A HOPF-LIKE FORMULA FOR MEAN-FIELD SPIN GLASS MODELS

VICTOR ISSA

ABSTRACT. We study mean-field spin glass models with general vector spins and convex covariance function. For those models, it is known that the limit of the free energy can be written as the supremum of a functional, this is the celebrated Parisi formula.

In this paper, we observe that the Parisi functional extends into a concave and Lipschitz functional on the set of signed measures. We use this fact and Fenchel-Moreau duality to derive an un-inverted version of the Parisi formula. Namely, we show that the limit of the free energy can be written as the infimum of a functional related to the Parisi functional.

This un-inverted formula can be interpreted as a Hopf-like formula for some Hamilton-Jacobi equation in Wasserstein space.

KEYWORDS AND PHRASES: mean-field spin glass, free energy, Parisi formula, convex analysis, Hopf formula, Hamilton-Jacobi equation.

MSC 2020: 60K35, 82B44, 82D30.

(Victor Issa) DEPARTMENT OF MATHEMATICS, ENS DE LYON, LYON, FRANCE
E-mail address: victor.issa@ens-lyon.fr.

Contents

1. Introduction

1.1. **Preamble.** Let $D \geqslant 1$ be an integer, and let $(H_N(\sigma))_{\sigma \in (\mathbb{R}^N)^D}$ be a centered Gaussian field such that, for every $\sigma = (\sigma_1, \ldots, \sigma_D)$ and $\tau = (\tau_1, \ldots, \tau_D) \in (\mathbb{R}^N)^D$, we have

$$\mathbb{E}\left[H_N(\sigma) H_N(\tau)\right] = N\xi\left(\frac{\sigma\tau^\intercal}{N}\right), \tag{1.1}$$

where $\xi \in C^\infty(\mathbb{R}^{D\times D}; \mathbb{R})$ is some fixed smooth function, and where $\sigma\tau^\intercal$ denotes the matrix of scalar products

$$\sigma\tau^\intercal = (\sigma_d \cdot \tau_{d'})_{1 \leqslant d, d' \leqslant D}. \tag{1.2}$$

We often identify $(\mathbb{R}^N)^D$ with the space $\mathbb{R}^{D\times N}$ of D-by-N matrices, which makes the notation in (1.2) natural. We assume that ξ has a convergent power-series expansion. We also give ourselves a reference measure P_1 on $\mathbb{R}^D$ of compact support, and we set $P_N = (P_1)^{\otimes N}$, which we think of as a probability measure on $\mathbb{R}^{D\times N}$. In other words, a sample $\sigma = (\sigma_{d,i})_{1\leqslant d \leqslant D, 1 \leqslant i \leqslant N}$ from P_N is such that the columns $(\sigma_{\cdot,i})_{1 \leqslant i \leqslant N}$ are independent with law P_1. We are interested in the study of the large-N behavior of the free energy

$$\overline{F}_N(t, \delta_0) = -\frac{1}{N}\mathbb{E}\log \int \exp\left(\sqrt{2t} H_N(\sigma) - tN\xi\left(\frac{\sigma\sigma^\intercal}{N}\right)\right) \mathrm{d}P_N(\sigma), \tag{1.3}$$

where $t \geqslant 0$. The term $\xi\left(\frac{\sigma\sigma^\intercal}{N}\right)$ in (1.3) is introduced as a convenience to simplify the expression of the limit; it is constant in classical cases of interest, such as when the coordinates of σ takes values in $\{-1, 1\}$ and ξ depends only on the diagonal entries of its argument. In general, the second argument of $\overline{F}_N$ can be any monotone probability measure on the space S^D_+ of positive semi-definite matrices, subject to a mild integrability requirement; the expression in (1.3) is with this argument chosen to be the Dirac mass at the origin. To explain what this space is, let us say that a path $\mathsf{q} : [0, 1) \to S^D_+$ is nondecreasing if for every $u \leqslant v \in [0, 1)$, we have $\mathsf{q}(v) - \mathsf{q}(u) \in S^D_+$. A probability measure on S^D_+ is said to be monotone if it is the image of the Lebesgue measure on $[0, 1]$ through a nondecreasing path from $[0, 1)$ to S^D_+. We write $\mathcal{P}^\uparrow(S^D_+)$ to denote the set of monotone measures on S^D_+, which is a subset of the set $\mathcal{P}(S^D_+)$ of probability measures on S^D_+. For every $p \in [1, +\infty]$, we also write $\mathcal{P}_p(S^D_+)$ to denote the subspace of $\mathcal{P}(S^D_+)$ of measures with finite p-th moment, with the understanding that $\mathcal{P}_\infty(S^D_+)$ is the subset of probability measures with compact support; we also write $\mathcal{P}^\uparrow_p(S^D_+) = \mathcal{P}^\uparrow(S^D_+) \cap \mathcal{P}_p(S^D_+)$. We postpone the precise definition of $\overline{F}_N(t, \mu)$ for arbitrary $\mu \in \mathcal{P}^\uparrow_1(S^D_+)$ to (2.5). In short, this quantity is obtained by adding an energy term in the exponential on the right side of (1.3) to encode the interaction of σ with an external magnetic field, and this external magnetic field has an ultrametric structure whose characteristics are encoded by the measure μ.

One can check [12, Proposition 3.1] that $\overline{F}_N(0,\cdot)$ does not depend on N; for every $\mu \in \mathcal{P}_1^{\uparrow}(S_+^D)$, we write

$$\psi(\mu) = \overline{F}_1(0,\mu) = \overline{F}_N(0,\mu).$$

This follows from the fact $P_N = P_1^{\otimes N}$ and that at $t = 0$ the N-body Hamiltonian has the same law as N copies of the 1-body Hamiltonian. When instead P_N is the uniform measure on the sphere of radius $\sqrt{N}$ centered at 0 in $(\mathbb{R}^D)^N$, $\overline{F}_N(0,\cdot)$ depends on N but converges to a smooth function of μ as $N \to +\infty$ [12, Proposition 3.1]. In what follows, we focus on models with $P_N = P_1^{\otimes N}$.

When ξ is convex on S_+^D, the limiting value of $\overline{F}_N(t,\delta_0)$ is known, this is the celebrated Parisi formula. The Parisi formula was first conjectured in [19] using a sophisticated non-rigorous argument now referred to as the replica method. The convergence of the free energy as $N \to +\infty$ was rigorously established in [10] in the case of the so-called Sherrington-Kirkpatrick model which corresponds to $D = 1$, $\xi(x) = x^2$ and $P_1 = \mathrm{Unif}(\{-1,1\})$. The Parisi formula for the Sherrington-Kirkpatrick model was then proven in [9, 21]. This was extended to the case $D = 1$, $P_1 = \mathrm{Unif}(\{-1,1\})$ and $\xi(x) = \sum_{p \geqslant 1} a_p x^p$ with $a_p \geqslant 0$ in [15]. Some models with $D > 1$ such as multispecies models, the Potts model, and a general class of models with vector spins were treated in [16, 17, 18], under the assumption that ξ is convex on $\mathbb{R}^{D\times D}$. Finally, the case $D > 1$ was treated in general in [4] assuming only that ξ is convex on the set of positive semi-definite matrices. The following version of the Parisi formula is [4, Corollary 8.2].

Theorem 1.1 ([4]). *If ξ is convex on S_+^D, then we have for every $t > 0$*

$$\lim_{N\to+\infty} \overline{F}_N(t,\delta_0) = \sup_{\mu \in \mathcal{P}_\infty^{\uparrow}(S_+^D)} \left\{ \psi(\mu) - t \int \xi^* \left(\frac{x}{t} \right) \mathrm{d}\mu(x) \right\}. \tag{1.4}$$

Here, ξ^* denotes the convex dual of ξ with respect to the cone S_+^D, it is the function $\mathbb{R}^{D\times D} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\xi^*(a) = \sup_{b \in S_+^D} \{ a \cdot b - \xi(b) \}. \tag{1.5}$$

1.2. **Main results.** In the classical version of the Parisi formula (1.4), the limit free energy is written as the supremum of a functional. In this paper, we manipulate the right-hand side of (1.4) to show that there exists a functional related to ψ whose infimum is equal to the limit of the free energy. Unless stated otherwise, we always work under the following assumptions, they are used to make sure that Theorem 1.1 holds.

(H1) For every $N \geqslant 1$, $P_N = P_1^{\otimes N}$.
(H2) The function ξ is convex on S_+^D.
(H3) The function ξ admits an absolutely convergent power series.

Here and throughout, we will use $\int h \mathrm{d}\mu$ as a shorthand for $\int_{S^D_+} h(x)\mathrm{d}\mu(x)$. When $D = 1$ we have that $\mathcal{P}^\uparrow(S^D_+) = \mathcal{P}(\mathbb{R}_+)$ is a convex set, and we have the result of [2] on the convexity of the Parisi functional, which is essentially the mapping $-\psi$, up to a linear term and a change of variables (see also [12]). This motivates the introduction of ψ_* the concave dual of ψ. The function ψ_* is defined for every Lipschitz function $\chi : \mathbb{R}_+ \to \mathbb{R}$ by

$$\psi_*(\chi) = \inf_{\mu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \psi(\mu) \right\}.$$

Given a Lipschitz function $\chi : \mathbb{R}_+ \to \mathbb{R}$, we also define for every $x \in \mathbb{R}_+$,

$$S_t\chi(x) = \sup_{y \in \mathbb{R}_+} \left\{ \chi(x+y) - t\xi^* \left(\frac{y}{t} \right) \right\}. \tag{1.6}$$

We recall that $\overline{F}_N(t,\mu)$ for $\mu \neq \delta_0$ is defined in (2.5).

Theorem 1.2. *Assume (H1), (H2) and (H3), also assume that $D = 1$, then for every $t \geqslant 0$ and $\mu \in \mathcal{P}_1(\mathbb{R}_+)$, we have*

$$\lim_{N \to +\infty} \overline{F}_N(t,\mu) = \inf_{\chi} \left\{ \int S_t\chi \mathrm{d}\mu - \psi_*(\chi) \right\}, \tag{1.7}$$

where the infimum is taken over the set of 1-Lipschitz, convex and nondecreasing functions $\chi : \mathbb{R}_+ \to \mathbb{R}$.

We add that when χ is convex, according to [5, Proposition 6.3], for every $x \in \mathbb{R}_+$, the quantity $S_t\chi(x)$ can also be represented in the following way,

$$S_t\chi(x) = \sup_{y \in \mathbb{R}_+} \left\{ xy - \chi^*(y) + t\xi(y) \right\}, \tag{1.8}$$

where $\chi^*(y) = \sup_{x \in \mathbb{R}_+} \{xy - \chi(x)\}$ is the convex dual of χ.

We equip S^D, the set of $D \times D$ symmetric matrices, with the Frobenius inner product,

$$x \cdot y = \sum_{d,d'=1}^{D} x_{dd'} y_{dd'}.$$

We let $|\cdot|$ denote the associated norm and $B(0,1)$ denotes the centered unit ball in S^D with respect to $|\cdot|$ and $\mathcal{K}_\xi = \nabla\xi(S^D_+ \cap B(0,1))$. When $D > 1$, the set $\mathcal{P}^\uparrow(S^D_+)$ is not convex. Due to this technical difficulty, we were unable to obtain the exact analog of (1.7) in this case. To circumvent this difficulty, we build a Lipschitz extension of ψ defined on the set of signed measures on $\mathcal{K}_\xi$. This yields a formula closely related to (1.7) in which ψ_* and $S_t\chi$ are replaced by the following

$$\psi^\xi_*(\chi) = \inf_{\mu \in \mathcal{P}^\uparrow(S^D_+) \cap \mathcal{P}(\mathcal{K}_\xi)} \left\{ \int \chi d\mu - \psi(\mu) \right\}, \tag{1.9}$$

$$\widetilde{S}_t\chi(x) = \sup_{y \in S^D_+ \cap B(0,1)} \left\{ \chi(x + t\nabla\xi(y)) - t\xi^*\left(\nabla\xi(y)\right) \right\}. \tag{1.10}$$

Theorem 1.3. *Assume (H1), (H2) and (H3), then for every $t \geqslant 0$ we have*

$$\lim_{N \to +\infty} \overline{F}_N(t, \delta_0) = \inf_{\chi} \left\{ \widetilde{S}_t \chi(0) - \psi_*^{\xi}(\chi) \right\}, \tag{1.11}$$

where the infimum is taken over the set of 1-Lipschitz functions $\chi : S_+^D \to \mathbb{R}$.

When mean-field spin glass models were first studied, it seemed natural to believe that the limit free energy could be written as the infimum of a functional. In this sense (1.4) which was first put forward in [19], can be referred to as an *inverted* variational formula. Below, we will therefore refer to formulas such as (1.7) and (1.11) as *un-inverted* variational formulas. We point out that a different un-inverted formula has already been obtained in [14, Theorem 1], where it is shown that the limit free energy can be written as an infimum over a set of martingales.

We finish this introduction by discussing some possible interpretations and generalizations of Theorems 1.2 and 1.3. We rely on the fact that, as pointed out in [12], the limit free energy is related to the viscosity solution of some Hamilton-Jacobi equation on $\mathcal{P}^{\uparrow}(S_+^D)$ (see Theorem 2.1 below).

One can show that any concave upper semicontinuous function on $\mathbb{R}^d$ coincides with the infimum of the family of affine functions that upper bound it, this is Fenchel-Moreau duality [23, Theorem 2.3.3]. The function $u(t,x) = a \cdot x + b + tH(a)$ can be interpreted as the unique solution (in the viscosity sense) of the Hamilton-Jacobi equation

$$\begin{cases} \partial_t u - H(\nabla u) = 0 \text{ on } (0, +\infty) \times \mathbb{R}^d \\ u(0, x) = a \cdot x + b. \end{cases}$$

Given a Hamilton-Jacobi equation on $\mathbb{R}^d$ with concave and upper semicontinuous initial condition φ, one can represent its unique solution (in the viscosity sense) as the infimum of the family of solutions with affine initial conditions that upper-bound φ, this is the Hopf representation [8, Theorem 3.13]. Note that for the Hopf representation to hold, it is *not* required to assume that $p \mapsto H(p)$ is convex.

For Hamilton-Jacobi equations on $\mathcal{P}^{\uparrow}(S_+^D)$, there is an important subtlety and more precision is needed to define concave and affine functions. There are two natural notions of geodesics on $\mathcal{P}^{\uparrow}(S_+^D)$. The first are the geodesics inherited from the inclusion of $\mathcal{P}^{\uparrow}(S_+^D)$ into the space of signed measures on S_+^D, those are the straight lines $\lambda \mapsto \lambda\mu + (1-\lambda)\mu'$. The second are the transport geodesics inherited from the inclusion of $\mathcal{P}^{\uparrow}(S_+^D)$ into $L^2([0,1), S_+^D)$ via the space of nondecreasing paths $\mathsf{q} : [0,1) \to S_+^D$, those are the images of the straight lines $\lambda \mapsto \lambda\mathsf{q} + (1-\lambda)\mathsf{q}'$ in $\mathcal{P}^{\uparrow}(S_+^D)$. This second kind of geodesics are *not* straight lines in $\mathcal{P}^{\uparrow}(S_+^D)$ and can be seen as the geodesics associated to the optimal transport geometry in $\mathcal{P}^{\uparrow}(S_+^D)$.

For the Hamilton-Jacobi equation on $\mathcal{P}^{\uparrow}(S_+^D)$ arising in the context of spin glasses (see (2.10) below), both of those geometries play a role. The transport derivative ∂_μ in (2.10) is defined using the transport geometry and tracks the infinitesimal slope of a function along the transport geodesics. On the other hand, the initial condition ψ in (2.10) is concave along straight lines $\lambda \mapsto \lambda\mu + (1-\lambda)\mu'$.

In [5, Theorem 1.1 (3)], it was shown that if the initial condition in (2.10) is concave along transport geodesics, then, perhaps unsurprisingly, the Hopf representation holds. But, it is known that this concavity assumption is not satisfied by the initial condition of (2.10) in general [11, Section 6].

At first glance, the two competing geometries in the formulation of (2.10) prevents the existence of a Hopf type representation for the solution. Indeed, the initial condition is concave along straight lines, so the relevant affine functions to consider for the Fenchel-Moreau duality are of the form $\mu \mapsto \int \chi(x)\mathrm{d}\mu(x)$. But, for a Hamilton-Jacobi equation on $\mathcal{P}^\uparrow(S^D_+)$ formulated with the transport derivative ∂_μ and with a fully general nonlinearity of the form $H(\partial_\mu f)$, there is no reason for the solution started with the initial condition $\mu \mapsto \int \chi(x)\mathrm{d}\mu(x)$ to be affine along straight lines. In the context of (2.10), there are additional structures to exploit in the nonlinear term $\int \xi(\partial_\mu f)\mathrm{d}\mu$ and this apparent incompatibility between the two geometries is resolved. We show in Theorem 4.2 that the solution of (2.10) with initial condition $\mu \mapsto \int \chi(x)\mathrm{d}\mu(x)$ is

$$(t,\mu) \mapsto \int S_t\chi(x)\mathrm{d}\mu(x),$$

where $S_t\chi$ is defined in (1.6). This proves that (2.10) does preserve affine functions (at least when $D = 1$).

Theorem 1.2 is in fact the statement that when $D = 1$, the solution of (2.10) is the infimum of the family of solutions with affine initial conditions that upper-bound ψ and (1.7) can therefore be interpreted as a Hopf-like formula for the solution of (2.10).

When ξ is not assumed to be convex on S^D_+, the Parisi formula breaks down and the value of the limit free energy is not known. In [12] it was conjectured that, in analogy with the convex case, the limit free energy should be related to the solution of (2.10). We believe that, at least under some additional assumptions on P_1 and ξ (see (H4) and (H5)), the Hopf-like representation derived in (1.7) should be available for the solution of (2.10) even when $D > 1$ and ξ is possibly non-convex on S^D_+. Together with [12, Conjecture 2.6], this yields a conjectural variational formula for the limiting value of the free energy when ξ is possibly non-convex on S^D_+ (see Conjecture 7.4).

1.3. **Organization of the paper.** We start by giving a proper definition of the enriched free energy in Section 2. In Section 3, we give a version of the Fenchel-Moreau theorem which holds for concave functions defined on the set of signed measures. In words, this result says that a concave function can be written as the infimum of the family of affine functions that upper bound it. In Section 4, using the Fenchel-Moreau theorem, we prove Theorem 1.2, the main argument is a $\sup - \inf$ interchange performed using [20]. To do so we rely crucially on the fact that $\mathcal{P}^{\uparrow}(\mathbb{R}_+) = \mathcal{P}(\mathbb{R}_+)$ is a convex set. In Section 5, to compensate for the lack of convexity of $\mathcal{P}^{\uparrow}(S^D_+)$ when $D > 1$, we show that any concave Lipschitz function $\psi : \mathcal{P}^{\uparrow}(S^D_+) \to \mathbb{R}$ can be extended to a concave Lipschitz function on $\mathcal{P}(S^D_+)$ (and even on the set of signed measures). The results of Section 5 allows us to prove Theorem 1.3 using similar arguments than in the proof of Theorem 1.2, this is done in Section 6. Finally, in Section 7 we explain the link between the un-inverted formulas and the Hopf-like representation for the viscosity solution of Hamilton-Jacobi equations.

Acknowledgement. I warmly thank Jean-Christophe Mourrat for motivating this project, as well as many helpful discussions and some decisive comments and ideas for Section 5.

2. The enriched free energy

Let $\mu \in \mathcal{P}^{\uparrow}_1(S^D_+)$ be a measure with finite support, we can write

$$\mu = \sum_{k=0}^{K} (\zeta_{k+1} - \zeta_k)\delta_{q_k}, \tag{2.1}$$

with $K \in \mathbb{N}$, $\zeta_0, \dots, \zeta_{K+1} \in \mathbb{R}$ satisfying

$$0 = \zeta_0 < \zeta_1 < \cdots < \zeta_K < \zeta_{K+1} = 1, \tag{2.2}$$

and $q_0, \dots, q_K \in S^D_+$ such that

$$0 = q_{-1} \leqslant q_0 < q_1 \cdots < q_{K-1} < q_K. \tag{2.3}$$

The definition of the enriched free energy will involve a probabilistic object called the Poisson-Dirichlet cascade. We will briefly recall some properties of this object, a full definition can be found in [15, Section 2.3]. We let

$$\mathcal{A} = \mathbb{N}^0 \cup \mathbb{N}^1 \cup \cdots \cup \mathbb{N}^K.$$

We think of $\mathcal{A}$ as a tree rooted at $\mathbb{N}^0 = \{\varnothing\}$, with depth K and such that each non-leaf vertex has countably infinite degree. For every $k \in \{0, \dots, K\}$, and $\alpha \in \mathbb{N}^k \subseteq \mathcal{A}$, we let $|\alpha| = k$ denote the depth of α. For every leaf $\alpha = (n_1, \dots, n_K) \in \mathbb{N}^K$, we let

$$\alpha\big|_k = (n_1, \dots, n_k),$$

with the understanding that $\alpha\big|_0$ is the root. For every $\alpha, \alpha' \in \mathcal{A}$ we define

$$\alpha \wedge \alpha' = \sup\left\{k \in \{0, \dots K\} \,\middle|\, \alpha\big|_k = \alpha'\big|_k\right\}.$$

A Poisson-Dirichlet cascade is the set $(v_\alpha)_{\alpha \in \mathbb{N}^K}$ of weights of a certain random probability measure on the set $\mathbb{N}^K$ of leaves of the tree $\mathcal{A}$. Those weights are constructed as follows. The children $\alpha 1, \alpha 2, \alpha 3, \dots$ of each vertex $\alpha \in \mathbb{N}^k$ for $k < K$ are decorated with the values $(u_{\alpha k})_{k \geqslant 1}$ (arranged in decreasing order) of an independent Poisson point process with intensity measure $\zeta_{k+1} x^{-(1+\zeta_{k+1})} dx$.

The weight of the leaf $\alpha \in \mathbb{N}^K$ is then calculated by taking the product of each of the weights associated to $\alpha\big|_k$ for $k \in \{1, \dots K\}$. Finally, the weights are normalized so that their sum is 1. Namely, if $\alpha \in \mathbb{N}^K$, we have

$$v_\alpha = \frac{w_\alpha}{\sum_{\beta \in \mathbb{N}^K} w_\beta},$$

where $w_\alpha = \prod_{k=1}^K u_{\alpha|_k}$.

We say that a random vector $z \in \mathbb{R}^{D\times N}$ is standard Gaussian when its coordinates in an orthonormal basis form a family of real independent standard Gaussian random variables. Let $(z_\alpha)_{\alpha \in \mathbb{N}^K}$ be a family of independent standard Gaussian vectors on $\mathbb{R}^{D\times N}$. For every $\sigma \in \mathbb{R}^{D\times N}$ and $\alpha \in \mathbb{N}^K$, we set

$$H_N^\mu(\sigma, \alpha) = \sum_{k=0}^K (2q_k - 2q_{k-1})^{1/2} z_{\alpha|_k} \cdot \sigma, \tag{2.4}$$

where $(2q_k - 2q_{k-1})^{1/2}$ should be understood as the square root of the symmetric positive semi-definite matrix $2q_k - 2q_{k-1}$. Here, $D \times D$ matrices act on $\mathbb{R}^{D\times N}$ via the standard multiplication of $D \times D$ matrices by $D \times N$ matrices. Alternatively, H_N^μ can be defined as the unique centered Gaussian process on $\mathbb{R}^D \times \mathbb{N}^K$ with the following covariance

$$\mathbb{E} H_N^\mu(\sigma, \alpha) H_N^\mu(\tau, \beta) = 2q_{\alpha\wedge\beta} \cdot \sigma\tau^\perp.$$

For every $t \geqslant 0$, we define the enriched free energy at (t, μ) by

$$F_N(t, \mu) = -\frac{1}{N} \log \int \sum_{\alpha \in \mathbb{N}^K} \exp\left(H_N^{t,\mu}(\sigma, \alpha)\right) v_\alpha \mathrm{d}P_N(\sigma), \tag{2.5}$$

where

$$H_N^{t,\mu}(\sigma, \alpha) = \sqrt{2t} H_N(\sigma) - Nt\xi\left(\frac{\sigma\sigma^\perp}{N}\right) + H_N^\mu(\sigma, \alpha) - q_K \cdot \sigma\sigma^\perp. \tag{2.6}$$

We let $\widetilde{\mathbb{E}}$ denote the expectation conditionally on the randomness coming from $(H_N(\sigma))_{\sigma \in \mathbb{R}^{D\times N}}$. We define the partially and fully averaged free energies

$$\widetilde{F}_N(t, \mu) = \widetilde{\mathbb{E}}[F_N(t, \mu)],$$
$$\overline{F}_N(t, \mu) = \mathbb{E}[F_N(t, \mu)].$$

For every $t \geqslant 0$, $\widetilde{F}_N(t, \cdot)$ is Lipschitz continuous on the set of finitely supported measures in $\mathcal{P}_\infty^\uparrow(S_+^D)$ [13, Proposition 3.1]. In particular $\overline{F}_N$ and $\widetilde{F}_N$ can be extended by continuity to $\mathbb{R}_+ \times \mathcal{P}_1^\uparrow(S_+^D)$. We let $\psi = F_1(0, \cdot)$, that is ψ is the unique Lipschitz continuous function on $\mathcal{P}_1^\uparrow(S_+^D)$ such that for every finitely supported μ as in (2.1), we have

$$\psi(\mu) = -\mathbb{E} \log \int \sum_{\alpha \in \mathbb{N}^K} \exp\left(H_1^\mu(\sigma, \alpha) - q_K \cdot \sigma\sigma^\perp\right) v_\alpha \mathrm{d}P_1(\sigma). \tag{2.7}$$

Let U be a uniform random variable on $[0,1)$. Given a monotonically coupled measure, $\mu \in \mathcal{P}_p^{\uparrow}(S_+^D)$, there exists a unique nondecreasing right continuous path with left limits $\mathsf{q}_\mu \in L^p([0,1),S^D)$ such that μ is the law of the random variable $X_\mu = \mathsf{q}_\mu(U)$. We let $\mathcal{Q}(S_+^D)$ denote the set of nondecreasing right continuous paths with left limits, and we define $\mathcal{Q}_p(S_+^D) = \mathcal{Q}(S_+^D) \cap L^p([0,1),S^D)$. The set $\mathcal{Q}_p(S_+^D)$ is a closed convex cone of $L^p([0,1),S^D)$, meaning that it is a closed convex subset and for every $\mathsf{q},\mathsf{q}' \in \mathcal{Q}_p(S_+^D)$, $t,t' \geqslant 0$ we have

$$t\mathsf{q} + t'\mathsf{q}' \in \mathcal{Q}_p(S_+^D).$$

The cone $\mathcal{Q}_2(S_+^D)$ is embedded in the Hilbert space $L^2([0,1),S^D)$, we can define its dual cone

$$\mathcal{Q}_2(S_+^D)^* = \left\{ \mathsf{p} \in L^2([0,1),S^D) \middle| \, \forall \mathsf{q} \in \mathcal{Q}_2(S_+^D), \ \langle \mathsf{p},\mathsf{q}\rangle_{L^2} \geqslant 0 \right\}.$$

Let $\kappa \in L^2$, according to [4, Lemma 3.5], we have $\kappa \in \mathcal{Q}_2(S_+^D)^*$ if and only if for every $t \in [0,1)$,

$$\int_t^1 \kappa(u)\mathrm{d}u \in S_+^D.$$

Given $\mathsf{q},\mathsf{q}' \in \mathcal{Q}_1(S_+^D)$, we write $\mathsf{q} \leqslant \mathsf{q}'$ whenever for every $t \in [0,1)$,

$$\int_t^1 \mathsf{q}'(u) - \mathsf{q}(u)\mathrm{d}u \in S_+^D.$$

When $\mathsf{q},\mathsf{q}' \in \mathcal{Q}_2(S_+^D)$, we have $\mathsf{q} \leqslant \mathsf{q}'$ if and only if $\mathsf{q}' - \mathsf{q} \in \mathcal{Q}_2(S_+^D)^*$. This notation can be extended to $\mathcal{P}_1^{\uparrow}(S_+^D)$ by setting $\mu \leqslant \mu'$ whenever $\mathsf{q}_\mu \leqslant \mathsf{q}_{\mu'}$. We say that a function defined on a subset of $\mathcal{P}_1^{\uparrow}(S_+^D)$ is nondecreasing when it is nondecreasing with respect to the order that we have just defined.

The dual of the closed convex cone $\mathcal{Q}_2(S_+^D)^*$ is the cone $\mathcal{Q}_2(S_+^D)$ itself, more precisely we have

$$\mathcal{Q}_2(S_+^D) = \left\{ \mathsf{q} \in L^2([0,1),S^D) \middle| \, \forall \mathsf{p} \in \mathcal{Q}_2(S_+^D)^*, \ \langle \mathsf{p},\mathsf{q}\rangle_{L^2} \geqslant 0 \right\}. \tag{2.8}$$

This property will be useful to show that certain functions defined on $\mathcal{P}_2^{\uparrow}(S_+^D)$ are nondecreasing.

The point of introducing the enriched free energy (2.5) is [4, Corollary 8.2] which we state as Theorem 2.1 below. Roughly speaking, Theorem 2.1 states that the limit of the enriched free energy is the unique solution to a partial differential equation. One can recover the classical Parisi formula (1.4) from Theorem 2.1 by simply setting $\mu = \delta_0$ in (2.9). Theorem 2.1 gives a new way of interpreting the Parisi formula, which does not solely rely on a variational representation. In particular, this point of view allows us to formulate a credible conjecture for the limit free energy of models whose covariance function ξ is nonconvex on S_+^D [12, Conjecture 2.6]. We refer to [4, 11, 13] for partial results in this direction.

Theorem 2.1 ([4])**.** *Suppose that the function ξ is convex over S^D_+. Then for every $t \geqslant 0$ and $\mu \in \mathcal{P}^\uparrow_1(S^D_+)$, we have*

$$\lim_{N\to+\infty} \overline{F}_N(t,\mu) = \sup_{\nu \in \mathcal{P}^\uparrow_\infty(S^D_+), \nu \geqslant \mu} \left\{ \psi(\nu) - t\mathbb{E}\left[\xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right] \right\}. \tag{2.9}$$

Moreover, denoting by $f(t,\mu)$ the expression above, we have that $f : \mathbb{R}_+ \times \mathcal{P}^\uparrow_2(S^D_+) \to \mathbb{R}$ solves the Hamilton-Jacobi equation

$$\begin{cases} \partial_t f - \int \xi(\partial_\mu f)\,\mathrm{d}\mu = 0 & \text{on } \mathbb{R}_+ \times \mathcal{P}^\uparrow_2(S^D_+) \\ f(0,\cdot) = \psi & \text{on } \mathcal{P}^\uparrow_2(S^D_+). \end{cases} \tag{2.10}$$

Remark 2.2. When $D = 1$, as pointed out in [6, Propostion A.3], we can drop the condition "$\nu \geqslant \mu$" in (2.9) and we have

$$\lim_{N\to+\infty} \overline{F}_N(t,\mu) = \sup_{\nu \in \mathcal{P}_\infty(\mathbb{R}_+)} \left\{ \psi(\nu) - t\mathbb{E}\left[\xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right] \right\}. \tag{2.11}$$

In (2.10), the symbol ∂_μ should be interpreted as a derivative of transport type. Informally, this means that given any sufficiently smooth function $g : \mathcal{P}^\uparrow_2(S^D_+) \to \mathbb{R}$, the symbol $\partial_\mu g(\mu)$ should be understood as an S^D-valued function defined on S^D_+ that is square integrable with respect to the measure μ and is such that the following holds as μ' converges to μ in $\mathcal{P}^\uparrow_2(S^D_+)$,

$$g(\mu') = g(\mu) + \mathbb{E}[\partial_\mu g(\mu)(X_\mu) \cdot (X_{\mu'} - X_\mu)] + o\left(\mathbb{E}\left[|X_{\mu'} - X_\mu|^2 \right]^{\frac{1}{2}} \right).$$

Definition 2.3. Let $\mathsf{g} : \mathcal{Q}_2(S^D_+) \to \mathbb{R}$ and $\mathsf{q} \in \mathcal{Q}_2(S^D_+)$. We say that g is Fréchet differentiable at q if there exists a unique $\mathsf{p} \in L^2([0,1), S^D)$ such that

$$\lim_{r\to 0} \sup_{\substack{\mathsf{q}' \in \mathcal{Q}_2(S^D_+) \\ 0 < |\mathsf{q}-\mathsf{q}'|_{L^2} \leqslant r}} \frac{\mathsf{g}(\mathsf{q}') - \mathsf{g}(\mathsf{q}) - \langle \mathsf{p}, \mathsf{q}' - \mathsf{q} \rangle_{L^2}}{|\mathsf{q} - \mathsf{q}'|_{L^2}} = 0.$$

In this case, we say that p is the Fréchet derivative of g at q and we denote it $\nabla \mathsf{g}(\mathsf{q})$.

The derivative ∂_μ can be reinterpreted as a Fréchet derivative in the following way. The map $\Omega : \mu \mapsto \mathsf{q}_\mu$ yields a nonlinear isometric bijection between $\mathcal{P}^\uparrow_2(S^D_+)$ and $\mathcal{Q}_2(S^D_+)$. Given a sufficiently smooth function $g : \mathcal{P}^\uparrow_2(S^D_+) \to \mathbb{R}$, we define $\mathsf{g} = g \circ \Omega^{-1}$. Then formally, we have for every $u \in [0,1)$

$$\partial_\mu g(\mu)(\mathsf{q}_\mu(u)) = \nabla \mathsf{g}(q_\mu)(u).$$

In particular, setting $\mathsf{f}(t,\mathsf{q}) = f(t, \Omega^{-1}\mathsf{q})$, the partial differential equation (2.10) can be seen as a special case of

$$\begin{cases} \partial_t \mathsf{f} - H(\nabla \mathsf{f}) = 0 \text{ on } (0,+\infty) \times \mathcal{Q}_2(S^D_+) \\ \mathsf{f}(0,\mathsf{q}) = \psi(\Omega^{-1}\mathsf{q}), \end{cases} \tag{2.12}$$

with $H(\mathsf{p}) = \int_0^1 \xi(\mathsf{p}(u))\mathrm{d}u$. This is the point of view adopted in [6] to prove that (2.10) is well-posed. Note that in (2.12) the nonlinearity H doesn't depend directly on q, while in (2.10) the nonlinearity depends on μ through the integration of $\xi(\partial_\mu f)$ with respect to μ. Despite this apparent simplification, some important properties of f and ψ can only be seen when they are considered as functions of μ and not q. For example $\psi : \mathcal{P}_2^\uparrow(S_+^D) \to \mathbb{R}$ is concave along straight lines [3], but this doesn't imply that $\mathsf{q} \mapsto \psi(\Omega^{-1}\mathsf{q})$ is concave, since in general,

$$\mathsf{q}_{(1-\lambda)\mu_0+\lambda\mu_1} \neq (1-\lambda)\mathsf{q}_{\mu_0} + \lambda \mathsf{q}_{\mu_1}.$$

Finally, note that the cone $\mathcal{Q}_2(S_+^D)$ has empty interior in L^2. We define $\mathcal{Q}_\uparrow(S_+^D)$ as the set of paths $\mathsf{q} \in \mathcal{Q}_2(S_+^D)$ such that there exists $c > 0$ satisfying the following for every $u \leqslant v$,

$$(\mathsf{q}(v) - cv\mathrm{Id}) - (\mathsf{q}(u) - cu\mathrm{Id}) \in S_+^D \text{ and } \mathrm{Ellipt}(\mathsf{q}(v) - \mathsf{q}(u)) \leqslant \frac{1}{c},$$

where for $m \in S^D$, $\mathrm{Ellipt}(m) \in [1, +\infty)$ denotes the ratio between the smallest and the largest eigenvalue of m. In practice $\mathcal{Q}_\uparrow(S_+^D)$ plays the role of the interior of $\mathcal{Q}_2(S_+^D)$ since for every $\mathsf{q} \in \mathcal{Q}_\uparrow(S_+^D)$ and every Lipschitz $\kappa : [0,1) \to S^D$, we have $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(S_+^D)$ for every $\varepsilon > 0$ small enough.

3. Fenchel-Moreau duality in Wasserstein space

Let $F \subseteq S_+^D$ be a closed set, and $x_0 \in F$. Let $\mathcal{L}(F, x_0)$ denote the set of Lipschitz functions $\chi : F \to \mathbb{R}$, equipped with the norm $|\chi|_{\mathcal{L}} = \max\{|\chi(x_0)|, |\chi|_{\mathrm{Lip}}\}$ where

$$|\chi|_{\mathrm{Lip}} = \sup_{\substack{x,y\in F \\ x\neq y}} \left\{ \frac{|\chi(x) - \chi(y)|}{|x-y|} \right\}.$$

When the point x_0 is clear from context, we simply write $\mathcal{L}(F)$. We let $\mathcal{L}_{\leqslant 1}(F)$ denote the unit ball of $\mathcal{L}(F)$, that is

$$\mathcal{L}_{\leqslant 1}(F) = \{\chi \in \mathcal{L}(F) \big| |\chi|_{\mathcal{L}} \leqslant 1\}.$$

Note that, according to the Arzelà-Ascoli theorem, $\mathcal{L}_{\leqslant 1}(F)$ is compact with respect to the topology of local uniform convergence. We also let $\mathcal{L}^0(F)$ denote the set of functions $\chi \in \mathcal{L}(F)$ satisfying $\chi(x_0) = 0$ and $\mathcal{L}^0_{\leqslant 1}(F) = \mathcal{L}^0(F) \cap \mathcal{L}_{\leqslant 1}(F)$.

Given a signed Borel measure ν on F with finite first moment, we define its Kantorovich-Rubinstein norm,

$$|\nu|_{\mathcal{M}} = \sup_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int_F \chi \mathrm{d}\nu \right\}.$$

We let $\mathcal{M}_1(F)$ denote the completion with respect to $|\cdot|_{\mathcal{M}}$ of the set of signed Borel measures on F having finite first moment. The closed linear span of $\{\delta_x \big| x \in F\}$ is $\mathcal{M}_1(F)$. Note that the distance $d(\nu, \nu') = |\nu - \nu'|_{\mathcal{M}}$ induced by the norm $|\cdot|_{\mathcal{M}}$ coincides with the optimal transport distance when restricted to $\mathcal{P}_1(F) \times \mathcal{P}_1(F)$ [22, Theorem 5.10], in particular for every $\nu, \nu' \in \mathcal{P}_1(F)$ we have

$$|\nu - \nu'|_{\mathcal{M}} = \inf_{\pi \in \Pi(\nu,\nu')} \left\{ \int_{F \times F} |x - y| \mathrm{d}\pi(x, y) \right\},$$

where $\Pi(\nu, \nu')$ denotes the set of probability measures $\pi \in \mathcal{P}_1(F \times F)$ such that $\pi_1 = \nu$ and $\pi_2 = \nu'$. Here and throughout, π_1 and π_2 denote the first and second marginal of the coupling π. More precisely, if $p_i : F \times F \to F$ denotes the projection on the i-th coordinate, we have $\pi_i = p_{i\#}\pi$

Proposition 3.1. *The continuous dual of* $(\mathcal{M}_1(F), |\cdot|_{\mathcal{M}})$ *is* $(\mathcal{L}(F), |\cdot|_{\mathcal{L}})$.

Proof. Let ℓ be a continuous linear form on $(\mathcal{M}(F), |\cdot|_{\mathcal{M}})$, for every $x \in F$ we let $\chi(x) = \ell(\delta_x)$. For every $x, y \in F$ we have

$$|\chi(x) - \chi(y)| = |\ell(\delta_x - \delta_y)| \leqslant c|\delta_x - \delta_y|_{\mathcal{M}} = c|x - y|,$$

therefore $\chi \in \mathcal{L}(F)$. For every $x_1, \ldots, x_K \in F$ and for every $\lambda_1, \ldots, \lambda_K \in \mathbb{R}$, we have

$$\ell\left(\sum_{k=1}^{K} \lambda_k \delta_{x_k}\right) = \sum_{k=1}^{K} \lambda_k \chi(x_k).$$

This means that for every finitely supported $\mu \in \mathcal{M}_1(F)$,

$$\ell(\mu) = \int \chi \mathrm{d}\mu.$$

Since $\chi \in \mathcal{L}(F)$, the map $\mu \mapsto \int \chi \mathrm{d}\mu$ defines a continuous linear form on $(\mathcal{M}_1(F), |\cdot|_{\mathcal{M}})$ that coincides with ℓ on the set of finitely supported measures. By density, we have for every $\mu \in \mathcal{M}_1(F)$,

$$\ell(\mu) = \int \chi \mathrm{d}\mu.$$

We have proven that $(\mathcal{M}_1(F), |\cdot|_{\mathcal{M}})^* = (\mathcal{L}(F), |\cdot|^*_{\mathcal{M}})$, where

$$|\chi|^*_{\mathcal{M}} = \sup_{|\mu|_{\mathcal{M}} \leqslant 1} \left\{ \int \chi \mathrm{d}\mu \right\}.$$

To conclude, let us show that for every $\chi \in \mathcal{L}(F)$, $|\chi|^*_{\mathcal{M}} = |\chi|_{\mathcal{L}}$. By construction, it is clear that $|\chi|^*_{\mathcal{M}} \leqslant |\chi|_{\mathcal{L}}$. To show the converse inequality, one can plug in $\mu = \pm(\delta_x - \delta_y)/|x - y|$ and $\mu = \pm\delta_{x_0}$ in the previous display to discover that

$$|\chi|^*_{\mathcal{M}} \geqslant \frac{|\chi(x) - \chi(y)|}{|x - y|},$$
$$|\chi|^*_{\mathcal{M}} \geqslant |\chi(x_0)|.$$

□

We say that a sequence of measures $(\mu_n)_n$ in $\mathcal{M}_1(F)$ weakly converges to μ in $\mathcal{M}_1(F)$ as $n \to +\infty$ when for every $\chi \in \mathcal{L}(F)$,

$$\lim_{n \to +\infty} \int \chi \mathrm{d}\mu_n = \int \chi \mathrm{d}\mu.$$

We say that a function $\varphi : \mathcal{M}_1(F) \to \mathbb{R} \cup \{-\infty\}$ is weakly upper semicontinuous when for every sequence $(\mu_n)_n$ that weakly converges to μ in $\mathcal{M}_1(F)$, we have

$$\limsup_{n \to +\infty} \varphi(\mu_n) \leqslant \varphi(\mu).$$

Every weakly upper semi-continuous function on $\mathcal{M}_1(F)$ is upper semi-continuous (that is, with respect to strong convergence under $|\cdot|_{\mathcal{M}}$). The converse is not true in general, but the following result holds.

Proposition 3.2. *Let $\varphi : \mathcal{M}_1(F) \to \mathbb{R}$ be a concave function, if φ is upper semi-continuous then φ is weakly upper semi-continuous.*

Proof. Since φ is concave and upper semi-continuous, the set

$$B = \{(x, \mu) \in \mathbb{R} \times \mathcal{M}_1(F) \big| \varphi(\mu) \geqslant x\},$$

of points below the graph of φ is closed and convex. In particular, it follows from the Hahn-Banach separation theorem, that B is weakly closed [7, Corollary 1.5]. This means that φ is weakly upper semi-continuous. □

Let us now give a statement of the Fenchel-Moreau duality in the context of the dual pair $(\mathcal{M}_1(F), \mathcal{L}(F))$. Usually, the Fenchel-Moreau duality is stated for convex functions, but it can be transformed into a statement about concave functions (and vice-versa) replacing each function φ by its opposite $-\varphi$. Here the functions we are ultimately interested in in Section 4 and 6 are concave, so we choose to state the Fenchel-Moreau duality as a result for concave functions.

Let $\varphi : \mathcal{M}_1(F) \to \mathbb{R} \cup \{-\infty\}$, we define its concave conjugate $\varphi_* : \mathcal{L}(F) \to \mathbb{R} \cup \{-\infty\}$ by

$$\varphi_*(\chi) = \inf_{\mu \in \mathcal{M}_1(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi(\mu) \right\}.$$

Similarly, given a function $\phi : \mathcal{L}(F) \to \mathbb{R} \cup \{-\infty\}$ we define its concave conjugate $\phi_* : \mathcal{M}_1(F) \to \mathbb{R} \cup \{-\infty\}$ by

$$\phi_*(\mu) = \inf_{\chi \in \mathcal{L}(F)} \left\{ \int \chi \mathrm{d}\mu - \phi(\chi) \right\}.$$

The following theorem is a translation of [23, Theorem 2.3.3] in our context.

Theorem 3.3 ([23])**.** *Let $\varphi : \mathcal{M}_1(F) \to \mathbb{R} \cup \{-\infty\}$ be a function that is not identically equal to $-\infty$. Then $\varphi_{**} = \varphi$ if and only if φ is concave and upper semicontinuous.*

The concave conjugate φ_* of φ is defined as the infimum of a family of affine functions, hence for some $\chi \in \mathcal{L}(F)$ we may have $\varphi_*(\chi) = -\infty$. We define,

$$\mathrm{dom}(\varphi_*) = \{\chi \in \mathcal{L}(F) \big| \varphi_*(\chi) > -\infty\}.$$

For every $\mu \in \mathcal{M}_1(F)$, we have

$$\inf_{\chi \in \mathcal{L}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\} = \inf_{\chi \in \mathrm{dom}(\varphi_*)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\}.$$

This alternative representation for φ_{**} can be useful when the functions in $\mathrm{dom}(\varphi_*)$ are shown to have special properties. The following proposition states that when φ is 1-Lipschitz, the set $\mathrm{dom}(\varphi_*)$ is contained in the unit ball of $\mathcal{L}(F)$.

Proposition 3.4. *Let $\varphi : (\mathcal{M}_1(F), |\cdot|_{\mathcal{M}}) \to \mathbb{R}$ be a 1-Lipschitz function. For every $\chi \in \mathcal{L}(F)$, if $|\chi|_{\mathcal{L}} > 1$, then*

$$\varphi_*(\chi) = -\infty.$$

Proof. Let $\chi \in \mathcal{L}(F)$.

Step 1. We show that if $|\chi|_{\mathcal{L}} > 1$, then there exists $\mu' \in \mathcal{M}_1(F)$ such that $|\mu'|_{\mathcal{M}} < 1$ and $\int \chi \mathrm{d}\mu' < -1$.

By contradiction, assume that for every $\mu \in \mathcal{M}_1(F)$, if $|\mu|_{\mathcal{M}} < 1$, then $\int \chi \mathrm{d}\mu \geqslant -1$. In this case, for every $\mu \in \mathcal{M}_1(F)$ and $\varepsilon > 0$, we have

$$\int \chi \mathrm{d}\left(\frac{-\mu}{(1+\varepsilon)|\mu|_{\mathcal{M}}} \right) \geqslant -1.$$

Letting $\varepsilon \to 0$ we obtain

$$\int \chi \mathrm{d}\mu \leqslant |\mu|_{\mathcal{M}}.$$

Now taking the supremum over $\mu \in \mathcal{M}_1(F)$, this yields $|\chi|_{\mathcal{L}} \leqslant 1$, a contradiction.

Step 2. Conclusion.

Let $0 \in \mathcal{M}_1(F)$ denote the null measure. For every $\mu \in \mathcal{M}_1(F)$, we have

$$\varphi(\mu) \geqslant \varphi(0) - |\mu|_{\mathcal{M}}.$$

It follows that

$$\varphi_*(\chi) \leqslant -\varphi(0) + \inf_{\mu \in \mathcal{M}(F)} \left\{ \int \chi \mathrm{d}\mu + |\mu|_{\mathcal{M}} \right\}.$$

Assume that $|\chi|_{\mathcal{L}} > 1$, let $\mu' \in \mathcal{M}_1(F)$ be such that $|\mu'|_{\mathcal{M}} < 1$ and $\int \chi \mathrm{d}\mu' \leqslant -1$ as in Step 1. We have

$$\lim_{t \to +\infty} \left(\int \chi \mathrm{d}(t\mu') + |t\mu'|_{\mathcal{M}} \right) = -\infty.$$

Thus,

$$\inf_{\mu \in \mathcal{M}_1(F)} \left\{ \int \chi \mathrm{d}\mu + |\mu|_{\mathcal{M}} \right\} = -\infty,$$

and it follows that $\varphi_*(\chi) = -\infty$. □

We use the notation $\min_{x \in X} f(x)$ to denote the value $\inf_{x \in X} f(x)$ when there exists $x_0 \in X$ such that $f(x_0) = \inf_{x \in X} f(x)$.

Corollary 3.5. *Let $\varphi : \mathcal{M}_1(F) \to \mathbb{R} \cup \{-\infty\}$ be a function that is not identically equal to $-\infty$. Assume that φ is 1-Lipschitz with respect to $|\cdot|_{\mathcal{M}}$ and concave, then for every $\mu \in \mathcal{M}_1(F)$ we have*

$$\varphi(\mu) = \min_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\}.$$

Of course, Corollary 3.5 is true if φ is only assumed to be L-Lipschitz for some $L \geqslant 0$, in this case we need to minimize over $\mathcal{L}_{\leqslant L}(F)$ rather than $\mathcal{L}_{\leqslant 1}(F)$.

Proof. Step 1. We show that

$$\varphi(\mu) = \inf_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\}.$$

According to Proposition 3.4, we have $\varphi_*(\chi) = -\infty$ whenever $\chi \notin \mathcal{L}_{\leqslant 1}(F)$, therefore

$$\inf_{\chi \in \mathcal{L}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\} = \inf_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\}.$$

Combining this with Theorem 3.3, we obtain

$$\varphi(\mu) = \inf_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\}.$$

Step 2. We show that the infimum in the variational formula of Step 1 is reached at some $\chi \in \mathcal{L}_{\leqslant 1}$.
For every $n \geqslant 1$, we let $\chi_n \in \mathcal{L}_{\leqslant 1}(F)$ be such that

$$\inf_{\chi \in \mathcal{L}_{\leqslant 1}(F)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*(\chi) \right\} \geqslant \int \chi_n \mathrm{d}\mu - \varphi_*(\chi_n) - \frac{1}{n}.$$

The sequence $(\chi_n)_n$ converges locally uniformly along a subsequence $(n_k)_k$ to some $\chi \in \mathcal{L}_{\leqslant 1}(F)$. In order to conclude it is enough to show that

$$\liminf_{k \to +\infty} \left\{ \int \chi_{n_k} \mathrm{d}\mu - \varphi_*(\chi_{n_k}) \right\} \geqslant \int \chi \mathrm{d}\mu - \varphi_*(\chi)$$

Let $\nu \in \mathcal{M}_1(F)$, we have $\chi_{n_k} \to \chi$ pointwise on F as $k \to +\infty$ and for every $x \in F$,

$$|\chi_{n_k}(x)| \leqslant 1 + |x - x_0|.$$

Therefore, by dominated convergence, we have

$$\lim_{k \to +\infty} \int \chi_{n_k} \mathrm{d}\nu = \int \chi \mathrm{d}\nu.$$

Since we have

$$\varphi_*(\chi) = \inf_{\nu \in \mathcal{M}_1(F)} \left\{ \int \chi \mathrm{d}\nu - \varphi(\nu) \right\},$$

it follows that $\limsup_{k \to +\infty} \varphi_*(\chi_{n_k}) \leqslant \varphi_*(\chi)$. Since we have

$$\lim_{k \to +\infty} \int \chi_{n_k} \mathrm{d}\mu = \int \chi \mathrm{d}\mu,$$

this concludes the proof. □

4. Scalar models

In this section, we focus on the case $D = 1$, and prove Theorem 1.2. Recall that in this setting $\mathcal{P}^{\uparrow}(\mathbb{R}_+) = \mathcal{P}(\mathbb{R}_+)$ is a convex set. Also recall that we have defined for every Lipschitz function $\chi : \mathbb{R}_+ \to \mathbb{R}$,

$$S_t\chi(x) = \sup_{y \in \mathbb{R}_+} \left\{ \chi(y) - t\xi^* \left(\frac{y-x}{t} \right) \right\}.$$

The function $(t,x) \mapsto S_t\chi(x)$ can be interpreted as the unique viscosity solution of

$$\begin{cases} \partial_t v - \xi(\nabla v) = 0 \text{ on } (0,+\infty) \times \mathbb{R}_+ \\ v(0,\cdot) = \chi. \end{cases} \tag{4.1}$$

4.1. Variational representation for the initial condition.

Let $\psi : \mathcal{P}_1(\mathbb{R}_+) \to \mathbb{R}$, for every Lipschitz function $\chi : \mathbb{R}_+ \to \mathbb{R}$, define

$$\psi_*(\chi) = \inf_{\mu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \psi(\mu) \right\}.$$

We recall that $\mathcal{L}(\mathbb{R}_+)$ denotes the set of Lipschitz functions $\chi : \mathbb{R}_+ \to \mathbb{R}$, we choose $x_0 = 0$ as the reference point for the norm $|\cdot|_{\mathcal{L}}$. We also recall that $\mathcal{L}_{\leqslant 1}(\mathbb{R}_+)$ denotes the unit ball of $\mathcal{L}(\mathbb{R}_+)$ and $\mathcal{L}^0(\mathbb{R}_+)$ denote the set of functions $\chi \in \mathcal{L}(\mathbb{R}_+)$ satisfying $\chi(0) = 0$. We let $\mathcal{X}$ denote the set of functions $\chi \in \mathcal{L}_{\leqslant 1}(\mathbb{R}_+)$ that are convex and nondecreasing, and $\mathcal{X}^0 = \mathcal{X} \cap \mathcal{L}^0$. Also recall that given $\mu, \nu \in \mathcal{P}_1(\mathbb{R}_+)$, we say that $\mu \leqslant \nu$ whenever for every $t \in [0,1)$,

$$\int_t^1 \mathsf{q}_\nu(u) - \mathsf{q}_\mu(u) \mathrm{d}u \geqslant 0, \tag{4.2}$$

and that $\psi : \mathcal{P}_1(\mathbb{R}_+) \to \mathbb{R}$ is said to be nondecreasing whenever we have for every $\mu, \nu \in \mathcal{P}_1(\mathbb{R}_+)$,

$$\mu \leqslant \nu \implies \psi(\mu) \leqslant \psi(\nu).$$

Finally, recall that we equip $\mathcal{M}_1(\mathbb{R}_+)$ with the norm $|\cdot|_{\mathcal{M}}$ and we denote by d the associated distance,

$$d(\nu, \nu') = |\nu - \nu'|_{\mathcal{M}}.$$

We recall that we use the notation $\min_{x \in X} f(x)$ to denote the value $\inf_{x \in X} f(x)$ when there exists $x_0 \in X$ such that $f(x_0) = \inf_{x \in X} f(x)$.

Lemma 4.1. *Let $\psi : \mathcal{P}_1(\mathbb{R}_+) \to \mathbb{R}$ be a 1-Lipschitz, concave and nondecreasing function. Then, for every $\mu \in \mathcal{P}_1(\mathbb{R}_+)$ we have*

$$\psi(\mu) = \min_{\chi \in \mathcal{X}^0} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\}.$$

In what follows, given $h : \mathbb{R}_+ \to \mathbb{R}$, we use $\int h(\mathsf{q})$ as a shorthand for $\int_0^1 h(\mathsf{q}(u))\mathrm{d}u$.

Proof of Lemma 4.1. For every $\mu \in \mathcal{M}_1(\mathbb{R}_+)$, define

$$\overline{\psi}(\mu) = \sup_{\nu \in \mathcal{P}_1(\mathbb{R}_+)} \{\psi(\nu) - d(\mu, \nu)\}.$$

The function $\overline{\psi} : \mathcal{M}_1(\mathbb{R}_+) \to \mathbb{R}$ is 1-Lipschitz as a supremum of Lipschitz functions and is concave as a supremum of a jointly concave functional. In addition, since ψ is 1-Lipschitz, for every $\mu \in \mathcal{P}_1(\mathbb{R}_+)$ we have $\overline{\psi}(\mu) = \psi(\mu)$. For every $\chi \in \mathcal{L}(\mathbb{R}_+)$, we let

$$\overline{\psi}_*(\chi) = \inf_{\mu \in \mathcal{M}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \overline{\psi}(\mu) \right\}.$$

Step 1. We show that for every $\chi \in \mathcal{L}_{\leqslant 1}(\mathbb{R}_+)$, $\overline{\psi}_*(\chi) = \psi_*(\chi)$.
Let $\chi \in \mathcal{L}_{\leqslant 1}(\mathbb{R}_+)$, the map $\mu \mapsto \int \chi \mathrm{d}\mu$ is 1-Lipschitz on $\mathcal{M}_1(\mathbb{R}_+)$ with respect to $|\cdot|_{\mathcal{M}}$. In particular, for every $\nu \in \mathcal{P}_1(\mathbb{R}_+)$, we have

$$\inf_{\mu \in \mathcal{M}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu + d(\nu, \mu) \right\} = \int \chi \mathrm{d}\nu.$$

Thus,

$$\begin{aligned}
\overline{\psi}_*(\chi) &= \inf_{\mu \in \mathcal{M}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \overline{\psi}(\mu) \right\} \\
&= \inf_{\mu \in \mathcal{M}_1(\mathbb{R}_+)} \inf_{\nu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \psi(\nu) + d(\mu, \nu) \right\} \\
&= \inf_{\nu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ -\psi(\nu) + \inf_{\mu \in \mathcal{M}_1(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu + d(\mu, \nu) \right\} \right\} \\
&= \inf_{\nu \in \mathcal{P}_1(\mathbb{R}_+)} \left\{ -\psi(\nu) + \int \chi \mathrm{d}\nu \right\} \\
&= \psi_*(\chi).
\end{aligned}$$

Step 2. We show that, for every $\mu \in \mathcal{P}_1(\mathbb{R}_+)$,

$$\psi(\mu) = \min_{\chi \in \mathcal{L}^0_{\leqslant 1}(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\}.$$

According to Corollary 3.5, we have for every $\mu \in \mathcal{M}_1(\mathbb{R}_+)$,

$$\overline{\psi}(\mu) = \min_{\chi \in \mathcal{L}_{\leqslant 1}(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \overline{\psi}_*(\chi) \right\}.$$

According to Step 1, for every $\chi \in \mathcal{L}_{\leqslant 1}(\mathbb{R}_+)$, $\overline{\psi}_*(\chi) = \psi_*(\chi)$. In addition, for every $c \in \mathbb{R}$, we have $\psi_*(\chi + c) = \psi_*(\chi) + c$. Therefore, since $\overline{\psi}$ is an extension of ψ, it follows from the previous display that for every $\mu \in \mathcal{P}_1(\mathbb{R}_+)$,

$$\psi(\mu) = \min_{\chi \in \mathcal{L}^0_{\leqslant 1}(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\}.$$

Step 3. We show that in the formula of Step 2, the minimum can be taken over the set of $\chi \in \mathcal{L}^0_{\leqslant 1}(\mathbb{R}_+)$ that are nondecreasing and convex.

Let λ denote the Lebesgue measure on $\mathbb{R}$, let $\mu \in \mathcal{P}_1(\mathbb{R}_+)$ be such that the associated path $\mathsf{q} = \mathsf{q}_\mu : [0,1) \to \mathbb{R}_+$ is surjective, belongs to $\mathcal{Q}_\uparrow(\mathbb{R}_+)$ and satisfies $\lambda(\mathsf{q}^{-1}(A)) = 0$ for any λ-negligible set $A \subseteq \mathbb{R}$. According to Step 2, there exists $\chi \in \mathcal{L}^0_{\leqslant 1}(\mathbb{R}_+)$ such that

$$\psi(\mu) = \int \chi \mathrm{d}\mu - \psi_*(\chi).$$

In addition, for every $\mu' \in \mathcal{P}_1(\mathbb{R}_+)$ we have,

$$\psi(\mu') \leqslant \int \chi \mathrm{d}\mu' - \psi_*(\chi).$$

Let $\mathsf{q}' \in \mathcal{Q}_2(\mathbb{R}_+)$ such that $\mathsf{q}' - \mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)^*$, and let $\mu' = \mathrm{Law}(\mathsf{q}'(U))$. By definition, we have $\mu' \geqslant \mu$ and,

$$\int \chi(\mathsf{q}') - \int \chi(\mathsf{q}) \geqslant \psi(\mu') - \psi(\mu) \geqslant 0.$$

Now let $\kappa \in \mathcal{Q}_2(\mathbb{R}_+)^*$ be a Lipschitz path, for $\varepsilon > 0$ small enough we have $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(\mathbb{R}_+)$, applying the previous display to $\mathsf{q}' = \mathsf{q} + \varepsilon\kappa$, we obtain

$$\int_0^1 \frac{\chi(\mathsf{q}(u) + \varepsilon\kappa(u)) - \chi(\mathsf{q}(u))}{\varepsilon} \mathrm{d}u \geqslant 0.$$

Since χ is Lipschitz, according to Rademacher's theorem, χ is differentiable almost everywhere, so for almost every $u \in [0,1)$, we have

$$\lim_{\varepsilon \to 0} \frac{\chi(\mathsf{q}(u) + \varepsilon\kappa(u)) - \chi(\mathsf{q}(u))}{\varepsilon} = \nabla\chi(\mathsf{q}(u))\kappa(u).$$

In addition,

$$\left| \frac{\chi(\mathsf{q}(u) + \varepsilon\kappa(u)) - \chi(\mathsf{q}(u))}{\varepsilon} \right| \leqslant |\kappa(u)|.$$

Therefore, by dominated convergence as $\varepsilon \to 0$, we have

$$\langle \nabla\chi \circ \mathsf{q}, \kappa \rangle_{L^2} \geqslant 0.$$

By density, the previous display holds for every $\kappa \in \mathcal{Q}_2(\mathbb{R}_+)^*$ and thus $\nabla\chi \circ \mathsf{q} \in \mathcal{Q}(\mathbb{R}_+)$. Since q is surjective, the function $\nabla\chi$ coincides almost everywhere with a nondecreasing and $\mathbb{R}_+$-valued function. In addition, since χ is Lipschitz, it is absolutely continuous, and we have for every $x < y$,

$$\frac{\chi(y) - \chi(x)}{y - x} = \frac{1}{y - x} \int_x^y \nabla\chi(z) \mathrm{d}z.$$

The right-hand side in the previous display is nonnegative and given $a < x < b$, the mean value of $\nabla\chi$ on $[a,x]$ is smaller than the mean value of $\nabla\chi$ on $[x,b]$. Thus, χ satisfies

$$0 \leqslant \frac{\chi(x) - \chi(a)}{x - a} \leqslant \frac{\chi(b) - \chi(x)}{b - x}.$$

This means that χ is convex and nondecreasing. This proves,

$$\psi(\mu) = \inf_{\chi \in \mathcal{X}^0} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\}.$$

In the previous display, the left and right-hand side are Lipschitz continuous and the equality holds for μ in a dense subset of $\mathcal{P}_1(\mathbb{R}_+)$. Therefore, by density, the previous display holds for every $\mu \in \mathcal{P}_1(\mathbb{R}_+)$. □

4.2. **Viscosity solution with linear initial condition.** The goal of this section is to show Theorem 4.2 below. The reader only interested in uninverted formulas for $\lim_{N\to+\infty} \overline{F}_N(t,\delta_0)$ but not for $\lim_{N\to+\infty} \overline{F}_N(t,\mu)$ can skip this subsection and directly go to Subsection 4.3, replacing the content of Theorem 4.2 by the following elementary formula,

$$\sup_{\mu \in \mathcal{P}_\infty(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\mu - t \int \xi^* \left(\frac{\cdot}{t} \right) \mathrm{d}\mu \right\} = \sup_{x \in \mathbb{R}_+} \left\{ \chi(x) - t\xi^* \left(\frac{x}{t} \right) \right\}. \tag{4.3}$$

Let $(\mathcal{H}, \langle \cdot,\cdot \rangle_{\mathcal{H}})$ be a Hilbert space. Given $\mathcal{C} \subseteq \mathcal{H}$, we say that $\mathcal{C}$ is a closed convex cone when $\mathcal{C}$ is a closed set in $\mathcal{H}$ and for every $x, x' \in \mathcal{C}$ and $t, t' \geqslant 0$, we have

$$tx + t'x' \in \mathcal{C}.$$

Given a closed convex cone $\mathcal{C}$ in $\mathcal{H}$, we define its dual cone by

$$\mathcal{C}^* = \{ y \in \mathcal{H} \,|\, \forall x \in \mathcal{C},\ \langle x, y \rangle_{\mathcal{H}} \geqslant 0 \}. \tag{4.4}$$

For $x, x' \in \mathcal{H}$, we say that $x \leqslant x'$ when $x' - x \in \mathcal{C}^*$ and we say that $\mathsf{g} : \mathcal{C} \to \mathbb{R}$ is $\mathcal{C}^*$-nondecreasing whenever for every $x, x' \in \mathcal{C}$, we have

$$x \leqslant x' \implies \mathsf{g}(x) \leqslant \mathsf{g}(x').$$

We define $\mathcal{V}(\mathcal{C})$ as the set of functions $V : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}$ such that, for some $L \geqslant 0$ the following holds. For every $t \geqslant 0$, $V(t,\cdot)$ is $\mathcal{C}^*$-nondecreasing and L-Lipschitz, and

$$\sup_{\substack{t>0 \\ x \in \mathcal{C}}} \left\{ \frac{V(t,x) - V(0,x)}{t} \right\} < +\infty.$$

Recall the notion of Fréchet derivative ∇ from Definition 2.3, also recall that given a Fréchet differentiable function $\mathsf{g} : \mathcal{Q}_2(\mathbb{R}_+) \to \mathbb{R}$, we have for every $\mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)$, $\nabla \mathsf{g}(\mathsf{q}) \in L^2([0,1),\mathbb{R})$. In particular expressions of the form $\int h(\nabla \mathsf{g}(\mathsf{q}))$ should be understood as $\int_0^1 h(\nabla \mathsf{g}(\mathsf{q})(u))\mathrm{d}u$. The notion of viscosity solution for (4.5) appearing in Theorem 4.2 below is introduced in details in [6, Definition 1.4].

Theorem 4.2. *Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ a Lipschitz, nondecreasing and convex function. The function $(t,\mathsf{q}) \mapsto \int S_t\chi(\mathsf{q})$ belongs to $\mathcal{V}(\mathcal{Q}_2(\mathbb{R}_+))$ and is the unique viscosity solution of*

$$\begin{cases} \partial_t V - \int \xi(\nabla V) = 0 \text{ on } (0,+\infty) \times \mathcal{Q}_2(\mathbb{R}_+) \\ V(0,\mathsf{q}) = \int \chi(\mathsf{q}) \text{ on } \mathcal{Q}_2(\mathbb{R}_+). \end{cases} \tag{4.5}$$

To prove Theorem 4.2, we are going to use a notion of differentiability that is weaker than Fréchet differentiability.

Definition 4.3. Let $\mathsf{g} : \mathcal{Q}_2(S^D_+) \to \mathbb{R}$ and $\mathsf{q} \in \mathcal{Q}_2(S^D_+)$. We say that g is Gateaux differentiable at q if the following conditions hold.

(1) For every $\kappa \in L^2([0,1), S^D)$ such that $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(S^D_+)$ for $\varepsilon > 0$ small enough, the following limit exists and is finite,

$$\mathsf{g}'(\mathsf{q}, \kappa) = \lim_{\varepsilon \to 0^+} \frac{\mathsf{g}(\mathsf{q} + \varepsilon\kappa) - \mathsf{g}(\mathsf{q})}{\varepsilon}.$$

(2) There exists a unique $\mathsf{p} \in L^2([0,1), S^D)$ such that for every $\kappa \in L^2([0,1), S^D)$ such that $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(S^D_+)$ for $\varepsilon > 0$ small enough, we have

$$\mathsf{g}'(\mathsf{q}, \kappa) = \langle \mathsf{p}, \kappa \rangle_{L^2}.$$

In this case, we say that p is the Gateaux derivative of g at q and we denote it $\nabla \mathsf{g}(\mathsf{q})$.

Since a Fréchet differentiable function is also Gateaux differentiable and both derivatives are equal, there is no harm in using the symbol ∇ to denote both the Gateaux and the Fréchet derivative. The Gateaux derivative allows us to characterize differentiable nondecreasing functions, roughly speaking a function $\mathsf{g} : \mathcal{Q}_2(S^D_+) \to \mathbb{R}$ is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing if and only if for every $\mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)$, we have $\nabla \mathsf{g}(\mathsf{q}) \in \mathcal{Q}_2(\mathbb{R}_+)$. In practice, this characterization is not exactly true at points in $\mathcal{Q}_2(\mathbb{R}_+) - \mathcal{Q}_\uparrow(\mathbb{R}_+)$ since the set of admissible directions at those points may not be rich enough.

Proposition 4.4. *Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be a Lipschitz function. The function $X : \mathsf{q} \mapsto \int_0^1 \chi(\mathsf{q}(u)) \mathrm{d}u$ is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing if and only if χ is nondecreasing and convex.*

Proof. We start by showing the desired equivalence under the additional assumption that χ is differentiable, and then we proceed by approximation.
Step 1. We assume that χ is differentiable on $\mathbb{R}_+$ and we show that X is Gateaux differentiable on $\mathcal{Q}_2(\mathbb{R}_+)$ and satisfies

$$\nabla X(\mathsf{q}) = \nabla \chi \circ \mathsf{q}.$$

Fix $\mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)$, let $\kappa \in L^2$ such that for $\varepsilon > 0$ small enough $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(\mathbb{R}_+)$. For every $u \in [0,1)$,

$$\left| \frac{\chi(\mathsf{q}(u) + \varepsilon\kappa(u)) - \chi(\mathsf{q}(u))}{\varepsilon} \right| \leqslant |\chi|_{\mathrm{Lip}} \kappa(u).$$

and we have as $\varepsilon \to 0$,

$$\frac{\chi(\mathsf{q}(u) + \varepsilon\kappa(u)) - \chi(\mathsf{q}(u))}{\varepsilon} \to \nabla \chi(\mathsf{q}(u))\kappa(u).$$

By dominated convergence, it follows that

$$X(\mathsf{q} + \varepsilon\kappa) = X(\mathsf{q}) + \varepsilon \int_0^1 \nabla \chi(\mathsf{q}(u))\kappa(u) \mathrm{d}u + o(\varepsilon).$$

So X is Gateaux differentiable at q and $\nabla X(\mathsf{q}) = \nabla \chi \circ \mathsf{q} \in L^\infty([0,1), \mathbb{R})$.

Step 2. We assume that χ is differentiable on $\mathbb{R}_+$ and we show that if χ is nondecreasing and convex on $\mathbb{R}_+$, then X is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing on $\mathcal{Q}_2(\mathbb{R}_+)$.

If χ is nondecreasing and convex, then $\nabla\chi : \mathbb{R} \to \mathbb{R}_+$ is nondecreasing and $\mathbb{R}_+$-valued. In particular, for every $\mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)$, we have

$$\nabla X(\mathsf{q}) = \nabla\chi \circ \mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+).$$

Thus, for every $\mathsf{q}, \mathsf{q}' \in \mathcal{Q}_2(\mathbb{R}_+)$ such that $\mathsf{q}' - \mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)^*$, we have

$$X(\mathsf{q}') - X(\mathsf{q}) = \int_0^1 \langle \nabla X(\lambda\mathsf{q}' + (1-\lambda)\mathsf{q}), \mathsf{q}' - \mathsf{q}\rangle_{L^2} \mathrm{d}\lambda \geqslant 0.$$

Step 3. We assume that χ is differentiable on $\mathbb{R}_+$ and we show that if X is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing on $\mathcal{Q}_2(\mathbb{R}_+)$, then χ is nondecreasing and convex on $\mathbb{R}_+$.

Let $\mathsf{q} \in \mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^2$, for every smooth function $\kappa \in \mathcal{Q}_2(\mathbb{R}_+)^*$ and every $\varepsilon > 0$ small enough we have $\mathsf{q} + \varepsilon\kappa \in \mathcal{Q}_2(\mathbb{R}_+)$ and

$$\frac{X(\mathsf{q} + \varepsilon\kappa) - X(\mathsf{q})}{\varepsilon} \geqslant 0.$$

Letting $\varepsilon \to 0$, we obtain

$$\langle \nabla\chi \circ \mathsf{q}, \kappa\rangle_{L^2} \geqslant 0.$$

By density, this last display is in fact valid for all $\kappa \in \mathcal{Q}_2^*$. Therefore $\nabla\chi \circ \mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)$. This ensures that $\nabla\chi$ is $\mathbb{R}_+$-valued and nondecreasing, therefore χ is convex and nondecreasing.

Step 4. Approximation by differentiable functions.

We let $\eta \in \mathcal{C}^\infty(\mathbb{R}_+, \mathbb{R})$ be a smooth function supported on $[1, 2]$ that takes nonnegative values and such that

$$\int_{\mathbb{R}_+} \eta(x)\mathrm{d}x = 1.$$

For every $\varepsilon > 0$, let $\eta_\varepsilon(x) = \frac{1}{\varepsilon}\eta\left(\frac{x}{\varepsilon}\right)$. We define,

$$\chi_\varepsilon(x) = \int_{\mathbb{R}_+} \chi(x + y)\eta_\varepsilon(y)\mathrm{d}y,$$

For every $x \in \mathbb{R}_+$, we have

$$\begin{aligned}
|\chi_\varepsilon(x) - \chi(x)| &= \left| \int_{\mathbb{R}_+} \chi(x+y)\eta_\varepsilon(y)\mathrm{d}y - \int_{\mathbb{R}_+} \chi(x)\eta_\varepsilon(y)\mathrm{d}y \right| \\
&\leqslant \int_{\mathbb{R}_+} |\chi(x+y) - \chi(x)|\eta_\varepsilon(y)\mathrm{d}y \\
&\leqslant \int_{\mathbb{R}_+} |\chi|_{\mathrm{Lip}} y \eta_\varepsilon(y)\mathrm{d}y \\
&= \varepsilon|\chi|_{\mathrm{Lip}} \int_{\mathbb{R}_+} y\eta(y)\mathrm{d}y.
\end{aligned}$$

Therefore as $\varepsilon \to 0$, $\chi_\varepsilon \to \chi$ uniformly on $\mathbb{R}_+$. In addition, for every $x \in \mathbb{R}_+$ and $h \in (-\varepsilon, \varepsilon)$,

$$\frac{\chi_\varepsilon(x+h) - \chi_\varepsilon(x)}{h} = \int_x^{x+2\varepsilon} \chi(z) \frac{\eta_\varepsilon(z-x-h) - \eta_\varepsilon(z-x)}{h} \mathrm{d}z.$$

We have, uniformly over $z \in [x, x+2\varepsilon]$,

$$\lim_{h \to 0} \frac{\eta_\varepsilon(z-x-h) - \eta_\varepsilon(z-x)}{h} = \nabla \eta_\varepsilon(z-x).$$

Therefore, χ_ε is differentiable at x and

$$\nabla \chi_\varepsilon(x) = \int_x^{x+2\varepsilon} \chi(z) \nabla \eta_\varepsilon(z-x) \mathrm{d}z.$$

Thus, the sequence $(\chi_\varepsilon)_\varepsilon$ is a sequence of Lipschitz differentiable functions on $\mathbb{R}_+$ that converge uniformly on $\mathbb{R}_+$ towards χ.

Step 5. Conclusion.

Let $(\chi_\varepsilon)_\varepsilon$ be the sequence built in Step 4. Set

$$X_\varepsilon(\mathsf{q}) = \int_0^1 \chi_\varepsilon(\mathsf{q}(u)) \mathrm{d}u.$$

We have,

$$X_\varepsilon(\mathsf{q}) = \int_{\mathbb{R}_+} X(\mathsf{q}+y) \eta_\varepsilon(y) \mathrm{d}y,$$

where $\mathsf{q}+y$ denotes the path obtained by adding the constant $y \in \mathbb{R}_+$ to the values of q. According to Step 4, χ_ε is differentiable on $\mathbb{R}_+$.

If χ is nondecreasing and convex, then so is χ_ε. According to Step 2, in this case X_ε is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing and letting $\varepsilon \to 0$, we discover that X is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing.

Conversely, if X is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing, then so is X_ε, using Step 3 we discover that χ_ε is convex and nondecreasing, and letting $\varepsilon \to 0$ we obtain that χ is convex and nondecreasing. □

Proposition 4.5. *Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be a Lipschitz, nondecreasing and convex function. For every $t \geqslant 0$, the function*

$$(t, \mathsf{q}) \mapsto \int_0^1 S_t \chi(\mathsf{q}(u)) \mathrm{d}u$$

belongs to $\mathcal{V}(\mathcal{Q}_2(\mathbb{R}_+))$.

Proof. Since χ is convex, $S_t\chi$ admits the Hopf representation [5, Proposition 6.3] that is, for every $x \in \mathbb{R}_+$, we have,

$$S_t\chi(x) = \sup_{y \in \mathbb{R}_+} \left\{ xy - \chi^*(y) + t\xi(y) \right\}.$$

In particular, $S_t\chi$ is the supremum of a family of affine functions of x, so $S_t\chi$ is convex on $\mathbb{R}_+$. By definition $S_t\chi \in \mathcal{V}(\mathbb{R}_+)$, so $S_t\chi$ is Lipschitz and nondecreasing. It then follows from Proposition 4.4 that $\mathsf{q} \mapsto \int_0^1 S_t\chi(\mathsf{q}(u))\mathrm{d}u$ is $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing and thus $(t, \mathsf{q}) \mapsto \int_0^1 S_t\chi(\mathsf{q}(u))\mathrm{d}u$ belongs to $\mathcal{V}(\mathcal{Q}_2(\mathbb{R}_+))$. □

For every $j \geqslant 1$, define

$$\mathcal{Q}^j(\mathbb{R}_+) = \left\{ (x_i)_{1 \leqslant i \leqslant j} \in \mathbb{R}_+^j \,\middle|\, x_i - x_{i-1} \in \mathbb{R}_+ \right\}. \tag{4.6}$$

The cone $\mathcal{Q}^j(\mathbb{R}_+)$ is the set of nondecreasing sequences of $\mathbb{R}_+$ with j terms. It is embedded in $\mathbb{R}^j$ and we equip it with the normalized Euclidean scalar product,

$$\langle x, y \rangle_j = \frac{1}{j} \sum_{i=1}^{j} x_i y_i.$$

One can think of $\mathcal{Q}^j(\mathbb{R}_+)$ as a finite dimensional approximation of $\mathcal{Q}_1(\mathbb{R}_+)$. Given $x \in \mathcal{Q}^j(\mathbb{R}_+)$, we define a path $l_j x \in \mathcal{Q}_1(\mathbb{R}_+)$ by setting

$$l_j x = \sum_{i=1}^{j} x_i \mathbf{1}_{\left[\frac{i-1}{j}, \frac{i}{j}\right)}.$$

Conversely, given a path $\mathsf{q} \in \mathcal{Q}_1(\mathbb{R}_+)$, we define $p_j \mathsf{q} \in \mathcal{Q}^j(\mathbb{R}_+)$ by setting

$$(p_j \mathsf{q})_i = j \int_{\frac{i-1}{j}}^{\frac{i}{j}} \mathsf{q}(u) \mathrm{d}u.$$

The linear maps l_j and p_j are adjoint in the following sense.

$$\langle \mathsf{q}, l_j x \rangle_{L^2} = \langle p_j \mathsf{q}, x \rangle_j .$$

It follows that for every $\mathcal{Q}_2(\mathbb{R}_+)^*$-nondecreasing function $\varphi : \mathcal{Q}_2(\mathbb{R}_+) \to \mathbb{R}$, the function $\varphi \circ l_j$ is $(\mathcal{Q}^j(\mathbb{R}_+))^*$-nondecreasing on $\mathcal{Q}^j(\mathbb{R}_+)$. We also define for every $x \in \mathbb{R}^j$,

$$H_j(x) = \frac{1}{j} \sum_{i=1}^{j} \xi(j x_i),$$

the function $H_j : \mathbb{R}^j \to \mathbb{R}$ is locally Lipschitz, $\mathbb{R}_+^j$-nondecreasing on $\mathbb{R}_+^j$ and $(\mathcal{Q}^j(\mathbb{R}_+))^*$-nondecreasing on $\mathcal{Q}^j(\mathbb{R}_+)$. Given $X \subseteq \mathbb{R}^j$, we let $\mathrm{int}(X)$ denote the interior of X.

Proposition 4.6. *Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be a Lipschitz, nondecreasing and convex function. The function $(t, x) \mapsto \frac{1}{j} \sum_{i=1}^{j} S_t \chi(x_i)$ belongs to $\mathcal{V}(\mathbb{R}_+^j)$ and is the unique viscosity solution of*

$$\begin{cases} \partial_t \overline{v}_j - H_j(\nabla \overline{v}_j) = 0 \ \text{on} \ (0, +\infty) \times \mathrm{int}(\mathbb{R}_+^j) \\ \overline{v}_j(0, x) = \frac{1}{j} \sum_{i=1}^{j} \chi(x_i) \ \text{on} \ \mathbb{R}_+^j. \end{cases} \tag{4.7}$$

Proof. The initial condition $\varphi_j : x \mapsto \frac{1}{j}\sum_{i=1}^j \chi(x_i)$ is convex on $\mathbb{R}_+^j$. We have,

$$\begin{aligned}\varphi_j^*(y) &= \sup_{x\in\mathbb{R}_+^j}\left\{\langle x,y\rangle_j - \frac{1}{j}\sum_{i=1}^j \chi(x_i)\right\}\\ &= \sup_{x_1\in\mathbb{R}_+}\dots\sup_{x_j\in\mathbb{R}_+}\left\{\frac{1}{j}\sum_{i=1}^j x_iy_i - \chi(x_i)\right\}\\ &= \frac{1}{j}\sum_{i=1}^j \sup_{x_i\in\mathbb{R}_+}\{x_iy_i - \chi(x_i)\}\\ &= \frac{1}{j}\sum_{i=1}^j \chi^*(y_i).\end{aligned}$$

According to [5, Proposition 6.3], the unique viscosity solution $\overline{v}_j$ of (4.7) admits the Hopf representation,

$$\overline{v}_j(t,x) = \sup_{y\in\mathbb{R}_+^j}\left\{\langle x,y\rangle_j - \varphi_j^*(y) + tH_j(y)\right\}.$$

Thus,

$$\begin{aligned}\overline{v}_j(t,x) &= \sup_{y\in\mathbb{R}_+^j}\left\{\frac{1}{j}\sum_{i=1}^j (x_iy_i - \chi^*(x_i) + t\xi(y_i))\right\}\\ &= \sup_{y_1\in\mathbb{R}_+}\dots\sup_{y_j\in\mathbb{R}_+}\left\{\frac{1}{j}\sum_{i=1}^j (x_iy_i - \chi^*(x_i) + t\xi(y_i))\right\}\\ &= \frac{1}{j}\sum_{i=1}^j \sup_{y_i\in\mathbb{R}_+}\{x_iy_i - \chi^*(x_i) + t\xi(y_i)\}\\ &= \frac{1}{j}\sum_{i=1}^j S_t\chi(x_i).\end{aligned}$$

By definition, $\overline{v}_j \in \mathcal{V}(\mathbb{R}_+^j)$, this concludes the proof. □

Proposition 4.7. *Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be a Lipschitz, nondecreasing and convex function. The function $(t,x) \mapsto \frac{1}{j}\sum_{i=1}^j S_t\chi(x_i)$ belongs to $\mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$ and is the unique viscosity solution of*

$$\begin{cases}\partial_t v_j - H_j(\nabla v_j) = 0 \text{ on } (0,+\infty)\times \mathrm{int}(\mathcal{Q}^j(\mathbb{R}_+))\\ u_j(0,x) = \frac{1}{j}\sum_{i=1}^j \chi(x_i) \text{ on } \mathcal{Q}^j(\mathbb{R}_+).\end{cases} \tag{4.8}$$

Proof. According to [5, Theorem 1.2], (4.8) admits a unique viscosity solution in $\mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$. Therefore, according to Proposition 4.6, it is enough to check that $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)}$ is a viscosity solution of (4.8) and belongs to $\mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$. Let $(t,x)\in\mathbb{R}_+\times\mathrm{int}(\mathcal{Q}^j(\mathbb{R}_+))$ and $\phi : (0,\infty)\times\mathrm{int}(\mathcal{Q}^j(\mathbb{R}_+))\to\mathbb{R}$ be a smooth function. Assume that $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)} - \phi$ has a local maximum at (t,x). Up to modifying ϕ outside a ball of small radius around (t,x), we may assume that ϕ is the restriction to $(0,\infty)\times\mathrm{int}(\mathcal{Q}^j(\mathbb{R}_+))$ of a smooth function $\overline{\phi}:(0,\infty)\times\mathbb{R}^j\to\mathbb{R}$. By construction $\overline{v}_j-\overline{\phi}$ coincide with $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)}-\phi$ on a neighborhood of (t,x) and thus has a local maximum at (t,x). Since $\overline{v}_j$ is the viscosity solution of (4.7), we have,

$$\left(\partial_t\overline{\phi}-H_j(\nabla\overline{\phi})\right)(t,x)\leqslant 0.$$

Since ϕ and $\overline{\phi}$ coincide on a neighborhood of (t,x), the previous display remains true if we replace $\overline{\phi}$ by ϕ. We reach similar conclusions if we assume that $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)}-\phi$ has local minimum at (t,x). This proves that $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)}$ is a viscosity solution of (4.8).

In addition, according to Proposition 4.6, we have for every $(t,x)\in\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)$,

$$\overline{v}_j(t,x)=\frac{1}{j}\sum_{i=1}^{j}S_t\chi(x_i)=\int S_t\chi(l_jx).$$

Therefore, according to Proposition 4.5, we have $\overline{v}_j|_{\mathbb{R}_+\times\mathcal{Q}^j(\mathbb{R}_+)}\in\mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$ which concludes. □

Proof of Theorem 4.2. According to Proposition 4.5, $(t,\mathsf{q})\mapsto\int S_t\chi(\mathsf{q})$ belongs to $\mathcal{V}(\mathcal{Q}_2(\mathbb{R}_+))$. Let $(t,\mathsf{q})\in\mathbb{R}_+\times\mathcal{Q}_2(\mathbb{R}_+)$, according to [6, Theorem 4.6 (1)], the value of the unique viscosity solution of (4.5) at (t,q) is the limit of $v_j(t,p_j\mathsf{q})$ as $j\to+\infty$, where v_j is the unique viscosity solution of (4.8). According to Proposition 4.7, we have

$$v_j(t,p_j\mathsf{q})=\int S_t\chi(l_jp_j\mathsf{q}).$$

Since $x\mapsto S_t\chi(x)$ is Lipschitz, we have

$$\left|\int S_t\chi(l_jp_j\mathsf{q})-\int S_t\chi(\mathsf{q})\right|\leqslant c|l_jp_j\mathsf{q}-\mathsf{q}|_{L^1}.$$

Finally, according to [6, Lemma 3.3 (7)] we have $\lim_{j\to+\infty}|l_jp_j\mathsf{q}-\mathsf{q}|_{L^1}=0$. This concludes the proof. □

4.3. **Proof of the un-inverted Parisi formula.**

Proof of Theorem 1.2. Let $(t,\mu)\in\mathbb{R}_+\times\mathcal{P}_\infty(\mathbb{R}_+)$, and

$$f(t,\mu)=\lim_{N\to+\infty}F_N(t,\mu).$$

According to Theorem 2.1 we have

$$f(t,\mu)=\sup_{\nu\in\mathcal{P}_\infty(\mathbb{R}_+)}\left\{\psi(\nu)-t\mathbb{E}\xi^*\left(\frac{X_\nu-X_\mu}{t}\right)\right\}.$$

We let $\Pi(\mu,\nu)$ denote the set of probability measures $\pi\in\mathcal{P}_\infty(\mathbb{R}_+\times\mathbb{R}_+)$ such that $\pi_1=\mu$ and $\pi_2=\nu$. According to [13, Proposition 2.5], we have

$$\mathbb{E}\xi^*\left(\frac{X_\nu-X_\mu}{t}\right)=\inf_{\pi\in\Pi(\nu,\mu)}\int\xi^*\left(\frac{y-x}{t}\right)\mathrm{d}\pi(x,y).\tag{4.9}$$

Define

$$\Pi(\mu,\cdot) = \bigcup_{\nu \in \mathcal{P}_\infty(\mathbb{R}_+)} \Pi(\mu,\nu).$$

It follows respectively from [4, Proposition 3.6], [4, Corollary 5.2] and [2], that ψ is nondecreasing, 1-Lipschitz and concave. Therefore, according to Lemma 4.1 we have

$$\begin{aligned} f(t,\mu) &= \sup_{\pi \in \Pi(\mu,\cdot)} \left\{ \psi(\pi_2) - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y) \right\} \\ &= \sup_{\pi \in \Pi(\mu,\cdot)} \inf_{\chi \in \mathcal{X}^0} \left\{ \int \chi \mathrm{d}\pi_2 - \psi_*(\chi) - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y) \right\} \\ &= \sup_{\pi \in \Pi(\mu,\cdot)} \inf_{\chi \in \mathcal{X}^0} \left\{ \int \chi \mathrm{d}\pi_2 - \psi_*(\chi) - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y) \right\}. \end{aligned}$$

The sets $\Pi(\mu,\cdot)$ and $\mathcal{X}^0$ are convex. In addition, according to the Arzelà-Ascoli theorem, $\mathcal{X}^0$ is compact with respect to the topology of local uniform convergence. For every $\pi \in \Pi(\mu,\cdot)$, the map $\chi \mapsto \int \chi \mathrm{d}\pi_2 - \psi_*(\chi)$ is lower semi-continuous on $\mathcal{X}^0$ with respect to the topology of local uniform convergence. Similarly since ξ^* is lower semi-continuous, according to Portmanteau's theorem for every $\chi \in \mathcal{X}^0$, the map $\pi \mapsto \int \chi \mathrm{d}\pi_2 - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y)$ is upper semi-continuous with respect to convergence under the optimal transport distance. Thus, according to [20, Corollary 3.4], we can interchange sup and inf in the previous display to obtain

$$\begin{aligned} f(t,\mu) &= \inf_{\chi \in \mathcal{X}^0} \left\{ -\psi_*(\chi) + \sup_{\pi \in \Pi(\mu,\cdot)} \left\{ \int \chi \mathrm{d}\pi_2 - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y) \right\} \right\} \\ &= \inf_{\chi \in \mathcal{X}^0} \left\{ -\psi_*(\chi) + \sup_{\pi \in \Pi(\mu,\cdot)} \left\{ \int \chi \mathrm{d}\pi_2 - t \int \xi^* \left(\frac{y-x}{t} \right) \mathrm{d}\pi(x,y) \right\} \right\} \\ &= \inf_{\chi \in \mathcal{X}^0} \left\{ -\psi_*(\chi) + \sup_{\nu \in \mathcal{P}_\infty(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\nu - t \mathbb{E} \xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right\} \right\} \end{aligned}$$

According to [6, Theorem 4.6 (2) and Proposition A.3], the viscosity solution of (4.5) has the Hopf-Lax representation and it follows from Theorem 4.2 that

$$\sup_{\nu \in \mathcal{P}_\infty(\mathbb{R}_+)} \left\{ \int \chi \mathrm{d}\nu - t \mathbb{E} \xi^* \left(\frac{X_\nu - X_\mu}{t} \right) \right\} = \int S_t \chi \mathrm{d}\mu.$$

Thus, from the two previous displays or simply from (4.3) if $\mu = \delta_0$, we have

$$f(t,\mu) = \inf_{\chi \in \mathcal{X}^0} \left\{ -\psi_*(\chi) + \int S_t \chi \mathrm{d}\mu \right\}.$$

Since $f(t,\cdot)$ is Lipschitz continuous on $\mathcal{P}_1(\mathbb{R}_+)$ and $\mathcal{P}_\infty(\mathbb{R}_+)$ is dense in $\mathcal{P}_1(\mathbb{R}_+)$, by density we can extend the equality of the last display to any $\mu \in \mathcal{P}_1(\mathbb{R}_+)$. □

5. Extension of concave functions

As discussed above, the main difficulty to generalize the results obtained in Section 4 to the case $D > 1$ is the fact that the set $\mathcal{P}^{\uparrow}(S^D_+)$ is nonconvex when $D > 1$. Indeed, this prevents us from performing the sup-inf interchange using [20] as in the proof of Theorem 1.2. To circumvent this difficulty, we show that the initial condition $\psi : \mathcal{P}^{\uparrow}(S^D_+) \to \mathbb{R}$ is the restriction of a concave and Lipschitz function defined on $\mathcal{P}(S^D_+)$. Using this extension, we show that at $\mu = \delta_0$ the supremum over $\mathcal{P}^{\uparrow}(S^D_+)$ in (2.9) can be rewritten as a supremum over $\mathcal{P}(S^D_+)$. Since the set $\mathcal{P}(S^D_+)$ is convex, we can then perform the sup-inf interchange and proceed as in the proof of Theorem 1.2. To do this properly, we will need to use some compacity properties, so in what follows S^D_+ is going to be replaced by a compact subset $\mathcal{K} \subseteq S^D_+$.

5.1. **Optimal transport of monotone probability measures.** Let $\mathcal{K} \subseteq S^D_+$ be a compact set and $\mathcal{P}(\mathcal{K})$ denote the set of Borel probability measures on $\mathcal{K}$. In this subsection, we study functions on possibly nonconvex subsets K of $\mathcal{P}(\mathcal{K})$. We prove that if such a function is concave on every convex subset of K, then it satisfies a Jensen-type inequality, provided that K is an extreme set in its convex hull (see Definition 5.5).

As previously, $\mathcal{L}$ denotes the set of Lipschitz functions $\chi : S^D_+ \to \mathbb{R}$, $\mathcal{L}_{\leqslant 1}$ denote the set of 1-Lipschitz functions $\chi : S^D_+ \to \mathbb{R}$, $\mathcal{L}^0$ denotes the set of functions $\chi \in \mathcal{L}$ satisfying $\chi(0) = 0$ and $\mathcal{L}^0_{\leqslant 1} = \mathcal{L}^0 \cap \mathcal{L}_{\leqslant 1}$. We equip $\mathcal{P}(\mathcal{K})$ with the optimal transport distance, for every $\mu, \nu \in \mathcal{P}(\mathcal{K})$,

$$d(\mu,\nu) = \inf_{\pi \in \Pi(\mu,\nu)} \left\{ \int_{\mathcal{K}\times\mathcal{K}} |x-y| \mathrm{d}\pi(x,y) \right\},$$

where $\Pi(\mu,\nu)$ denotes the set of probability measures $\pi \in \mathcal{P}(\mathcal{K}\times\mathcal{K})$ such that $\pi_1 = \mu$ and $\pi_2 = \nu$. The distance d also admits the following dual representation [22, Theorem 5.10],

$$d(\mu,\nu) = \sup_{\chi \in \mathcal{L}_{\leqslant 1}} \left\{ \int_{S^D_+} \chi \mathrm{d}(\mu-\nu) \right\}. \tag{5.1}$$

In what follows, we will consider a compact set $K \subseteq \mathcal{P}(\mathcal{K})$. The set $\mathcal{P}(K)$ of probability measures on K will appear, we will equip this set with the optimal transport distance d derived from the optimal transport distance d on K, that is for every $\eta, \eta' \in \mathcal{P}(K)$,

$$\mathsf{d}(\eta,\eta') = \inf_{\pi \in \Pi(\eta',\eta')} \left\{ \int_{K\times K} d(\nu,\nu') \mathrm{d}\pi(\nu,\nu') \right\},$$

where $\Pi(\eta,\eta')$ denotes the set of probability measures $\pi \in \mathcal{P}(K\times K)$ such that $\pi_1 = \eta$ and $\pi_2 = \eta'$. As usual, the optimal transport distance d admits the dual representation [22, Theorem 5.10],

$$\mathsf{d}(\eta,\eta') = \sup_X \left\{ \int X(\nu)\mathrm{d}\eta(\nu) - \int X(\nu)\mathrm{d}\eta'(\nu) \right\}, \tag{5.2}$$

where the supremum is taken over the set of 1-Lipschitz functions $X : K \to \mathbb{R}$.

Proposition 5.1. *Let E be a compact Polish space, we equip the set $\mathcal{P}(E)$ of Borel probability measures on E with the optimal transport distance. Let $\mu \in \mathcal{P}(E)$, let $(X_i)_{i\geqslant 1} : \Omega \to E^{\mathbb{N}}$ be independent and identically distributed random variables with law μ, then almost surely*

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i} \underset{n\to\infty}{\longrightarrow} \mu.$$

Remark 5.2. By Caratheodory's extension theorem, given a Borel probability measure μ on a Polish space E, we can always construct a sequence of independent and identically distributed random variables with law μ.

Proof. Let d denote the optimal transport distance on $\mathcal{P}(E)$. According to [22, Theorem 6.9], since the space E is compact, given a sequence $(\mu_n)_{n\geqslant 1}$ in $\mathcal{P}(E)$, we have $\mu_n \to \mu$ in $(\mathcal{P}(E), d)$ if and only if for every continuous function $h : E \to \mathbb{R}$,

$$\lim_{n\to+\infty}\int h\mathrm{d}\mu_n = \int h\mathrm{d}\mu.$$

Let $\mathcal{C}(E)$ denote the set of real-valued continuous functions on E. Since E is compact and metrizable, $(\mathcal{C}(E), |\cdot|_\infty)$ is separable [7, Theorem 6.6]. Let $(h_p)_p$ be a dense sequence in $(\mathcal{C}(E), |\cdot|_\infty)$, to check convergence in $(\mathcal{P}(E), d)$, it is sufficient to check convergence against h_p for every $p \geqslant 1$. By the law of large numbers, for every $p \geqslant 1$ the following holds almost surely

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}h_p(X_i) = \int_E h_p\mathrm{d}\mu.$$

Let $\widetilde{\mu}_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i} \in \mathcal{P}(E)$, since $(h_p)_p$ is countable, almost surely we have for every $p \geqslant 1$,

$$\lim_{n\to\infty}\int_E h_p\mathrm{d}\widetilde{\mu}_n = \int_E h_p\mathrm{d}\mu.$$

That is, almost surely the sequence $(\widetilde{\mu}_n)_n$ converges to μ in $(\mathcal{P}(E), d)$. □

Given $\eta \in \mathcal{P}(K)$, we define the barycenter $\mathrm{Bar}(\eta) \in \mathcal{P}(\mathcal{K})$ by

$$\mathrm{Bar}(\eta)(A) = \int_K \nu(A)\mathrm{d}\eta(\nu).$$

The barycenter of η is the unique probability measure on $\mathcal{K}$ satisfying for every continuous function $h : \mathcal{K} \to \mathbb{R}$,

$$\int_K\int_{\mathcal{K}} h(x)\mathrm{d}\nu(x)\mathrm{d}\eta(\nu) = \int_{\mathcal{K}} h(x)\mathrm{d}\mathrm{Bar}(\eta)(x).$$

Heuristically, we can think of $\mathrm{Bar}(\eta)$ as the first moment of η.

Remark 5.3. When η is a finitely supported probability measure, say of the form $\sum_{i=1}^{p} c_i\delta_{\nu_i}$, we have $\mathrm{Bar}(\eta) = \sum_{i=1}^{p} c_i\nu_i$.

Proposition 5.4. *The barycenter map* $Bar : (\mathcal{P}(K), \mathsf{d}) \to (\mathcal{P}(\mathcal{K}), d)$ *is Lipschitz continuous and for every* $\eta, \eta' \in \mathcal{P}(K)$ *and* $\lambda \in [0,1]$, *we have*

$$Bar(\lambda\eta + (1-\lambda)\eta') = \lambda Bar(\eta) + (1-\lambda) Bar(\eta').$$

Proof. Let $\chi \in \mathcal{L}_{\leqslant 1}$, the map $\nu \mapsto \int \chi \mathrm{d}\nu$ is 1-Lipschitz on K. Let $\eta, \eta' \in \mathcal{P}(K)$, we have

$$\begin{aligned} d(\mathrm{Bar}(\eta), \mathrm{Bar}(\eta')) &= \sup_{\chi \in \mathcal{L}_{\leqslant 1}} \left\{ \int_{\mathcal{K}} \chi \mathrm{dBar}(\eta) - \int_{\mathcal{K}} \chi \mathrm{dBar}(\eta') \right\} \\ &= \sup_{\chi \in \mathcal{L}_{\leqslant 1}} \left\{ \int_K \int_{\mathcal{K}} \chi \mathrm{d}\nu \mathrm{d}\eta(\nu) - \int_K \int_{\mathcal{K}} \chi \mathrm{d}\nu \mathrm{d}\eta'(\nu) \right\} \\ &\leqslant \sup_X \left\{ \int_K X d\eta - \int_K X \mathrm{d}\eta' \right\} \\ &= \mathsf{d}(\eta, \eta'). \end{aligned}$$

□

Let C denote the closed convex hull of K in $\mathcal{P}(\mathcal{K})$.

Definition 5.5. We say that K is an extreme set in C when for every $\eta \in \mathcal{P}(C)$, if $\mathrm{Bar}(\eta) \in K$ then $\eta \in \mathcal{P}(K)$.

For example, if K is an extreme set in C, then taking $\eta = \lambda\delta_\mu + (1-\lambda)\delta_\nu$, where $\mu, \nu \in C$ and $\lambda \in (0,1)$, we have that

$$\lambda\mu + (1-\lambda)\nu \in K \implies \mu, \nu \in K.$$

Proposition 5.6. *The closed convex hull* C *of* K *satisfies*

$$C = \left\{ Bar(\eta) \middle| \eta \in \mathcal{P}(K) \right\}. \tag{5.3}$$

Proof. According to Proposition 5.4, the set defined in (5.3) is closed and convex. Let $C' \subseteq \mathcal{P}(\mathcal{K})$ be a closed convex set containing K, let $\eta \in \mathcal{P}(K)$. According to Proposition 5.1, there exists a sequence of finitely supported measures $(\eta_n)_n$ such that $\eta_n \to \eta$ as $n \to +\infty$. Since C' is convex, according to Remark 5.3, we have $\mathrm{Bar}(\eta_n) \in C'$. In addition, since C' is closed, letting $n \to \infty$ and using Proposition 5.4 we have $\mathrm{Bar}(\eta) \in C'$. We have proven that $\left\{\mathrm{Bar}(\eta) \middle| \eta \in \mathcal{P}(K)\right\} \subseteq C'$, therefore $\left\{\mathrm{Bar}(\eta) \middle| \eta \in \mathcal{P}(K)\right\}$ is the closed convex hull of K. □

Proposition 5.7. *Assume that* K *is an extreme set in its convex hull. Let* $\eta \in \mathcal{P}(K)$ *be such that* $Bar(\eta) \in K$. *There exists a sequence* $\eta_n \in \mathcal{P}(K)$ *of finitely supported probability measures such that* $Bar(\eta_n) \in K$ *and* $\eta_n \to \eta$ *as* $n \to +\infty$.

Proof. Let $(\nu_i)_{i \geqslant 1}$ be a sequence of independent and identically distributed random measures in K with law η, define

$$\widetilde{\eta}_n = \frac{1}{n} \sum_{i=1}^n \delta_{\nu_i}.$$

According to Proposition 5.1, we have $\widetilde{\eta}_n \to \eta$ almost surely. Let us show that almost surely for every $n \geqslant 1$, $\text{Bar}(\widetilde{\eta}_n) \in K$. Define,

$$\mathcal{O} = \{\omega \in \Omega \big| \, \forall n \geqslant 1, \ \text{Bar}(\widetilde{\eta}_n(\omega)) \in K\}.$$

Let $P_n \in \mathcal{P}(C)$ denote the law of the random probability measure $\text{Bar}(\widetilde{\eta}_n) \in C$. The probability measure P_n is the only probability measure in $\mathcal{P}(C)$ that satisfies for every $h : C \to \mathbb{R}$,

$$\int_C h(\nu) \mathrm{d}P_n(\nu) = \int_{K^n} h\left(\frac{1}{n}\sum_{i=1}^n \nu_i\right) \prod_{i=1}^n \mathrm{d}\eta(\nu_i).$$

For every $A \subseteq C$, we have,

$$\begin{aligned} \text{Bar}(P_n)(A) &= \int_C \nu(A) \mathrm{d}P_n(\nu) \\ &= \int_{K^n} \frac{1}{n}\sum_{i=1}^n \nu_i(A) \prod_{i=1}^n \mathrm{d}\eta(\nu_i) \\ &= \int_K \nu(A) \mathrm{d}\eta(\nu) \\ &= \text{Bar}(\eta)(A). \end{aligned}$$

So $\text{Bar}(P_n) = \text{Bar}(\eta) \in K$, since K is an extreme set in its closed convex hull, we have $P_n \in \mathcal{P}(K)$ and almost surely $\text{Bar}(\widetilde{\eta}_n) \in K$. Thus $\mathbb{P}(\mathcal{O}) = 1$ and for almost all $\omega \in \Omega$, the measures $\eta_n = \widetilde{\eta}_n(\omega)$ satisfy the desired properties. □

Proposition 5.8. *Assume that K is an extreme set in its closed convex hull, let $\varphi : K \to \mathbb{R}$ be a bounded upper semicontinuous function. Assume that for every $\nu, \nu' \in K$, $\lambda \in [0,1]$ such that $\lambda\nu + (1-\lambda)\nu' \in K$ we have*

$$\lambda\varphi(\nu) + (1-\lambda)\varphi(\nu') \leqslant \varphi\left(\lambda\nu + (1-\lambda)\nu'\right).$$

Then φ satisfies Jensen's inequality, that is for every $\eta \in \mathcal{P}(K)$ such that $Bar(\eta) \in K$, we have

$$\int_K \varphi(\nu) \mathrm{d}\eta(\nu) \leqslant \varphi\left(Bar(\eta)\right). \tag{5.4}$$

Proof. Assume that η is finitely supported, then $\eta = \sum_{i=1}^p c_i \delta_{\nu_i}$ for some $\nu_i \in K$ and $c_i > 0$ such that $\sum_{i=1}^p c_i = 1$ and $\sum_{i=1}^p c_i \nu_i \in K$. Since K is an extreme set in its convex hull C, for every nonempty $I \subseteq \{1, \dots, p\}$, we have

$$\frac{\sum_{i \in I} c_i \nu_i}{\sum_{i \in I} c_i} \in K. \tag{5.5}$$

It follows by induction on $p \geqslant 1$, that

$$\sum_{i=1}^p c_i \varphi(\nu_i) \leqslant \varphi\left(\sum_{i=1}^p c_i \nu_i\right).$$

Therefore, (5.4) holds when η is finitely supported. Assume now that $\eta \in \mathcal{P}(K)$, is any probability measure such that $\text{Bar}(\eta) \in K$. According to Proposition 5.7, the probability measure η can be approximated, by finitely supported probability measures $\eta_n \in \mathcal{P}(K)$ with barycenters in K. Using Proposition 5.4, we can pass to the limit in

$$\int_K \varphi(\nu) d\eta_n(\nu) \leqslant \varphi\left(\mathrm{Bar}(\eta_n)\right),$$

and use the upper semicontinuity of φ to obtain (5.4). □

5.2. **Monotone probability measures on a compact set.** Let $\mathcal{K}$ be a compact subset of S^D_+, define

$$\mathcal{P}^{\uparrow}(\mathcal{K}) = \mathcal{P}^{\uparrow}(S^D_+) \cap \mathcal{P}(\mathcal{K}). \tag{5.6}$$

In this section, we show that $K = \mathcal{P}^{\uparrow}(\mathcal{K})$ is a compact set and that K is extreme in its convex hull. This means that Proposition 5.8 can be applied to functions on $\mathcal{P}^{\uparrow}(\mathcal{K})$.

Recall that we have defined a partial order on S^D by setting $x \leqslant x'$ whenever $x' - x \in S^D_+$. We say that a subset $S \subseteq S^D_+$ is totally ordered when for every $x, x' \in S$ we have $x \leqslant x'$ or $x' \leqslant x$.

Proposition 5.9. *The set $\mathcal{P}^{\uparrow}(\mathcal{K})$ is a closed subset of $(\mathcal{P}(\mathcal{K}), d)$, in particular $\mathcal{P}^{\uparrow}(\mathcal{K})$ is compact with respect to the topology induced by d. In addition, for every $\mu \in \mathcal{P}(\mathcal{K})$, the support of μ is totally ordered if and only if $\mu \in \mathcal{P}^{\uparrow}(\mathcal{K})$.*

Proof. Let $\mu_n \in \mathcal{P}^{\uparrow}(\mathcal{K})$ be a sequence that converges to some $\mu \in \mathcal{P}(\mathcal{K})$. The sequence $(\mathsf{q}_{\mu_n})_n$ is a Cauchy sequence in $L^1([0,1), \mathcal{K})$, let $\mathsf{q} \in L^1([0,1), \mathcal{K})$ denote its limit. There exists a subsequence $(n_k)_k$ such that $\mathsf{q}_{\mu_{n_k}} \to \mathsf{q}$ almost everywhere, so there exists an almost everywhere representative of q in $\mathcal{Q}_1(S^D_+)$, we also denote this representant by q and we have $\mu = \mathrm{Law}(\mathsf{q}(U)) \in \mathcal{P}^{\uparrow}(\mathcal{K})$.

Let $\mu \in \mathcal{P}^{\uparrow}(\mathcal{K})$ then for every $x, y \in \mathrm{supp}(\mu)$ there exists $u, v \in [0,1)$ such that $x = \mathsf{q}_\mu(u)$ and $y = \mathsf{q}_\mu(v)$. We must have $u \leqslant v$ or $v \leqslant u$, and since q_μ is nondecreasing, this implies $x \leqslant y$ or $y \leqslant x$ and $\mathrm{supp}(\mu)$ is totally ordered. Conversely, assume that $\mathrm{supp}(\mu)$ is totally ordered, let $(X_i)_i$ be a sequence of independent and identically distributed random variables with law μ, set

$$\widetilde{\mu}_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}.$$

According to Proposition 5.1, $\widetilde{\mu}_n \to \mu$ almost surely. Let us fix $\omega \in \Omega$ such that $\widetilde{\mu}_n(\omega) \to \mu$ and define $\mu_n = \widetilde{\mu}_n(\omega)$. Since $\mathrm{supp}(\mu)$ is totally ordered, for every $n \geqslant 1$ there exists a permutation $s_n \in \mathcal{S}_n$ such that the sequence $(X_{s_n(i)}(\omega))_{1 \leqslant i \leqslant n}$ is nondecreasing. Let

$$\mathsf{q}_n = \sum_{i=1}^n X_{s_n(i)}(\omega) \mathbf{1}_{[\frac{i-1}{n}, \frac{i}{n})},$$

we have $\mathsf{q}_n \in \mathcal{Q}(S^D_+) \cap L^\infty$ and $\mu_n = \mathrm{Law}(\mathsf{q}_n(U)) \in \mathcal{P}^{\uparrow}(\mathcal{K})$. Since $\mu_n \to \mu$ and $\mathcal{P}^{\uparrow}(\mathcal{K})$ is closed, we have $\mu \in \mathcal{P}^{\uparrow}(\mathcal{K})$. □

Proposition 5.10. *The closed convex hull of* $\mathcal{P}^{\uparrow}(\mathcal{K})$ *is* $\mathcal{P}(\mathcal{K})$ *and* $\mathcal{P}^{\uparrow}(\mathcal{K})$ *is an extreme set in* $\mathcal{P}(\mathcal{K})$.

Proof. According to Proposition 5.6, to show that the closed convex hull of $\mathcal{P}^{\uparrow}(\mathcal{K})$ is $\mathcal{P}(\mathcal{K})$, it suffices to show that for every $\mu \in \mathcal{P}(\mathcal{K})$, there exists $\eta \in \mathcal{P}(\mathcal{P}^{\uparrow}(\mathcal{K}))$ such that $\mu = \mathrm{Bar}(\eta)$.

According to Proposition 5.1, there exists a sequence of finitely supported measures $(\mu_n)_n$ such that $\mu_n \to \mu$. For every $n \geqslant 1$, there exists p_n, c_i^n, x_i^n such that $\mu_n = \sum_{i=1}^{p_n} c_i^n \delta_{x_i^n}$. Define,

$$\eta_n = \sum_{i=1}^{p_n} c_i^n \delta_{\delta_{x_i^n}},$$

we have $\mu_n = \mathrm{Bar}(\eta_n)$. Up to extraction, the sequence $(\eta_n)_n$ converges to some $\eta \in \mathcal{P}(\mathcal{P}^{\uparrow}(\mathcal{K}))$. Therefore, according to Proposition 5.4, we have $\mu = \mathrm{Bar}(\eta)$. This proves that the closed convex hull of $\mathcal{P}^{\uparrow}(\mathcal{K})$ is $\mathcal{P}(\mathcal{K})$.

Let η be a probability measure on $\mathcal{P}(\mathcal{K})$, such that $\mathrm{Bar}(\eta) \in \mathcal{P}^{\uparrow}(\mathcal{K})$. Let A denote the support of $\mathrm{Bar}(\eta)$, we have,

$$0 = \mathrm{Bar}(\eta)(A^c) = \int_{\mathcal{P}(\mathcal{K})} \nu(A^c) d\eta(\nu).$$

So η-almost surely $\nu(A^c) = 0$. The set A is a closed and according to Proposition 5.9, A is totally ordered. In particular, η-almost surely, the support of the measure ν is contained in A, therefore $\nu \in \mathcal{P}^{\uparrow}(\mathcal{K})$. Thus, $\eta \in \mathcal{P}(\mathcal{P}^{\uparrow}(\mathcal{K}))$, which concludes the proof. □

5.3. **Building the extension.** In this section, we assume that the function $\xi : \mathbb{R}^{D\times D} \to \mathbb{R}$ is strictly convex on S^D_+. As usual U denotes a uniform random variable in $[0,1)$, let $L^\infty_{\leqslant 1}$ denote the set of functions $[0,1) \to S^D_+$ that are essentially bounded by 1 in norm and define

$$\mathcal{P}^{\uparrow}_{\xi} = \left\{ \mathrm{Law}(\nabla\xi(\mathsf{q}(U))) \middle| \mathsf{q} \in \mathcal{Q}(S^D_+) \cap L^\infty_{\leqslant 1} \right\} \subseteq \mathcal{P}^{\uparrow}(S^D_+). \tag{5.7}$$

Recall that $B(0,1)$ denotes the unit ball centered at 0 in S^D, we let

$$\mathcal{K}_\xi = \nabla\xi(B(0,1) \cap S^D_+).$$

The goal of this subsection is to prove that any Lipschitz and concave function on the nonconvex set $\mathcal{P}^{\uparrow}_{\xi}$ extends into a Lipschitz and concave function on $\mathcal{M}_1(\mathcal{K}_\xi)$ with the same Lipschitz constant.

Note that we have $\mathcal{P}^{\uparrow}_{\xi} \subseteq \mathcal{P}^{\uparrow}(\mathcal{K}_\xi)$, but in general the inclusion is strict when $D > 1$. The set $\mathcal{P}^{\uparrow}_{\xi}$ is the image of $\mathcal{P}^{\uparrow}(B(0,1) \cap S^D_+)$ by the mapping $\mu \mapsto \nabla\xi_{\#}\mu$. The closed convex envelope and the closed linear span of $\mathcal{P}^{\uparrow}_{\xi}$ are respectively $\mathcal{P}(\mathcal{K}_\xi)$ and $\mathcal{M}_1(\mathcal{K}_\xi)$. Note that since $\mathcal{K}_\xi$ is compact, $\mathcal{M}_1(\mathcal{K}_\xi)$ is in fact equal to the set of signed measures on $\mathcal{K}_\xi$ with finite mass.

Definition 5.11. We say that a function $\varphi : \mathcal{P}_\xi^\uparrow \to \mathbb{R}$ is pre-concave, when for every $\mu, \nu \in \mathcal{P}_\xi^\uparrow$ and $\lambda \in [0,1]$,

$$\lambda\mu + (1-\lambda)\nu \in \mathcal{P}_\xi^\uparrow \implies \lambda\varphi(\mu) + (1-\lambda)\varphi(\nu) \leqslant \varphi(\lambda\mu + (1-\lambda)\nu).$$

As in Section 3, we equip $\mathcal{M}_1(\mathcal{K}_\xi)$ with the norm $|\cdot|_{\mathcal{M}}$ and we denote by d the distance induced by $|\cdot|_{\mathcal{M}}$. The restriction of the distance d to $\mathcal{P}(\mathcal{K}_\xi)$ is the optimal transport distance. Given a continuous function $\varphi : \mathcal{P}_\xi^\uparrow \to \mathbb{R}$, for every $\mu' \in \mathcal{M}_1(\mathcal{K}_\xi)$, we define

$$\overline{\varphi}(\mu') = \sup_{\eta \in \mathcal{P}\left(\mathcal{P}_\xi^\uparrow\right)} \left\{ \int \varphi d\eta - d(\mathrm{Bar}(\eta), \mu') \right\}. \tag{5.8}$$

For every function $\chi \in \mathcal{L}(\mathcal{K}_\xi)$, we also define

$$\varphi_*^\xi(\chi) = \inf_{\mu \in \mathcal{P}_\xi^\uparrow} \left\{ \int \chi d\mu - \varphi(\mu) \right\}. \tag{5.9}$$

We choose $x_0 = \nabla\xi(0)$ as a reference point for the norm $|\cdot|_{\mathcal{L}}$ in $\mathcal{L}(\mathcal{K}_\xi)$. we equip $\mathcal{P}(\mathcal{P}_\xi^\uparrow)$ with the optimal transport distance inherited from the restriction of the distance d to $\mathcal{P}_\xi^\uparrow$. That is for every $\eta, \eta' \in \mathcal{P}(\mathcal{P}_\xi^\uparrow)$,

$$\mathsf{d}(\eta, \eta') = \inf_{\pi \in \Pi(\eta, \eta')} \left\{ \int_{\mathcal{P}_\xi^\uparrow} d(\nu, \nu') \mathrm{d}\pi(\nu, \nu') \right\}. \tag{5.10}$$

Here $\Pi(\eta, \eta')$ denotes the set of measures $\pi \in \mathcal{P}(\mathcal{P}_\xi^\uparrow \times \mathcal{P}_\xi^\uparrow)$ with marginals $\pi_1 = \eta$ and $\pi_2 = \eta'$. As usual, the optimal transport distance d admits the dual representation [22, Theorem 5.10],

$$\mathsf{d}(\eta, \eta') = \sup_X \left\{ \int X(\nu) \mathrm{d}\eta(\nu) - \int X(\nu) \mathrm{d}\eta'(\nu) \right\}, \tag{5.11}$$

where the supremum is taken over the set of 1-Lipschitz functions $X : \mathcal{P}_\xi^\uparrow \to \mathbb{R}$.

The goal of this subsection is to show the following extension theorem, which will be applied to the function ψ defined in (2.7) in the next section.

Theorem 5.12. *For every 1-Lipschitz and pre-concave function $\varphi : \mathcal{P}_\xi^\uparrow \to \mathbb{R}$, the function $\overline{\varphi} : \mathcal{M}_1(\mathcal{K}_\xi) \to \mathbb{R}$ defined by (5.8) is a 1-Lipschitz and concave extension of φ. In addition, for every $\mu' \in \mathcal{M}(\mathcal{K}_\xi)$, we have*

$$\overline{\varphi}(\mu') = \min_{\chi \in \mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)} \left\{ \int \chi \mathrm{d}\mu - \varphi_*^\xi(\chi) \right\}. \tag{5.12}$$

Note that given $S \subseteq S_+^D$ and $\chi : S \to \mathbb{R}$, by letting

$$\overline{\chi}(x) = \inf_{y \in S} \left\{ \chi(y) + |\chi|_{\mathrm{Lip}} |x - y| \right\},$$

we define a Lipschitz extension of χ to S_+^D with the same Lipschitz constant. Therefore, in Theorem 5.12, (5.12) remains true if we take the minimum over $\mathcal{L}_{\leqslant 1}(S_+^D)$ rather than $\mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)$.

Proposition 5.13. *Let $\xi \in \mathcal{C}^\infty(\mathbb{R}^{D\times D}, \mathbb{R})$ be a strictly convex function on S^D_+, then the function $\nabla \xi : S^D_+ \to \mathbb{R}^{D\times D}$ is injective.*

Proof. Let $a \neq b$ in S^D_+, the function $\gamma : t \mapsto \xi(ta + (1-t)b)$ is strictly convex on $[0,1]$, so its derivative is strictly increasing on $(0,1)$ and for every $s < t$ we have

$$\nabla\xi(sa + (1-s)b) \cdot (a-b) < \nabla\xi(ta + (1-t)b) \cdot (a-b). \tag{5.13}$$

Letting $s \to 0$ and $t \to 1$, we obtain

$$\xi(b) \cdot (a-b) \leqslant \gamma'(1/3) < \gamma'(2/3) \leqslant \xi(a) \cdot (a-b). \tag{5.14}$$

This justifies $\nabla\xi(a) \neq \nabla\xi(b)$. □

Proposition 5.14. *The set $\mathcal{P}^\uparrow_\xi$ is compact and is an extreme set in its closed convex hull $\mathcal{P}(\mathcal{K}_\xi)$.*

Proof. The set $\mathcal{P}^\uparrow_\xi$ is the image of $\mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$ by the continuous map $\nu \mapsto \nabla\xi_\# \nu$. According to Proposition 5.9, $\mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$ is compact, so $\mathcal{P}^\uparrow_\xi$ is compact. Let $\eta \in \mathcal{P}(\mathcal{P}(\mathcal{K}_\xi))$ be such that $\mathrm{Bar}(\eta) \in \mathcal{P}^\uparrow_\xi$ and let $\rho \in \mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$ such that $\mathrm{Bar}(\eta) = \nabla\xi_\#\rho$. In particular, we have $\mathrm{Bar}(\eta) \in \mathcal{P}^\uparrow(\mathcal{K}_\xi)$, since $\mathcal{P}^\uparrow(\mathcal{K}_\xi)$ is an extreme set in its convex hull according to Proposition 5.10, it follows that η is supported on $\mathcal{P}^\uparrow(\mathcal{K}_\xi)$. Therefore, for every μ in the support of η, we have $\mu \in \mathcal{P}^\uparrow(\mathcal{K}_\xi)$ so the path $\mathsf{q}_\mu \in \mathcal{Q}_\infty(S^D_+)$ is well-defined and valued in $\mathcal{K}_\xi$. In particular given μ in the support of η, there exists a, possibly non-monotone, map $\mathsf{p}_\mu : [0,1) \to B(0,1) \cap S^D_+$ such that for every $u \in [0,1)$,

$$\mathsf{q}_\mu(u) = \nabla\xi(\mathsf{p}_\mu(u)).$$

Moreover, we have $\mu = \nabla\xi_\# \mathrm{Law}(\mathsf{p}_\mu(U))$, this implies

$$\nabla\xi_\# \int \mathrm{Law}(\mathsf{p}_\mu(U)) \mathrm{d}\eta(\mu) = \nabla\xi_\#\rho. \tag{5.15}$$

According to Proposition 5.13, $\nabla\xi : S^D_+ \to S^D_+$ is injective, so we have

$$\int \mathrm{Law}(\mathsf{p}_\mu(U)) \mathrm{d}\eta(\mu) = \rho.$$

Finally, since $\rho \in \mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$, and $\mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$ is an extreme set in its convex hull according to Proposition 5.10, we deduce that η-almost surely $\mathrm{Law}(\mathsf{p}_\mu(U)) \in \mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$. Hence, η-almost surely $\mu \in \mathcal{P}^\uparrow_\xi$ which concludes. □

Proposition 5.15. *For any bounded function $\varphi : \mathcal{P}^\uparrow_\xi \to \mathbb{R}$, the function $\overline{\varphi}$ defined by (5.8) is concave and 1-Lipschitz on $\mathcal{M}_1(\mathcal{K}_\xi)$.*

Proof. The function $(\eta, \mu') \mapsto \int \varphi \mathrm{d}\eta - d(\mathrm{Bar}(\eta), \mu')$ is concave on $\mathcal{P}\left(\mathcal{P}^\uparrow_\xi\right) \times \mathcal{M}_1(\mathcal{K}_\xi)$, so $\overline{\varphi}$ is concave as the supremum of a jointly concave functionnal. For every $\eta \in \mathcal{P}\left(\mathcal{P}^\uparrow_\xi\right)$, the function $\mu' \mapsto \int \varphi \mathrm{d}\eta - d(\mathrm{Bar}(\eta), \mu')$ is 1-Lipschitz, so $\overline{\varphi}$ is 1-Lipschitz as the pointwise supremum of a family of 1-Lipschitz functions. □

Proposition 5.16. *Assume that $\varphi : \mathcal{P}^{\uparrow}_{\xi} \to \mathbb{R}$ is 1-Lipschitz and pre-concave, then the function $\overline{\varphi}$ defined by (5.8) is an extension of φ.*

Proof. We fix $\mu' \in \mathcal{P}^{\uparrow}_{\xi}$ and our goal is to show that $\overline{\varphi}(\mu') = \varphi(\mu')$. Choosing $\eta = \delta_{\mu'}$ in (5.8), we obtain $\overline{\varphi}(\mu') \geqslant \varphi(\mu')$, so we only need to show the other bound.

Step 1. We show that for every $\eta \in \mathcal{P}(\mathcal{P}^{\uparrow}_{\xi})$, there exists $\eta' \in \mathcal{P}(\mathcal{P}^{\uparrow}_{\xi})$ such that $\mathrm{Bar}(\eta') = \mu'$ and $\mathsf{d}(\eta, \eta') \leqslant d(\mathrm{Bar}(\eta), \mu')$.

Let π be an optimal coupling between $\mathrm{Bar}(\eta)$ and μ', the existence of such a π is guaranteed by [22, Theorem 4.1]. According to [1, Theorem 5.3.1], there exists a family of probability measures $(\mu_x)_{x \in \mathcal{K}_\xi}$ such that,

$$\int_{\mathcal{K}_\xi \times \mathcal{K}_\xi} h(x,y) \mathrm{d}\pi(x,y) = \int_{\mathcal{K}_\xi} \int_{\mathcal{K}_\xi} h(x,y) \mathrm{d}\mu_x(y) \mathrm{d}\mathrm{Bar}(\eta)(x).$$

Given a random probability measure $\nu \in \mathcal{P}^{\uparrow}_{\xi}$ sampled from η, we define $\nu' \in \mathcal{P}(\mathcal{K}_\xi)$ as the unique probability measure satisfying

$$\int_{\mathcal{K}_\xi} h(y) \mathrm{d}\nu'(y) = \int_{\mathcal{K}_\xi} \int_{\mathcal{K}_\xi} h(y) \mathrm{d}\mu_x(y) \mathrm{d}\nu(x).$$

We let $\eta' \in \mathcal{P}(\mathcal{P}(\mathcal{K}_\xi))$ be the law of the random variable $\nu' \in \mathcal{P}(\mathcal{K}_\xi)$. We have,

$$\begin{aligned}
\int_{\mathcal{K}_\xi} h(y) \mathrm{d}\mathrm{Bar}(\eta')(y) &= \int_{\mathcal{P}^{\uparrow}_{\xi}} \int_{\mathcal{K}_\xi} h(y) \mathrm{d}\nu'(y) \mathrm{d}\eta(\nu) \\
&= \int_{\mathcal{P}^{\uparrow}_{\xi}} \int_{\mathcal{K}_\xi} \int_{\mathcal{K}_\xi} h(y) \mathrm{d}\mu_x(y) \mathrm{d}\nu(x) \mathrm{d}\eta(\nu) \\
&= \int_{\mathcal{K}_\xi} \int_{\mathcal{K}_\xi} h(y) \mathrm{d}\mu_x(y) \mathrm{d}\mathrm{Bar}(\eta)(x) \\
&= \int_{\mathcal{K}_\xi \times \mathcal{K}_\xi} h(y) \mathrm{d}\pi(x,y) \\
&= \int_{\mathcal{K}_\xi} h(y) \mathrm{d}\mu'(y).
\end{aligned}$$

Therefore, $\mathrm{Bar}(\eta') = \mu' \in \mathcal{P}^{\uparrow}_{\xi}$. According to Proposition 5.14, this imposes that η' is supported on $\mathcal{P}^{\uparrow}_{\xi}$. Finally, by definition of the optimal transport distance, we have,

$$\begin{aligned}
\mathsf{d}(\eta,\eta') &\leqslant \int_{\mathcal{P}_\xi^\uparrow} \mathrm{d}(\nu,\nu')\mathrm{d}\eta(\nu) \\
&\leqslant \int_{\mathcal{P}_\xi^\uparrow}\int_{\mathcal{K}_\xi}\int_{\mathcal{K}_\xi} \mathrm{d}(x,y)\mathrm{d}\mu_x(y)\mathrm{d}\nu(x)\mathrm{d}\eta(\nu) \\
&= \int_{\mathcal{K}_\xi}\int_{\mathcal{K}_\xi} d(x,y)\mathrm{d}\mu_x(y)\mathrm{d}\mathrm{Bar}(\eta)(x) \\
&= \int_{\mathcal{K}_\xi}\int_{\mathcal{K}_\xi} d(x,y)\mathrm{d}\pi(x,y) \\
&= d(\mathrm{Bar}(\eta),\mu').
\end{aligned}$$

This concludes Step 1.
Step 2. We show that $\overline{\varphi}(\mu') \leqslant \varphi(\mu')$.
Let $\eta \in \mathcal{P}(\mathcal{P}_\xi^\uparrow)$ and let $\eta' \in \mathcal{P}(\mathcal{P}_\xi^\uparrow)$ be as built in Step 1. We have,

$$\int \varphi \mathrm{d}\eta - d(\mathrm{Bar}(\eta),\mu') = \int \varphi \mathrm{d}\eta' + \int \varphi \mathrm{d}(\eta - \eta') - d(\mathrm{Bar}(\eta),\mu').$$

By Proposition 5.8, $\int \varphi \mathrm{d}\eta' \leqslant \varphi(\mu')$. Since φ is 1-Lipschitz, according to (5.11), we have

$$\int \varphi \mathrm{d}(\eta - \eta') \leqslant \mathsf{d}(\eta,\eta').$$

In Step 1, we have built η' so that $-d(\mathrm{Bar}(\eta),\mu') \leqslant -\mathsf{d}(\eta,\eta')$, therefore, it follows from the two previous displays that

$$\int \varphi \mathrm{d}\eta - d(\mathrm{Bar}(\eta),\mu') \leqslant \varphi(\mu').$$

By taking the supremum over $\eta \in \mathcal{P}(\mathcal{P}_\xi^\uparrow)$ on the left-hand side, we can conclude. □

Proposition 5.17. *Let $\varphi : \mathcal{P}_\xi^\uparrow \to \mathbb{R}$, the function $\overline{\varphi}$ defined in (5.8) satisfies for every $\chi_0 \in \mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)$, $\overline{\varphi}_*(\chi_0) = \varphi_*^\xi(\chi_0)$.*

Proof. Observe that when $\chi_0 \in \mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)$, the map $\mu \mapsto \int \chi_0 \mathrm{d}\mu$ is 1-Lipschitz on $\mathcal{M}_1(\mathcal{K}_\xi)$ with respect to the distance d. In particular, for every $\nu \in \mathcal{P}_\xi^\uparrow$, we have

$$\inf_{\mu \in \mathcal{M}_1(\mathcal{K}_\xi)} \left\{ \int \chi_0 \mathrm{d}\mu + d(\nu,\mu) \right\} = \int \chi_0 \mathrm{d}\nu.$$

Thus,

$$\begin{aligned}
\overline{\varphi}_*(\chi_0) &= \inf_{\mu\in\mathcal{M}_1(\mathcal{K}_\xi)}\left\{\int \chi_0 \mathrm{d}\mu - \overline{\varphi}(\mu)\right\} \\
&= \inf_{\mu\in\mathcal{M}_1(\mathcal{K}_\xi)}\inf_{\eta\in\mathcal{P}(\mathcal{P}^\uparrow_\xi)}\left\{\int \chi_0\mathrm{d}\mu - \int \varphi\mathrm{d}\eta + d(\mathrm{Bar}(\eta),\mu)\right\} \\
&= \inf_{\eta\in\mathcal{P}(\mathcal{P}^\uparrow_\xi)}\left\{-\int \varphi \mathrm{d}\eta + \inf_{\mu\in\mathcal{M}_1(\mathcal{K}_\xi)}\left\{\int \chi_0\mathrm{d}\mu + d(\mathrm{Bar}(\eta),\mu)\right\}\right\} \\
&= \inf_{\eta\in\mathcal{P}(\mathcal{P}^\uparrow_\xi)}\left\{-\int\varphi\mathrm{d}\eta + \int \chi_0 \mathrm{dBar}(\eta)\right\} \\
&= \inf_{\eta\in\mathcal{P}(\mathcal{P}^\uparrow_\xi)}\left\{\int\left(-\varphi(\nu) + \int \chi_0\mathrm{d}\nu\right)\mathrm{d}\eta(\nu)\right\} \\
&= \inf_{\nu\in P^\uparrow_\xi}\left\{-\varphi(\nu) + \int\chi_0\mathrm{d}\nu\right\} \\
&= \varphi^\xi_*(\chi_0).
\end{aligned}$$

□

Proof of Theorem 5.12. The function $\overline{\varphi}$ is 1-Lipschitz and concave on the set $\mathcal{M}_1(\mathcal{K}_\xi)$ according to Proposition 5.15. It follows from Corollary 3.5, that for every $\mu' \in \mathcal{M}_1(\mathcal{K}_\xi)$,

$$\overline{\varphi}(\mu') = \min_{\chi\in\mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)}\left\{\int \chi\mathrm{d}\mu - \overline{\varphi}_*(\chi)\right\}.$$

According to Proposition 5.17, we have for every $\chi\in\mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)$, $\overline{\varphi}_*(\chi) = \varphi^\xi_*(\chi)$, thus

$$\overline{\varphi}(\mu') = \min_{\chi\in\mathcal{L}_{\leqslant 1}(\mathcal{K}_\xi)}\left\{\int \chi\mathrm{d}\mu - \varphi^\xi_*(\chi)\right\}.$$

Finally, we have shown in Proposition 5.16 that $\overline{\varphi}$ is an extension of φ. □

6. Vector models

We now adapt the arguments of the proof of Theorem 1.2 to prove Theorem 1.3. To do so we will use Theorem 5.12, which is proven to be valid only when ξ is strictly convex on S^D_+. Therefore, the proof we give is only valid when ξ is assumed to be strictly convex on S^D_+. We will fix this issue in Section 6.2 using a continuity argument.

6.1. **Strictly convex models.**

Let $\underline{\xi} = \xi + I_{B(0,1)\cap S^D_+}$ denote the function that coincides with ξ on $B(0,1)\cap S^D_+$ and is equal to $+\infty$ outside of $B(0,1)\cap S^D_+$. We define

$$\theta(x) = x\cdot\nabla\xi(x) - \xi(x).$$

Proposition 6.1. *The function $\underline{\xi}^* : S^D_+ \to \mathbb{R}$ is 1-Lipschitz and for every $x \in B(0,1) \cap S^D_+$, we have*

$$\theta(x) = \underline{\xi}^*(\nabla\xi(x)).$$

Proof. For every $y \in \mathbb{R}^{D\times D}$, we have

$$\underline{\xi}^*(y) = \sup_{x \in B(0,1)\cap S^D_+} \left\{ x \cdot y - \xi(x) \right\}.$$

Therefore, $\underline{\xi}^*$ is 1-Lipschitz as the supremum of a family of 1-Lipschitz functions.

Let $x \in B(0,1) \cap S^D_+$, from the definition of θ, it is clear that $\theta(x) \leqslant \xi^*(\nabla\xi(x))$. Given $x' \in S^D_+$, by convexity of ξ on S^D_+, we have for every $\lambda \in (0,1]$,

$$\frac{\xi(\lambda x' + (1-\lambda)x) - \xi(x)}{\lambda} \leqslant \xi(x') - \xi(x).$$

Letting $\lambda \to 0$ in the previous display, we obtain

$$x \cdot \nabla\xi(x) - \xi(x) \geqslant x' \cdot \nabla\xi(x) - \xi(x').$$

This proves that,

$$\theta(x) = \sup_{x' \in S^D_+} \left\{ x' \cdot \nabla\xi(x) - \xi(x') \right\} = \xi^*(\nabla\xi(x)).$$

In particular, the supremum in the previous display is reached at $x' = x \in S^D_+ \cap B(0,1)$ and we have,

$$\theta(x) = \sup_{x' \in S^D_+ \cap B(0,1)} \left\{ x' \cdot \nabla\xi(x) - \xi(x') \right\} = \underline{\xi}^*(\nabla\xi(x)).$$ □

Recall that for every Lipschitz function $\chi : S^D_+ \to \mathbb{R}$, we have defined for every $x \in S^D_+$,

$$\widetilde{S}_t\chi(x) = \sup_{y \in S^D_+ \cap B(0,1)} \left\{ \chi(x + t\nabla\xi(y)) - t\xi^*\left(\nabla\xi(y)\right) \right\}.$$

Theorem 6.2. *Assume that ξ is strictly convex on S^D_+, then Theorem 1.3 holds.*

Proof. According to [4, Corollary 8.7 (2)], the following refinement of (2.9) holds,

$$\lim_{N\to+\infty} \overline{F}_N(t) = \sup_{\mu \in \mathcal{P}^{\uparrow}_{\xi}} \left\{ \psi(\mu) - t \int \xi^*\left(\frac{\cdot}{t}\right) \mathrm{d}\mu \right\}.$$

We have, $(t\xi)^* = t\xi^*\left(\frac{\cdot}{t}\right)$ so up to replacing ξ by $t\xi$, we may assume without loss of generality that $t = 1$.

Step 1. We show that

$$\lim_{N\to+\infty} \overline{F}_N(1) = \sup_{\mu \in \mathcal{P}(\mathcal{K}_{\xi})} \left\{ \overline{\psi}(\mu) - \int \underline{\xi}^* \mathrm{d}\mu \right\}.$$

We have

$$\lim_{N\to+\infty} \overline{F}_N(1) = \sup_{\mu\in\mathcal{P}^\uparrow(B(0,1)\cap S^D_+)} \left\{\psi(\nabla\xi_\#\mu) - \int \theta \mathrm{d}\mu\right\}.$$

According Proposition 6.1, we have $\int \theta(x)\mathrm{d}\mu(x) = \int \underline{\xi}^*(\nabla\xi(x))\mathrm{d}\mu(x)$, replacing μ by $\nabla\xi_\#\mu$, we obtain

$$\lim_{N\to+\infty} \overline{F}_N(1) = \sup_{\mu\in\mathcal{P}^\uparrow_\xi} \left\{\psi(\mu) - \int \underline{\xi}^* \mathrm{d}\mu\right\}.$$

In addition, it follows from Theorem 5.12 that $\overline{\psi}$ is an extension of ψ, therefore the previous display implies

$$\lim_{N\to+\infty} \overline{F}_N(1) \leqslant \sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \left\{\overline{\psi}(\mu) - \int \underline{\xi}^* \mathrm{d}\mu\right\}. \tag{6.1}$$

Conversely, let us fix $\mu\in\mathcal{P}(\mathcal{K}_\xi)$, for every $\varepsilon > 0$, there exists $\eta \in \mathcal{P}(\mathcal{P}^\uparrow_\xi)$, such that

$$\overline{\psi}(\mu) \leqslant \varepsilon + \int \psi \mathrm{d}\eta - d(\mathrm{Bar}(\eta), \mu).$$

Since $\underline{\xi}^*$ is 1-Lipschitz according to Proposition 6.1 by definition of the distance d, it follows that,

$$\begin{aligned}
\overline{\psi}(\mu) - \int \underline{\xi}^*\mathrm{d}\mu &\leqslant \varepsilon + \int \psi\mathrm{d}\eta - \int \underline{\xi}^*\mathrm{d}\mu - d(\mathrm{Bar}(\eta),\mu) \\
&\leqslant \varepsilon + \int \left(\psi(\nu) - \int \underline{\xi}^*\mathrm{d}\nu\right)\mathrm{d}\eta(\nu) \\
&\qquad\qquad + \int \underline{\xi}^*\mathrm{d}(\mathrm{Bar}(\eta) - \mu) - d(\mathrm{Bar}(\eta),\mu) \\
&\leqslant \varepsilon + \int \left(\psi(\nu) - \int \underline{\xi}^*\mathrm{d}\nu\right)\mathrm{d}\eta(\nu) \\
&\leqslant \varepsilon + \sup_{\nu\in\mathcal{P}^\uparrow_\xi} \left\{\psi(\nu) - \int \underline{\xi}^*\mathrm{d}\nu\right\} \\
&= \varepsilon + \lim_{N\to+\infty} \overline{F}_N(1).
\end{aligned}$$

Taking the supremum over $\mu\in\mathcal{P}(\mathcal{K}_\xi)$, this yields

$$\sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \left\{\overline{\psi}(\mu) - \int \underline{\xi}^*\mathrm{d}\mu\right\} \leqslant \varepsilon + \lim_{N\to+\infty} \overline{F}_N(1).$$

Finally, since $\varepsilon > 0$ is arbitrary, we obtain the desired result by letting $\varepsilon \to 0$.
Step 2. We show that

$$\lim_{N\to+\infty} \overline{F}_N(1) = \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \left\{\int \chi\mathrm{d}\mu - \psi^\xi_*(\chi) - \int \underline{\xi}^*\mathrm{d}\mu\right\}.$$

According to Theorem 5.12, we have for every $\mu\in\mathcal{P}(\mathcal{K}_\xi)$,

$$\overline{\psi}(\mu) = \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \left\{\int \chi\mathrm{d}\mu - \psi^\xi_*(\chi)\right\}.$$

It thus follows from Step 1 that

$$\lim_{N\to+\infty} \overline{F}_N(1) = \sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \mathcal{G}(\chi,\mu),$$

where,

$$\mathcal{G}(\chi,\mu) = \int \chi \mathrm{d}\mu - \psi^\xi_*(\chi) - \int \underline{\xi}^* \mathrm{d}\mu.$$

The set $\mathcal{P}(\mathcal{K}_\xi)$ is convex compact with respect to the topology of the distance d. For every $\chi \in \mathcal{L}^0_{\leqslant 1}$ the map $\mathcal{G}(\chi,\cdot)$ is concave and Lipschitz continuous with respect to d. For every $\mu \in \mathcal{P}(\mathcal{K}_\xi)$, the map $\mathcal{G}(\cdot,\mu)$ is convex and lower semi continuous with respect to the topology of local uniform convergence. Therefore, according to [20, Corollary 3.4] we can perform a sup-inf interchange in the previous display to obtain,

$$\lim_{N\to+\infty} \overline{F}_N(1) = \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \mathcal{G}(\chi,\mu).$$

Step 3. Conclusion.
Using Step 2, we have

$$\begin{aligned}
\lim_{N\to+\infty} \overline{F}_N(1) &= \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \left\{ -\psi^\xi_*(\chi) + \sup_{\mu\in\mathcal{P}(\mathcal{K}_\xi)} \left\{ \int \chi d\mu - \int \underline{\xi}^* \mathrm{d}\mu \right\} \right\} \\
&= \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \left\{ -\psi^\xi_*(\chi) + \sup_{x\in B(0,1)\cap S^D_+} \left\{ \chi(\nabla\xi(x)) - \underline{\xi}^*(\nabla\xi(x)) \right\} \right\} \\
&= \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}} \left\{ \widetilde{S}_t\chi(0) - \psi^\xi_*(\chi) \right\}.
\end{aligned}$$

□

6.2. **Convex models.** In this subsection, we prove Theorem 1.3 when ξ is only assumed to be convex but not necessarily strictly convex. To do so, we will use the fact that for every $\alpha > 0$, $x \mapsto \xi(x) + \alpha|x|^2$ is strictly convex. This allows us to establish (1.11) with $\alpha > 0$ using Theorem 6.2, and we can then let $\alpha \to 0$ to conclude.

We start by defining a family of Gaussian processes $(H^\alpha_N)_{N\geqslant 1}$ with covariance function $\xi_\alpha(x) = \xi(x) + \alpha|x|^2$. For every $\sigma \in (\mathbb{R}^D)^N$, we let

$$H^{\mathrm{Potts}}_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^N J_{ij}\sigma_i \cdot \sigma_j,$$

where the J_{ij}'s are independent standard Gaussian random variables. We choose $(J_{ij})_{i,j\geqslant 1}$ independent of $(H_N(\sigma))_{\sigma\in\mathbb{R}^{D\times N}}$. For every $\alpha \geqslant 0$, let

$$H^\alpha_N(\sigma) = H_N(\sigma) + \sqrt{\alpha} H^{\mathrm{Potts}}_N(\sigma).$$

We have for every $\sigma,\tau \in (\mathbb{R}^D)^N$,

$$\mathbb{E}\left[H^\alpha_N(\sigma)H^\alpha_N(\tau)\right] = N\xi_\alpha\left(\frac{\sigma\tau^\perp}{N}\right).$$

We let $\overline{F}^\alpha_N(t)$ denote the free energy of H^α_N, more precisely

$$\overline{F}^\alpha_N(t) = -\frac{1}{N}\mathbb{E}\log\int \exp\left(\sqrt{2t}H^\alpha_N(\sigma) - Nt\xi_\alpha\left(\frac{\sigma\sigma^\perp}{N}\right)\right)\mathrm{d}P_N(\sigma).$$

Note that at $\alpha = 0$, we have $\xi_0 = \xi$, $H^0_N = H_N$ and $\overline{F}^0_N(t) = \overline{F}_N(t)$. We let $\langle\cdot\rangle_\alpha$ denote the Gibbs measure associated to $\overline{F}^\alpha_N(t)$, it is a random probability measure on $\mathbb{R}^{D\times N}$ defined by

$$d\langle\cdot\rangle_\alpha(\sigma) \propto \exp\left(\sqrt{2t}H^\alpha_N(\sigma) - Nt\xi_\alpha\left(\frac{\sigma\sigma^\perp}{N}\right)\right)\mathrm{d}P_N(\sigma).$$

Proposition 6.3. *For every $t \geqslant 0$, there exists a constant $C \geqslant 0$ such that for every $N \in \mathbb{N}$ and every $\alpha, \alpha' \geqslant 0$,*

$$|\overline{F}^\alpha_N(t) - \overline{F}^{\alpha'}_N(t)| \leqslant C|\alpha - \alpha'|.$$

Proof. Without loss of generality, we may assume that $t = 1/2$. The function $\alpha \mapsto \overline{F}^\alpha_N(1/2)$ is continuous on $\mathbb{R}_+$ and differentiable on $(0, +\infty)$, and we have

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}F^\alpha_N(1/2) = -\frac{1}{N}\mathbb{E}\left\langle \frac{1}{2\sqrt{\alpha}}H^{\mathrm{Potts}}_N(\sigma) - \frac{N}{2}\left|\frac{\sigma\sigma^\perp}{N}\right|^2\right\rangle_\alpha.$$

Using the Gaussian integration by part formula [15, Lemma 1.1], it follows that,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}F^\alpha_N(1/2) = \frac{1}{2}\mathbb{E}\left\langle\left|\frac{\sigma\tau^\perp}{N}\right|^2\right\rangle_\alpha.$$

Since the reference measure P_1 is compactly supported, there exists a constant $c \geqslant 0$ such that $\mathbb{E}\langle\cdot\rangle_\alpha$-almost surely $\left|\frac{\sigma\tau^\perp}{N}\right|^2 \leqslant D^2c^2$, the result follows. □

For every $\chi \in \mathcal{L}^0_{\leqslant 1}$, we let

$$\psi^\alpha_*(\chi) = \inf_{\mu\in\mathcal{P}^\uparrow_{\xi_\alpha}}\left\{\int \chi\mathrm{d}\mu - \psi(\mu)\right\},$$

and

$$\widetilde{S}^\alpha_t\chi = \sup_{x\in B(0,1)\cap S^D_+}\left\{\chi(t\nabla\xi_\alpha(x)) - t\theta_\alpha(x)\right\},$$

where,

$$\theta_\alpha(x) = x\cdot\nabla\xi_\alpha(x) - \xi_\alpha(x) = \theta(x) + \alpha|x|^2.$$

Proposition 6.4. *For every $t \geqslant 0$, the function*

$$\alpha \mapsto \inf_{\chi\in\mathcal{L}^0_{\leqslant 1}}\left\{\widetilde{S}^\alpha_t\chi(0) - \psi^\alpha_*(\chi)\right\},$$

is Lipschitz on $\mathbb{R}_+$.

Proof. If for every $\chi \in \mathcal{L}^0_{\leqslant 1}$, the functions $\alpha \mapsto \widetilde{S}^\alpha_t \chi(0)$ and $\alpha \mapsto \psi^\alpha_*(\chi)$ are Lipschitz functions on $\mathbb{R}_+$ with Lipschitz constant independent of χ, then the result follows because in this case $\alpha \mapsto \inf_{\chi \in \mathcal{L}^0_{\leqslant 1}} \left\{ \widetilde{S}^\alpha_t \chi(0) - \psi^\alpha_*(\chi) \right\}$ is the infimum of a family of uniformly Lipschitz functions.
Step 1. We show that for every $\chi \in \mathcal{L}^0_{\leqslant 1}$, $\alpha \mapsto \widetilde{S}^\alpha_t \chi(0)$ is $3t$–Lipschitz.
Let $x \in B(0,1) \cap S^D_+$, the function

$$\alpha \mapsto \chi(t\nabla\xi(x) + 2t\alpha x) - t\theta(x) - t\alpha|x|^2$$

is $3t$-Lipschitz on $\mathbb{R}_+$ and we have

$$\widetilde{S}^\alpha_t \chi = \sup_{x \in B(0,1)} \left\{ \chi(t\nabla\xi(x) + 2t\alpha x) - t\theta(x) - t\alpha|x|^2 \right\}.$$

Therefore, $\alpha \mapsto \widetilde{S}^\alpha_t \chi(0)$ is $3t$-Lipschitz on $\mathbb{R}_+$ as the supremum of a family of $3t$-Lipschitz functions on $\mathbb{R}_+$.
Step 2. We show that for every $\chi \in \mathcal{L}^0_{\leqslant 1}$, $\alpha \mapsto \psi^\alpha_*(\chi)$ is 4–Lipschitz on $\mathbb{R}_+$.
Let $\alpha, \alpha' \geqslant 0$, we have for every $\mu \in \mathcal{P}^\uparrow(B(0,1) \cap S^D_+)$

$$\begin{aligned} \left| \int \chi(\nabla\xi(x) + 2\alpha x)\mathrm{d}\mu(x) - \int \chi(\nabla\xi(x) + 2\alpha' x)\mathrm{d}\mu(x) \right| &\leqslant \int 2|\alpha - \alpha'||x|\mathrm{d}\mu \\ &\leqslant 2|\alpha - \alpha'|. \end{aligned}$$

In addition, since ψ is 1-Lipschitz with respect to the optimal transport distance d, we have

$$\begin{aligned} &\left| \psi\left((\nabla\xi + 2\alpha \mathrm{id}_{S^D_+})_{\#}\mu \right) - \psi\left((\nabla\xi + 2\alpha' \mathrm{id}_{S^D_+})_{\#}\mu \right) \right| \\ &\leqslant d((\nabla\xi + 2\alpha \mathrm{id}_{S^D_+})_{\#}\mu, (\nabla\xi + 2\alpha' \mathrm{id}_{S^D_+})_{\#}\mu) \\ &\leqslant \sup_{\chi_0 \in \mathcal{L}^0_{\leqslant 1}} \left\{ \int \chi_0(\nabla\xi(x) + 2\alpha x)\mathrm{d}\mu(x) - \int \chi_0(\nabla\xi(x) + 2\alpha' x)\mathrm{d}\mu(x) \right\} \\ &\leqslant \sup_{\chi_0 \in \mathcal{L}^0_{\leqslant 1}} \left\{ \int 2|\alpha - \alpha'||x|\mathrm{d}\mu(x) \right\} \\ &\leqslant 2|\alpha - \alpha'|. \end{aligned}$$

Finally, since

$$\psi^\alpha_*(\chi) = \inf_{\mu \in \mathcal{P}^\uparrow(B(0,1) \cap S^D_+)} \left\{ \int \chi(\nabla\xi(x) + 2\alpha x)\mathrm{d}\mu(x) - \psi\left((\nabla\xi + 2\alpha \mathrm{id}_{S^D_+})_{\#}\mu \right) \right\},$$

we deduce that $\alpha \mapsto \psi^\alpha_*(\chi)$ is 4-Lipschitz as the infimum of a family of 4-Lipschitz functions. □

Proof of Theorem 1.3. Let $\alpha > 0$, ξ_α is strictly convex on S^D_+ so according to Theorem 6.2, we have

$$\lim_{N \to +\infty} \overline{F}^\alpha_N(t) = \inf_{\chi \in \mathcal{L}_{\leqslant 1}} \left\{ \widetilde{S}^\alpha_t \chi(0) - \psi^\alpha_*(\chi) \right\}.$$

According to Propositions 6.3 and 6.4, we can let $\alpha \to 0$ in the previous display to obtain

$$\lim_{N\to+\infty} \overline{F}_N(t) = \inf_{\chi\in\mathcal{L}_{\leqslant 1}} \left\{ \widetilde{S}_t\chi(0) - \psi_*^\xi(\chi) \right\}.$$

□

7. Interpretations of the main results

This last section has a more speculative flavor, its aim is to give an interpretation of the main results of this paper through the lens of Hamilton-Jacobi equations. We start by explaining why we think that (1.7) can be interpreted as a Hopf-like formula. Assuming that this interpretation can be made rigorous in the case $D > 1$, we construct a conjectural variational formula for the limit free energy when ξ is not assumed to be convex on S_+^D.

7.1. Hopf and Hopf-like formulas.

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function and let $H : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function. It is well known that the Hamilton-Jacobi equation

$$\begin{cases} \partial_t u - H(\nabla u) = 0 \text{ on } (0,+\infty)\times\mathbb{R}^n \\ u(0,\cdot) = \psi \end{cases} \tag{7.1}$$

admits a unique viscosity solution u and that when ψ is concave, u admits the Hopf representation [8, Theorem 3.13]. We recall this in the next proposition.

Proposition 7.1 ([8]). *Assume that $\psi : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz and concave, assume that $H : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz. Then, the unique viscosity solution of (7.1) satisfies*

$$u(t,x) = \inf_{y\in\mathbb{R}^n} \left\{ x\cdot y - \psi_*(y) + tH(y) \right\}, \tag{7.2}$$

where $\psi_*(y) = \inf_{x\in\mathbb{R}^n} \{x \cdot y - \psi(y)\}$.

Intuitively, the Hopf formula follows from the following observations. At time $t = 0$, $u(0,\cdot)$ coincides with the infimum of the family of values taken by affine functions that upper bound ψ (this is Fenchel-Moreau duality),

$$u(0,x) = \inf_{y\in\mathbb{R}^n} \left\{ x\cdot y - \psi_*(y) \right\}. \tag{7.3}$$

For every $y \in \mathbb{R}^n$ the viscosity solution of (7.1) with the affine initial condition $x \mapsto x \cdot y - \psi_*(y)$ is

$$(t,x) \mapsto x\cdot y - \psi_*(y) + tH(y).$$

Provided that we can interchange the semigroup of (7.1) with the infimum in (7.3), we obtain (7.2).

Notice that in Proposition 7.1, the geodesics used to define the notion of concavity and the notion of derivative (the symbol ∇ in (7.1)) are the same, both notions are defined using straight lines in $\mathbb{R}^n$.

For the rest of this subsection we assume that $D = 1$. Let $\psi : \mathcal{P}_1(\mathbb{R}_+) \to \mathbb{R}$ be a nondecreasing Lipschitz function, we have explained that (2.10) could be reformulated into (2.12). Now observe that in (2.12) the notion of derivative (the symbol ∇ in (2.12)) is the Fréchet derivative in L^2, it is defined using straight lines in L^2. Hence, as might be expected, it was shown in [6, Theorem 4.6 (3)] that the Hopf representation is also available for viscosity solutions of (2.10), provided that the function $\mathsf{q} \mapsto \psi(\Omega^{-1}\mathsf{q})$ is concave. In this case, f the unique solution of (2.12), satisfies

$$\mathsf{f}(t, \mathsf{q}) = \inf_{\mathsf{p} \in \mathcal{Q}_\infty(\mathbb{R}_+)} \left\{ \langle \mathsf{q}, \mathsf{p} \rangle_{L^2} - \psi_\circ(\mathsf{p}) + t \int \xi(\mathsf{p}) \right\}, \tag{7.4}$$

where

$$\psi_\circ(\mathsf{p}) = \inf_{\mathsf{q} \in \mathcal{Q}_2(\mathbb{R}_+)} \left\{ \langle \mathsf{q}, \mathsf{p} \rangle_{L^2} - \psi(\Omega^{-1}\mathsf{q}) \right\}.$$

Even thought (7.4) is the natural generalization of (7.2), this formula seems to be of limited use in the context of spin glasses since for the function ψ defined by (2.7), the function $\mathsf{q} \mapsto \psi(\Omega^{-1}q)$ is not concave (nor convex) in general as shown in [11, Section 6].

We now assume that ψ is given by (2.7). According to [2], the function $\mu \mapsto \psi(\mu)$ is concave on $\mathcal{P}_2(\mathbb{R}_+)$. When considering (2.10) in this case, the geodesics used to define concavity and derivatives are not the same. We use straight lines in $\mathcal{P}_2(\mathbb{R}_+)$ to define concavity and transport geodesics in $\mathcal{P}_2(\mathbb{R}_+)$ (that is straight lines in L^2) to define the symbol ∂_μ appearing in (2.10). Nonetheless, the formula (1.7) seems to be the natural adaptation of (7.4) under this slightly unusual setup. Indeed, let f be the viscosity solution of (2.10), recall that according to Lemma 4.1, we have

$$f(0, \mu) = \inf_{\chi \in \mathcal{X}} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\},$$

where $\mathcal{X}$ denotes the set of Lipschitz functions $\chi : \mathbb{R}_+ \to \mathbb{R}$ which are nondecreasing and convex. Theorem 1.2 states that

$$f(t, \mu) = \inf_{\chi \in \mathcal{X}} \left\{ \int S_t \chi \mathrm{d}\mu - \psi_*(\chi) \right\}, \tag{7.5}$$

and according to Theorem 4.2, $(t, \mu) \mapsto \int S_t \chi \mathrm{d}\mu - \psi_*(\chi)$ is the unique viscosity solution of

$$\begin{cases} \partial_t u - \int \xi(\partial_\mu u) \mathrm{d}\mu = 0 \text{ on } (0, +\infty) \times \mathcal{P}_2(\mathbb{R}_+) \\ u(0, \mu) = \int \chi \mathrm{d}\mu - \psi_*(\chi). \end{cases}$$

Therefore (7.5) and (7.4) have the same structure; those two formulas express the viscosity solution as the infimum of a family of viscosity solutions started at an affine initial condition. The only difference is that for (7.4) the relevant linear initial conditions are of the form $\mathsf{q} \mapsto \langle \mathsf{p}, \mathsf{q} \rangle_{L^2}$ when for (7.5) the relevant linear initial conditions are of the form $\mu \mapsto \int \chi \mathrm{d}\mu$. This is why we refer to (1.7) as a Hopf-like formula.

7.2. **Some conjectures.** In this subsection, we do not assume that $D = 1$. When ξ is not assumed to be convex on S^D_+, the Parisi formula breaks down and there is no known generalization of Theorem 1.1. In [11, 12, 13] it was conjectured that results such as Theorem 2.1 should generalize to the setting where the function ξ is not assumed to be convex on S^D_+. Namely, it should hold under (H1) and (H3) only that the free energy converges as $N \to +\infty$ to the viscosity solution of (2.10). When ξ is not assumed to be convex the Hopf-Lax representation for the viscosity solutions of (2.10) is not available and variational representations for the limit free energy such as (2.9) are proven to be false (see [11, Section 6]).

As explained in the previous subsection, for Hamilton-Jacobi equations with possibly nonconvex nonlinearity, if the initial condition is concave, a variational representation is available for the viscosity solution, and it seems that the un-inverted formula (1.7) can be interpreted as an instance of such a variational representation. One of the main ingredients for proving (1.7) and coincidently also for the Hopf representation to hold is Fenchel-Moreau duality (which manifests through Lemma 4.1 here). If one wishes to generalize (1.7) to models with $D > 1$, a seemingly crucial step is thus to generalize Lemma 4.1. The attentive reader will notice that a version of Lemma 4.1 follows from Theorem 5.12, this allows us to write ψ as an infimum over Lipschitz functions $\chi : S^D_+ \to \mathbb{R}$. But this version of Lemma 4.1 is not strong enough for our purpose. Indeed, to guarantee that (7.12) below is well-posed, one needs that the initial condition χ is nondecreasing [5, Theorem 1.2 (2)]. The argument we have used in the proof of Lemma 4.1 to show that the infimum could in fact be taken over nondecreasing χ's doesn't seem to be easily applicable when $D > 1$. Indeed, we have used crucially the fact that the set of surjective paths $\mathsf{q} \in \mathcal{Q}(\mathbb{R}_+)$ is dense in $\mathcal{Q}(\mathbb{R}_+)$. When $D > 1$, the set of surjective paths $\mathsf{q} \in \mathcal{Q}(S^D_+)$ is empty, as any surjective function $[0, 1) \to S^D_+$ must be non-monotonous. In order to obtain a suitable generalization of Lemma 4.1 to the case $D > 1$, we thus make the following additional assumptions.

- (H4) There exists compactly supported probability measures $\pi_1, \dots, \pi_D \in \mathcal{P}_\infty(\mathbb{R})$ such that $P_1 = \pi_1 \otimes \dots \otimes \pi_D$.
- (H5) The function ξ only depends on the diagonal coefficients of its argument. That is, there exists a function $\overline{\xi} : \mathbb{R}^D \to \mathbb{R}$ such that $\xi(A) = \overline{\xi}((A_{dd})_{1 \leqslant d \leqslant D})$.

Thanks to (H5) we can encode the limit free energy with a partial differential equation on $\mathcal{P}^\uparrow_2(\mathbb{R}^D_+)$ rather than $\mathcal{P}^\uparrow_2(S^D_+)$, and using (H4) we can adapt Lemma 4.1. Let us explain this in more details.

The map $x \mapsto \mathrm{diag}(x)$ which maps each vector $x \in \mathbb{R}^D_+$ to the matrix in S^D_+ whose diagonal coefficients are $x_1, \dots, x_D$, defines an injection from $\mathbb{R}^D_+$ to S^D_+. In particular each $\mu \in \mathcal{P}^\uparrow_1(\mathbb{R}^D_+)$ can be interpreted as a probability measure in $\mathcal{P}^\uparrow_1(S^D_+)$. This means that the quantity $\overline{F}_N(t, \mu)$ is also defined when $\mu \in \mathcal{P}_1(\mathbb{R}_+^D)$. We can then easily adapt the arguments of [4, Section 8] to prove the following theorem.

Theorem 7.2 (limit free energy for convex diagonal models). *Assume that (H1), (H2), (H3) and (H5) hold, then for every $t \geqslant 0$ and $\mu \in \mathcal{P}_1^{\uparrow}(\mathbb{R}_+^D)$, we have*

$$\lim_{N \to +\infty} \overline{F}_N(t,\mu) = \sup_{\nu \in \mathcal{P}_\infty^{\uparrow}(\mathbb{R}_+^D), \nu \geqslant \mu} \left(\psi(\nu) - t\mathbb{E}\left[\overline{\xi}^* \left(\frac{X_\nu - X_\mu}{t} \right) \right] \right). \tag{7.6}$$

Moreover, denoting by $f(t,\mu)$ the expression above, we have that $f : \mathbb{R}_+ \times \mathcal{P}_2^{\uparrow}(\mathbb{R}_+^D) \to \mathbb{R}$ solves the Hamilton-Jacobi equation

$$\begin{cases} \partial_t f - \int \overline{\xi}(\partial_\mu f)\, \mathrm{d}\mu = 0 & \text{on } \mathbb{R}_+ \times \mathcal{P}_2^{\uparrow}(\mathbb{R}_+^D), \\ f(0,\cdot) = \psi & \text{on } \mathcal{P}_2^{\uparrow}(\mathbb{R}_+^D). \end{cases} \tag{7.7}$$

Let ψ be the functional defined in (2.7) and let $\psi_d : \mathcal{P}_1(\mathbb{R}_+) \to \mathbb{R}$ denote the function obtained from (2.7) when $D = 1$ and the reference probability measure is π_d. Given $\mu \in \mathcal{P}(\mathbb{R}_+^D)$ let $\mu_1, \dots, \mu_D \in \mathcal{P}(\mathbb{R}_+)$ denote its marginals. Assumption (H4) yields the following simplification, for every $\mu \in \mathcal{P}(\mathbb{R}_+^D)$,

$$\psi(\mu) = \sum_{d=1}^{D} \psi_d(\mu_d). \tag{7.8}$$

Let $\mathcal{X}_D$ denote the set of functions $\chi : \mathbb{R}_+^D \to \mathbb{R}$ which satisfy

$$\chi(x) = \sum_{d=1}^{D} \chi_d(x_d), \tag{7.9}$$

with $\chi_1, \dots, \chi_D : \mathbb{R}_+ \to \mathbb{R}$ 1-Lipschitz, nondecreasing and convex. The set $\mathcal{X}_D$ is contained in the set of Lipschitz, nondecreasing and convex functions on $\mathbb{R}_+^D$. In particular if we combine (7.8) with Lemma 4.1 applied to each ψ_d, we obtain the following for every $\mu \in \mathcal{P}_1^{\uparrow}(\mathbb{R}_+^D)$,

$$\psi(\mu) = \inf_{\chi \in \mathcal{X}_D} \left\{ \int \chi \mathrm{d}\mu - \psi_*(\chi) \right\}, \tag{7.10}$$

where, as usual,

$$\psi_*(\chi) = \inf_{\mu \in \mathcal{P}_1^{\uparrow}(\mathbb{R}_+^D)} \left\{ \int \chi \mathrm{d}\mu - \psi(\mu) \right\}. \tag{7.11}$$

With (7.10) in mind, we can expect that the mechanism leading to (7.5) in the case $D = 1$, leads to the following result when $D \geqslant 1$.

Conjecture 7.3. *Assume that (H1), (H3), (H4) and (H5) hold but with $\overline{\xi}$ possibly non-convex on $\mathbb{R}_+^D$, then f the unique viscosity solution of (7.7) satisfies*

$$f(t,\mu) = \inf_{\chi \in \mathcal{X}_D} \left\{ \int S_t \chi \mathrm{d}\mu - \psi_*(\chi) \right\},$$

where $S_t\chi(x) = \sup_{y\in\mathbb{R}^D_+}\left\{x\cdot y - \chi^*(y) + t\overline{\xi}(y)\right\}$ *is the unique viscosity solution of*

$$\begin{cases}\partial_t u - \overline{\xi}(\nabla u) = 0 \ \text{on} \ (0,+\infty)\times\mathbb{R}^D_+\\ u(0,\cdot) = \chi.\end{cases}$$

Combining Conjecture 7.3 with [12, Conjecture 2.6], we obtain the following conjectural variational formula for the limit free energy of nonconvex models.

Conjecture 7.4. *Assume that (H1), (H3), (H4) and (H5) hold but with* $\overline{\xi}$ *possibly non-convex on* $\mathbb{R}^D_+$, *then*

$$\lim_{N\to+\infty}\overline{F}_N(t,\delta_0) = \inf_{\chi\in\mathcal{X}_D}\left\{S_t\chi(0) - \psi_*(\chi)\right\},$$

where $\mathcal{X}_D$ *is defined in* (7.9), $S_t\chi(0) = \sup_{y\in\mathbb{R}^D_+}\left\{-\chi^*(y) + t\overline{\xi}(y)\right\}$ *is the value at* $(t,0)$ *of the unique viscosity solution of*

$$\begin{cases}\partial_t u - \overline{\xi}(\nabla u) = 0 \ \text{on} \ (0,+\infty)\times\mathbb{R}^D_+\\ u(0,\cdot) = \chi,\end{cases} \tag{7.12}$$

and where

$$\psi_*(\chi) = \inf_{\mu\in\mathcal{P}_1^\uparrow(\mathbb{R}^D_+)}\left\{\int \chi \,\mathrm{d}\mu - \psi(\mu)\right\}. \tag{7.13}$$

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Basel, 2008.

[2] Antonio Auffinger and Wei-Kuo Chen. The Parisi formula has a unique minimizer. *Comm. Math. Phys.*, 335(3):1429–1444, 2015.

[3] Hong-Bin Chen. Parisi PDE and convexity for vector spins. Preprint, arXiv:2311.10446.

[4] Hong-Bin Chen and Jean-Christophe Mourrat. On the free energy of vector spin glasses with non-convex interactions. *Prob. Math. Phys.*, to appear, 2023.

[5] Hong-Bin Chen and Jiaming Xia. Hamilton-Jacobi equations with monotone nonlinearities on convex cones. *Ann. Fac. Sci. Toulouse Math.*, to appear, 2022.

[6] Hong-Bin Chen and Jiaming Xia. Hamilton-Jacobi equations from mean-field spin glasses. Preprint, arXiv:2201.12732.

[7] John B. Conway. *A course in functional analysis*. Graduate Text in Mathematics. Springer, 1990.

[8] 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.

[9] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[10] Francesco Guerra and Fabio Lucio Toninelli. The thermodynamic limit in mean-field spin glass models. *Comm. Math. Phys.*, 230:71–79, 2002.

[11] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.

[12] Jean-Christophe Mourrat. The Parisi formula is a Hamilton-Jacobi equation in Wasserstein space. *Canad. J. Math.*, 74(3):607–629, 2022.

[13] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(3):1143–1182, 2023.

[14] Jean-Christophe Mourrat. Un-inverting the Parisi formula. *Ann. Henri Poincaré Probab. Stat.*, to appear, 2023.

[15] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, 2013.

[16] Dmitry Panchenko. The free energy in a multi-species Sherrington-Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.

[17] Dmitry Panchenko. Free energy in the Potts spin glass. *Ann. Probab.*, 46(2):829–864, 2018.

[18] Dmitry Panchenko. Free energy in the mixed p-spin models with vector spins. *Ann. Probab.*, 46(2):865–896, 2018.

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

[20] Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.

[21] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

[22] Cédric Villani. *Optimal transport, old and new*. Grundlehren der Mathematischen Wissenschaften. Springer, 2009.

[23] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific Publishing, 2002.