in \Delta_{m, N}, \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| = 0.$$

For every $N \in \mathbb{N}$, $\nu \in \Delta_{m,N}$, $t \in [0, T]$, $a \in \mathcal{Q}^K$ with $|a|_{\ell^\infty} \leq Q$, and $\varepsilon > 0$, we have

$$(5.10) \quad \begin{aligned} |\bar{F}_N^{\nu'}(t, L_K a) - f^{\nu,K}(t, a)| &\leq \sup_{\substack{\nu' \in \Delta_{m,N}, \ p \in \mathcal{Q}_\infty \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| \\ &\quad + \varepsilon \sqrt{D} + t \left(C(N\varepsilon)^{-1/3} + \frac{M_1 \sqrt{D}}{K} \right). \end{aligned}$$

Proof. By (2.17), (4.2), and (2.18), for every $\nu' \in \Delta_{m,N}$ and $p \in \mathcal{Q}_\infty$ with $|p|_{L^\infty} \leq Q'$, we have

$$|\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| \leq |\bar{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu')}(0, p) - g_p(\nu')| + \frac{m \log(N+1)}{N}.$$

Taking suprema and using the uniform initial convergence in Proposition 4.1 prove (5.9).

Step 1: Estimate for $\partial_t V_N - \xi_K(\nabla_K V_N)$. Fix $\nu \in \Delta_{m,N}$ and define $V_N : \mathbb{R}_+ \times \mathcal{Q}^K \rightarrow \mathbb{R}$ by

$$V_N(t, a) := \bar{F}_N^{\nu}(t, L_K a + \varepsilon \ell).$$

Write $q_a := L_K a + \varepsilon \ell$ and $p_N := \partial_q \bar{F}_N^{\nu}(t, q_a)$. Lemma 2.1 gives $p_N \in \mathcal{Q}_\infty$ and $|p_N|_{L^\infty} \leq 1$. For a direction $b \in E_K$, the chain rule and the definition of P_K give

$$\langle \nabla_K V_N(t, a), b \rangle_K = \langle p_N, L_K b \rangle_{L^2} = \langle P_K p_N, b \rangle_K.$$

Thus $\nabla_K V_N(t, a) = P_K p_N \in \mathcal{Q}^K$, and $|\nabla_K V_N(t, a)|_{\ell^\infty} \leq 1$. In particular, $V_N(t, \cdot)$ is 1-Lipschitz in $|\cdot|_{\ell_K^1}$ and $(\mathcal{Q}^K)^*$ -increasing.

For $t > 0$ and $a \in \text{int } \mathcal{Q}^K$, Lemma 5.4 permits the use of (3.1) at $q = q_a$. Since $\xi_K(P_K p_N) = \int_0^1 \xi((L_K P_K p_N)(s)) ds$ by (5.2), we have

$$(5.11) \quad \begin{aligned} |\partial_t V_N - \xi_K(\nabla_K V_N)| &\leq \left| \partial_t \bar{F}_N^{\nu}(t, q_a) - \int_0^1 \xi(p_N(s)) ds \right| + \left| \int_0^1 \xi(p_N(s)) ds - \xi_K(P_K p_N) \right| \\ &\stackrel{\text{P. 3.1, L. 5.3}}{\leq} C(N\varepsilon)^{-1/3} + \frac{M_1 \sqrt{D}}{K}. \end{aligned}$$

Step 2: Comparison with $f^{\nu,K}$. Let $\gamma_{N,K,\varepsilon} := C(N\varepsilon)^{-1/3} + M_1 \sqrt{D}/K$. By (5.11), the functions

$$V_N^-(t, a) := V_N(t, a) - \gamma_{N,K,\varepsilon} t \quad \text{and} \quad V_N^+(t, a) := V_N(t, a) + \gamma_{N,K,\varepsilon} t$$

are respectively a viscosity subsolution and a viscosity supersolution of (5.3) on the interior of $\mathcal{Q}^K$. For each fixed $t \in \mathbb{R}_+$, the functions $V_N^-(t, \cdot)$ and $V_N^+(t, \cdot)$ remain 1-Lipschitz in $|\cdot|_{\ell_K^1}$ and $(\mathcal{Q}^K)^*$ -increasing. By (2.4), their time increments are bounded by $(M_0 + \gamma_{N,K,\varepsilon})|t - t'|$, uniformly in a , as required in Lemma 5.5. The comparison lemma requires these PDE inequalities only on the interior.

For every $b \in \mathcal{Q}^K$ with $|b|_{\ell^\infty} \leq Q'$, the initial condition for $f^{\nu,K}$, Lemma 2.1, and $|\ell|_{L^1} = \frac{1}{2} \sqrt{D}$ yield

$$(5.12) \quad |V_N(0, b) - f^{\nu,K}(0, b)| \leq \sup_{\substack{\nu' \in \Delta_{m,N}, \ p \in \mathcal{Q}_\infty \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| + \frac{1}{2} \varepsilon \sqrt{D}.$$

Apply Lemma 5.5 to $(u, v) = (V_N^-, f^{\nu,K})$ and then to $(u, v) = (f^{\nu,K}, V_N^+)$. Since $Q' = Q + M_1 T + 1$, equation (5.12) gives, for $0 \leq t \leq T$ and $|a|_{\ell^\infty} \leq Q$,

$$(5.13) \quad \begin{aligned} |V_N(t, a) - f^{\nu,K}(t, a)| &\leq \sup_{\substack{\nu' \in \Delta_{m,N}, \ p \in \mathcal{Q}_\infty \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| \\ &\quad + \frac{1}{2} \varepsilon \sqrt{D} + \gamma_{N,K,\varepsilon} t. \end{aligned}$$

Finally, the L^1 -Lipschitz estimate (2.4) gives

$$|\bar{F}_N^{\nu}(t, L_K a) - V_N(t, a)| \leq \varepsilon |\ell|_{L^1} = \frac{1}{2} \varepsilon \sqrt{D}.$$

Combining this with (5.13) proves (5.10). $\square$

We are now ready to rigorously verify that $\bar{F}_N^\nu \simeq f^\nu$ when N is large. Recall that at the formal level this follows from the fact that $\bar{F}_N^\nu$ approximately satisfies the same Hamilton–Jacobi equation as f^ν .

Theorem 5.7. *Suppose that μ has finite support. For every $T, Q \in \mathbb{R}_+$, we have*

$$(5.14) \quad \lim_{N \rightarrow \infty} \sup_{\nu \in \Delta_{m,N}} \sup_{\substack{0 \leq t \leq T, \\ |q|_{L^\infty} \leq Q}} \sup_{q \in \mathcal{Q}_\infty} |\bar{F}_N^\nu(t, q) - f^\nu(t, q)| = 0.$$

For every $\nu \in \mathcal{P}(\text{supp } \mu)$ and every sequence $\nu_N \in \Delta_{m,N}$ satisfying $\lim_{N \rightarrow \infty} W_1(\nu_N, \nu) = 0$, we also have

$$(5.15) \quad \lim_{N \rightarrow \infty} \bar{F}_N^{\nu_N}(t, q) = f^\nu(t, q), \quad \text{at every } (t, q) \in \mathbb{R}_+ \times \mathcal{Q}_2.$$

Proof. Fix $T, Q \in \mathbb{R}_+$ and put $Q' := Q + M_1 T + 1$. Let $N \in \mathbb{N}$, $\varepsilon > 0$, $\eta \in \Delta_{m,N}$, $\nu \in \mathcal{P}(\text{supp } \mu)$, $t \in [0, T]$, and $q \in \mathcal{Q}_\infty$ with $|q|_{L^\infty} \leq Q$. For each K , Lemma 5.5 in both directions, using (4.16), gives

$$|f^{\eta,K}(t, P_K q) - f^{\nu,K}(t, P_K q)| \leq 2Q' W_1(\eta, \nu).$$

Combine this with Proposition 5.6, the L^1 regularity of $\bar{F}_N^\eta$, and (5.7):

$$\begin{aligned} |\bar{F}_N^\eta(t, q) - f^\nu(t, q)| &\leq \sup_{\substack{\nu' \in \Delta_{m,N}, \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| \\ &\quad + \varepsilon \sqrt{D} + Ct(N\varepsilon)^{-1/3} + 2Q' W_1(\eta, \nu) \\ &\quad + \frac{\sqrt{D}(Q + tM_1)}{K} + |f^{\nu,K}(t, P_K q) - f^\nu(t, q)|. \end{aligned}$$

Take $K \rightarrow \infty$ with all other parameters fixed and use (5.4). We obtain

$$(5.16) \quad \begin{aligned} |\bar{F}_N^\eta(t, q) - f^\nu(t, q)| &\leq \sup_{\substack{\nu' \in \Delta_{m,N}, \\ |p|_{L^\infty} \leq Q'}} |\bar{F}_N^{\nu'}(0, p) - \varphi^{\nu'}(p)| \\ &\quad + \varepsilon \sqrt{D} + Ct(N\varepsilon)^{-1/3} + 2Q' W_1(\eta, \nu). \end{aligned}$$

Choose $\varepsilon = N^{-1/4}$. With $\eta = \nu \in \Delta_{m,N}$, the bound tends to zero uniformly over these measures and $t \in [0, T]$, $|q|_{L^\infty} \leq Q$, proving (5.14). With $\eta = \nu_N \rightarrow \nu$ in W_1 , it proves (5.15) for bounded paths.

For $q \in \mathcal{Q}_2$, put $q^{(s)}(r) := q(r \wedge s)$, $s < 1$. These paths are bounded and $|q - q^{(s)}|_{L^1} \rightarrow 0$ as $s \uparrow 1$. Equations (2.4) and (5.5) give

$$|\bar{F}_N^{\nu_N}(t, q) - f^\nu(t, q)| \leq 2|q - q^{(s)}|_{L^1} + |\bar{F}_N^{\nu_N}(t, q^{(s)}) - f^\nu(t, q^{(s)})|.$$

Let $N \rightarrow \infty$ and then $s \uparrow 1$ to finish the sequential assertion. $\square$

To convert the previous theorem into a proof of (1.6), we need one last result. Namely, we need to verify that f^ν is continuous in ν . This way we can transfer our control for $\nu \in \Delta_{m,N}$ to a control for every $\nu \in \mathcal{P}(\text{supp } \mu)$.

Lemma 5.8. *Suppose that μ has finite support. For every $\nu, \nu' \in \mathcal{P}(\text{supp } \mu)$, $t \in \mathbb{R}_+$, and $q \in \mathcal{Q}_\infty$,*

$$|f^\nu(t, q) - f^{\nu'}(t, q)| \leq 2(tM_1 + |q|_{L^\infty}) W_1(\nu, \nu').$$

Proof. Choose $\nu_N, \nu'_N \in \Delta_{m,N}$ converging to ν, ν' in W_1 . Lemma 2.5 gives the same bound for $\bar{F}_N^{\nu_N}$ and $\bar{F}_N^{\nu'_N}$. Pass to the limit using (5.15) and continuity of W_1 . For any two finitely supported target measures, apply Theorem 5.7 with a reference probability measure whose support is the union of their supports; the initial data and solutions depend only on the target measure. $\square$

Combining all the previous results of the section, we obtain the following weak-version of the main result. This is only a weaker version because it assumes that μ has a finite support. We explain in the next section how to deduce the main result from this weak version.

Corollary 5.9. *If μ has finite support, then (1.6) holds. For every $T, Q \in \mathbb{R}_+$, the convergence is uniform over $t \in [0, T]$ and $q \in \mathcal{Q}_{\infty, \leq Q}$.*

Proof. Fix $T, Q \in \mathbb{R}_+$ and assume $\text{supp } \mu = \{a_1, \dots, a_m\}$. We identify the limit uniformly over $t \in [0, T]$ and $q \in \mathcal{Q}_{\infty, \leq Q}$.

Step 1: Reduction to a minimum over $\Delta_{m,N}$. Let $s_{0,N} := \min\{1/2, \sqrt{\log(N+1)/N}\}$. Then we have

$$\lim_{N \rightarrow \infty} s_{0,N} = 0 \quad \text{and} \quad \lim_{N \rightarrow \infty} \frac{\log(N+1)}{Ns_{0,N}} = 0.$$

By Lemma 3.6, the difference between $\bar{F}_{N, \mu^{\otimes N}}(t, q)$ and the minimum of $\bar{F}_{N, \mu^{\otimes N}, \Sigma_N(\nu)}(t, q)$ is between zero and $Cs_{0,N} + m \log(N+1)/(Ns_{0,N})$. By (2.17) and (2.18), we therefore have

$$\begin{aligned} & \left| \bar{F}_{N, \mu^{\otimes N}}(t, q) - \min_{\nu \in \Delta_{m,N}} \{ \text{Ent}(\nu | \mu) + f^\nu(t, q) \} \right| \\ & \leq Cs_{0,N} + \frac{m \log(N+1)}{Ns_{0,N}} + \frac{m \log(N+1)}{N} + \sup_{\nu \in \Delta_{m,N}} |\bar{F}_N^\nu(t, q) - f^\nu(t, q)|. \end{aligned}$$

By (5.14) in Theorem 5.7, the right-hand side tends to zero uniformly over $t \in [0, T]$ and $q \in \mathcal{Q}_{\infty, \leq Q}$.

Step 2: Passing from $\Delta_{m,N}$ to $\mathcal{P}(\text{supp } \mu)$. For $t \in [0, T]$ and $q \in \mathcal{Q}_{\infty, \leq Q}$, define $G_{t,q} : \mathcal{P}(\text{supp } \mu) \rightarrow \mathbb{R}$ by

$$G_{t,q}(\nu) := \text{Ent}(\nu | \mu) + f^\nu(t, q).$$

By Lemma 5.8, we have

$$|G_{t,q}(\nu) - G_{t,q}(\nu')| \leq |\text{Ent}(\nu | \mu) - \text{Ent}(\nu' | \mu)| + 2(Q + M_1 T)W_1(\nu, \nu').$$

The entropy is uniformly continuous on the compact set $\mathcal{P}(\text{supp } \mu)$. Define $\nu^{[N]} : \mathcal{P}(\text{supp } \mu) \rightarrow \Delta_{m,N}$ by taking $\lfloor N\nu_j \rfloor$ spins equal to a_j for $j \in \{1, \dots, m-1\}$ and assigning the remaining spins to a_m . For every $\nu \in \mathcal{P}(\text{supp } \mu)$, we then have

$$\sum_{j=1}^m |\nu_j - \nu^{[N]}(\nu)(\{a_j\})| \leq \frac{2(m-1)}{N}.$$

Since the total variation distance between two measures on $\text{supp } \mu$ is half the sum of the differences of their atom masses, a coupling gives

$$\sup_{\nu \in \mathcal{P}(\text{supp } \mu)} W_1(\nu, \nu^{[N]}(\nu)) \leq \frac{1}{N}(m-1) \text{diam}(\text{supp } \mu).$$

Since $\Delta_{m,N} \subseteq \mathcal{P}(\text{supp } \mu)$, the preceding two estimates give

$$\lim_{N \rightarrow \infty} \sup_{\substack{0 \leq t \leq T \\ q \in \mathcal{Q}_{\infty, \leq Q}}} \left| \min_{\nu \in \Delta_{m,N}} G_{t,q}(\nu) - \min_{\nu \in \mathcal{P}(\text{supp } \mu)} G_{t,q}(\nu) \right| = 0.$$

Combining Steps 1 and 2 proves (1.6) and its uniformity for the finite-support case. For $q \in \mathcal{Q}_2$, the conclusion follows by approximating q in L^1 by bounded paths and using (2.4) and (5.5). $\square$

6. EXTENSION TO EVERY REFERENCE MEASURE ON THE UNIT BALL

Throughout this section, μ is an arbitrary probability measure supported on $\mathbb{B}$. The initial data and solutions were defined in (1.3) and Definition 5.2. We approximate them and the reference measure by finite discretizations to prove Theorem 1.1.

Proposition 6.1. *For every $\nu, \nu' \in \mathcal{P}(\text{supp } \mu)$, $t \in \mathbb{R}_+$, and $q \in \mathcal{Q}_\infty$, the solutions of Definition 5.2 satisfy*

$$(6.1) \quad |f^\nu(t, q) - f^{\nu'}(t, q)| \leq 2(tM_1 + |q|_{L^\infty})W_1(\nu, \nu').$$

In addition, for every $(t, q) \in \mathbb{R}_+ \times \mathcal{Q}_2$, the map $\nu \mapsto f^\nu(t, q)$ is continuous in W_1 .

Proof. The regularity of every φ^ν follows from Lemma 2.2. Since $\text{supp } \mu$ is compact, choose for each $\delta > 0$ a finite δ -net and a Borel nearest-point map $\mathcal{T}_\delta : \text{supp } \mu \rightarrow \text{supp } \mu$. Then

$$(6.2) \quad |\mathcal{T}_\delta(a) - a| \leq \delta, \quad a \in \text{supp } \mu.$$

Put $\nu_\delta := (\mathcal{T}_\delta)_\# \nu$, so $W_1(\nu_\delta, \nu) \leq \delta$.

Fix $q \in \mathcal{Q}_\infty$ and $\delta > 0$. Write $\mathcal{T}_\delta(\text{supp } \mu) = \{a_1, \dots, a_m\}$ and let $C_j := \mathcal{T}_\delta^{-1}(\{a_j\})$ for $j \in \{1, \dots, m\}$. Fix $h \in \mathcal{B}_b(\text{supp } \mu)$, and define $h_\delta : \{a_1, \dots, a_m\} \rightarrow \mathbb{R}$ by

$$h_\delta(a_j) := \begin{cases} \log \left(\frac{1}{\nu(C_j)} \int_{C_j} e^{h(a)} \nu(da) \right), & \nu(C_j) > 0, \\ 0, & \nu(C_j) = 0. \end{cases}$$

Extend h_δ to a bounded Borel function on $\text{supp } \mu$. By construction, we have

$$(\mathcal{T}_\delta)_\#(e^h \nu) = e^{h_\delta} \nu_\delta \quad \text{and} \quad \int e^h d\nu = \int e^{h_\delta} d\nu_\delta.$$

Recall from (1.3) that $\psi_{\nu, h}(q) = \bar{F}_{1, e^h \nu}(0, q)$ and $\psi_{\nu_\delta, h_\delta}(q) = \bar{F}_{1, e^{h_\delta} \nu_\delta}(0, q)$. The two finite reference measures $e^h \nu$ and $e^{h_\delta} \nu_\delta$ have a common positive mass. Normalize them by this mass and couple the resulting probability measures through $\mathcal{T}_\delta$. Lemma 2.3 with $N = 1$ and $t = 0$ bounds the difference of their free energies; restoring the common mass changes each free energy by the same constant. Hence, we have

$$(6.3) \quad |\psi_{\nu, h}(q) - \psi_{\nu_\delta, h_\delta}(q)| \leq 2|q|_{L^\infty} \delta.$$

Jensen's inequality on each C_j gives

$$\int_{C_j} h d\nu \leq \nu(C_j) h_\delta(a_j), \quad j \in \{1, \dots, m\}.$$

Combining this inequality with (6.3) and taking the supremum over h in (1.3), we obtain

$$\varphi^\nu(q) \leq \varphi^{\nu_\delta}(q) + 2|q|_{L^\infty} \delta.$$

Conversely, for any bounded h_δ on $\mathcal{T}_\delta(\text{supp } \mu)$, choose $h := h_\delta \circ \mathcal{T}_\delta$. In this case $\int h d\nu = \int h_\delta d\nu_\delta$, and (6.3) applies. Taking the supremum over h_δ gives the reverse inequality, and therefore

$$(6.4) \quad |\varphi^\nu(q) - \varphi^{\nu_\delta}(q)| \leq 2|q|_{L^\infty} \delta, \quad q \in \mathcal{Q}_\infty.$$

For $b \in \mathcal{Q}^K$, (6.4) gives

$$|\varphi^\nu(L_K b) - \varphi^{\nu_\delta}(L_K b)| \leq 2|b|_{\ell^\infty} \delta.$$

Apply Lemma 5.5 in both directions to the already-defined solutions $f^{\nu, K}$ and $f^{\nu_\delta, K}$. For $t \leq T$ and $|a|_{\ell^\infty} \leq Q$,

$$(6.5) \quad |f^{\nu, K}(t, a) - f^{\nu_\delta, K}(t, a)| \leq 2(Q + M_1 T + 1)\delta.$$

Taking $a = P_K q$ and letting $K \rightarrow \infty$ in (5.4), we can deduce from the previous display that $f^{\nu_\delta} \rightarrow f^\nu$ uniformly on $[0, T] \times \mathcal{Q}_{\infty, \leq Q}$. This means that if we approximate the measures ν and ν' by ν_δ and ν'_δ , apply Lemma 5.8, and let $\delta, \delta' \downarrow 0$, we obtain (6.1).

For $q \in \mathcal{Q}_2$, put $q^{(s)}(r) := q(r \wedge s)$ for $s < 1$. If $|q|_{L^2} \leq R$, monotonicity and the Cauchy-Schwarz inequality give

$$(6.6) \quad |q^{(s)}|_{L^\infty} \leq \frac{R}{\sqrt{1-s}}, \quad |q - q^{(s)}|_{L^1} \leq 2R\sqrt{1-s}.$$

Indeed, $(1-s)|q(s)|^2 \leq \int_s^1 |q(r)|^2 dr$, and $|q - q^{(s)}|_{L^1} \leq \int_s^1 |q(r)| dr + (1-s)|q(s)|$. Apply (5.5) at q and $q^{(s)}$, and (6.1) at the bounded path. We obtain, for all target measures,

$$(6.7) \quad \sup_{\substack{0 \leq t \leq T, \ q \in \mathcal{Q}_2 \\ |q|_{L^2} \leq R}} |f^\nu(t, q) - f^{\nu'}(t, q)| \leq 4R\sqrt{1-s} + 2 \left(TM_1 + \frac{R}{\sqrt{1-s}} \right) W_1(\nu, \nu').$$

First choose s close to 1 and then let $W_1(\nu, \nu') \rightarrow 0$. This proves continuity and also uniform convergence of the finite discretizations on bounded subsets of $\mathcal{Q}_2$. $\square$

For the remainder of the section, fix $\delta > 0$ and the map $\mathcal{T}_\delta$ satisfying (6.2), and let $\mu_\delta := (\mathcal{T}_\delta)_\# \mu$. Its support is finite and contained in $\text{supp } \mu$. Applying Lemma 2.3 to the product coupling induced by $\mathcal{T}_\delta$ gives, for $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_\infty$,

$$(6.8) \quad \sup_{N \in \mathbb{N}} |\bar{F}_{N, \mu^{\otimes N}}(t, q) - \bar{F}_{N, \mu_\delta^{\otimes N}}(t, q)| \leq 2\delta(tM_1 + |q|_{L^\infty}).$$

Indeed, the joint law of $(\sigma, \mathcal{T}_\delta(\sigma))$ under $\mu^{\otimes N}$ has marginals $\mu^{\otimes N}$ and $\mu_\delta^{\otimes N}$, and every pair in its support satisfies $N^{-1} \sum_{i=1}^N |\sigma_{\cdot, i} - \mathcal{T}_\delta(\sigma_{\cdot, i})| \leq \delta$. Thus (6.8) is independent of the number of points in $\text{supp } \mu_\delta$.

Recall the relative entropy $\text{Ent}(\cdot | \mu)$ introduced in (1.5). Define $J, J_\delta : \mathbb{R}_+ \times \mathcal{Q}_2 \rightarrow \mathbb{R}$ by

$$(6.9) \quad J(t, q) := \inf_{\nu \in \mathcal{P}(\text{supp } \mu)} \{ \text{Ent}(\nu | \mu) + f^\nu(t, q) \} \quad \text{and} \quad J_\delta(t, q) := \inf_{\eta \in \mathcal{P}(\text{supp } \mu_\delta)} \{ \text{Ent}(\eta | \mu_\delta) + f^\eta(t, q) \}.$$

Both quantities in (6.9) are finite. Indeed, entropy is nonnegative, and (5.5) and Lemma 2.2 give $|f^\nu(t, q)| \leq M_0 t + |q|_{L^1}$ uniformly in ν ; the choices $\nu = \mu$ and $\eta = \mu_\delta$ have zero entropy.

Lemma 6.2. *For every $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_\infty$, we have*

$$(6.10) \quad |J(t, q) - J_\delta(t, q)| \leq 2(tM_1 + |q|_{L^\infty})\delta.$$

Proof. Write $\text{supp } \mu_\delta = \{a_1, \dots, a_m\}$ and define the cells $C_j := \mathcal{T}_\delta^{-1}(\{a_j\})$. Then $\mu(C_j) = \mu_\delta(\{a_j\}) > 0$, and these cells cover $\text{supp } \mu$ up to a μ -null set.

Entropy contraction. Fix $\nu \in \mathcal{P}(\text{supp } \mu)$ with $\text{Ent}(\nu | \mu) < \infty$ and let $\eta := (\mathcal{T}_\delta)_\# \nu$. The coupling $(a, \mathcal{T}_\delta(a))$ gives $W_1(\nu, \eta) \leq \delta$. In addition, relative entropy decreases under the measurable map $\mathcal{T}_\delta$:

$$(6.11) \quad \text{Ent}(\eta | \mu_\delta) \leq \text{Ent}(\nu | \mu).$$

Indeed, write $r := d\nu/d\mu$. On every cell with $\mu(C_j) > 0$, Jensen's inequality for $x \mapsto x \log x$ gives

$$\int_{C_j} r \log r \, d\mu \geq \nu(C_j) \log \frac{\nu(C_j)}{\mu(C_j)}.$$

Summing these inequalities proves (6.11). Combining it with (6.1), we obtain

$$\text{Ent}(\eta | \mu_\delta) + f^\eta(t, q) \leq \text{Ent}(\nu | \mu) + f^\nu(t, q) + 2(tM_1 + |q|_{L^\infty})\delta.$$

Taking the infimum over such ν proves the first inequality.

Entropy-preserving lift. Fix $\eta \in \mathcal{P}(\text{supp } \mu_\delta)$ and let $\beta_j := \eta(\{a_j\})$. Define $\tilde{\nu} \in \mathcal{P}(\text{supp } \mu)$ by

$$(6.12) \quad \tilde{\nu} := \sum_{j=1}^m \beta_j \mu(\cdot | C_j).$$

The pushforward $(\mathcal{T}_\delta)_\# \tilde{\nu}$ equals η , so $W_1(\tilde{\nu}, \eta) \leq \delta$. Moreover, the density of the measure in (6.12) with respect to μ equals $\beta_j/\mu(C_j)$ on C_j . Consequently, we have

$$(6.13) \quad \text{Ent}(\tilde{\nu} | \mu) = \sum_{j=1}^m \beta_j \log \frac{\beta_j}{\mu(C_j)} = \text{Ent}(\eta | \mu_\delta).$$

Using (6.13) and (6.1), we therefore obtain

$$\text{Ent}(\tilde{\nu} | \mu) + f^{\tilde{\nu}}(t, q) \leq \text{Ent}(\eta | \mu_\delta) + f^\eta(t, q) + 2(tM_1 + |q|_{L^\infty})\delta.$$

Taking the infimum over η completes (6.10). $\square$

Proof of Theorem 1.1. Step 1: Uniform convergence on bounded paths. Fix $T, Q \in \mathbb{R}_+$. Since μ_δ has finite support, Corollary 5.9 gives

$$(6.14) \quad \lim_{N \rightarrow \infty} \sup_{0 \leq t \leq T, \, q \in \mathcal{Q}_{\infty, \leq Q}} |\bar{F}_{N, \mu_\delta^{\otimes N}}(t, q) - J_\delta(t, q)| = 0.$$

For every $(t, q) \in [0, T] \times \mathcal{Q}_{\infty, \leq Q}$, the triangle inequality gives

$$\begin{aligned} |\bar{F}_{N, \mu^{\otimes N}}(t, q) - J(t, q)| &\leq |\bar{F}_{N, \mu^{\otimes N}}(t, q) - \bar{F}_{N, \mu_\delta^{\otimes N}}(t, q)| \\ &\quad + |\bar{F}_{N, \mu_\delta^{\otimes N}}(t, q) - J_\delta(t, q)| + |J_\delta(t, q) - J(t, q)|. \end{aligned}$$

By (6.8), Lemma 6.2, and (6.14), we therefore obtain

$$(6.15) \quad \limsup_{N \rightarrow \infty} \sup_{0 \leq t \leq T, q \in \mathcal{Q}_{\infty, \leq Q}} |\bar{F}_{N, \mu^{\otimes N}}(t, q) - J(t, q)| \leq 4\delta(TM_1 + Q).$$

Since $\delta > 0$ is arbitrary, sending $\delta \downarrow 0$ in (6.15) proves uniform convergence on $[0, T] \times \mathcal{Q}_{\infty, \leq Q}$.

Step 2: Extension to every $q \in \mathcal{Q}_2$. Fix $t \in \mathbb{R}_+$ and $q \in \mathcal{Q}_2$. For $s \in [0, 1]$, take $q^{(s)}$ from (6.6). Lemma 2.1 and (5.5) imply that both $q \mapsto \bar{F}_{N, \mu^{\otimes N}}(t, q)$ and $q \mapsto J(t, q)$ are 1-Lipschitz in L^1 ; for J , this follows by taking the infimum of functions with the same Lipschitz constant. Hence

$$|\bar{F}_{N, \mu^{\otimes N}}(t, q) - J(t, q)| \leq 2|q - q^{(s)}|_{L^1} + |\bar{F}_{N, \mu^{\otimes N}}(t, q^{(s)}) - J(t, q^{(s)})|.$$

The second term tends to zero as $N \rightarrow \infty$ for fixed s , by Step 1. By (6.6), the first term tends to zero as $s \uparrow 1$. This proves (1.6) with the infimum in place of the minimum.

Step 3: Attainment of the infimum. The space $\mathcal{P}(\text{supp } \mu)$ is compact in W_1 , since $\text{supp } \mu$ is compact. Relative entropy is lower semicontinuous in W_1 . By (6.7), the map $\nu \mapsto f^\nu(t, q)$ is continuous in W_1 for every $q \in \mathcal{Q}_2$. Therefore, the objective

$$\nu \mapsto \text{Ent}(\nu | \mu) + f^\nu(t, q)$$

is lower semicontinuous on $\mathcal{P}(\text{supp } \mu)$. It is finite at $\nu = \mu$, since $\text{Ent}(\mu | \mu) = 0$ and $f^\mu(t, q)$ is finite. Compactness therefore yields a minimizer, proving (1.6) as stated. $\square$

APPENDIX A. PROOF OF THE FINITE-DIMENSIONAL COMPARISON PRINCIPLE

We give a proof of Lemma 5.5 using the weighted norms of Section 5. In particular, the argument uses only the monotonicity and regularity of ξ_K on $\mathcal{Q}^K$ and does not require a boundary condition on $\partial\mathcal{Q}^K$.

Proof of Lemma 5.5. Fix $K \in \mathbb{N}$. Use the weighted inner product $\langle \cdot, \cdot \rangle_K$ and let $|a|_K := \langle a, a \rangle_K^{1/2}$. Then

$$|a|_{\ell_K^1} \leq |a|_K \leq |a|_{\ell^\infty}, \quad a \in E_K,$$

and $|\cdot|_{\ell_K^1}$ and $|\cdot|_{\ell^\infty}$ are dual norms. By assumption, for some $C \in \mathbb{R}_+$,

$$(A.1) \quad |u(t, a) - u(0, a)| \vee |v(t, a) - v(0, a)| \leq Ct, \quad (t, a) \in \mathbb{R}_+ \times \mathcal{Q}^K.$$

Localization and the boundary barrier. Put $R := Q + M_1T + 1$ and

$$S_R := \sup_{a \in \mathcal{Q}^K: |a|_{\ell^\infty} \leq R} (u(0, a) - v(0, a)).$$

Radial projection onto this ball, together with the spatial Lipschitz bounds, gives

$$(A.2) \quad u(0, a) - v(0, a) \leq S_R + 2(|a|_{\ell^\infty} - R)_+, \quad a \in \mathcal{Q}^K.$$

For $\kappa \in (0, 1/4)$, define the smooth function

$$(A.3) \quad \rho_\kappa(a) := \sqrt{|a_K|^2 + \kappa^2}.$$

On $\mathcal{Q}^K$, positivity of the increments implies $|a|_{\ell^\infty} = |a_K|$, so

$$|a|_{\ell^\infty} \leq \rho_\kappa(a) \leq |a|_{\ell^\infty} + \kappa, \quad a \in \mathcal{Q}^K.$$

Moreover, $\nabla_K \rho_\kappa$ has only its last block nonzero, equal to $Ka_K/\rho_\kappa(a)$, and hence $|\nabla_K \rho_\kappa|_{\ell_K^1} \leq 1$ on E_K . Choose a smooth nondecreasing $\theta_\kappa: \mathbb{R} \rightarrow \mathbb{R}_+$ with $0 \leq \theta'_\kappa \leq 1$, $\theta_\kappa(r) = 0$ for $r \leq 0$, and $\theta_\kappa(r) \geq (r - \kappa)_+$. Let

$$(A.4) \quad \chi(t, a) := 3\theta_\kappa(\rho_\kappa(a) + M_1t - R).$$

The preceding bounds give

$$(A.5) \quad \begin{aligned} \partial_t \chi &\geq M_1 |\nabla_K \chi|_{\ell_K^1}, \\ \chi(t, a) &= 0, \quad 0 \leq t \leq T, \quad a \in \mathcal{Q}^K, \quad |a|_{\ell^\infty} \leq Q, \\ \chi(t, a) &\geq 3(|a|_{\ell^\infty} - R - 2\kappa)_+, \quad (t, a) \in \mathbb{R}_+ \times \mathcal{Q}^K. \end{aligned}$$

In particular,

$$(A.6) \quad \sup_{a \in \mathcal{Q}^K} (u(0, a) - v(0, a) - \chi(0, a)) \leq S_R + 6\kappa.$$

For $a \in \text{int } \mathcal{Q}^K$, put $a_0 := 0$ and define

$$(A.7) \quad b(a) := \sum_{k=1}^K \text{tr}[(a_k - a_{k-1})^{-1}].$$

This smooth positive function diverges at the boundary. For $c \in \mathcal{Q}^K$, differentiation gives

$$(A.8) \quad \langle \nabla_K b(a), c \rangle_K = - \sum_{k=1}^K \text{tr}[(a_k - a_{k-1})^{-2} (c_k - c_{k-1})] \leq 0, \quad c_0 := 0.$$

Thus $-\nabla_K b(a) \in (\mathcal{Q}^K)^*$.

Doubled variables and interior contact points. Suppose the comparison fails. There are $t_* \in (0, T]$ and $a_* \in \mathcal{Q}^K$, with $|a_*|_{\ell^\infty} \leq Q$, such that $u(t_*, a_*) - v(t_*, a_*) > S_R$. Fix $T_* > T$. Choose κ , and then $\gamma, \beta > 0$, small enough that

$$u(t_*, a_*) - v(t_*, a_*) - \chi(t_*, a_*) - \gamma t_* - \frac{2\beta}{T_* - t_*} > S_R + 12\kappa.$$

By continuity and density of $\text{int } \mathcal{Q}^K$, choose $\widehat{a}_* \in \text{int } \mathcal{Q}^K$ close to a_* , and then $\sigma > 0$ small enough that

$$(A.9) \quad u(t_*, \widehat{a}_*) - v(t_*, \widehat{a}_*) - \chi(t_*, \widehat{a}_*) - \gamma t_* - \frac{2\beta}{T_* - t_*} - \sigma b(\widehat{a}_*) > S_R + 12\kappa.$$

Keep these parameters fixed. For $\alpha > 1$, define on $[0, T_*)^2 \times \text{int } \mathcal{Q}^K \times \mathcal{Q}^K$

$$(A.10) \quad \begin{aligned} \Psi_\alpha(t, t', a, a') &:= u(t, a) - v(t', a') - \frac{\alpha}{2}(|t - t'|^2 + |a - a'|_K^2) \\ &\quad - \chi(t, a) - \gamma t - \frac{\beta}{T_* - t} - \frac{\beta}{T_* - t'} - \sigma b(a). \end{aligned}$$

The diagonal point from (A.9) gives

$$(A.11) \quad \sup \Psi_\alpha > S_R + 12\kappa.$$

The time-growth and spatial Lipschitz assumptions imply

$$(A.12) \quad u(t, a) - v(t', a') \leq S_R + 2(|a|_{\ell^\infty} - R)_+ + 2CT_* + |a - a'|_{\ell_K^1}.$$

The cutoff with coefficient $3 > 2$ controls a , the quadratic penalty controls a' , the time penalties exclude T_* , and $\sigma b(a)$ excludes $\partial \mathcal{Q}^K$. Hence Ψ_α attains a maximum at $(t_\alpha, t'_\alpha, a_\alpha, a'_\alpha)$. Equations (A.12) and (A.11), using $x - \alpha x^2/2 \leq 1/(2\alpha)$, give a constant $C_* < \infty$, independent of $\alpha > 1$, such that

$$(A.13) \quad |a_\alpha|_K \vee |a'_\alpha|_K \leq C_*, \quad \alpha(|t_\alpha - t'_\alpha|^2 + |a_\alpha - a'_\alpha|_K^2) + b(a_\alpha) \leq C_*.$$

Since each inverse trace is bounded by C_* , every increment of a_α has smallest eigenvalue at least C_*^{-1} . Quadratic closeness gives $|a_\alpha - a'_\alpha|_K \rightarrow 0$; for fixed K , Weyl's inequality then shows $a'_\alpha \in \text{int } \mathcal{Q}^K$ for all sufficiently large α .

If either contact time is zero, the time-growth assumption gives

$$u(t_\alpha, a_\alpha) - v(t'_\alpha, a'_\alpha) \leq u(0, a_\alpha) - v(0, a_\alpha) + C|t_\alpha - t'_\alpha| + |a_\alpha - a'_\alpha|_K.$$

Since $\chi(t_\alpha, a_\alpha) \geq \chi(0, a_\alpha)$, completing squares and using (A.6) yields

$$\Psi_\alpha(t_\alpha, t'_\alpha, a_\alpha, a'_\alpha) \leq S_R + 6\kappa + \frac{C^2 + 1}{2\alpha},$$

contradicting (A.11) for large α . Fix such an α , so both contact points lie in $(0, T_*) \times \text{int } \mathcal{Q}^K$.

Hamiltonian comparison and viscosity inequalities. The test gradients for u and v are

$$(A.14) \quad p^v := \alpha(a_\alpha - a'_\alpha), \quad p^u := p^v + \nabla_K \chi(t_\alpha, a_\alpha) + \sigma \nabla_K b(a_\alpha).$$

At interior contact points, the spatial Lipschitz and $(\mathcal{Q}^K)^*$ -monotonicity assumptions give

$$(A.15) \quad p^u, p^v \in \mathcal{Q}^K, \quad |p^u|_{\ell^\infty} \vee |p^v|_{\ell^\infty} \leq 1.$$

For example, test monotonicity in a direction $c \in (\mathcal{Q}^K)^*$ at $a_\alpha + hc$ for the upper test and at $a'_\alpha - hc$ for the lower test. Letting $h \downarrow 0$ gives nonnegative pairing with every $c \in (\mathcal{Q}^K)^*$, hence membership in $((\mathcal{Q}^K)^*)^* = \mathcal{Q}^K$. Testing arbitrary directions gives the norm bounds.

The monotonicity of ξ_K and the definition of M_1 imply

$$(A.16) \quad \nabla_K \xi_K(p) \in \mathcal{Q}^K, \quad |\nabla_K \xi_K(p)|_{\ell^\infty} \leq M_1, \quad p \in \mathcal{Q}^K, \quad |p|_{\ell^\infty} \leq 1.$$

The first inclusion follows by differentiating monotonicity in directions of $(\mathcal{Q}^K)^*$ at interior points and extending by continuity; the norm bound follows from $(\nabla_K \xi_K(p))_k = \nabla \xi(p_k)$. By (A.8),

$$(A.17) \quad r := -\sigma \nabla_K b(a_\alpha) \in (\mathcal{Q}^K)^*, \quad p^u - p^v = \nabla_K \chi(t_\alpha, a_\alpha) - r.$$

Integrating the derivative of ξ_K along the segment between p^v and p^u , which lies in the set in (A.15), gives

$$(A.18) \quad \begin{aligned} \xi_K(p^u) - \xi_K(p^v) &= \int_0^1 \langle \nabla_K \xi_K(p^v + s(p^u - p^v)), \nabla_K \chi - r \rangle_K ds \\ &\leq M_1 |\nabla_K \chi(t_\alpha, a_\alpha)|_{\ell_K^1}. \end{aligned}$$

The time derivatives of the upper and lower tests are respectively

$$\alpha(t_\alpha - t'_\alpha) + \partial_t \chi(t_\alpha, a_\alpha) + \gamma + \frac{\beta}{(T_* - t_\alpha)^2}, \quad \alpha(t_\alpha - t'_\alpha) - \frac{\beta}{(T_* - t'_\alpha)^2}.$$

Subtracting the viscosity supersolution inequality from the subsolution inequality, and discarding the positive time-penalty terms, yields

$$\gamma \leq \xi_K(p^u) - \xi_K(p^v) - \partial_t \chi(t_\alpha, a_\alpha) \leq M_1 |\nabla_K \chi|_{\ell_K^1} - \partial_t \chi \leq 0,$$

by (A.5). This contradicts $\gamma > 0$ and proves the comparison. $\square$

REFERENCES

- [1] Elena Agliari, Adriano Barra, Raffaella Burioni, and Aldo Di Biasio. Notes on the p-spin glass studied via Hamilton–Jacobi and smooth-cavity techniques. *J. Math. Phys.*, 53(6):063304, 29, 2012.
- [2] P. M. Aronow and Patrick Lopatto. Quantitative Parisi formulas and fluctuations in the Sherrington–Kirkpatrick model, 2026.
- [3] Antonio Auffinger and Wei-Kuo Chen. On properties of Parisi measures. *Probab. Theory Related Fields*, 161(3):817–850, 2015.
- [4] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.
- [5] Jean Barbier, Dmitry Panchenko, and Manuel Sáenz. Strong replica symmetry for high-dimensional disordered log-concave Gibbs measures. *Information and Inference: A Journal of the IMA*, 11(3):1079–1108, 2022.
- [6] 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.
- [7] Adriano Barra, Gino Del Ferraro, and Daniele Tantari. Mean field spin glasses treated with PDE techniques. *Eur. Phys. J. B*, 86(7):Art. 332, 10, 2013.
- [8] Adriano Barra, Aldo Di Biasio, and Francesco Guerra. Replica symmetry breaking in mean-field spin glasses through the Hamilton–Jacobi technique. *J. Stat. Mech. Theory Exp.*, pages P09006, 22, 2010.
- [9] Erik Bates and Youngtak Sohn. Free energy in multi-species mixed p-spin spherical models. *Electron. J. Probab.*, 27:Paper No. 52, 75, 2022.
- [10] Erik Bates and Youngtak Sohn. Parisi formula for balanced Potts spin glass. *Comm. Math. Phys.*, 405(10):Paper No. 228, 68, 2024.
- [11] Erik Bates and Youngtak Sohn. Balanced multi-species spin glasses. *Preprint, arXiv:2507.06522*, 2025.
- [12] Anton Bovier and Anton Klimovsky. The Aizenman–Sims–Starr and Guerra’s schemes for the SK model with multidimensional spins. *Electron. J. Probab.*, 14:no. 8, 161–241, 2009.
- [13] 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(4):366–389, 1976.

- [14] Francesco Camilli, Pierluigi Contucci, Federica Gerace, and Emanuele Mingione. Variational bounds for perceptron learning from structured data. *Preprint, arXiv:2608.04882*, 2026.
- [15] Hong-Bin Chen. Parisi PDE and convexity for vector spins. *Preprint, arXiv:2311.10446*, 2023.
- [16] Hong-Bin Chen. On the self-overlap in vector spin glasses. *Journal of Mathematical Physics*, 65(3):031901, 2024.
- [17] 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.
- [18] Hong-Bin Chen and Jean-Christophe Mourrat. Simultaneous replica-symmetry breaking for vector spin glasses. *Preprint, arXiv:2411.14105*, 2024.
- [19] 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.
- [20] Hong-Bin Chen and Jean-Christophe Mourrat. Free energy of non-convex multi-species spherical spin glasses. *Preprint, arXiv:2609.32710*, 2026.
- [21] Hong-Bin Chen and Jiaming Xia. Hamilton–Jacobi equations from mean-field spin glasses. *Probab. Theory Related Fields*, 192(3):803–873, 2025.
- [22] Wei-Kuo Chen. Variational representations for the Parisi functional and the two-dimensional Guerra–Talagrand bound. *Ann. Probab.*, 45(6A):3929–3966, 2017.
- [23] 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.
- [24] Rick Durrett. *Probability: Theory and Examples*. Cambridge University Press, 5 edition, 2019.
- [25] Ronen Eldan. A simple approach to chaos for p -spin models. *J. Stat. Phys.*, 181:1266–1276, 2020.
- [26] Jürg Fröhlich and Bogusław Zegarliński. Spin glasses and other lattice systems with long range interactions. *Comm. Math. Phys.*, 120(4):665–688, 1989.
- [27] Giuseppe Genovese and Adriano Barra. A mechanical approach to mean field spin models. *J. Math. Phys.*, 50(5):053303, 2009.
- [28] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.
- [29] Fu-Hsuan Ho. Convergence of the free energy of vector spin glasses with non-convex interactions. *Preprint, arXiv:2609.39307*, 2026.
- [30] Aukosh Jagannath and Ian Tobasco. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.*, 144(7):3135–3150, 2016.
- [31] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics (113) (Book 113). Springer New York, 1991.
- [32] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Preprint, arXiv:2010.09114*, 2020.
- [33] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.
- [34] Jean-Christophe Mourrat. The Parisi formula is a Hamilton–Jacobi equation in Wasserstein space. *Canad. J. Math.*, 74(3):607–629, 2022.
- [35] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(3):1143–1182, 2023.
- [36] Jean-Christophe Mourrat and Dmitry Panchenko. Extending the Parisi formula along a Hamilton–Jacobi equation. *Electron. J. Probab.*, 25:Paper No. 23, 17, 2020.
- [37] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.
- [38] Dmitry Panchenko. *The Sherrington–Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.
- [39] Dmitry Panchenko. The free energy in a multi-species Sherrington–Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.
- [40] Dmitry Panchenko. Free energy in the mixed p -spin models with vector spins. *Ann. Probab.*, 46(2):865–896, 2018.
- [41] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754, 1979.
- [42] Giorgio Parisi. A sequence of approximated solutions to the SK model for spin glasses. *J. Phys. A: Math. Gen.*, 13(4):L115–L121, 1980.
- [43] Michel Talagrand. Free energy of the spherical mean field model. *Probab. Theory Related Fields*, 134(3):339–382, 2006.
- [44] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

(Hong-Bin Chen) NYU-ECNU INSTITUTE OF MATHEMATICAL SCIENCES, NYU SHANGHAI, CHINA
Email address: hongbin.chen@nyu.edu

(Fu-Hsuan Ho) WEIZMANN INSTITUTE OF SCIENCE, REHOVOT, ISRAEL
Email address: fu-hsuan.ho@weizmann.ac.il

(Victor Issa) LPSM, SORBONNE UNIVERSITY, PARIS, FRANCE
Email address: victor.issa@ens-lyon.fr