EXISTENCE AND UNIQUENESS OF PERMUTATION-INVARIANT OPTIMIZERS FOR PARISI FORMULA

VICTOR ISSA

Abstract. It has recently been shown in [4] that, upon constraining the system to stay in a balanced state, the Parisi formula for the mean-field Potts model can be written as an optimization problem over permutation-invariant functional order parameters.

In this paper, we focus on permutation-invariant mean-field spin glass models. After introducing a correction term in the definition of the free energy and without constraining the system, we show that the limit free energy can be written as an optimization problem over permutation-invariant functional order parameters. We also show that for some models this optimization problem admits a unique optimizer. In the case of Ising spins, the correction term can be easily removed, and those results transfer to the uncorrected limit free energy.

We also derive an upper bound for the limit free energy of some nonconvex permutation-invariant models. This upper bound is expressed as a variational formula and is related to the solution of some Hamilton-Jacobi equation. We show that if no first order phase transition occurs, then this upper bound is equal to the lower bound derived in [17]. We expect that this hypothesis holds at least in the high temperature regime.

Our method relies on the fact that the free energy of any convex mean-field spin glass model can be interpreted as the strong solution of some Hamilton-Jacobi equation.

Keywords and phrases: multi-species spin glass, Hamilton-Jacobi equations, free energy, Parisi formula.

MSC 2020: 35D40, 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 $D > 1$ be an integer, that we will keep fixed throughout the paper. We study Gaussian processes $(H_N(\sigma))_{\sigma\in\mathbb{R}^{D\times N}}$, whose covariance is of the form

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

Here $\sigma\tau^* = (\sigma_d \cdot \tau_{d'})_{1\leqslant d,d'\leqslant D}$, $x \cdot y$ is the standard scalar product on $\mathbb{R}^N$ and $\xi : \mathbb{R}^{D\times D} \to \mathbb{R}$ is a function given by an absolutely convergent power series. Unless stated otherwise, we will always assume that ξ is convex on the set of positive semi-definite matrices. We give ourselves for each N a reference probability measure P_N on $\mathbb{R}^{D\times N}$ of the form $P_N = P_1^{\otimes N}$ where P_1 is a compactly supported probability measure on $\mathbb{R}^D$. For every $t \geqslant 0$, One can associate a random probability measure on $\mathbb{R}^{D\times N}$ to the process H_N called the Gibbs measure and denoted by $\langle\cdot\rangle$. It is defined by

$$\langle h(\sigma)\rangle = \frac{\int h(\sigma)\exp\left(\sqrt{2t}H_N(\sigma) - Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)\right)\mathrm{d}P_N(\sigma)}{\int \exp\left(\sqrt{2t}H_N(\sigma) - Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)\right)\mathrm{d}P_N(\sigma)}.$$

An important step, in understanding the Gibbs measure associated to the family of processes $(H_N)_N$ is the computation of the large N limit of the free energy

$$\overline{F}_N(t) = -\frac{1}{N}\mathbb{E}\log\int \exp\left(\sqrt{2t}H_N(\sigma) - Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)\right)\mathrm{d}P_N(\sigma). \tag{1.2}$$

For the models of interest here, a variational formula for the limiting value of $\overline{F}_N(t)$ is known, this is the celebrated Parisi formula. The Parisi formula was first conjectured in [24] 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 [14] 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 [13, 26]. 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 [20]. Some models with $D > 1$ such as multispecies models, the Potts model, and a general class of models with vector spins were treated in [21, 22, 23], under the assumption that ξ is convex on $\mathbb{R}^{D\times D}$. Finally, the case $D > 1$ was treated in general in [7] assuming only that that ξ is convex on the set of positive semi-definite matrices. The following version of the Parisi formula is [7, Corollary 8.2].

Theorem 1.1 ([7]). *If ξ is convex on S^D_+, then for every $t > 0$,*

$$\lim_{N\to+\infty} \overline{F}_N(t) = \sup_{\mathfrak{q}}\left\{\psi(\mathfrak{q}) - t\int_0^1 \xi^*\left(\frac{\mathfrak{q}(u)}{t}\right)\mathrm{d}u\right\}. \tag{1.3}$$

Here S^D_+ denotes the set of positive semi-definite symmetric matrices in $\mathbb{R}^{D\times D}$. We equip $\mathbb{R}^{D\times D}$ with the order $A \leqslant B$ if and only $B - A \in S^D_+$. The supremum in (1.3) is taken over the set of bounded functions in

$$\mathcal{Q}(S^D_+) = \{\mathfrak{q} : [0,1) \to S^D_+ \big| \, \mathfrak{q} \text{ is càdlàg and nondecreasing}\}. \tag{1.4}$$

The function ψ is the cascade transform of the measure P_1 and ξ^* is the convex conjugate of ξ with respect to the cone S^D_+. We postpone the precise definition of these objects to (2.5) and (3.5) respectively.

Let $\mathscr{S}_D$ denote the group of permutations of $\{1,\dots,D\}$, we say that ξ is permutation-invariant and P_1 is permutation-invariant when for every permutation $s \in \mathscr{S}_D$,

$$\forall R \in \mathbb{R}^{D\times D},\ \xi(R) = \xi\left((R_{s(d),s(d')})_{1\leqslant d,d'\leqslant D}\right), \tag{1.5}$$

$$\forall \chi \in \mathcal{C}_b(\mathbb{R}^D),\ \int \chi(x_1,\dots,x_D)\mathrm{d}P_1(x) = \int \chi(x_{s(1)},\dots,x_{s(D)})\mathrm{d}P_1(x). \tag{1.6}$$

We will show that when ξ and P_1 are permutation-invariant, the supremum in (1.3) can be taken over the set of permutation-invariant paths in $\mathcal{Q}(S^D_+)$. This result can be interpreted as an absence of breaking of permutation-invariance by the system. This persistence of permutation-invariance was predicted to happen for the Potts model (see (1.7) below) in [11]. It has been rigorously proven that the Potts model does not break its permutation invariance, but only after constraining the system to stay in a balanced state [4] or by introducing a correction term of the form $-Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)$ [5]. In (1.2) the correction term of [5] is also present, but our result is general enough to cover models where P_1 is supported on $\{-1,1\}^D$ and ξ only depends on the diagonal coefficients of its argument, for those models the correction term is constant and our main results transfer to the uncorrected free energy.

1.2. **Main results.** Let $\mathcal{Q}(\mathbb{R}_+)$ denote the set of càdlàg and nondecreasing functions $[0,1) \to \mathbb{R}_+$. Let U denote a uniform random variable in $[0,1)$, given $p_1, p_2 \in \mathcal{Q}(\mathbb{R}_+)$, we let $(p_1,p_2)^\perp : [0,1) \to S^D_+$ be defined by

$$(p_1,p_2)^\perp = p_1\left(\mathrm{id}_D - \frac{\mathbf{1}_D}{D}\right) + p_2\frac{\mathbf{1}_D}{D},$$

where id_D denotes the $D\times D$ identity matrix and $\mathbf{1}_D$ the $D\times D$ matrix whose coefficients are all equal to 1. As will be proven in Section 4, every permutation-invariant path in $\mathcal{Q}(S^D_+)$ is of the form $(p_1,p_2)^\perp$.

Heuristically, given a maximizing path $\mathfrak{q}$ in (1.3), the law of the matrix $\mathfrak{q}(U)$ is the limiting distribution of $\frac{\sigma\tau^*}{N}$ where σ and τ are two independent random variables with law $\langle\cdot\rangle$. For some specific models, the distribution of the overlap matrix has some additional properties, and those additional properties allow us to write the limit free energy as an optimization over a smaller set of paths. For example, consider the Potts model which corresponds to the family of processes

$$H_N^{\mathrm{Potts}}(\sigma) = \frac{1}{\sqrt{N}}\sum_{i,j=1}^N J_{ij}\sigma_i\cdot\sigma_j. \tag{1.7}$$

Here $(J_{ij})_{i,j\geqslant 1}$ denotes a family of independent standard Gaussian random variables. The process H_N^{Potts} satisfies (1.1) with $\xi(R) = \sum_{d,d'=1}^D R_{dd'}^2$. Often, the Potts model is considered with reference measure $P_1 = \mathrm{Unif}\{e_1,\dots,e_D\}$, where $(e_1,\dots,e_D)$ denote the canonical basis of $\mathbb{R}^D$. With this assumption, the terms of the form $\sigma_i\cdot\sigma_j$ appearing in (1.7) only take the value 0 or 1. In this case, if we sample two independent random variables $\sigma,\tau\in\mathbb{R}^{D\times N}$ with law $\langle\cdot\rangle$, their overlap matrix $R = \frac{\sigma\tau^*}{N}$ satisfies $\sum_{d,d'=1}^D R_{dd'} = 1$ almost surely and for every $s\in\mathscr{S}_D$, the matrices R and $(R_{s(d)s(d')})_{1\leqslant d,d'\leqslant D}$ are equal in law under $\mathbb{E}\langle\cdot\rangle$. This means that if $\mathfrak{q}$ is a maximizing path for the Parisi formula of the Potts model, we should expect that

$$\mathfrak{q} = p\left(\mathrm{id}_D - \frac{\mathbf{1}_D}{D}\right) + \frac{\mathbf{1}_D}{D^2} = \left(p,\frac{1}{D}\right)^\perp$$

for some $p\in\mathcal{Q}(\mathbb{R}_+)$. This observation on the set of optimal paths in (1.3) for the Potts model was leveraged in [5, 4]. In [4], the authors show that when the system is constrained to stay in a balanced state, the limit free energy of the Potts model can be written a supremum over $\mathcal{Q}(\mathbb{R}_+)$. In [5], the author does not constrain the system but introduces a correction term of the form $-Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)$ like in (1.2) and obtains results similar to [4].

In this paper, we show that similar results can be obtained in different settings. We will focus on permutation-invariant models, that is models with ξ and P_1 satisfying (1.5) and (1.6). We will show that for those models, the following variational formula holds.

Theorem 1.2. *Assume that ξ is convex on S^D_+ and permutation-invariant, assume that P_1 is permutation-invariant. Then, for every $t\geqslant 0$,*

$$\begin{aligned}\lim_{N\to+\infty}\overline{F}_N(t) = \sup_{(p_1,p_2)}\inf_{(r_1,r_2)}\Bigg\{&\psi((p_1,p_2)^\perp)\\ &-\langle p_1,r_1\rangle_{L^2} - \langle p_2,r_2\rangle_{L^2} + t\int_0^1 \xi\left(\left(\frac{p_1(u)}{D-1},p_2(u)\right)^\perp\right)\mathrm{d}u\Bigg\}.\end{aligned}\tag{1.8}$$

Where the supremum and the infimum are taken over $(\mathcal{Q}(\mathbb{R}_+)\cap L^\infty)^2$.

Similarly to the Potts model, at the heuristic level (1.8) follows from the fact that given two independent random variables σ and τ under $\langle\cdot\rangle$, for every $s\in\mathscr{S}_D$, the matrices $\frac{\sigma\tau^*}{N}$ and $\left(\frac{\sigma_{s(d)}\cdot\tau_{s(d')}}{N}\right)_{1\leqslant d,d'\leqslant D}$ are equal in law under $\mathbb{E}\langle\cdot\rangle$. This means that every maximizing path should be of the form $\mathfrak{q} = (p_1,p_2)^\perp$ where $p_1,p_2\in\mathcal{Q}(\mathbb{R}_+)$ and the limit free energy should be a supremum over $\mathcal{Q}(\mathbb{R}_+)^2$.

We will also consider permutation-invariant models, where the interaction function ξ is further assumed to only depend on $\frac{\sigma_1\cdot\tau_1}{N},\dots,\frac{\sigma_D\cdot\tau_D}{N}$. In this case, we identify the map $R\mapsto\xi(R)$ defined on $\mathbb{R}^{D\times D}$ with the map $x\mapsto\xi(\mathrm{diag}(x))$ defined on $\mathbb{R}^D$. Since ξ only depends on the diagonal coefficients of its argument, we only need to keep track of the diagonal of the overlap matrix and the limit free energy can be written as a supremum over $\mathcal{Q}(\mathbb{R}_+)^D$. Performing the same heuristic computation as above, we then expect that under our permutation invariance assumptions, the limit free energy can in fact be written as a supremum over $\mathcal{Q}(\mathbb{R}_+)$. We will show that this is indeed the case.

Given a maximizing path r of the Parisi formula, the law of the random variable $r(U)$ is called a Parisi measure. When $D = 1$, it is known that there exists a unique Parisi measure [1]. In this case, the proof relies on a strict concavity property of the Parisi functional. However, when $D > 1$, for technical reasons this strict concavity property does not carry over well to $\mathcal{Q}(S^D_+)$ or $\mathcal{Q}(\mathbb{R}_+)^D$. But, in (1.9) below, the Parisi formula is written as a maximization over $\mathcal{Q}(\mathbb{R}_+)$, the set of 1-dimensional paths. Thanks to this, we can proceed as in [1] to prove uniqueness of Parisi measures.

Theorem 1.3. *Assume that ξ is convex on S^D_+, permutation-invariant and only depends on the diagonal coefficients of its argument, assume that P_1 is permutation-invariant. Then, for every $t > 0$,*

$$\lim_{N\to+\infty} \overline{F}_N(t) = \sup_{p\in\mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{\psi(p id_D) - t\int_0^1 \xi^*\left(\frac{p(u) id_D}{t}\right)\mathrm{d}u\right\}. \tag{1.9}$$

In addition, the supremum in (1.9) is reached at a unique $p^* \in \mathcal{Q}(\mathbb{R}_+)\cap L^\infty$.

Let us now comment on the inclusion of the term $-Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)$ in the definition of the free energy (1.2). Consider again the Potts model, that is $\xi(R) = \sum_{d,d'=1}^D R^2_{dd'}$ and $P_1 = \mathrm{Unif}\{e_1,\dots,e_D\}$. For any $\sigma \in \{e_1,\dots,e_D\}^N$, we have $\frac{\sigma\sigma^*}{N} = \mathrm{diag}(\alpha_1,\dots,\alpha_D)$ where

$$\alpha_d = \frac{1}{N}\#\{i \leqslant N, \sigma_i = e_d\}.$$

So, $\xi(\sigma\sigma^*/N) = \sum_{d=1}^D \alpha_d^2$ with $\sum_{d=1}^D \alpha_d = 1$ and $\alpha_d \geqslant 0$. This means that for the Potts model, the value of the correction term $-Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)$ is minimal on configurations of the form $\sigma = (e_d,\dots,e_d)$ and maximal on configurations satisfying $\alpha_d = \frac{1}{D}$. In particular, at least in the case of the Potts model, even though we are not constraining the system to stay in a balanced state, the correction term favors configurations $\sigma \in \mathbb{R}^{D\times N}$ which are balanced. Thankfully, this is not always the case and for some models the correction term $-Nt\xi\left(\frac{\sigma\sigma^*}{N}\right)$ can be removed. For example, if we assume that ξ only depends on the diagonal coefficients of its argument and P_1 is supported on $\{-1,1\}^D$, then the correction term is constant. In this case, the variational formula (1.9) can be rewritten as a variational formula for the uncorrected free energy.

Let $\alpha \geqslant 1$, for $\sigma \in \mathbb{R}^{2\times N}$ define

$$H_N^{\mathrm{BP+SK}}(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^N J_{ij}^{11} \sigma_{1i}\sigma_{1j} + \frac{1}{\sqrt{N}} \sum_{i,j=1}^N J_{ij}^{22} \sigma_{2i}\sigma_{2j} + \frac{1}{\sqrt{N}} \sum_{i,j=1}^N J_{ij}^{12} \sigma_{1i}\sigma_{2j},$$

where $J^{11} = (J^{11}_{ij})_{i,j\geqslant 1}$ and $J^{22} = (J^{22}_{ij})_{i,j\geqslant 1}$ are independent families of independent centered Gaussian random variables with variance $\alpha/2$ and $(J^{12}_{ij})_{i,j\geqslant 1}$ is a family of independent Gaussian random variables with variance 1 and independent of J^{11} and J^{22}. The process $H_N^{\mathrm{BP+SK}}$ satisfies (1.1) with,

$$\xi^{BP+SK}(R) = \frac{\alpha}{2} R_{11}^2 + \frac{\alpha}{2} R_{22}^2 + R_{11}R_{22}.$$

Choosing $P_1 = \mathrm{Unif}(\{-1,1\}^2)$ as the reference probability measure, Theorem 1.3 can be applied to $H_N^{\mathrm{BP+SK}}$ to discover that the system does not break its permutation invariance. One can wonder what happens when $0 \leqslant \alpha < 1$, in this case ξ^{BP+SK} is nonconvex on S^2_+ and there is no known generalization of Theorem 1.1. In Section 7, we will discuss some results that are applicable in this case.

When ξ is not assumed to be convex on S^D_+, to the best of our knowledge, there is no proof of the fact that $\overline{F}_N(t)$ converges as $N \to +\infty$. Building upon an interpolation argument developed in [2, Section 2.1], we will show that for permutation-invariant models where ξ only depends on the diagonal coefficients of its argument and is nonconvex on S^D_+, the right-hand side in (1.9) is an upper bound on the lim sup of the free energy.

Theorem 1.4. *Assume that ξ is permutation-invariant and only depends on the diagonal coefficients of its argument, assume that P_1 is permutation-invariant. Then, even when ξ is nonconvex, we have for every $t > 0$,*

$$\limsup_{N\to+\infty} \overline{F}_N(t) \leqslant \sup_{p\in\mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ \psi(p id_D) - t\int_0^1 \Xi^*\left(\frac{p(u) id_D}{t}\right) \mathrm{d}u \right\},$$

where $\Xi(x) = \frac{1}{D}\sum_{d=1}^D \xi(x_d, \dots, x_d)$.

In addition, we show that if no first order phase transition occurs, then the upper bound of Theorem 1.4 matches the lower bound derived in [17, 15]. The no first order phase transition hypothesis is formally written in (H_{ξ,P_1}) below, note that this hypothesis is true when ξ is convex. We refer to Theorem 7.10 below for a formal version of Theorem 1.5.

Theorem 1.5. *Assume that the hypotheses of Theorem 1.4 hold and further assume that no first order phases transition occurs, then for every $t > 0$, $\overline{F}_N(t)$ converges as $N \to +\infty$ and*

$$\lim_{N\to+\infty} \overline{F}_N(t) = \sup_{p\in\mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ \psi(p id_D) - t\int_0^1 \Xi^*\left(\frac{p(u) id_D}{t}\right) \mathrm{d}u \right\},$$

where $\Xi(x) = \frac{1}{D}\sum_{d=1}^D \xi(x_d, \dots, x_d)$.

As explained in Remark 7.12 below, it is plausible that there exists $t_c > 0$ such that no first order phase transition occurs on $[0, t_c)$. In this case, the argument used to prove Theorem 1.5 allows to identify the limit free energy on $[0, t_c]$.

1.3. Motivations.

Most of the proof we derive below rely on the following result, which we state informally.

Theorem 1.6 ([16, 7]). *Assume that ξ is convex on S^D_+, then for every $t \geqslant 0$,*

$$\lim_{N \to +\infty} \overline{F}_N(t) = f(t, 0),$$

where $f : \mathbb{R}_+ \times \mathcal{Q}(S^D_+) \to \mathbb{R}$ is the unique solution of

$$\begin{cases} \partial_t f - \int \xi(\nabla f) = 0 \text{ on } \mathbb{R}_+ \times \mathcal{Q}(S^D_+) \\ f(0, \cdot) = \psi \text{ on } \mathcal{Q}(S^D_+). \end{cases} \tag{1.10}$$

In Section 2 we will give a precise definition of f and ψ, we will also explain the precise meaning of (1.10). In the context of Theorem 1.2 and Theorem 1.3, this approach trough partial differential equations does not seem to possess any clear advantage when compared with the methods developed in [5]. The true power of Theorem 1.6 is revealed when dealing with nonconvex models. When ξ is nonconvex, the Parisi formula completely breaks down. To the best of our knowledge, until recently, it seems that there was no clear conjecture on what the limit of the free energy should be in this case. In [16, Conjecture 2.6], it is proposed that results such as Theorem 1.6 should generalize to nonconvex models. Further development in [17, 15] have led to the following lower bound for the free energy of nonconvex models.

Theorem 1.7 ([17, 15]). *For every $t \geqslant 0$,*

$$\liminf_{N \to +\infty} \overline{F}_N(t) \geqslant f(t, 0),$$

where $f : \mathbb{R}_+ \times \mathcal{Q}(S^D_+) \to \mathbb{R}$ is the unique solution of

$$\begin{cases} \partial_t f - \int \xi(\nabla f) = 0 \text{ on } \mathbb{R}_+ \times \mathcal{Q}(S^D_+) \\ f(0, \cdot) = \psi \text{ on } \mathcal{Q}(S^D_+). \end{cases} \tag{1.11}$$

To identify the limit free energy of nonconvex models, a possible approach is to prove an upper bound for $\limsup_{N \to +\infty} \overline{F}_N(t)$ and compare it with $f(t, 0)$. In Section 7, we will show that

$$f(t, 0) \leqslant \liminf_{N \to +\infty} \overline{F}_N(t) \leqslant \limsup_{N \to +\infty} \overline{F}_N(t) \leqslant g(t, 0),$$

where g is the solution of an equation similar to (1.10) (see (7.11) below). We are able to phrase the upper bound this way thanks to the fact that the proofs of Theorem 1.2 and Theorem 1.3 are formalized in the language of partial differential equations. In particular, we will show that under some unproven regularity assumption on f, $f(t, 0)$ and $g(t, 0)$ are equal. We refer to Section 7.4 for a more precise discussion.

1.4. **Organization of the paper.** We only consider permutation-invariant models. That is, models with ξ permutation-invariant and $P_N = P_1^{\otimes N}$ with P_1 permutation-invariant and compactly supported. Without loss of generality, we will assume that P_1 is supported inside the unit ball of $\mathbb{R}^D$. Excluding Section 7, ξ is always assumed to be convex on S^D_+.

In Section 2 we will give a precise definition of the functions f and ψ appearing in Theorem 1.6, we will also explain the precise meaning of (1.10). In Section 3 we will define an appropriate notion of solution for (1.10) following [10]. In particular, in Section 2 we introduce a variant of the free energy $\overline{F}_N(t)$ which will depend on $t \in \mathbb{R}_+$ and an extra parameter $\mathfrak{q} \in \mathcal{Q}(S^D_+)$. This enriched version of the free energy will be denoted $\overline{F}_N(t,\mathfrak{q})$ and satisfies $\overline{F}_N(t,0) = \overline{F}_N(t)$. It is through the introduction of this extra parameter that it is possible to obtain a partial differential equation for the limiting objects. In section 4 we will introduce some basic notations on permutation-invariant matrices and paths that will be useful for the later sections. We make the elementary but important remark that the set of permutation-invariant matrices is isomorphic to $\mathbb{R}^2$ and the set of permutation-invariant vectors is isomorphic to $\mathbb{R}$. The bulk of our analysis is in Section 5, in this section we start by showing that the enriched free energy is permutation-invariant, namely for every permutation $s \in \mathscr{S}_D$,

$$\overline{F}_N(t,(\mathfrak{q}_{dd'})_{1\leqslant d,d\leqslant D}) = \overline{F}_N(t,(\mathfrak{q}_{s(d)s(d')})_{1\leqslant d,d\leqslant D}).$$

Thanks to this property and (1.10), we can write a partial differential equation satisfied by $(t,\mathfrak{q}) \mapsto \lim_{N\to+\infty} \overline{F}_N(t,\mathfrak{q})$ on the set of permutation-invariant paths. Since the set of permutation-invariant paths in $\mathcal{Q}(S^D_+)$ is isomorphic to $\mathcal{Q}(\mathbb{R}^2_+)$, we will deduce Theorem 1.2. Moreover, when ξ is further assumed to only depend on the diagonal coefficient of its argument, it is even possible to write a partial differential equation on the set of permutation-invariant paths in $\mathcal{Q}(\mathbb{R}_+)^D$ which is isomorphic to $\mathcal{Q}(\mathbb{R}_+)$, this will yield the variational formula (1.9) of Theorem 1.3. Using a strict concavity property of the Parisi functional, in Section 6 we will deduce the existence and the uniqueness of an optimizer in the variational formula (1.9). Finally, in Section 7 we assume that ξ is permutation-invariant and depends only on the diagonal coefficients of its argument, but we do not assume that ξ is convex. We will show that under those hypotheses if ξ admits an absolutely convergent power series and satisfies (1.1), then

$$\forall x \in \mathbb{R}^D_+,\ \xi(x) \leqslant \frac{1}{D}\sum_{d=1}^{D} \xi(x_d,\dots,x_d).$$

With this inequality, we will prove Theorem 1.4 via a simple interpolation argument. To conclude, we will explore how this upper bound is related to the lower bound for the free energy of nonconvex models derived in [17, 15] and we will prove a rigorous version of Theorem 1.5, namely Theorem 7.10.

1.5. **Acknowledgement.** I warmly thank Jean-Christophe Mourrat for countless fascinating discussions on spin glasses and many helpful comments and insights during this project.

2. The enriched free energy

The method we propose relies on the fact that the limit free energy $f(t) = \lim \overline{F}_N(t)$ solves a partial differential equation. At this stage, f only depends on one parameter, in order to obtain a partial differential equation, we are going to introduce an extra parameter. The nature of this parameter will not be the same for models that satisfy (1.1) and models that satisfy (1.1) under the additional assumption that ξ only depends on the diagonal coefficients of its argument. We will provide the construction for the first class of models and then explain the additional adjustments that can be made when ξ only depends on the diagonal coefficients of its argument.

Here S^D_+ denotes the set of positive semi-definite symmetric matrices in $\mathbb{R}^{D\times D}$ and S^D_{++} the set of positive symmetric matrices in $\mathbb{R}^{D\times D}$. We equip $R^{D\times D}$ with the order $A \leqslant B$ if and only $B - A \in S^D_+$. The extra parameter we introduce in the definition of the free energy belongs to the space

$$\mathcal{Q}(S^D_+) = \{\mathfrak{q} : [0,1) \to S^D_+ \big| \, \mathfrak{q} \text{ is càdlàg and nondecreasing}\}. \tag{2.1}$$

In general paths in $\mathcal{Q}(S^D_+)$ are not defined at $u = 1$, but if $\mathfrak{q} \in \mathcal{Q}(S^D_+) \cap L^\infty$ we can set

$$\mathfrak{q}(1) = \lim_{u \uparrow 1} \mathfrak{q}(u) \in S^D_+.$$

We let $\mathfrak{R}$ be a Poisson Dirichlet cascade, briefly $\mathfrak{R}$ is a specific random probability measure on the unit sphere of a separable Hilbert space $(\mathfrak{H}, \wedge)$, such that given two independent random variables $\alpha, \alpha' \in \mathfrak{H}$ sampled from $\mathfrak{R}$, the random variable $\alpha \wedge \alpha'$ is uniformly distributed in $[0,1]$ (see [20, Section 2] and [7, Section 4]). We let $\langle \cdot \rangle_{\mathfrak{R}}$ denote the expectation with respect to $\mathfrak{R}^{\otimes \mathbb{N}}$ and $\mathfrak{U}$ the support of $\mathfrak{R}$. By construction of $\mathfrak{R}$, almost surely $\mathfrak{U}$ is an ultrametric set, that is almost surely for every $\alpha^1, \alpha^2, \alpha^3 \in \mathfrak{U}$,

$$\alpha^1 \wedge \alpha^3 \geqslant \min\left(\alpha^1 \wedge \alpha^2, \alpha^2 \wedge \alpha^3\right).$$

Let $\mathfrak{q} \in \mathcal{Q}(S^D_+) \cap L^\infty$, according to [7, Propsoition 4.1] $\mathfrak{R}$-almost surely, there exists a $\mathbb{R}^D$-valued centered Gaussian process $(w^{\mathfrak{q}}(\alpha))_{\alpha \in \mathfrak{U}}$ such that

$$\mathbb{E}\left[w^{\mathfrak{q}}(\alpha)\left(w^{\mathfrak{q}}(\alpha')\right)^*\right] = \mathfrak{q}(\alpha \wedge \alpha').$$

Let $N \geqslant 1$, and consider $(W^{\mathfrak{q}}_N(\alpha))_{\alpha \in \mathfrak{U}}$ be a $(\mathbb{R}^D)^N$-valued random process whose coordinates are independent and have the same law as $(w^{\mathfrak{q}}(\alpha))_{\alpha \in \mathfrak{U}}$. We set

$$F_N(t, \mathfrak{q}) = -\frac{1}{N} \log \iint \exp(H^{t,\mathfrak{q}}_N(\sigma, \alpha)) dP_N(\sigma) d\mathfrak{R}(\alpha), \tag{2.2}$$

where

$$H_N^{t,\mathfrak{q}}(\sigma,\alpha) = \sqrt{2t}H_N(\sigma) - tN\xi\left(\frac{\sigma\sigma^*}{N}\right) + \sqrt{2}W_N^{\mathfrak{q}}(\alpha)\cdot\sigma - \mathfrak{q}(1)\cdot\sigma\sigma^*.$$

We let $\overline{F}_N(t,\mathfrak{q}) = \mathbb{E}F_N(t,\mathfrak{q})$, the function $\overline{F}_N : \mathbb{R}_+ \times \left(\mathcal{Q}(S^D_+)\cap L^\infty\right) \to \mathbb{R}$ is Lipschitz, more precisely according to [7, Proposition 5.1], we have

$$|\overline{F}_N(t,\mathfrak{q}) - \overline{F}_N(t',\mathfrak{q}')| \leqslant |\mathfrak{q}-\mathfrak{q}'|_{L^1} + |t-t'|\sup_{|a|\leqslant 1}|\xi(a)|. \tag{2.3}$$

Here, $|a|$ denotes the Frobenius norm of $a \in \mathbb{R}^{D\times D}$, that is the norm associated to the scalar product

$$a\cdot b = \sum_{d,d'=1}^{D} a_{dd'}b_{dd'}.$$

Using (2.3), we can then extend $\overline{F}_N$ by continuity to $\mathbb{R}_+ \times \left(\mathcal{Q}(S^D_+)\cap L^1\right)$. Note that if 0 denotes the null path, we have $\overline{F}_N(t,0) = \overline{F}_N(t)$. So the function we have defined is indeed an extension of the free energy.

Let us now introduce the notion of differentiability that we will use to formulate the partial differential equation (1.10). Given $\mathfrak{q} \in \mathcal{Q}(S^D_+)\cap L^2$, we let

$$\mathrm{Adm}(\mathfrak{q}) = \{\kappa \in L^2 \big| \ \exists r > 0,\ \forall t \in [0,r],\ \mathfrak{q} + t\kappa \in \mathcal{Q}(S^D_+)\cap L^2\}.$$

The set $\mathrm{Adm}(\mathfrak{q})$ is the set of admissible directions at $\mathfrak{q}$. Let $g : \mathcal{Q}(S^D_+)\cap L^2 \to \mathbb{R}$, we say that g is Gateaux differentiable at $\mathfrak{q} \in \mathcal{Q}(S^D_+)\cap L^2$ when for every $\kappa \in \mathrm{Adm}(\mathfrak{q})$, the following limit exists

$$g'(\mathfrak{q};\kappa) = \lim_{t\downarrow 0}\frac{g(\mathfrak{q}+t\kappa) - g(\mathfrak{q})}{t},$$

and there exists a unique $d \in L^2$ such that for every $\kappa \in \mathrm{Adm}(\mathfrak{q})$,

$$g'(\mathfrak{q};\kappa) = \langle d,\kappa\rangle_{L^2}.$$

In this case, the vector $d \in L^2$ is denoted $\nabla g(\mathfrak{q})$ and we call it the Gateaux derivative of g at $\mathfrak{q}$. Note that $\nabla g(\mathfrak{q}) \in L^2$, so we can make sense of expressions of the form $\int_0^1 \xi(\nabla g(\mathfrak{q})(u))\mathrm{d}u$, provided that the integral converges. Such expressions will be simply abbreviated by $\int \xi(\nabla g(\mathfrak{q}))$ in what follows.

We define $\mathcal{Q}_\uparrow(S^D_+)$ as the set of paths $\mathfrak{q} \in \mathcal{Q}(S^D_+)\cap L^2$ such that $\mathfrak{q}(0) = 0$, and there exists a constant $c > 0$ such that $u \mapsto \mathfrak{q}(u) - cu\mathrm{id}_D$ is nondecreasing and for all $u < v$,

$$\mathrm{Ellipt}(\mathfrak{q}(v) - \mathfrak{q}(u)) \leqslant \frac{1}{c}.$$

Here, for $m \in S^D_{++}$, $\mathrm{Ellipt}(m)$ denotes the ratio of the biggest and smallest eigenvalue of m. When ξ is convex, which will always be the case except in Section 7, it can be shown that $\overline{F}_N$ converges pointwise to some Lipschitz function $f : \mathbb{R}_+ \times \left(\mathcal{Q}(S^D_+)\cap L^1\right) \to \mathbb{R}$. Furthermore, according to [7, Propositions 7.2 & 8.6], $f : \mathbb{R}_+ \times (\mathcal{Q}(S^D_+)\cap L^2) \to \mathbb{R}$ is Gateaux differentiable on

$(0,+\infty) \times (\mathcal{Q}_\uparrow(S^D_+) \cap L^\infty)$ and satisfies

$$\begin{cases} \partial_t f - \int \xi(\nabla f) = 0 & \text{on } (0,+\infty) \times (\mathcal{Q}_\uparrow(S^D_+) \cap L^\infty) \\ f(0,\cdot) = \psi & \text{on } \mathcal{Q}_\uparrow(S^D_+) \cap L^\infty, \end{cases} \tag{2.4}$$

where we have defined $\psi = \lim_{N\to+\infty} \overline{F}_N(0,\cdot)$. In fact, since we have assumed that $P_N = P_1^{\otimes N}$ we simply have

$$\psi = \overline{F}_1(0,\cdot). \tag{2.5}$$

The fact that f is a solution of (2.4) at every point where it is differentiable is a consequence of the cavity computations conducted in [7] (see [7, Propostion 7.2]), this result is robust and holds true even when ξ is nonconvex. The real miracle is the fact that f is differentiable at every (nice) points, this is proven in [7, Propostion 8.6] and relies strongly on the fact that ξ is convex on S^D_+. We stress that in general, nonlinear partial differential equations such as (2.4) do not always have a differentiable solution. Therefore, to study (2.4) and other similar equations, we will need to appeal to the notion of viscosity solution. this notion will be introduced precisely in Section 3. Also note, that according to [7, Proposition 3.6], we have $\nabla f(t,\mathfrak{q}) \in \mathcal{Q}(S^D_+) \cap L^2$. This means that the behavior of (2.4) is only governed by the restriction of ξ to S^D_+, this is why we can assume only that ξ is convex on S^D_+.

In this framework, the Parisi formula can be seen as a consequence of a variational representation for the viscosity solution of (2.4). According to [9, Theorem 4.6 (2)], for $(t,\mathfrak{q}) \in \mathbb{R}_+ \times \mathcal{Q}(S^D_+) \cap L^2$, we have

$$f(t,\mathfrak{q}) = \sup_{\mathfrak{p}\in\mathcal{Q}(S^D_+)\cap L^\infty} \inf_{\mathfrak{r}\in\mathcal{Q}(S^D_+)\cap L^\infty} \left\{ \psi(\mathfrak{q}+\mathfrak{p}) - \langle \mathfrak{p},\mathfrak{r}\rangle_{L^2} + t\int_0^1 \xi(\mathfrak{r}(u))\mathrm{d}u \right\}.$$

At $\mathfrak{q} = 0$ this yields, $\lim_{N\to+\infty} \overline{F}_N(t) = \sup_{\mathfrak{p}\in\mathcal{Q}(S^D_+)} \mathcal{P}_{t,\xi,P_1}(\mathfrak{p})$, where

$$\mathcal{P}_{t,\xi,P_1}(\mathfrak{p}) = \psi(\mathfrak{p}) - \sup_{\mathfrak{r}\in\mathcal{Q}(S^D_+)\cap L^\infty} \left\{ \langle \mathfrak{p},\mathfrak{r}\rangle_{L^2} - t\int_0^1 \xi(\mathfrak{r}(u))\mathrm{d}u \right\}. \tag{2.6}$$

As explained in the proof of [7, Proposition 8.1], one can recover the "usual" Parisi functional from $\mathcal{P}_{t,\xi,P_1}$ by plugging in paths of the form $\nabla\xi\circ\mathfrak{p}$. If we define $\theta(a) = a\cdot\nabla\xi(a) - \xi(a)$, we have

$$\mathcal{P}_{t,\xi,P_1}(t\nabla\xi\circ\mathfrak{p}) = \psi(t\nabla\xi\circ\mathfrak{p}) - t\int_0^1 \theta(\mathfrak{p}(u))\mathrm{d}u.$$

Before moving to the next section, we explain the adjustments that can be made to obtain a simpler Hamilton-Jacobi equation when ξ is assumed to only depend on the diagonal coefficients of its argument. Henceforth, we will refer to models with this additional property as diagonal models. For those models, we identify the function $A \mapsto \xi(A)$ defined on $\mathbb{R}^{D\times D}$ and the function $x \mapsto \xi(\mathrm{diag}(x))$ defined on $\mathbb{R}^D$. Recall that heuristically, the paths $\mathfrak{q}$ are to be understood as encoding the limiting distribution of the overlap matrix, $\frac{\sigma\tau^*}{N}$. When the model is diagonal, we do not need to keep track of the full overlap matrix and encoding the distribution of the diagonal coefficients is enough. To this end, let us introduce another space of paths,

$$\mathcal{Q}(\mathbb{R}_+^D) = \{q : [0,1) \to \mathbb{R}_+^D \,\big|\, q \text{ is càdlàg nondecreasing}\}.$$

Here, by q is nondecreasing we mean that for every $u \leqslant v$, $q(v) - q(u) \in \mathbb{R}_+^D$. Note that $\mathcal{Q}(\mathbb{R}_+^D)$ is isomorphic to the subset of $\mathcal{Q}(S_+^D)$ composed of the paths that are valued in the set of diagonal matrices. If $q \in \mathcal{Q}(\mathbb{R}_+^D)$ we denote by $\mathrm{diag}(q)$ the associated diagonal matrix valued path. We then define $\mathcal{Q}_\uparrow(\mathbb{R}_+^D)$ as the set of paths $q \in \mathcal{Q}(\mathbb{R}_+^D)$ such that the path $\mathrm{diag}(q)$ belongs to $\mathcal{Q}_\uparrow(S_+^D)$. In order to not get mixed up with the previous definition of enriched free energy, we will denote by $\overline{F}_N^{\mathrm{diag}}$ the restriction of $\overline{F}_N$ to $\mathcal{Q}(\mathbb{R}_+^D)$. That is, for every $(t,q) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}_+^D) \cap L^1)$,

$$\overline{F}_N^{\mathrm{diag}}(t,q) = \overline{F}_N(t, \mathrm{diag}(q)).$$

As previously, the sequence $(\overline{F}_N^{\mathrm{diag}}(t,q))_N$ converges to $f(t,\mathrm{diag}(q))$. We let $f^{\mathrm{diag}}(t,q) = f(t,\mathrm{diag}(q))$, the function f^{diag} is Gateaux differentiable on $(0,+\infty) \times \left(\mathcal{Q}_\uparrow(\mathbb{R}_+^D) \cap L^\infty\right)$ and solves

$$\begin{cases} \partial_t f^{\mathrm{diag}} - \int \xi(\nabla f^{\mathrm{diag}}) = 0 & \text{on } (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}_+^D) \cap L^\infty) \\ f^{\mathrm{diag}}(0,\cdot) = \psi^{\mathrm{diag}} & \text{on } \mathcal{Q}_\uparrow(\mathbb{R}_+^D) \cap L^\infty, \end{cases} \tag{2.7}$$

where we have defined $\psi^{\mathrm{diag}}(q) = \psi(\mathrm{diag}(q))$. This can be checked by differentiating f^{diag} and using (2.4). The point of this derivation is that we can now express the limit free energy of diagonal models as the value at $(t,0)$ of the solution of a Hamilton-Jacobi equation on $\mathcal{Q}(\mathbb{R}_+^D)$ rather than $\mathcal{Q}(S_+^D)$. This yields a different variational representation for the limit free energy.

3. Hamilton-Jacobi equations on closed convex cones

As explained above, Hamilton-Jacobi equations on closed convex cones, such as (2.4) and (2.7), will play an important role in this paper. Here, we summarize some known results on those equations. We will focus on the theory of Hamilton-Jacobi equations on closed convex cones in finite dimensional Hilbert spaces. But, most of the result exposed below remain valid in infinite dimensional Hilbert spaces [9]. We will use both settings in the following sections.

Let $(\mathcal{H}, \langle \cdot,\cdot \rangle_{\mathcal{H}})$ be a finite dimensional Hilbert space. Let $\mathcal{C} \subseteq \mathcal{H}$, we say that $\mathcal{C}$ is a closed convex cone when $\mathcal{C}$ is a closed set and for every $s,t \geqslant 0$, $x,y \in \mathcal{C}$, we have

$$sx + ty \in \mathcal{C}.$$

In what follows, we fix a nonempty closed convex cone $\mathcal{C} \subseteq \mathcal{H}$. We will assume that the interior of $\mathcal{C}$ is nonempty and that the only vector belonging to $\mathcal{C}$ and $-\mathcal{C}$ simultaneously is 0. When dealing with infinite dimensional Hilbert spaces like in [9], $\mathcal{C}$ is allowed to have empty interior, but for simplicity we do not consider this possibility in this section.

Let $\psi : \mathcal{C} \to \mathbb{R}$ be a Lipschitz function and $H : \mathcal{H} \to \mathbb{R}$ be such that $H\big|_{\mathcal{C}}$ is locally Lipschitz. We are interested in equations of the form

$$\begin{cases} \partial_t v - H(\nabla v) = 0 \text{ on } (0,+\infty) \times \mathring{\mathcal{C}} \\ v(0,\cdot) = \psi \text{ on } \mathcal{C}. \end{cases} \tag{3.1}$$

When $\mathcal{C} = \mathbb{R}^d$, there is already a rich theory for equations like (3.1) [12, Section 10]. One of the main accomplishments of this theory is the introduction of a suitable notion of solution. Equation (3.1) may not have any differentiable solution [12, Section 3.3 Example 6]. But, by introducing the notion of viscosity solutions (see Definition 3.1 below), it is possible to guarantee existence and uniqueness of solutions under some mild conditions, provided that we allow non-differentiable functions to solve (3.1).

Most of the theory of viscosity solutions on $\mathbb{R}^d$ can be adapted for Hamilton-Jacobi equations on closed convex cones [10]. In principle, when $\mathbb{R}^d$ is replaced by $\mathcal{C}$, some boundary conditions should be enforced to guarantee that (3.1) admits a unique solution. But this requirement can be bypassed provided that some monotony conditions on ψ and H hold.

Given $\mathcal{D} \subseteq \mathcal{C}$, a function $\phi : (0,+\infty) \times \mathcal{D} \to \mathbb{R}$ is said to be differentiable at $(t,x) \in (0,+\infty) \times \mathcal{D}$ when there exists a unique $(a,p) \in \mathbb{R} \times \mathcal{H}$ such that the following holds as $(s,y) \in (0,+\infty) \times \mathcal{D}$ tends to (t,x),

$$\phi(s,y) = \phi(t,x) + (s-t)a + \langle x-y, p\rangle_{\mathcal{H}} + O\left(|s-t|^2 + |x-y|^2_{\mathcal{H}}\right).$$

In this case, (a,p) is denoted $(\partial_t \phi(t,x), \nabla \phi(t,x))$. When ϕ is differentiable at every point in $(0,+\infty) \times \mathcal{D}$ and the function $(t,x) \mapsto (\partial_t \phi(t,x), \nabla \phi(t,x))$ is continuous, we say that ϕ is smooth. Note that the notions of differentiability and smoothness defined here make sense even when x does not belong to the interior of $\mathcal{C}$ and does not requires $\mathcal{D}$ to be an open set.

Definition 3.1 (Viscosity solutions)**.**

(1) An upper semi-continuous function $v : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}$ is a viscosity subsolution of (3.1) when for every $(t,x) \in (0,+\infty) \times \mathring{\mathcal{C}}$ and every smooth function $\phi : (0,+\infty) \times \mathring{\mathcal{C}} \to \mathbb{R}$ such that $v - \phi$ has a local maximum at (t,x), we have

$$\partial_t \phi(t,x) - H(\nabla \phi(t,x)) \leqslant 0.$$

(2) A lower semi-continuous function $v : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}$ is a viscosity supersolution of (3.1) when for every $(t,x) \in (0,+\infty) \times \mathring{\mathcal{C}}$ and every smooth function $\phi : \mathbb{R}_+ \times \mathring{\mathcal{C}} \to \mathbb{R}$ such that $v - \phi$ has a local minimum at (t,x), we have

$$\partial_t \phi(t,x) - H(\nabla \phi(t,x)) \geqslant 0.$$

(3) A continuous function $v : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}$ is a viscosity solution when it is both a viscosity subsolution and a viscosity supersolution.

We define the dual cone $\mathcal{C}^*$ of the cone $\mathcal{C}$ by setting,

$$\mathcal{C}^* = \{x \in \mathcal{H} \big| \, \forall y \in \mathcal{C}, \; \langle x, y\rangle_{\mathcal{H}} \geqslant 0\}. \tag{3.2}$$

The dual cone $\mathcal{C}^*$ is a closed convex cone and satisfies $(\mathcal{C}^*)^* = \mathcal{C}$. Given $\mathcal{A}, \mathcal{B} \subseteq \mathcal{H}$, we say that $g : \mathcal{A} \to \mathbb{R}$ is $\mathcal{B}$-nondecreasing when for every $x, x' \in \mathcal{A}$, if $x - x' \in \mathcal{B}$ then $g(x) \geqslant g(x')$. Note that, if $h : \mathcal{C} \to \mathbb{R}$ is differentiable, then h is $\mathcal{C}^*$-nondecreasing if and only if $\nabla h(\mathcal{C}) \subseteq \mathcal{C}$. We define $\mathcal{V}(\mathcal{C})$ to be the set of functions $v : \mathbb{R}_+ \times \mathcal{C} \to \mathbb{R}$, such that for every $t \geqslant 0$, $v(t, \cdot)$ is $\mathcal{C}^*$-nondecreasing and such that the following estimates hold

$$\sup_{\substack{t>0 \\ x\in\mathcal{C}}} \frac{|v(0,x) - v(t,x)|}{t} < +\infty \text{ and } \sup_{t>0} \|v(t,\cdot)\|_{\mathrm{Lip}} < +\infty. \tag{3.3}$$

Here, $\|v(t,\cdot)\|_{\mathrm{Lip}}$ denotes the optimal Lipschitz constant of $v(t,\cdot)$. The following theorem is extracted from [10, Theorem 1.2].

Theorem 3.2 ([10])**.** *Let $\psi : \mathcal{C} \to \mathbb{R}$ be a Lipschitz and $\mathcal{C}^*$-nondecreasing function and let $H : \mathcal{H} \to \mathbb{R}$ be such that $H\big|_{\mathcal{C}}$ is $\mathcal{C}^*$-nondecreasing and locally Lipschitz. Then, the Hamilton-Jacobi equation*

$$\begin{cases} \partial_t v - H(\nabla v) = 0 \;\; \text{on } (0,+\infty) \times \mathring{\mathcal{C}} \\ v(0,\cdot) = \psi \;\; \text{on } \mathcal{C}, \end{cases} \tag{3.4}$$

admits a unique viscosity solution in $\mathcal{V}(\mathcal{C})$ that will be denoted v.

Theorem 3.2 ensures that (3.1) is well posed as long as ψ and H are $\mathcal{C}^*$-nondecreasing. When H is convex and some additional requirements are put on the cone $\mathcal{C}$ it is possible to obtain an explicit representation for the unique viscosity solution. We refer to this representation as the Hopf-Lax representation of the viscosity solution.

Given $g : \mathcal{C} \to \mathbb{R} \cup \{+\infty\}$, we define the (monotone) convex conjugate of g over $\mathcal{C}$ by

$$g^*(y) = \sup_{x\in\mathcal{C}} \{\langle x, y\rangle_{\mathcal{H}} - g(x)\}. \tag{3.5}$$

When $\mathcal{C} = \mathbb{R}^d$, the functions that satisfy $g^{**} = g$ are exactly the lower semicontinuous convex functions [25, Section 12]. If we let

$$g^{**}(x) = \sup_{y\in\mathcal{C}^*} \{\langle x, y\rangle_{\mathcal{H}} - g(y)\}$$

denote the convex conjugate of g^* over $\mathcal{C}^*$, we have that g^{**} is lower semicontinuous, convex and $\mathcal{C}$-nondecreasing. So, the functions satisfying $g = g^{**}$ must be lower semicontinuous, convex and $\mathcal{C}$-nondecreasing.

Definition 3.3. We say that a closed convex cone $\mathcal{C}$ has the Fenchel-Moreau property when for every $g : \mathcal{C} \to \mathbb{R} \cup \{+\infty\}$ not identically equal to $+\infty$, we have $g^{**} = g$ if and only if g is lower semicontinuous, convex and $\mathcal{C}$-nondecreasing.

Theorem 3.4 ([10])**.** *Assume that $\mathcal{C}$ has the Fenchel-Moreau property. Let $\psi : \mathcal{C} \to \mathbb{R}$ be a Lipschitz and $\mathcal{C}^*$-nondecreasing function and let $H : \mathcal{H} \to \mathbb{R}$ be such that $H\big|_{\mathcal{C}}$ is $\mathcal{C}^*$-nondecreasing, locally Lipschitz and* convex*. Then, the unique viscosity solution v of (3.4) admits the Hopf-Lax representation, that is for every $(t,x) \in \mathbb{R}_+ \times \mathcal{C}$,*

$$v(t,x) = \sup_{y \in \mathcal{C}} \inf_{z \in \mathcal{C}} \left\{ \psi(x+y) - \langle y, z \rangle_{\mathcal{H}} + tH(z) \right\}.$$

Most of the cones that we will encounter in this document have the Fenchel-Moreau property. Examples of cones with the Fenchel-Moreau property include $\mathbb{R}^D_+$ and S^D_+. We refer to [8] for an in-depth study of Fenchel-Moreau cones. Note that the cones $\mathcal{Q}(S^D_+)$, $\mathcal{Q}(\mathbb{R}^D_+)$ and $\mathcal{Q}^j(S^D_+)$, $\mathcal{Q}^j(\mathbb{R}^D_+)$ defined below are all Fenchel-Moreau cones, for those cones the Hopf-Lax representation will therefore be available. Note that when ψ is assumed to be convex and H is possibly nonconvex, another variational representation for the viscosity solution is available [10, Proposition 6.3]. Finally, we point out that, as proven in [10, Theorem 1.2 (1)], the comparison principle also remains valid for (3.1). Note that crucially, to use the comparison principle, the condition to be a viscosity solution needs only to be checked on the interior of $\mathcal{C}$. This means that in practice when dealing with solutions of (3.1), the boundary points of $\mathcal{C}$ do not need to be considered.

Theorem 3.5 ([10])**.** *Let $u, v \in \mathcal{V}(\mathcal{C})$, assume that u is a viscosity subsolution and v is a viscosity supersolution. Then,*

$$\sup_{(t,x) \in \mathbb{R}_+ \times \mathcal{C}} \left\{ u(t,x) - v(t,x) \right\} = \sup_{x \in \mathcal{C}} \left\{ u(0,x) - v(0,x) \right\}.$$

We conclude this section by using the comparison principle to prove a stability result for approximate strong solutions of (3.1) on $\mathring{\mathcal{C}}$.

Proposition 3.6. *Let $\psi : \mathcal{C} \to \mathbb{R}$ be a Lipschitz and $\mathcal{C}^*$-nondecreasing function, and let $H : \mathcal{H} \to \mathbb{R}$ be such that $H\big|_{\mathcal{C}}$ is $\mathcal{C}^*$-nondecreasing and locally Lipschitz. Let $g \in \mathcal{V}(\mathcal{C})$, assume that $g(0,\cdot) = \psi$, that g is differentiable on $(0,+\infty) \times \mathring{\mathcal{C}}$ and that there exists a constant $c > 0$ such that*

$$\sup_{t>0, x \in \mathring{\mathcal{C}}} \left| \partial_t g(t,x) - H(\nabla g(t,x)) \right| \leqslant c. \tag{3.6}$$

Then, for every $t \geqslant 0$, $\sup_{x \in \mathcal{C}} |v(t,x) - g(t,x)| \leqslant ct$. Where v is the unique viscosity solution of (3.4).

Proof. Define $\widetilde{g}(t,x) = g(t,x) + ct$, we have $\widetilde{g} \in \mathcal{V}(\mathcal{C})$ and $\widetilde{g}(0,\cdot) = \psi$. In addition, $\widetilde{g}$ is differentiable on $(0,+\infty) \times \mathring{\mathcal{C}}$ and

$$\partial_t \widetilde{g} - H(\nabla \widetilde{g}) \geqslant 0 \text{ on } (0,+\infty) \times \mathring{\mathcal{C}}.$$

Therefore $\widetilde{g}$ is supersolution of (3.4) on $\mathring{\mathcal{C}}$. According to the comparison principle [10, Theorem 1.2 (1)], we have $v - \widetilde{g} \leqslant 0$ on $\mathbb{R}_+ \times \mathcal{C}$. So for every $x \in \mathcal{C}$,

$$v(t,\cdot) - g(t,\cdot) \leqslant ct.$$

Similarly, we show that $(t,x) \mapsto g(t,x) - ct$ is a subsolution and use the comparison principle to deduce $g(t,\cdot) - v(t,\cdot) \leqslant ct$ on $\mathbb{R}_+ \times \mathcal{C}$. Combining those two bounds, we obtain the desired result. □

4. Permutation-invariant objects

In this section, we introduce basic notations and results for matrix valued paths. Those results will be useful in later sections for the analysis of Hamilton-Jacobi equations.

4.1. **Permutation-invariant matrices.** Let $m \in \mathbb{R}^{D\times D}$, for every $s \in \mathscr{S}_D$ we define $m^s = (m_{s(d),s(d')})_{1\leqslant d,d'\leqslant D}$. When for every $s \in \mathscr{S}_D$, $m^s = m$ we say that m is permutation-invariant. An in depth study of permutation-invariant matrices is conducted in [18], for the sake of completeness we extract the following result. Recall that id_D denotes the $D \times D$ identity matrix and $\mathbf{1}_D$ denotes the $D \times D$ matrix whose coefficients are all equal to 1.

Proposition 4.1 ([18])**.** *Let $m \in \mathbb{R}^{D\times D}$ be permutation-invariant. We have*

$$m = \begin{pmatrix} a & t & \cdots & \cdots & t \\ t & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & t \\ t & \cdots & \cdots & t & a \end{pmatrix} = a id_D + t\left(\mathbf{1}_D - id_D\right), \tag{4.1}$$

where $a = m_{11}$ and $t = m_{12}$. The matrix m admits two eigenvalues $\lambda_1 = a - t$ and $\lambda_2 = a - t + Dt$ (or 1 eigenvalue when $t = 0$) and can be expressed in terms of its eigenvalues via

$$\begin{aligned} m &= \frac{1}{D}\begin{pmatrix} \lambda_2 + (D-1)\lambda_1 & \lambda_2 - \lambda_1 & \cdots & \cdots & \lambda_2 - \lambda_1 \\ \lambda_2 - \lambda_1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \lambda_2 - \lambda_1 \\ \lambda_2 - \lambda_1 & \cdots & \cdots & \lambda_2 - \lambda_1 & \lambda_2 + (D-1)\lambda_1 \end{pmatrix} \\ &= \lambda_1\left(id_D - \frac{\mathbf{1}_D}{D}\right) + \lambda_2\frac{\mathbf{1}_D}{D}. \end{aligned} \tag{4.2}$$

Proof. Let $m \in \mathbb{R}^{D\times D}$ be permutation-invariant, and let $a = m_{11}$ and $t = m_{12}$. Let $d, d' \in \{1,\dots,D\}$, if $d \neq d'$ there exists $s \in \mathscr{S}_D$, such that $(s(d), s(d')) = (1,2)$, since m is permutation-invariant, this yields $m_{d,d'} = t$. Otherwise, $d = d'$ and there exists $s \in \mathscr{S}_D$, such that $(s(d), s(d')) = (1,1)$, from which it follows that $m_{d,d} = a$. This proves that $m = aI_D + t\left(\mathbf{1}_D - \mathrm{id}_D\right)$. Moreover, the characteristic polynomial of the matrix $\mathbf{1}_D$ is $\chi_{\mathbf{1}_D}(X) = (X-D)X^{D-1}$. The eigenvalues of m can be deduced from the fact that if $t \neq 0$, $\chi_m(X) = t^D\chi_{\mathbf{1}_D}\left(\frac{X-a+t}{t}\right)$. □

In what follows, for every $\lambda_1, \lambda_2 \in \mathbb{R}$, we will denote by $m(\lambda_1,\lambda_2)$ the matrix $\lambda_1\left(\mathrm{id}_D - \frac{\mathbf{1}_D}{D}\right) + \lambda_2\frac{\mathbf{1}_D}{D}$. According to Proposition 4.1, $m(\lambda_1,\lambda_2)$ is the only $D \times D$ permutation-invariant matrix whose eigenvalues are λ_1 and λ_2. Recall that $\mathbb{R}^{D\times D}$ is equipped with the Frobenius inner product

$$m \cdot n = \sum_{d,d'=1}^{D} m_{dd'} n_{dd'}.$$

In particular, we have $m(\lambda) \cdot m(\mu) = (D-1)\lambda_1\mu_1 + \lambda_2\mu_2$. As per (3.5), we let ξ^* denote the convex conjugate of ξ with respect to the cone S^D_+, that is

$$\xi^*(m) = \sup_{n \in S^D_+} \{m \cdot n - \xi(n)\}.$$

For every $\lambda_1, \lambda_2 \in \mathbb{R}$, we define

$$\xi_\perp(\lambda_1,\lambda_2) = \xi\left(m\left(\frac{\lambda_1}{D-1}, \lambda_2\right)\right). \tag{4.3}$$

We equip $\mathbb{R}^D$ with the standard inner product $x \cdot y = \sum_{d=1}^D x_d y_d$, and let $\xi_\perp^*$ denote the convex conjugate of $\xi_\perp$ with respect to $\mathbb{R}^2_+$,

$$\xi_\perp^*(\lambda) = \sup_{\mu \in \mathbb{R}^2_+} \{\lambda \cdot \mu - \xi_\perp(\mu)\}.$$

Proposition 4.2. *For every $\lambda_1, \lambda_2 \in \mathbb{R}$, we have*

$$\xi_\perp^*(\lambda_1,\lambda_2) = \xi^*(m(\lambda_1,\lambda_2)).$$

Proof. Clearly, $\xi^*(m(\lambda_1,\lambda_2)) \geqslant (\xi_\perp)^*(\lambda_1,\lambda_2)$. Conversely, given $n \in S^D_+$, we define $n_0 = \frac{1}{D!}\sum_{s\in\mathscr{S}_D} n^s$. The matrix n_0 is permutation-invariant and positive semi-definite, by Proposition 4.1, there exists $\mu_1, \mu_2 \in \mathbb{R}^2_+$ such that $n_0 = m(\mu_1/(D-1), \mu_2)$. Furthermore, since ξ is convex and permutation-invariant, we have

$$\xi(n_0) \leqslant \frac{1}{D!}\sum_{s\in\mathscr{S}_D} \xi(n^s) = \xi(n).$$

Therefore,

$$\begin{aligned} m(\lambda_1,\lambda_2)\cdot n - \xi(n) &= m(\lambda_1,\lambda_2)\cdot n_0 - \xi(n) \\ &= \lambda \cdot \mu - \xi(n) \\ &\leqslant \lambda\cdot\mu - \xi\left(m\left(\frac{\mu_1}{D-1}, \mu_2\right)\right) \\ &\leqslant (\xi_\perp)^*(\lambda_1,\lambda_2). \end{aligned}$$

Taking the supremum over $n \in S^D_+$, we obtain the desired inequality. □

4.2. **Permutation-invariant paths.** Let $V = S^D$ or $V = \mathbb{R}^D$, we equip V with any norm $|\cdot|$. For $p \in [1,+\infty)$, we denote by $L^p([0,1), V)$ or simply L^p the set of functions $h : [0,1) \to V$ such that $|h|^p$ is integrable. As usual, functions in L^p are to be understood as equivalence classes of functions modulo equality almost everywhere. We define the p-norm on L^p, by

$$|h|_{L^p} = \left(\int_0^1 |h(u)|^p \mathrm{d}u\right)^{1/p}.$$

The set L^∞ is the set of essentially bounded functions, and we equip it with the sup norm

$$|h|_{L^\infty} = \text{ess-sup}_{u\in[0,1)}|h(u)|.$$

When V is equipped with a scalar product $x \cdot y$ and $|\cdot|$ is the associated Euclidean norm, the norm $|\cdot|_{L^2}$ is Hilbertian and comes from the scalar product

$$\langle h, k\rangle_{L^2} = \int_0^1 h(u)\cdot k(u)\mathrm{d}u.$$

At fixed $p \in [1,+\infty]$, the different norms $|\cdot|_{L^p}$ obtained by changing the norm $|\cdot|$ on V are all equivalent. Therefore, the statement $h \in L^p$ does not depend on the particular norm $|\cdot|$ we choose, but the precise value of $|h|_{L^p}$ does. Given $q \in \mathcal{Q}(\mathbb{R}^2_+)$, we let $q^\perp \in \mathcal{Q}(S^D_+)$ be the path defined by $q^\perp(u) = m(q(u))$. For $\mathfrak{q} \in \mathcal{Q}(S^D_+)$ and $s \in \mathscr{S}_D$, we define $\mathfrak{q}^s \in \mathcal{Q}(S^D_+)$ by $\mathfrak{q}^s(u) = (\mathfrak{q}(u))^s = (\mathfrak{q}_{s(d),s(d')}(u))_{1\leqslant d,d'\leqslant D}$. As previously, we equip S^D with the Frobenius inner product and $\mathbb{R}^2$ with the standard inner product.

Proposition 4.3. *Let $\mathfrak{q} \in \mathcal{Q}(S^D_+)$, assume that for every $s \in \mathscr{S}_D$, we have $\mathfrak{q}^s = \mathfrak{q}$, then there exists $q = (q_1, q_2) \in \mathcal{Q}(\mathbb{R}^2_+)$ such that $\mathfrak{q} = q^\perp$. In addition, for every $p \in [1,+\infty]$, $\mathfrak{q} \in L^p$ if and only if $q \in L^p$. Moreover, when $\mathfrak{q} \in L^2$, we have for every $q' = (q'_1, q'_2) \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^2$,*

$$\langle \mathfrak{q}, (q')^\perp\rangle_{L^2} = \langle q_1, (D-1)q'_1\rangle_{L^2} + \langle q_2, q'_2\rangle_{L^2}.$$

Proof. Let $u \in [0,1)$, for every $s \in \mathscr{S}_D$, the matrix $\mathfrak{q}(u) \in S^D_+$ satisfies $(\mathfrak{q}(u))^s = \mathfrak{q}(u)$. According to Proposition 4.1, there exists $q(u) \in \mathbb{R}^2_+$ such that $\mathfrak{q}(u) = m(q(u))$. If $u \leqslant v$, we have

$$m(q(v) - q(u)) = \mathfrak{q}(v) - \mathfrak{q}(u) \in S^D_+,$$

so $q_1(v) - q_1(u) \geqslant 0$ and $q_2(v) - q_2(u) \geqslant 0$.

In addition, the map $m(\lambda_1, \lambda_2) \mapsto (\lambda_1, \lambda_2)$ is continuous according to Proposition 4.1. This justifies that $q \in \mathcal{Q}(\mathbb{R}^2_+)$ and by definition $q^\perp = \mathfrak{q}$. Let $|x|_p = \left(\frac{1}{D}\sum_{d=1}^D |x_d|^p\right)^{1/p}$ denote the normalized p-norm on $\mathbb{R}^D$ and $|x|_\infty = \sup_{1\leqslant d\leqslant D}|x_d|$. We equip S^D with the norm

$$N_2(m) = \sup_{|x|_2=1} |mx|_2.$$

When $m \in S^D_+$, $N_2(m)$ is equal to the biggest eigenvalue of m. Therefore, for every $\lambda_1, \lambda_2 \in \mathbb{R}_+$,

$$N_2(m(\lambda_1, \lambda_2)) = \max\{\lambda_1, \lambda_2\}.$$

Thus, for every $u \in [0,1)$, $N_2(\mathfrak{q}(u)) = |q(u)|_\infty$, in particular $\mathfrak{q} \in L^p$ if and only if $q \in L^p$.

Finally, assume that $\mathfrak{q} \in L^2$, then by the previous argument $q \in L^2$. Recall that we have $m(\lambda) \cdot m(\mu) = (D-1)\lambda_1\mu_1 + \lambda_2\mu_2$, for every $q' \in \mathbb{R}_+^2$, we have

$$\begin{aligned}\langle \mathfrak{q}, (q')^\perp \rangle_{L^2} &= \int_0^1 m(q(u)) \cdot m(q'(u)) \mathrm{d}u \\ &= \int_0^1 (D-1) q_1(u) q_1'(u) + q_2(u) q_2'(u) \mathrm{d}u \\ &= \langle q_1, (D-1) q_1' \rangle_{L^2} + \langle q_2, q_2' \rangle_{L^2}.\end{aligned}$$

□

4.3. **Piecewise linear approximations of paths.** Given a finite dimensional vector space V and a cone $\mathcal{C} \subseteq V$. For every $j \geqslant 1$, we define

$$\mathcal{Q}^j(\mathcal{C}) = \{x \in \mathcal{C}^j \big| \ \forall i \in \{1, \dots, j\},\ x_i - x_{i-1} \in \mathcal{C}\}. \tag{4.4}$$

As previously, we only use $(V, \mathcal{C}) = (S^D, S^D_+)$ or $(V, \mathcal{C}) = (\mathbb{R}^D, \mathbb{R}^D_+)$. We will always adopt the convention $x_0 = 0$. Given $x \in \mathcal{Q}^j(\mathbb{R}^D_+)$, we define $\Lambda^j x \in \mathcal{Q}(\mathbb{R}^D_+)$ to be the path that linearly interpolates between the values $(0,0), (\frac{1}{j}, x_1), (\frac{2}{j}, x_2), \dots, (1, x_j)$. More precisely, for every $u \in [0,1)$,

$$\Lambda^j x(u) = \sum_{i=1}^{j} \mathbf{1}_{\left[\frac{i-1}{j}, \frac{i}{j}\right)}(u) \left(x_{i-1} + j \left(u - \frac{i-1}{j} \right) (x_i - x_{i-1}) \right). \tag{4.5}$$

We will write $\mathbf{1}_i$ instead of $\mathbf{1}_{\left[\frac{i-1}{j}, \frac{i}{j}\right)}$ in what follows, with the understanding that $\mathbf{1}_{j+1} = 0$. We also define

$$\widetilde{\mathbf{1}}_i(u) = j \left(u - \frac{i-1}{j} \right) \mathbf{1}_i(u) + j \left(\frac{i+1}{j} - u \right) \mathbf{1}_{i+1}(u).$$

Note that we have $\Lambda^j x = \sum_{i=1}^j x_i \widetilde{\mathbf{1}}_i$.

We equip $\mathbb{R}^2$ with the normalized ℓ^1-norm, $|v| = \frac{1}{2}(|v_1| + |v_2|)$ and the standard normalized scalar product $v \cdot w = \frac{1}{2}(v_1 w_1 + v_2 w_2)$. For every $p \in [1, +\infty)$, we equip $(\mathbb{R}^2)^j$ with the norm $|\cdot|_p$, defined by

$$|x|_p = \left(\frac{1}{j} \sum_{i=1}^{j} |x_i|^p \right)^{\frac{1}{p}}. \tag{4.6}$$

We also equip $(\mathbb{R}^2)^j$ with the normalized scalar product,

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

Given a path, $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^1$ we set

$$\Lambda_j p(u) = (\langle p(u), j \widetilde{\mathbf{1}}_i \rangle_{L^2})_{1 \leqslant i \leqslant j},$$

this defines a path $\Lambda_j p \in \mathcal{Q}^j(\mathbb{R}_+)$. For every $q \in \mathcal{Q}(\mathbb{R}^D_+)$, if we write $q(u) = (q_d(u))_{1 \leqslant d \leqslant D}$, then $q_d \in \mathcal{Q}(\mathbb{R}_+)$ and we define

$$\Lambda_j q(u) = (\Lambda_j q_1(u), \dots, \Lambda_j q_D(u)).$$

The linear maps $\Lambda^j : \mathcal{Q}^j(\mathbb{R}^2_+) \to \mathcal{Q}(\mathbb{R}^2_+) \cap L^1$ and $\Lambda_j : \mathcal{Q}(\mathbb{R}^2_+) \cap L^1 \to \mathcal{Q}^j(\mathbb{R}^2_+)$ form an adjoint pair in the following sense, for every $p \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^1$ and $x \in \mathcal{Q}^j(\mathbb{R}^2_+)$, we have

$$\langle \Lambda^j x, p \rangle_{L^2} = \langle x, \Lambda_j p \rangle_j .$$

Proposition 4.4. *There exists a constant $c > 0$ such that for every $j \geqslant 1$ and every $q = (q_1, q_2) \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty$, we have*

$$|q - \Lambda^j \Lambda_j q|_{L^1} \leqslant \frac{c|q|_{L^\infty}}{j}.$$

Proof. Let $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$, to lighten notations, we let $p_i = p\left(\frac{i}{j}\right)$ and define $\overline{p} = (p_i)_{1 \leqslant i \leqslant j} \in \mathcal{Q}^j(\mathbb{R}_+)$. For every $i \in \{1, \dots, j\}$, we have

$$\frac{p_i + p_{i-1}}{2} \leqslant \langle p, j\widetilde{\mathbf{1}}_i \rangle_{L^2} \leqslant \frac{p_i + p_{i+1}}{2}.$$

Therefore $|jp \cdot \widetilde{\mathbf{1}}_i - p_i| \leqslant \max\left\{\frac{p_{i+1} - p_i}{2}, \frac{p_i - p_{i-1}}{2}\right\} \leqslant \frac{p_{i+1} - p_{i-1}}{2}$. Summing over i, we obtain $|\Lambda_j p - \overline{p}|_1 \leqslant \frac{|p|_{L^\infty}}{j}$. In addition, given $x \in (\mathbb{R}_+)^j$, we have

$$|\Lambda^j(x)|_{L^1} \leqslant \sum_{i=1}^{j} |x_i| |\widetilde{\mathbf{1}}_i|_{L^1} = |x|_1.$$

Therefore,

$$|\Lambda^j \Lambda_j p - \Lambda^j(\overline{p})|_{L^1} \leqslant |\Lambda_j p - \overline{p}|_1 \leqslant \frac{|p|_{L^\infty}}{j}.$$

Observe, that for $u \in \left[\frac{i-1}{j}, \frac{i}{j}\right)$, $|\Lambda^j \overline{p}(u) - p(u)| \leqslant p_i - p_{i-1}$, thus

$$|\Lambda^j(\overline{p}) - p|_{L^1} \leqslant \frac{|p|_{L^\infty}}{j}.$$

Combining the previous two displays, we obtain

$$|\Lambda^j \Lambda_j p - p|_1 \leqslant |\Lambda^j \Lambda_j p - \Lambda^j(\overline{p})|_1 + |\Lambda^j(\overline{p}) - p|_1 \leqslant \frac{|p|_{L^\infty}}{j} + \frac{|p|_{L^\infty}}{j} = \frac{2|p|_{L^\infty}}{j}.$$

Finally, given $q \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty$ if we write $q(u) = (q_1(u), q_2(u))$, then $q_1, q_2 \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$. Applying the bound in the previous display to q_1 and q_2, we obtain

$$\begin{aligned} |\Lambda^j \Lambda_j q - q|_{L^1} &= \frac{1}{2}\left(|\Lambda^j \Lambda_j q_1 - q_1|_{L^1} + |\Lambda^j \Lambda_j q_2 - q_2|_{L^1}\right) \\ &\leqslant \frac{|q_1|_{L^\infty} + |q_2|_{L^\infty}}{j} \\ &\leqslant \frac{2|q|_{L^\infty}}{j}. \end{aligned}$$

□

Recall that given a closed convex cone $\mathcal{C}$ we have defined its dual cone $\mathcal{C}^*$ in (3.2).

Proposition 4.5. *For every $x \in (\mathbb{R}^2_+)^j$, we have*

$$|\Lambda^j(x)|_{L^1} \leqslant |x|_1.$$

In addition, if $x \in (\mathcal{Q}^j(\mathbb{R}^2_+))^*$, then $(\Lambda^j(x))^\perp \in (\mathcal{Q}(S^D_+) \cap L^2)^*$.

Proof. Let $x = (x_1, x_2) \in (\mathbb{R}^2_+)^j$, we have $|\widetilde{\mathbf{1}}_i|_{L^1} = \frac{1}{j}$. So,

$$|\Lambda^j x_1|_{L^1} = \left| \sum_{i=1}^j x_{1i} \widetilde{\mathbf{1}}_i \right|_{L^1} \leqslant \sum_{i=1}^j x_{1i} \left|\widetilde{\mathbf{1}}_i\right|_{L^1} = \frac{1}{j} \sum_{i=1}^j |x_{1i}|.$$

It follows that,

$$|\Lambda^j x|_{L^1} = \frac{|\Lambda^j x_1|_{L^1} + |\Lambda^j x_2|_{L^1}}{2} \leqslant \frac{|x_1|_1 + |x_2|_1}{2} = |x|_1.$$

This proves the first part of the proposition, let us now further assume that $x \in (\mathcal{Q}^j(\mathbb{R}^2_+))^*$. Let $\mathfrak{q} \in \mathcal{Q}(S^D_+) \cap L^2$, the path $\frac{1}{D!} \sum_{s \in \mathscr{S}_D} \mathfrak{q}^s$ is permutation-invariant. According to Proposition 4.3, there exists $q \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^2$ such that $\frac{1}{D!} \sum_{s \in \mathscr{S}_D} \mathfrak{q}^s = q^\perp$. We have

$$\begin{aligned}
\langle \mathfrak{q}, (\Lambda^j x)^\perp \rangle_{L^2} &= \left\langle \mathfrak{q}, \frac{1}{D!} \sum_{s \in \mathscr{S}_D} (\Lambda^j x)^{\perp, s^{-1}} \right\rangle_{L^2} \\
&= \left\langle \frac{1}{D!} \sum_{s \in \mathscr{S}_D} \mathfrak{q}^s, (\Lambda^j x)^\perp \right\rangle_{L^2} \\
&= \left\langle q^\perp, (\Lambda^j x)^\perp \right\rangle_{L^2} \\
&= \left\langle (D-1) q_1, \Lambda^j x_1 \right\rangle_{L^2} + \left\langle q_2, \Lambda^j x_2 \right\rangle_{L^2} \\
&= \left\langle (D-1) \Lambda_j q_1, x_1 \right\rangle_j + \left\langle \Lambda_j q_2, x_2 \right\rangle_j.
\end{aligned}$$

We have $((D-1)\Lambda_j q_1, \Lambda_j q_2) \in \mathcal{Q}^j(\mathbb{R}^2_+)$, so by definition of $(\mathcal{Q}^j(\mathbb{R}^2_+))^*$, the last line in the previous display is $\geqslant 0$. This justifies $(\Lambda^j(x))^\perp \in (\mathcal{Q}(S^D_+) \cap L^2)^*$. □

Recall the definition of $\xi_\perp$ from (4.3). Also recall that, given $r \in \mathcal{Q}$ we use the notation $\int h(r)$ as a shorthand for $\int_0^1 h(r(u)) \mathrm{d}u$.

Proposition 4.6. *Let $q \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^1$, for every $j \geqslant 1$, we have*

$$\int \xi_\perp(\Lambda^j \Lambda_j q) \leqslant \int \xi_\perp(q).$$

Proof. We let $\rho_1, \dots, \rho_j \in \mathbb{R}^2_+$ denote the coordinates of $\Lambda_j q$. By convexity of ξ, we have

$$\begin{aligned}
\int \xi_\perp(\Lambda^j \Lambda_j q) &= \int_0^1 \xi_\perp(\Lambda^j \Lambda_j q(u)) \mathrm{d}u \\
&= \sum_{i=1}^{j} \int_{\left[\frac{i-1}{j}, \frac{i}{j}\right)} \xi_\perp\left(\rho_{i-1} + j\left(u - \frac{i-1}{j}\right)(\rho_i - \rho_{i-1})\right) \mathrm{d}u \\
&= \sum_{i=1}^{j} \int_{\left[\frac{i-1}{j}, \frac{i}{j}\right)} \xi_\perp\left(j\left(\frac{i}{j} - u\right)\rho_{i-1} + j\left(u - \frac{i-1}{j}\right)\rho_i\right) \mathrm{d}u \\
&\leqslant \sum_{i=1}^{j} \int_{\left[\frac{i-1}{j}, \frac{i}{j}\right)} j\left(\frac{i}{j} - u\right)\xi_\perp(\rho_{i-1}) + j\left(u - \frac{i-1}{j}\right)\xi_\perp(\rho_i) \,\mathrm{d}u \\
&= \sum_{i=1}^{j} \frac{\xi_\perp(\rho_{i-1}) + \xi_\perp(\rho_i)}{2j}.
\end{aligned}$$

Using Jensen's inequality, it follows that $\xi_\perp(\rho_i) \leqslant \langle \xi_\perp(q), j\widetilde{\mathbf{1}}_i \rangle_{L^2}$. Finally, since $\sum_{i=1}^{j} \widetilde{\mathbf{1}}_i = 1$, we obtain

$$\int \xi_\perp(\Lambda^j \Lambda_j q) \leqslant \int \xi_\perp(q).$$

□

5. The free energy and Hamilton-Jacobi equations

We recall that $f(t, \mathfrak{q}) = \lim_{N \to +\infty} \overline{F}_N(t, \mathfrak{q})$. In this section, we use the fact that f solves (2.4) and the permutation invariance of the model to show that $\lim \overline{F}_N(t, 0)$ can be expressed as the value at $(t, 0)$ of the solution of a reduced Hamilton-Jacobi equation.

5.1. **The equation for models with matrix valued paths.**

Given $q \in \mathcal{Q}(\mathbb{R}^2_+)$, recall that we have defined $q^\perp(u) = m(q(u))$. The goal of this section is to show that the functions defined by $f^\perp(t, q) = f(t, q^\perp)$ is the viscosity solution of some Hamilton-Jacobi on $\mathcal{Q}(\mathbb{R}^2_+)$. Since we know that f is Gateaux differentiable and solves (2.4), this will basically only amount to computing the Gateaux derivatives of $f^\perp$.

Proposition 5.1. *For every $t \geqslant 0$, $\mathfrak{q} \in \mathcal{Q}(S^D_+) \cap L^2$ and $s \in \mathscr{S}_D$, we have*

$$f(t, \mathfrak{q}^s) = f(t, \mathfrak{q}). \tag{5.1}$$

Proof. Let $\mathfrak{q} \in \mathcal{Q}(S^D_+)$ and $t \geqslant 0$, recall from (2.2) that we have

$$F_N(t, \mathfrak{q}) = -\frac{1}{N} \mathbb{E} \log \iint \exp(H^{t,\mathfrak{q}}_N(\sigma, \alpha)) dP_N(\sigma) d\mathfrak{R}(\alpha),$$

where

$$H^{t,\mathfrak{q}}_N(\sigma, \alpha) = \sqrt{2t} H_N(\sigma) - tN\xi\left(\frac{\sigma\sigma^*}{N}\right) + \sqrt{2} W^{\mathfrak{q}}_N(\alpha) \cdot \sigma - \mathfrak{q}(1) \cdot \sigma\sigma^*.$$

The process $(H_N^{t,\mathfrak{q}}(\sigma,\alpha))_{\sigma\in\mathbb{R}^D,\alpha\in\mathfrak{U}}$ is a Gaussian process with the following mean and covariance,

$$\mathbb{E}H_N^{t,\mathfrak{q}}(\sigma,\alpha) = -tN\xi\left(\frac{\sigma\sigma^*}{N}\right) - \mathfrak{q}(1)\cdot\sigma\sigma^*,$$

$$\mathrm{Cov}\left(H_N^{t,\mathfrak{q}}(\sigma,\alpha), H_N^{t,\mathfrak{q}}(\tau,\beta)\right) = 2N\left(t\xi\left(\frac{\sigma\tau^*}{N}\right) + \frac{1}{N}\sum_{i=1}^N \sigma_i\cdot\mathfrak{q}(\alpha\wedge\beta)\tau_i\right).$$

Let us show that for every $s\in\mathscr{S}_D$, we have

$$(H_N^{t,\mathfrak{q}^{s^{-1}}}(\sigma,\alpha))_{\sigma\in\mathbb{R}^D,\alpha\in\mathfrak{U}} \overset{(d)}{=} (H_N^{t,\mathfrak{q}}(\sigma^s,\alpha))_{\sigma\in\mathbb{R}^D,\alpha\in\mathfrak{U}}.$$

To proceed, we compute the covariance and the mean of those two Gaussian processes and discover that they are equal. Let $\sigma,\tau\in\mathbb{R}^D$ and $\alpha,\beta\in\mathfrak{U}$, since ξ is permutation-invariant, we have $\xi\left(\frac{(\sigma^s)(\tau^s)^*}{N}\right) = \xi\left(\frac{\sigma\tau^*}{N}\right)$. Therefore,

$$\begin{aligned}
\mathbb{E}H_N^{t,\mathfrak{q}^{s^{-1}}}(\sigma,\alpha)H_N^{t,\mathfrak{q}^{s^{-1}}}(\tau,\beta) &= 2N\left(t\xi\left(\frac{\sigma\tau^*}{N}\right) + \frac{1}{N}\sum_{i=1}^N \sigma_i\cdot\mathfrak{q}^{s^{-1}}(\alpha\wedge\beta)\tau_i\right)\\
&= 2N\left(t\xi\left(\frac{\sigma\tau^*}{N}\right) + \frac{1}{N}\sum_{i=1}^N \sigma_i^s\cdot\mathfrak{q}(\alpha\wedge\beta)\tau_i^s\right)\\
&= 2N\left(t\xi\left(\frac{(\sigma^s)(\tau^s)^*}{N}\right) + \frac{1}{N}\sum_{i=1}^N \sigma_i^s\cdot\mathfrak{q}(\alpha\wedge\beta)\tau_i^s\right)\\
&= \mathbb{E}H_N^{t,\mathfrak{q}}(\sigma^s,\alpha)H_N^{t,\mathfrak{q}}(\tau^s,\beta).
\end{aligned}$$

In addition,

$$\begin{aligned}
\mathbb{E}H_N^{t,\mathfrak{q}^{s^{-1}}}(\sigma,\alpha) &= -Nt\xi\left(\frac{\sigma\sigma^*}{N}\right) - \mathfrak{q}^{s^{-1}}(1)\cdot\sigma\sigma^*\\
&= -Nt\xi\left(\left(\frac{\sigma\sigma^*}{N}\right)^s\right) - \mathfrak{q}(1)\cdot(\sigma\sigma^*)^s\\
&= -Nt\xi\left(\frac{\sigma^s(\sigma^s)^*}{N}\right) - \mathfrak{q}(1)\cdot\sigma^s(\sigma^s)^*\\
&= \mathbb{E}H_N^{t,\mathfrak{q}}(\sigma^s,\alpha).
\end{aligned}$$

The desired equality in law follows. We now show that $f(t,\mathfrak{q}^s) = f(t,\mathfrak{q})$. According to the previous result and the permutation invariance of P_N, we have

$$\begin{aligned}
F_N(t,\mathfrak{q}^s) &= -\frac{1}{N}\mathbb{E}\log\iint \exp(H_N^{t,\mathfrak{q}^s}(\sigma,\alpha))dP_N(\sigma)d\mathfrak{R}(\alpha)\\
&= -\frac{1}{N}\mathbb{E}\log\iint \exp(H_N^{t,\mathfrak{q}}(\sigma^{s^{-1}},\alpha))dP_N(\sigma)d\mathfrak{R}(\alpha)\\
&= -\frac{1}{N}\mathbb{E}\log\iint \exp(H_N^{t,\mathfrak{q}}(\sigma,\alpha))dP_N(\sigma)d\mathfrak{R}(\alpha)\\
&= F_N(t,\mathfrak{q}).
\end{aligned}$$

Letting $N\to+\infty$, we obtain the desired result.

□

Proposition 5.2. *For every* $(t,q) \in (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty)$*,* $f(t,\cdot)$ *is Gateaux differentiable at* $q^\perp$ *and the path* $\nabla f(t,q^\perp) \in \mathcal{Q}(S^D_+)$ *is permutation-invariant that is, for every* $u \in [0,1)$ *and* $s \in \mathscr{S}_D$*,*

$$(\nabla f(t,q^\perp)(u))^s = \nabla f(t,q^\perp)(u).$$

Proof. According to [7, Proposition 8.6], $f(t,\cdot)$ is Gateaux differentiable at $q^\perp$. Let $\kappa \in L^2([0,1), S^D)$ such that for $\varepsilon > 0$ small enough $q^\perp + \varepsilon\kappa \in \mathcal{Q}(S^D_+)$. According to Proposition 5.1, we have for every $s \in \mathscr{S}_D$,

$$\frac{f(t,q^\perp + \varepsilon\kappa^s) - f(t,q^\perp)}{\varepsilon} = \frac{f(t,q^\perp + \varepsilon\kappa) - f(t,q^\perp)}{\varepsilon}.$$

Letting $\varepsilon \to 0$, we obtain that $\langle \nabla f(t,q^\perp), \kappa\rangle_{L^2} = \langle \nabla f(t,q^\perp), \kappa^s\rangle_{L^2}$. So,

$$\langle \nabla f(t,q^\perp), \kappa\rangle_{L^2} = \langle (\nabla f(t,q^\perp))^{s^{-1}}, \kappa\rangle_{L^2}.$$

This means that $(\nabla f(t,q^\perp))^{s^{-1}} = \nabla f(t,q^\perp)$. □

For $q \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^2$, we define $f^\perp(t,q) = f(t,q^\perp)$, $\psi^\perp(q) = \psi(q^\perp)$. Recall the definition of $\xi_\perp$ in (4.3).

Proposition 5.3. *The function* $f^\perp$ *is Gateaux differentiable at every* $(t,q) \in (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty)$*, and it satisfies*

$$\begin{cases} \partial_t f^\perp - \int \xi_\perp(\nabla f^\perp) = 0 \text{ on } (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty) \\ f^\perp(0,\cdot) = \psi^\perp. \end{cases} \tag{5.2}$$

Proof. According to [7, Propositions 7.2 & 8.6], f is Gateaux differentiable at every $(t,\mathfrak{q}) \in (0,+\infty) \times (\mathcal{Q}_\uparrow(S^D_+) \cap L^\infty)$ and we have

$$\partial_t f(t,\mathfrak{q}) = \int \xi(\nabla f(t,\mathfrak{q})).$$

Let $t > 0$ and $q = (q_1,q_2) \in \mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty$ and let $\kappa = (\kappa_1,\kappa_2) \in L^2([0,1),\mathbb{R}^2)$ such that for $\varepsilon > 0$ small enough $q + \varepsilon\kappa \in \mathcal{Q}_\uparrow(\mathbb{R}^2_+)$. Passing to the limit as $\varepsilon \to 0$ in

$$\frac{f^\perp(t,q+\varepsilon\kappa) - f^\perp(t,q)}{\varepsilon} = \frac{f(t,q^\perp + \varepsilon\kappa^\perp) - f(t,q^\perp)}{\varepsilon},$$

yields $\langle \nabla f^\perp(t,q), \kappa\rangle_{L^2} = \langle \nabla f(t,q^\perp), \kappa^\perp\rangle_{L^2}$. According to Proposition 5.2, the path $\nabla f(t,q^\perp)$ is permutation-invariant. According to Proposition 4.3, there exists $r = (r_1,r_2) \in \mathcal{Q}(\mathbb{R}^2_+)$ such that, $\nabla f(t,q^\perp) = r^\perp$ and we have

$$\langle (\nabla f^\perp(t,q))_1, \kappa_1\rangle_{L^2} + \langle (\nabla f^\perp(t,q))_2, \kappa_2\rangle_{L^2} = \langle (D-1)r_1, \kappa_1\rangle_{L^2} + \langle r_2, \kappa_2\rangle_{L^2}.$$

So, $\nabla f^\perp(t,q) = ((D-1)r_1, r_2)$. In particular, for every $u \in [0,1)$,

$$\xi_\perp(\nabla f^\perp(t,q)) = \xi(r^\perp) = \xi(\nabla f(t,q^\perp)).$$

Finally, we have

$$\partial_t f^\perp(t,q) = \partial_t f(t,q^\perp) = \int \xi(\nabla f(t,q^\perp)) = \int \xi_\perp(\nabla f^\perp(t,q)).$$

□

If (5.2) was written on $\mathcal{Q}(\mathbb{R}^2_+) \cap L^2$ rather than $\mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty$, then we could immediately conclude that $f^\perp$ is a solution in the viscosity sense and (1.8) would directly follow from the Hopf-Lax representation. This is indeed how are going to argue, but to do so we need to show that we can neglect boundary points. That is, paths in $\mathcal{Q}(\mathbb{R}^2_+) \cap L^2$ that do not belong to $\mathcal{Q}_\uparrow(\mathbb{R}^2_+) \cap L^\infty$. This can be done using the content of Section 3. To connect the two settings, we will need to consider finite dimensional approximations of (5.2). Recall the definition of the lift and projection maps Λ^j and Λ_j from Section 4.3. Given $(t,x) \in \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}^2_+)$, we define $f^{\perp,j}(t,x) = f^\perp(t, \Lambda^j x)$ and $H^j_\perp(x) = \int \xi_\perp(\Lambda^j x)$.

Proposition 5.4. *There exists a constant $c > 0$ such that the following holds. For every $j \geqslant 1$, the function $f^{\perp,j}$ is differentiable on $(0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}^2_+)$ and for every $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}^2_+)$,*

$$|\partial_t f^{\perp,j}(t,x) - H^j_\perp(\nabla f^{\perp,j}(t,x))| \leqslant \frac{c}{j}.$$

Furthermore, $f^{\perp,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}^2_+))$.

Proof. Step 1. We show that $f^{\perp,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}^2_+))$.

According to [7, Proposition 5.1], $f : \mathbb{R}_+ \times (\mathcal{Q}(S^D_+) \cap L^1) \to \mathbb{R}$ is Lipschitz. More precisely, we have for every $(t,\mathfrak{q}), (t',\mathfrak{q}') \in \mathbb{R}_+ \times (\mathcal{Q}(S^D_+) \cap L^1)$,

$$|f(t,\mathfrak{q}) - f(t',\mathfrak{q}')| \leqslant |\mathfrak{q} - \mathfrak{q}'|_{L^1} + |t - t'| \sup_{|a| \leqslant 1} |\xi(a)|.$$

Let $|\cdot|$ denote the Frobenius norm on S^D. There exists a constant $c > 0$ depending only on $|\cdot|$ such that the following holds. For every $x \in \mathcal{Q}^j(\mathbb{R}^2_+)$, we have

$$\begin{aligned}
|(\Lambda^j x)^\perp|_{L^1} &= \int_0^1 |m(\Lambda^j x(u))| \mathrm{d}u \\
&= \int_0^1 |\Lambda^j x_1(u)| \left| \mathrm{id}_D - \frac{\mathbf{1}_D}{D} \right| + |\Lambda^j x_2(u)| \left| \frac{\mathbf{1}_D}{D} \right| \mathrm{d}u \\
&\leqslant c \int_0^1 \frac{|\Lambda^j x_1(u)| + |\Lambda^j x_2(u)|}{2} \mathrm{d}u \\
&= c |\Lambda^j x|_{L^1} \\
&\leqslant c |x|_1,
\end{aligned}$$

where the last line is a consequence of Proposition 4.5. Since $f^{\perp,j}(t,x) = f(t, (\Lambda^j x)^\perp)$, it follows that $f^{\perp,j} : \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}^2_+) \to \mathbb{R}$ is Lipschitz. In particular,

$$\sup_{t > 0, x \in \mathcal{Q}^j(\mathbb{R}^2_+)} \frac{|f^{\perp,j}(t,x) - f^{\perp,j}(0,x)|}{t} < +\infty \text{ and } \sup_{t > 0} \|f^{\perp,j}(t,\cdot)\|_{\mathrm{Lip}} < +\infty.$$

Furthermore, according to [7, Proposition 3.6], for every $t \geqslant 0$, $f(t,\cdot)$ is $(\mathcal{Q}(S^D_+) \cap L^2)^*$-nondecreasing. Let $x, y \in \mathcal{Q}^j(\mathbb{R}^2_+)$, such that $y - x \in (\mathcal{Q}^j(\mathbb{R}^2_+))^*$, we have $(\Lambda^j x)^\perp, (\Lambda^j y)^\perp \in \mathcal{Q}(S_+^D) \cap L^\infty$ and according to Proposition 4.5, $(\Lambda^j(y-x))^\perp \in (\mathcal{Q}(S_+^D) \cap L^2)^*$. Therefore,

$$f^{\perp,j}(t,y) - f^{\perp,j}(t,x) = f(t,(\Lambda^j y)^\perp) - f(t,(\Lambda^j x)^\perp) \geqslant 0.$$

Thus, $f^{\perp,j}(t,\cdot)$ is $(\mathcal{Q}^j(\mathbb{R}_+^2))^*$-nondecreasing and we have proven that $f^{\perp,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+^2))$. This concludes Step 1.

Step 2. We show that there exists $c > 0$ such that for every $j \geqslant 1$ and every $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+^2)$, $f^{\perp,j}$ is differentiable at (t,x) and

$$|\partial_t f^j(t,x) - H_\perp^j(\nabla f^j(t,x))| \leqslant \frac{c}{j}.$$

For every $x \in \mathcal{Q}^j(\mathring{\mathbb{R}}_+^2)$, $\Lambda^j x \in \mathcal{Q}_\uparrow(\mathbb{R}_+^2) \cap L^\infty$. Using Proposition 5.3, we deduce that $f^{\perp,j}$ is differentiable on $(0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+^2)$ and for every $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+^2)$ we have $\nabla f^{\perp,j}(t,x) = \Lambda_j(\nabla f^\perp(t,\Lambda^j x))$. In addition, we have

$$\begin{aligned}
\partial_t f^{\perp,j}(t,x) &= \partial_t f^\perp(t,\Lambda^j x) \\
&= \int \xi_\perp(\nabla f^\perp(t,\Lambda^j x)) \\
&= H_\perp^j(\nabla f^{\perp,j}(t,x)) + \int \left(\xi_\perp(\nabla f^\perp(t,\Lambda^j x)) - \xi_\perp(\Lambda^j \Lambda_j \nabla f^\perp(t,\Lambda^j x))\right).
\end{aligned}$$

Since $f^\perp$ is 1-Lipschitz with respect to L^1-norm, we have $|\nabla f^\perp(t,\Lambda^j x))|_{L^\infty} \leqslant 1$. Now, using Step 1 and Proposition 4.4 we discover that

$$\begin{aligned}
\left|\partial_t f^{\perp,j}(t,x) - H_\perp^j(\nabla f^{\perp,j}(t,x))\right| &= |\xi_\perp(\nabla f^\perp(t,\Lambda^j x)) - \xi_\perp(\Lambda^j\Lambda_j \nabla f^\perp(t,\Lambda^j x))|_1 \\
&\leqslant \ell_\perp |\nabla f^\perp(t,\Lambda^j x) - \Lambda^j \Lambda_j \nabla f^\perp(t,\Lambda^j x)|_{L^1} \\
&\leqslant \frac{c}{j},
\end{aligned}$$

where $\ell_\perp = \sup_{|a|\leqslant 1} |\xi_\perp(a)|$ and $c \geqslant 0$ is some constant depending on ξ and $|\cdot|$. □

Theorem 5.5. *The function $f^\perp$ is the unique viscosity solution of*

$$\begin{cases} \partial_t u^\perp - \int \xi_\perp(\nabla u^\perp) = 0 \ \text{on} \ (0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+^2) \cap L^2) \\ u^\perp(0,\cdot) = \psi^\perp. \end{cases} \tag{5.3}$$

In addition, for every $(t,q) \in (0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+^2) \cap L^2)$, $f^\perp$ admits the Hopf-Lax representation at (t,q),

$$f^\perp(t,q) = \sup_{p \in \mathcal{Q}(\mathbb{R}_+^2) \cap L^\infty} \inf_{r \in \mathcal{Q}(\mathbb{R}_+^2) \cap L^\infty} \left\{ \psi^\perp(p+q) - \langle p, r\rangle_{L^2} + t \int_0^1 \xi_\perp(r) \right\}. \tag{5.4}$$

Remark 5.6. Taking $q = 0$ in (5.4), we obtain Theorem 1.2.

Proof. According to [9, Theorem 4.6], (5.3) has a unique Lipschitz viscosity solution $u^\perp$ and it is given by the variational formula (5.4). Let us show that $f^\perp$ and $u^\perp$ coincide on $[0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+^2) \cap L^2)$.

Let $j \geqslant 1$, define Let $\psi^{\perp,j} = f^{\perp,j}(0,\cdot)$. According to Proposition 5.4, $\psi^{\perp,j}$ is $(\mathcal{Q}^j(\mathbb{R}^2_+))^*$-nondecreasing and Lipschitz. In addition, $H^j_\perp|_{\mathcal{Q}^j(\mathbb{R}^2_+)}$ is $(\mathcal{Q}^j(\mathbb{R}^2_+))^*$-nondecreasing and locally Lipschitz. According to Theorem 3.2, the following Hamilton-Jacobi equation is well posed

$$\begin{cases} \partial_t u^{\perp,j} - H^j_\perp(\nabla u^{\perp,j}) = 0 \text{ on } (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}^2_+) \\ u^{\perp,j}(0,\cdot) = \psi^{\perp,j}. \end{cases}$$

We let $u^{\perp,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}^2_+))$ be its unique viscosity solution.
Step 1. We show that, for every $(t,q) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty)$,

$$f^\perp(t,q) = \lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q).$$

According to Proposition 3.6 and Proposition 5.4, there exists $c > 0$ such that for every $j \geqslant 1$ and $(t,x) \in \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}^2_+)$, we have

$$|f^{\perp,j}(t,x) - u^{\perp,j}(t,x)| \leqslant \frac{ct}{j}. \tag{5.5}$$

Let $q \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty$, according to Proposition 4.4 as $j \to +\infty$ we have $\Lambda^j \Lambda_j q \to q$ in L^1. By Lipschitz continuity of $f^\perp$ we have

$$f^\perp(t,q) = \lim_{j\to+\infty} f^{\perp,j}(t,\Lambda_j q) = \lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q).$$

Step 2. We show that, for every $(t,q) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty)$,

$$\lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q) = u^\perp(t,q).$$

According to [10, Theorem 1.2 (2) (d)], $u^{\perp,j}$ admits the Hopf-Lax representation. That is, for every $(t,x) \in \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}^2_+)$,

$$u^{\perp,j}(t,x) = \sup_{y\in\mathcal{Q}^j(\mathbb{R}^2_+)} \inf_{z\in\mathcal{Q}^j(\mathbb{R}^2_+)} \left\{ \psi^{\perp,j}(x+y) - \langle y,z\rangle_j + t\int \xi_\perp(\Lambda^j z) \right\}.$$

Similarly, from [9, Theorem 1.1], we know that for every $(t,q) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}^2_+) \cap L^2)$ we have

$$u^\perp(t,q) = \sup_{p\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \left\{ \psi^\perp(q+p) - \langle p,r\rangle_{L^2} + t\int \xi_\perp(r) \right\}.$$

Step 2.1 We show that, $\lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q) \leqslant u^\perp(t,q)$.
Observe that $\mathcal{Q}^j(\mathbb{R}^2_+) = \{\Lambda_j p, p \in \mathcal{Q}(\mathbb{R}^2_+) \cap L^\infty\}$. So,

$$u^{\perp,j}(t,\Lambda_j q) = \sup_{p\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \Big\{ \psi^\perp(\Lambda^j\Lambda_j q + \Lambda^j\Lambda_j p) - \langle \Lambda_j p, \Lambda_j r\rangle_j + t\int \xi_\perp(\Lambda^j\Lambda_j r) \Big\}.$$

Since, $\langle \Lambda_j p, \Lambda_j r\rangle_j = \langle \Lambda^j \Lambda_j p, r\rangle_{L^2}$ and $\{\Lambda^j\Lambda_j p, p \in \mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty\} \subseteq \mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty$. We have

$$u^{\perp,j}(t,\Lambda_j q) \leqslant \sup_{p\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \Big\{ \psi^\perp(\Lambda^j\Lambda_j q + p) - \langle p, r\rangle_{L^2} + t\int \xi_\perp(\Lambda^j\Lambda_j r)\Big\}.$$

Finally, according to Proposition 4.6, we have $\int \xi_\perp(\Lambda^j\Lambda_j r) \leqslant \int \xi_\perp(r)$, so

$$u^{\perp,j}(t,\Lambda_j q) \leqslant \sup_{p\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \Big\{ \psi^\perp(\Lambda^j\Lambda_j q + p) - \langle p, r\rangle_{L^2} + t\int \xi_\perp(r)\Big\}.$$

Using the Lipschitz continuity of $\psi^\perp$, we discover that

$$\begin{aligned} u^{\perp,j}(t,\Lambda_j q) \leqslant & |\Lambda^j\Lambda_j q - q|_1 \\ & + \sup_{p\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \Big\{ \psi^\perp(q + p) - \langle p, r\rangle_{L^2} + t\int \xi_\perp(r)\Big\}. \end{aligned}$$

Using Proposition 4.4, we obtain $\lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q) \leqslant u^\perp(t,q)$.
Step 2.2 We show that, $\lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q) \geqslant u^\perp(t,q)$.
For every $\varepsilon > 0$, there exists $p \in \mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty$, such that

$$u^\perp(t,q) \leqslant \varepsilon + \psi^\perp(p+q) + \inf_{r\in\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty} \Big\{ -\langle p, r\rangle_{L^2} + t\int \xi_\perp(r)\Big\}.$$

We have $\{\Lambda^j z, z \in \mathcal{Q}^j(\mathbb{R}^2_+)\} \subseteq \mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty$. So,

$$u^\perp(t,q) \leqslant \varepsilon + \psi^\perp(p+q) + \inf_{z\in\mathcal{Q}^j(\mathbb{R}^2_+)} \Big\{ -\langle p, \Lambda^j z\rangle_{L^2} + t\int \xi_\perp(\Lambda^j z)\Big\}.$$

Using the Lipschitz continuity of $\psi^\perp$, we have $\psi^\perp(p+q) \leqslant \psi^{\perp,j}(\Lambda_j p + \Lambda_j q) + |\Lambda^j\Lambda_j(p+q) - (p+q)|_{L^1}$. And we have $\langle p, \Lambda^j z\rangle_{L^2} = \langle \Lambda_j p, z\rangle_j$. So,

$$\begin{aligned} u^\perp(t,q) \leqslant \varepsilon + |\Lambda^j\Lambda_j(p+q) - (p+q)|_{L^1} + \psi^{\perp,j}(\Lambda_j p + \Lambda_j q) \\ + \inf_{z\in\mathcal{Q}^j(\mathbb{R}^2_+)} \Big\{ -\langle \Lambda_j p, z\rangle_j + t\int \xi_\perp(\Lambda^j z)\Big\}. \end{aligned}$$

Since, $\{\Lambda_j p, p \in \mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty\} \subseteq \mathcal{Q}^j(\mathbb{R}^2_+)$, obtain

$$u^\perp(t,q) \leqslant \varepsilon + |\Lambda^j\Lambda_j(p+q) - (p+q)|_{L^1} + u^{\perp,j}(t,\Lambda_j q).$$

Using Proposition 4.4 and letting $j\to+\infty$, we obtain

$$u^\perp(t,q) \leqslant \varepsilon + \lim_{j\to+\infty} u^{\perp,j}(t,\Lambda_j q),$$

since $\varepsilon > 0$ is arbitrary this concludes Step 2.
Step 3. Conclusion.

From Step 1 and Step 2, we deduce that $f^\perp$ and $u^\perp$ coincide on $\mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty)$. Since $\mathcal{Q}(\mathbb{R}^2_+)\cap L^\infty$ is dense in $\mathcal{Q}(\mathbb{R}^2_+)\cap L^2$ with respect to L^1 convergence and $f^\perp$ and $u^\perp$ are both Lipschitz continuous with respect to $|\cdot|_{L^1}$, we conclude that $f^\perp$ and $u^\perp$ coincide on $\mathbb{R}_+\times(\mathcal{Q}(\mathbb{R}^2_+)\cap L^2)$. ☐

5.2. **The equation for models with vector valued paths.** In this section, we repeat the analysis of Section 5.1 under the additional assumption that the interaction function ξ only depends on the diagonal coefficients of the overlap matrix. In this case, we can use the fact the value of the limit free energy is encoded by a Hamilton-Jacobi equation on $\mathcal{Q}(\mathbb{R}_+^D)$, as pointed out in (2.7). In this setting, permutation-invariant paths $q \in \mathcal{Q}(\mathbb{R}_+^D)$ are of the form $q = (p, \dots, p)$ with $p \in \mathcal{Q}(\mathbb{R}_+)$. Leveraging symmetries of the problem, this will allow us to show that the limit free energy is in fact encoded by a Hamilton-Jacobi equation on $\mathcal{Q}(\mathbb{R}_+)$. Since the proofs are very similar, we will not write them out in as many details as in the previous section.

For $p \in \mathcal{Q}(\mathbb{R}_+)$, we let $\mathfrak{q} = \mathrm{diag}(p, \dots, p) \in \mathcal{Q}(S_+^D)$ denote the path such that $\mathfrak{q}(u)$ is the diagonal matrix with diagonal coefficients $(p(u), \dots, p(u))$. We define $f^\dagger(t,p) = f(t, \mathrm{diag}(p, \dots, p))$, $\psi^\dagger(p) = \psi(\mathrm{diag}(p, \dots, p))$. Recall that here, we identify the function $A \mapsto \xi(A)$ defined on $\mathbb{R}^{D\times D}$ and the function $x \mapsto \xi(\mathrm{diag}(x))$ defined on $\mathbb{R}^D$. With this in mind, we set $\xi_\dagger(\lambda) = \xi(\lambda, \dots, \lambda)$.

Proposition 5.7. *The function $f^\dagger$ is Gateaux differentiable at every $(t,p) \in (0, +\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty)$ and satisfies*

$$\begin{cases} \partial_t f^\dagger - \int \xi_\dagger\left(\frac{\nabla f^\dagger}{D}\right) = 0 \ \text{on } (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty) \\ f^\dagger(0, \cdot) = \psi^\dagger \ \text{on } \mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty. \end{cases}$$

Proof. We reproduce the proof of Proposition 5.3, replacing f by f^{diag}, the function defined by $f^{\mathrm{diag}}(t,q) = f(t, \mathrm{diag}(q))$ for $q \in \mathcal{Q}(\mathbb{R}_+^D)$. Let $(t,p) \in \mathbb{R}_+ \times (\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty)$ according to Proposition 5.2, f is Gateaux differentiable at $(t, \mathrm{diag}(p, \dots, p))$ and its gradient is a permutation-invariant path. Therefore, f^{diag} is Gateaux differentiable at $(t, (p, \dots, p))$ and its gradient $\nabla f^{\mathrm{diag}}(t, (p, \dots, p)) \in \mathcal{Q}(\mathbb{R}_+^D)$ satisfies $\nabla f^{\mathrm{diag}}(t, (p, \dots, p)) = (r, \dots, r)$ for some $r \in \mathcal{Q}(\mathbb{R}_+)$. It follows that $f^\dagger$ is Gateaux differentiable at (t,p) and its Gateaux derivative satisfies $\nabla f^\dagger(t,p) = Dr$. According to (2.7), we have

$$\partial_t f^\dagger(t,p) = \partial_t f^{\mathrm{diag}}(t,p) = \int \xi(\nabla f^{\mathrm{diag}}(t, (p, \dots, p))) = \int \xi_\dagger\left(\frac{\nabla f^\dagger(t,p)}{D}\right).$$

□

Recall from (4.4) that

$$\mathcal{Q}^j(\mathbb{R}_+) = \{x \in \mathbb{R}_+^j \big| \, x_1 \leqslant \dots \leqslant x_j\}.$$

Also recall that given $x \in \mathcal{Q}^j(\mathbb{R}_+)$, $\Lambda^j x$ denote the path which linearly interpolates between the values $(0,0), (\frac{1}{j}, x_1), \dots, (1, x_j)$ and that given $p \in \mathcal{Q}(\mathbb{R}_+)$ the vector $\Lambda_j p \in \mathcal{Q}^j(\mathbb{R}_+)$ is defined so that Λ_j and Λ^j form an adjoint pair. As previously, we define $f^{\dagger,j}(t,x) = f^\dagger(t, \Lambda^j x)$ and $H_\dagger^j(x) = \int_0^1 \xi_\dagger(\Lambda^j x)$.

Proposition 5.8. *There exists a constant $c > 0$ such that the following holds. For every $j \geqslant 1$, the function $f^{\dagger,j}$ is differentiable on $(0, +\infty) \times \mathring{\mathcal{Q}}^j(\mathbb{R}_+)$ and* *for every* $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+)$,

$$\left| \partial_t f^{\dagger,j}(t,x) - H^j_\dagger \left(\frac{\nabla f^{\dagger,j}(t,x)}{D} \right) \right| \leqslant \frac{c}{j}.$$

Furthermore, $f^{\dagger,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$.

Proof. The proof will follow from Proposition 5.4 and the fact that $f^\dagger(t,p) = f^\perp(t,(p,p))$.

Step 1. We show that $f^{\dagger,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$.

According to Proposition 5.4, $f^{\perp,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}^2_+))$. In particular,

$$\begin{aligned}
\sup_{\substack{t>0 \\ x\in\mathcal{Q}^j(\mathbb{R}_+)}} \frac{|f^{\dagger,j}(t,x) - f^{\dagger,j}(0,x)|}{t} &= \sup_{\substack{t>0 \\ x\in\mathcal{Q}^j(\mathbb{R}_+)}} \frac{|f^{\perp,j}(t,(x,x)) - f^{\perp,j}(0,(x,x))|}{t} \\
&\leqslant \sup_{\substack{t>0 \\ y\in\mathcal{Q}^j(\mathbb{R}^2_+)}} \frac{|f^{\perp,j}(t,y) - f^{\perp,j}(0,y)|}{t} \\
&< +\infty.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\sup_{t>0} \|f^{\dagger,j}(t,\cdot)\|_{\mathrm{Lip}} &= \sup_{t>0} \|x \mapsto f^{\perp,j}(t,(x,x))\|_{\mathrm{Lip}} \\
&\leqslant \sup_{t>0} \|f^{\perp,j}(t,\cdot)\|_{\mathrm{Lip}} \\
&< +\infty.
\end{aligned}$$

Furthermore, given $x, x' \in \mathcal{Q}^j(\mathbb{R}_+)$, it is clear that if $x' - x \in (\mathcal{Q}^j(\mathbb{R}_+))^*$, then $(x',x') - (x,x) \in (\mathcal{Q}^j(\mathbb{R}^2_+))^*$. So,

$$f^{\dagger,j}(t,x') - f^{\dagger,j}(t,x) = f^{\perp,j}(t,(x',x')) - f^{\perp,j}(t,(x,x)) \geqslant 0.$$

Thus, $f^{\dagger,j}(t,\cdot)$ is $(\mathcal{Q}^j(\mathbb{R}_+))^*$-nondecreasing. We have proven that $f^{\dagger,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}^2_+))$, this concludes Step 1.

Step 2. We show that there exists $c > 0$ such that for every $j \geqslant 1$ and every $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+)$, $f^{\dagger,j}$ is differentiable at (t,x) and

$$\left| \partial_t f^{\dagger,j}(t,x) - H^j_\dagger \left(\frac{\nabla f^{\dagger,j}(t,x)}{D} \right) \right| \leqslant \frac{c}{j}.$$

For every $x \in \mathcal{Q}^j(\mathring{\mathbb{R}}_+)$, $(x,x) \in \mathcal{Q}^j(\mathring{\mathbb{R}}^2_+)$. Using Proposition 5.4, we deduce that $f^{\dagger,j}$ is differentiable on $(0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+)$ and for every $(t,x) \in (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+)$, we have

$$\partial_t f^{\dagger,j}(t,x) - H^j_\dagger \left(\frac{\nabla f^{\dagger,j}(t,x)}{D} \right) = \partial_t f^{\perp,j}(t,(x,x)) - H^j_\perp(\nabla f^{\perp,j}(t,(x,x))).$$

The result then follows from Proposition 5.4. □

Theorem 5.9. *The function $f^\dagger$ is the unique viscosity solution of,*

$$\begin{cases} \partial_t u^\dagger - \int \xi_\dagger\left(\frac{\nabla u^\dagger}{D}\right) = 0 \text{ on } (0,+\infty) \times \left(\mathcal{Q}(\mathbb{R}_+) \cap L^2\right) \\ u^\dagger(0,\cdot) = \psi^\dagger. \end{cases} \tag{5.6}$$

In addition, for every $(t,r) \in [0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+) \cap L^2)$, $f^\dagger$ admits the Hopf-Lax representation,

$$f^\dagger(t,r) = \sup_{p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \left\{ \psi^\dagger(p) - t \int_0^1 \xi^*\left(\frac{p-r}{t}, \dots, \frac{p-r}{t}\right) \right\}. \tag{5.7}$$

Remark 5.10. Taking $r = 0$ in (5.7), we obtain the variational formula (1.9) of Theorem 1.3.

Remark 5.11. Here, since (5.6) is written over the set of 1-dimensional paths, a different version of the Hopf-Lax representation is available [9, Proposition A.3]. This allows us to write the viscosity solution as a “sup” rather than a “sup inf”.

Remark 5.12. Reproducing the proof of Proposition 4.2, we have for every $\lambda \geqslant 0$,

$$\xi^*(\lambda, \dots, \lambda) = \left(\xi_\dagger\left(\frac{\cdot}{D}\right)\right)^*(\lambda).$$

Proof. According to [9, Theorem 4.6] and [9, Proposition A.3], (5.6) has a unique Lipschitz viscosity solution $u^\dagger$ and it is given by the variational formula (5.7). Let us show that $f^\dagger$ and $u^\dagger$ coincide on $[0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+) \cap L^2)$.

Let $j \geqslant 1$, define $\psi^{\dagger,j} = f^{\dagger,j}(0,\cdot)$, according to Proposition 5.8, $\psi^{\dagger,j}$ is $(\mathcal{Q}^j(\mathbb{R}_+))^*$-nondecreasing and Lipschitz. In addition, $H^j_\dagger|_{\mathcal{Q}^j(\mathbb{R}_+)}$ is $(\mathcal{Q}^j(\mathbb{R}_+))^*$-nondecreasing and locally Lipschitz. Let $u^{\dagger,j} \in \mathcal{V}(\mathcal{Q}^j(\mathbb{R}_+))$ be the unique viscosity solution of

$$\begin{cases} \partial_t u^{\dagger,j} - H^j_\dagger\left(\frac{\nabla u^{\dagger,j}}{D}\right) = 0 \text{ on } (0,+\infty) \times \mathcal{Q}^j(\mathring{\mathbb{R}}_+) \\ u^{\dagger,j}(0,\cdot) = \psi^{\dagger,j}. \end{cases}$$

Step 1. We show that, for every $(t,p) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}_+) \cap L^\infty)$,

$$f^\dagger(t,p) = \lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p).$$

According to Proposition 3.6 and Proposition 5.8, there exists $c > 0$ such that for every $j \geqslant 1$ and $(t,x) \in \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}_+)$, we have

$$|f^{\dagger,j}(t,x) - u^{\dagger,j}(t,x)| \leqslant \frac{ct}{j}.$$

Let $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$, according to Proposition 4.4, as $j \to +\infty$ we have $\Lambda^j \Lambda_j p \to p$ in L^1. By Lipschitz continuity of $f^\dagger$ we have

$$f^\dagger(t,p) = \lim_{j \to +\infty} f^{\dagger,j}(t, \Lambda_j p) = \lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p).$$

Step 2. We show that, for every $(t,p) \in \mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}_+) \cap L^\infty)$,

$$\lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p) = u^\dagger(t,p).$$

According to [10, Theorem 1.2 (2) (d)], $u^{\dagger,j}$ admits the Hopf-Lax representation. That is, for every $(t,x) \in \mathbb{R}_+ \times \mathcal{Q}^j(\mathbb{R}_+)$,

$$u^{\dagger,j}(t,x) = \sup_{y \in \mathcal{Q}^j(\mathbb{R}_+)} \inf_{z \in \mathcal{Q}^j(\mathbb{R}_+)} \left\{ \psi^{\dagger,j}(x+y) - \langle y, z \rangle_j + t \int \xi_\dagger \left(\frac{\Lambda^j z}{D} \right) \right\}.$$

Similarly, from [9, Theorem 4.6], we know that for every $(t,p) \in \mathbb{R}_+ \times (Q(\mathbb{R}_+) \cap L^2)$ we have

$$u^\dagger(t,p) = \sup_{r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \inf_{s \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \left\{ \psi^\dagger(p+r) - \langle r, s \rangle_{L^2} + t \int \xi_\dagger \left(\frac{s}{D} \right) \right\}.$$

Step 2.1 We show that, $\lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j q) \leqslant u^\dagger(t,q)$.
Observe that $\mathcal{Q}^j(\mathbb{R}_+) = \{\Lambda_j p, p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty\}$. So,

$$\begin{aligned} u^{\dagger,j}(t, \Lambda_j p) = \sup_{r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \inf_{s \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \Bigg\{ & \psi^\dagger(\Lambda^j \Lambda_j p + \Lambda^j \Lambda_j r) \\ & - \langle \Lambda_j r, \Lambda_j s \rangle_j + t \int \xi(\Lambda^j \Lambda_j s) \Bigg\}. \end{aligned}$$

Since, $\langle \Lambda_j r, \Lambda_j s \rangle_j = \langle \Lambda^j \Lambda_j r, s \rangle_{L^2}$ and $\{\Lambda^j \Lambda_j r, r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty\} \subseteq \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$. We have

$$\begin{aligned} u^{\dagger,j}(t, \Lambda_j p) \leqslant \sup_{r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \inf_{s \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \Bigg\{ & \psi^\dagger(\Lambda^j \Lambda_j p + r) \\ & - \langle r, s \rangle_{L^2} + t \int \xi_\dagger \left(\frac{\Lambda^j \Lambda_j s}{D} \right) \Bigg\}. \end{aligned}$$

Finally, according to Proposition 4.6, we have $\int \xi_\dagger(\Lambda^j \Lambda_j s / D) \leqslant \int \xi_\dagger(s/D)$, so

$$\begin{aligned} u^{\dagger,j}(t, \Lambda_j p) \leqslant \sup_{p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \inf_{r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \Bigg\{ & \psi^\dagger(\Lambda^j \Lambda_j p + r) \\ & - \langle r, s \rangle_{L^2} + t \int \xi_\dagger \left(\frac{s}{D} \right) \Bigg\}. \end{aligned}$$

Using the Lipschitz continuity of $\psi^\dagger$, we discover that

$$\begin{aligned} u^{\dagger,j}(t, \Lambda_j p) \leqslant & |\Lambda^j \Lambda_j p - p|_1 \\ & + \sup_{p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \inf_{r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \left\{ \psi^\dagger(p+r) - \langle r, s \rangle_{L^2} + t \int \xi_\dagger \left(\frac{s}{D} \right) \right\}. \end{aligned}$$

Using Proposition 4.4, we finally obtain $\lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p) \leqslant u^\dagger(t,p)$.
Step 2.2 We show that, $\lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p) \geqslant u^\dagger(t,p)$.

For every $\varepsilon > 0$, there exists $r \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$, such that

$$u^\dagger(t,p) \leqslant \varepsilon + \psi^\dagger(p+r) + \inf_{s \in \mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ -\langle r, s\rangle_{L^2} + t \int \xi_\dagger \left(\frac{s}{D} \right) \right\}.$$

We have $\{\Lambda^j z, z \in \mathcal{Q}^j(\mathbb{R}_+)\} \subseteq \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$. So,

$$u^\dagger(t,p) \leqslant \varepsilon + \psi^\dagger(p+r) + \inf_{z \in \mathcal{Q}^j(\mathbb{R}_+)} \left\{ -\langle r, \Lambda^j z\rangle_{L^2} + t \int \xi_\dagger \left(\frac{\Lambda^j z}{D} \right) \right\}.$$

Using the Lipschitz continuity of $\psi^\dagger$, we have $\psi^\dagger(p+r) \leqslant \psi^{\dagger,j}(\Lambda_j p + \Lambda_j r) + |\Lambda^j \Lambda_j (p+r) - (p+r)|_{L^1}$. And we have $\langle r, \Lambda^j z\rangle_{L^2} = \langle \Lambda_j r, z\rangle_j$. So,

$$\begin{aligned} u^\dagger(t,p) \leqslant \varepsilon + |\Lambda^j \Lambda_j (p+r) - (p+r)|_{L^1} + \psi^{\dagger,j}(\Lambda_j p + \Lambda_j r) \\ + \inf_{z \in \mathcal{Q}^j(\mathbb{R}_+)} \left\{ -\langle \Lambda_j r, z\rangle_j + t \int \xi_\dagger \left(\frac{\Lambda^j z}{D} \right) \right\}. \end{aligned}$$

Since, $\Lambda_j r \in \mathcal{Q}^j(\mathbb{R}_+)$, obtain

$$u^\dagger(t,p) \leqslant \varepsilon + |\Lambda^j \Lambda_j (p+r) - (p+r)|_{L^1} + u^{\dagger,j}(t, \Lambda_j p).$$

Using Proposition 4.4 and letting $j \to +\infty$, we obtain

$$u^\dagger(t,p) \leqslant \varepsilon + \lim_{j \to +\infty} u^{\dagger,j}(t, \Lambda_j p),$$

since $\varepsilon > 0$ is arbitrary this concludes Step 2.

Step 3. Conclusion.

From Step 1 and Step 2, we deduce that $f^\dagger$ and $u^\dagger$ coincide on $\mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}_+) \cap L^\infty)$. Since $\mathcal{Q}(\mathbb{R}_+) \cap L^\infty$ is dense in $\mathcal{Q}(\mathbb{R}_+) \cap L^2$ with respect to L^1 convergence and $f^\dagger$ and $u^\dagger$ are both Lipschitz continuous with respect to L^1, we conclude that $f^\dagger$ and $u^\dagger$ coincide on $\mathbb{R}_+ \times (\mathcal{Q}(\mathbb{R}_+) \cap L^2)$. □

6. Uniqueness of the optimizer for diagonal models

Let $\mathcal{P}^\uparrow(\mathbb{R}^D_+)$ denote the set of probability measures on $\mathbb{R}^D_+$ of the form $\mathrm{Law}(q(U))$ with $q \in \mathcal{Q}(\mathbb{R}^D_+)$ and U a uniform random variable on $[0,1)$. The map $q \mapsto \mathrm{Law}(q(U))$ is an isometric bijection from $\mathcal{Q}(\mathbb{R}^D_+)$ to $\mathcal{P}^\uparrow(\mathbb{R}^D_+)$. Therefore, we can also think of the Parisi functional

$$q \mapsto \psi(q) - t \int_0^1 \xi^* \left(\frac{q}{t} \right),$$

as a functional depending on $\mu \in \mathcal{P}^\uparrow(\mathbb{R}^D_+)$ rather than $q \in \mathcal{Q}(\mathbb{R}^D_+)$. In a sense, replacing $\mathcal{Q}(\mathbb{R}^D_+)$ by $\mathcal{P}^\uparrow(\mathbb{R}^D_+)$ changes the geometry. Given $\mu_0, \mu_1 \in \mathcal{P}^\uparrow(\mathbb{R}^D_+)$ and $q_0, q_1 \in \mathcal{Q}(\mathbb{R}^D_+)$ such that $\mu_i = \mathrm{Law}(q_i(U))$, in general, for $\lambda \in (0,1)$ we have

$$\lambda \mu_1 + (1-\lambda)\mu_0 \neq \mathrm{Law}(\lambda q_1(U) + (1-\lambda) q_0(U)).$$

When $D = 1$, this allows us to reveal a hidden concavity property of the Parisi functional, according to [1, Theorem 2], the Parisi functional is strictly concave on $\mathcal{P}^\uparrow(\mathbb{R}_+) = \mathcal{P}(\mathbb{R}_+)$. This is the key property used to establish the uniqueness of Parisi measures for models with $D = 1$ [1, Corollary 1].

In the setting, $D > 1$ this approach seems to fail. A difficulty is that the convexity of the set $\mathcal{P}^{\uparrow}(\mathbb{R}_+)$ is an exception rather than the rule. For every $D > 1$, the set $\mathcal{P}^{\uparrow}(\mathbb{R}_+^D)$ is not convex, since for example, $\frac{\delta_{(0,1)}+\delta_{(1,0)}}{2} \notin \mathcal{P}^{\uparrow}(\mathbb{R}_+^2)$. This is problematic, as given two possible maximizing probability measures $\mu_0 \neq \mu_1$ in $\mathcal{P}^{\uparrow}(\mathbb{R}_+^D)$, the Parisi functional may not even be defined at $\frac{\mu_0+\mu_1}{2}$.

In [6, Theorem 1.2], it was pointed out that, that the restriction of the Parisi functional to any "one-dimensional subspace of $\mathcal{P}^{\uparrow}(\mathbb{R}_+^D)$" is strictly concave. Thanks to Theorem 1.3, the Parisi formula can be written as a supremum over probability measures in $\mathcal{P}^{\uparrow}(\mathbb{R}_+^D)$ supported on $\{(\lambda, \dots, \lambda), \lambda \geqslant 0\}$. This space is of course a "one-dimensional subspace" in the sense of [6, Theorem 1.2], this allows us to prove uniqueness of Parisi measures.

Let $\mathcal{P}_1(\mathbb{R}_+)$ denote the set of Borel probability measures on $\mathbb{R}_+$ with finite first moment. Given $\mu \in \mathcal{P}_1(\mathbb{R}_+)$, for every $u \in [0,1)$, we define

$$p_\mu(u) = \inf\{x \geqslant 0 \,|\, \mu([0,x]) > u\}.$$

The path $p_\mu \in \mathcal{Q}(\mathbb{R}_+) \cap L^1$ is the quantile function of the probability measure μ and classically, $\mathrm{Law}(p_\mu(U)) = \mu$.

Proposition 6.1 ([6]). *The map $\mu \mapsto \psi^{\dagger}(p_\mu)$ is strictly concave on $\mathcal{P}_1(\mathbb{R}_+)$.*

Proof. First, note that in [6] the Parisi functional depends on paths in,

$$\Pi(S_+^D) = \{\pi : [0,1] \to S_+^D \,|\, \pi \text{ is left-continuous and nondecreasing}\}.$$

The Parisi functional is continuous with respect to $|\cdot|_{L^1}$, so one can replace paths $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$ with their left continuous version without affecting the value of the Parisi functional. Hence, the distinction between Π and $\mathcal{Q}$ is simply a matter of taste. Let $\widetilde{\xi} : \mathbb{R}^{D\times D} \to \mathbb{R}$ be a convex function satisfying (1.1) for some Gaussian process $\widetilde{H}_N$. Presumably, as a consequence of [6, Theorem 1.2], choosing $\Psi = s\mathrm{id}_D$, we only have that the function $\mu \mapsto \psi(\nabla\widetilde{\xi} \circ \Psi \circ p_\mu)$ is strictly concave on $\mathcal{P}_1(\mathbb{R}_+)$. But, ψ is independent of the choice of $\widetilde{\xi}$, and with $\widetilde{\xi}(R) = \frac{1}{2}\sum_{d,d'=1}^D R_{dd'}^2$ we obtain that $\mu \mapsto \psi^{\dagger}(p_\mu)$ is strictly concave on $\mathcal{P}_1(\mathbb{R}_+)$. □

Proof of Theorem 1.3. The variational formula (1.9) follows from Theorem 5.9, we now prove existence and uniqueness of an optimizer in (1.9). In the first three steps of the proof, we explain that in (1.9), the supremum can be taken over a subset of paths bounded in L^∞, this yields existence of an optimizer. The last step is dedicated to showing uniqueness of the optimizer through Proposition 6.1.

Step 1. We show that there exists $c \geqslant 0$, such that for every $\lambda > \lambda' \geqslant c$ we have

$$\frac{\xi_{\dagger}^*(\lambda) - \xi_{\dagger}^*(\lambda')}{\lambda - \lambda'} \geqslant 1.$$

Let $a > 0$, there exists $b \geqslant 0$ such that $\xi_{\dagger}(\lambda) \leqslant b$ on $[0,a]$. Let $I_{[0,a]}$ denote the convex function taking the value 0 on $[0,a]$ and $+\infty$ otherwise. We have

$\xi_\dagger \leqslant b + I_{[0,a]}$ on $\mathbb{R}_+$. Thus, for $\lambda \geqslant 0$,

$$\begin{aligned}\xi_\dagger^*(\lambda) &= \sup_{\lambda' \geqslant 0} \left\{ \lambda\lambda' - \xi_\dagger(\lambda) \right\} \\ &\geqslant \sup_{\lambda' \geqslant 0} \left\{ \lambda\lambda' - b - I_{[0,a]} \right\} \\ &= \sup_{\lambda' \in [0,a]} \left\{ \lambda\lambda' - b \right\} \\ &= a\lambda - b.\end{aligned}$$

This imposes $\liminf_{\lambda \to +\infty} \frac{\xi_\dagger^*(\lambda)}{\lambda} \geqslant a$. Since $a > 0$ is arbitrary, we have

$$\liminf_{\lambda \to +\infty} \frac{\xi_\dagger^*(\lambda)}{\lambda} = +\infty.$$

Let $c \geqslant 0$ be so that for every $\lambda \geqslant c$,

$$\frac{\xi_\dagger^*(\lambda) - \xi_\dagger^*(0)}{\lambda} \geqslant 1.$$

Then, since $\xi_\dagger^*$ is convex, its slope increases and for every $\lambda > \lambda' \geqslant c$, we have

$$\frac{\xi_\dagger^*(\lambda) - \xi_\dagger^*(\lambda')}{\lambda - \lambda'} \geqslant \frac{\xi_\dagger^*(\lambda) - \xi_\dagger^*(0)}{\lambda} \geqslant 1.$$

This concludes Step 1.

Step 2. Let $L^\infty_{\leqslant tc}$ denote the ball of center 0 and radius tc in L^∞. We show that

$$\sup_{p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty} \left\{ \psi^\dagger(p) - \int \xi_\dagger^* \left(\frac{p}{t} \right) \right\} = \sup_{p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty_{\leqslant tc}} \left\{ \psi^\dagger(p) - \int \xi_\dagger^* \left(\frac{p}{t} \right) \right\}.$$

Let LHS and RHS denote the left-hand side and right-hand side in the previous display. It is clear that LHS $\geqslant$ RHS, we only prove the other inequality. Let $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$, and let $\widetilde{p} = p \wedge (t\bar{c})$ be the path which coincides with p until $u^* = \inf\{u \in [0,1), p(u) > tc\}$ and is constant $= tc$ on $[u^*, 1)$. We have

$$\begin{aligned}&\left(\psi^\dagger(p) - t\int_0^1 \xi_\dagger^* \left(\frac{p}{t} \right) \right) - \left(\psi^\dagger(\widetilde{p}) - t\int_0^1 \xi_\dagger^* \left(\frac{\widetilde{p}}{t} \right) \right) \\ &\leqslant \int_0^1 |p - \widetilde{p}| - t\int_0^1 \left(\xi_\dagger^* \left(\frac{p}{t} \right) - \xi_\dagger^* \left(\frac{\widetilde{p}}{t} \right) \right) \\ &= \int_{u^*}^1 p - tc - t\left(\xi_\dagger^* \left(\frac{p}{t} \right) - \xi_\dagger^*(c) \right) \\ &= \int_{u^*}^1 (p - tc) \left(1 - \frac{\xi_\dagger^* \left(\frac{p}{t} \right) - \xi_\dagger^*(c)}{\frac{p}{t} - c} \right) \\ &\leqslant 0.\end{aligned}$$

Since $|\widetilde{p}|_{L^\infty} \leqslant tc$, this concludes Step 2.

Step 3. We show that the sup in the variational formula of Theorem 1.3 is reached at some $p^* \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$.

The functional $p \mapsto \psi^\dagger(p) - t \int_0^1 \xi_\dagger^* \left(\frac{p}{t}\right)$ is continuous with respect to the L^1-norm. So, according to Step 2 it is enough to prove that the set $\mathcal{Q}(\mathbb{R}_+) \cap L^\infty_{\leqslant tc}$ is compact with respect to L^1 convergence. Let $(p_n)_n$ be a sequence of paths in $\mathcal{Q}(\mathbb{R}_+)$ which is bounded by tc in $|\cdot|_{L^\infty}$. By a diagonal argument, up to extraction (not relabelled) we may assume that for every $u \in \mathbb{Q} \cap [0,1)$, we have $\lim_{n \to +\infty} p_n(u) = p(u)$ for some $p(u) \in \mathbb{R}_+$. The map $u \mapsto p(u)$ is nondecreasing and bounded by tc on $\mathbb{Q} \cap [0,1)$. For every $v \in [0,1)$, the quantity $p(u)$ converges as $u \to v$ in $\mathbb{Q} \cap [v,1)$, we denote by $p(v)$ its limiting value. The function $p : [0,1) \to \mathbb{R}_+$ thus defined is càdlàg, nondecreasing and bounded by tc. Since p is monotone, the set of discontinuities of p is countable. Let $v \in [0,1)$ be a point at which p is continuous, let $u \in \mathbb{Q} \cap [v,1)$ we have $p_n(v) \leqslant p_n(u)$, letting $n \to +\infty$ we obtain $\limsup_{n \to +\infty} p_n(v) \leqslant p(u)$ and since p is continuous at v, letting $u \to v$ yields $\limsup_{n \to +\infty} p_n(v) \leqslant p(v)$. By considering $u \in \mathbb{Q} \cap [0,v]$, we can repeat the same argument to discover that $\liminf_{n \to +\infty} p_n(v) \geqslant p(v)$. In conclusion, $p_n \to p$ pointwise on $[0,1)$ outside a countable set of points. Since $|p_n|_{L^\infty} \leqslant tc$, by dominated convergence, it follows that $p_n \to p$ in L^1.

Step 4. We show that the sup in the variational formula of Theorem 1.3 is reached at most at one $p^* \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$.

By contradiction, assume that there are two maximizers $p_0 \neq p_1$ of (1.9) in $\mathcal{Q}(\mathbb{R}_+) \cap L^\infty$. Let $\mu_i \in \mathcal{P}(\mathbb{R}_+)$ be the law of the random variable $p_i(U)$ where U is a uniform random variable in $[0,1)$. Let $p = p_\mu \in \mathcal{Q}(\mathbb{R}_+) \cap L^\infty$ denote the quantile function of $\mu = \frac{\mu_0 + \mu_1}{2} \in \mathcal{P}(\mathbb{R}_+)$. According to Proposition 6.1, we have

$$\psi^\dagger(p) > \frac{\psi^\dagger(p_0) + \psi^\dagger(p_1)}{2}.$$

By definition

$$\int \xi_\dagger^* \left(\frac{p}{t}\right) = \frac{1}{2} \int \xi_\dagger^* \left(\frac{p_0}{t}\right) + \frac{1}{2} \int \xi_\dagger^* \left(\frac{p_1}{t}\right).$$

Therefore,

$$\psi^\dagger(p) - t \int \xi_\dagger^* \left(\frac{p}{t}\right) > \frac{1}{2}\left(\psi^\dagger(p_0) - t \int \xi_\dagger^* \left(\frac{p_0}{t}\right)\right) + \frac{1}{2}\left(\psi^\dagger(p_1) - t \int \xi_\dagger^* \left(\frac{p_1}{t}\right)\right).$$

This is a contradiction since in the previous display the left-hand side is upper-bounded by $\lim_{N \to +\infty} \overline{F}_N(t)$ and the right-hand side is equal to $\lim_{N \to +\infty} \overline{F}_N(t)$.

□

7. Upper-bound for nonconvex models

In this section, we assume that the interaction function ξ only depends on the diagonal coefficients of its argument, and we do *not* assume that ξ is convex. We give a proof of Theorem 1.4. We believe that this Theorem is related to the approach developed in [17, 15] where the limit free energy of nonconvex models is lower bounded in terms of the value at $(t,0)$ of the viscosity solution of some Hamilton-Jacobi equation. Theorem 1.4 will follow from a classic interpolation argument after observing that ξ satisfies the following inequality

$$\forall x \in \mathbb{R}_+^D,\ \xi(x) \leqslant \frac{1}{D} \sum_{d=1}^{D} \xi(x_d, \dots, x_d).$$

7.1. **An inequality for permutation-invariant covariance functions.**

Proposition 7.1. *Let $i_1, \dots, i_D \in \mathbb{N}$ and $I = \sum_{d=1}^{D} i_d$, for every $x_1, \dots, x_D \in \mathbb{R}_+$, we have*

$$\frac{1}{D!} \sum_{s \in \mathscr{S}_D} \prod_{d=1}^{D} x_{s(d)}^{i_d} \leqslant \frac{1}{D} \sum_{d=1}^{D} x_d^I. \tag{7.1}$$

In addition, if I is even then (7.1) holds for $x_1, \dots, x_D \in \mathbb{R}$.

Proof. If $I = 0$ then the inequality is clear, otherwise by the inequality of arithmetic and geometric means we have

$$\prod_{d=1}^{D} x_{s(d)}^{i_d} \leqslant \frac{1}{I} \sum_{\delta=1}^{D} i_\delta x_{s(\delta)}^I.$$

Summing over $s \in \mathscr{S}_D$, we obtain

$$\frac{1}{D!} \sum_{s \in \mathscr{S}_D} \prod_{d=1}^{D} x_{s(d)}^{i_d} \leqslant \frac{1}{I} \sum_{\delta=1}^{D} i_\delta \frac{1}{D!} \sum_{s \in \mathscr{S}_D} x_{s(\delta)}^I.$$

Observing that $\frac{1}{D!} \sum_{s \in \mathscr{S}_D} x_{s(\delta)}^I = \frac{1}{D} \sum_{d=1}^{D} x_d^I$ is independent of δ, we obtain (7.1). Assume now that I is even, and let $y \in \mathbb{R}^D$, define $x_d = |y_d|$. Let E be the set of $d \in \{1, \dots, D\}$ such that $y_d < 0$. We have

$$\prod_{d=1}^{D} y_{s(d)}^{i_d} = (-1)^{\sum_{s(d) \in E} i_d} \prod_{d=1}^{D} x_{s(d)}^{i_d} \leqslant \prod_{d=1}^{D} x_{s(d)}^{i_d}.$$

Summing over $s \in \mathscr{S}_D$ and applying (7.1) to x, we obtain

$$\begin{aligned} \frac{1}{D!} \sum_{s \in \mathscr{S}_D} \prod_{d=1}^{D} y_{s(d)}^{i_d} &\leqslant \frac{1}{D!} \sum_{s \in \mathscr{S}_D} \prod_{d=1}^{D} x_{s(d)}^{i_d} \\ &\leqslant \frac{1}{D} \sum_{d=1}^{D} x_d^I \\ &= \frac{1}{D} \sum_{d=1}^{D} y_d^I. \end{aligned}$$

Where the last line follows from the fact that $x_d = \pm y_d$ and I is even. □

Recall that we have defined for every $\lambda \in \mathbb{R}$, $\xi_\dagger(\lambda) = \xi(\lambda, \dots, \lambda)$ and we say that ξ is permutation-invariant when for every $s \in \mathscr{S}_D$ and $x \in \mathbb{R}^D$, $\xi(x^s) = \xi(x)$. Let $A \in \mathbb{R}^{D\times D}$ be a $D \times D$ matrix and $p \in \mathbb{N}^*$, we denote

$$A^{\otimes p} = (A_{i_1,j_1} \times \dots \times A_{i_p,j_p})_{1\leqslant i_1,\dots,i_p,j_1,\dots,j_p\leqslant D} \in \mathbb{R}^{D^p\times D^p}$$

the p-fold tensor product of A with itself. For every $C, C' \in \mathbb{R}^{D^p\times D^p}$ we define

$$C \cdot C' = \sum_{1\leqslant i_1,\dots,i_p,j_1,\dots,j_p\leqslant D} C_{(i_1,j_1),\dots,(i_p,j_p)} C'_{(i_1,j_1),\dots,(i_p,j_p)}.$$

Proposition 7.2. *Let $\xi : \mathbb{R}^D \to \mathbb{R}$, such that ξ is permutation-invariant and admits an absolutely convergent power series. Assume that for every, $N \geqslant 1$ there exists a Gaussian process $(H_N(\sigma))_{\sigma\in\mathbb{R}^{D\times N}}$ such that for every $\sigma, \tau \in \mathbb{R}^{D\times N}$,*

$$\mathbb{E}[H_N(\sigma)H_N(\tau)] = N\xi\left(\frac{\sigma_1\cdot\tau_1}{N}, \dots, \frac{\sigma_D\cdot\tau_D}{N}\right).$$

Then, $\xi_\dagger$ is convex on $\mathbb{R}_+$ and for every $x \in \mathbb{R}^D_+$, we have

$$\xi(x) \leqslant \frac{1}{D}\sum_{d=1}^{D} \xi_\dagger(x_d). \tag{7.2}$$

Remark 7.3. For some ξ, (7.2) is not satisfied on $\mathbb{R}^D$. For example, assume that $D = 2$ and consider $\xi(x_1, x_2) = x_1x_2(x_1 + x_2)$, we have $\xi_\dagger(\lambda) = 2\lambda^3$, and

$$\xi(-2, 1) = 2 > -7 = \frac{\xi_\dagger(-2) + \xi_\dagger(1)}{2}.$$

Proof. Given $\sigma, \tau \in \mathbb{R}^{D\times N}$, we consider their overlap matrix

$$\frac{\sigma\tau^*}{N} = \left(\frac{\sigma_d\cdot\tau_{d'}}{N}\right)_{1\leqslant d,d'\leqslant D},$$

the diagonal of the overlap matrix $\frac{\sigma\tau^*}{N}$ is the overlap vector

$$R(\sigma,\tau) = \left(\frac{\sigma_1\cdot\tau_1}{N}, \dots, \frac{\sigma_D\cdot\tau_D}{N}\right).$$

For every $A \in \mathbb{R}^{D\times D}$, we define

$$\overline{\xi}(A) = \xi(A_{11}, \dots, A_{dd}).$$

We have $\overline{\xi}(\frac{\sigma\tau^*}{N}) = \xi\left(\frac{\sigma_1\cdot\tau_1}{N}, \dots, \frac{\sigma_D\cdot\tau_D}{N}\right)$. According to [17, Poposition 6.6], there exists a sequence of matrices $(C^{(p)})_{p\geqslant1}$ such that $C^{(p)} \in S^{D^p}_+$ and, for every $A \in \mathbb{R}^{D\times D}$,

$$\overline{\xi}(A) = \sum_{p\geqslant1} C^{(p)} \cdot A^{\otimes p}.$$

Since $\overline{\xi}(A)$ only depends on the diagonal of A, the matrices $C^{(p)}$ must be diagonal with nonnegative coefficients. In particular, there exists nonnegative real numbers $(a_{i_1,\ldots,i_D})_{i_1,\ldots,i_D\geqslant 0}$, such that for every $x\in\mathbb{R}^D$,

$$\xi(x) = \sum_{i_1,\ldots,i_D\geqslant 0} a_{i_1,\ldots,i_D}\prod_{d=1}^D x_d^{i_d}.$$

Therefore, for every $\lambda\in\mathbb{R}$,

$$\xi_\dagger(\lambda) = \sum_{I\geqslant 0}\lambda^I \sum_{\substack{i_1,\ldots,i_D\geqslant 0\\ \sum_{\delta=1}^D i_\delta = I}} a_{i_1,\ldots,i_D}$$

This shows that $\xi_\dagger$ has a power series expansion with nonnegative coefficients, so $\xi_\dagger$ is convex on $\mathbb{R}_+$. In addition, since ξ is permutation-invariant, given $x\in\mathbb{R}_+^D$ we have

$$\begin{aligned}\xi(x) &= \frac{1}{D!}\sum_{s\in\mathscr{S}_D}\xi(x^s)\\ &= \sum_{i_1,\ldots,i_D\geqslant 0} a_{i_1,\ldots,i_D}\frac{1}{D!}\sum_{s\in\mathscr{S}_D}\prod_{d=1}^D x_{s(d)}^{i_d}.\end{aligned}$$

According to Proposition 7.1, we thus have

$$\begin{aligned}\xi(x) &\leqslant \sum_{i_1,\ldots,i_D\geqslant 0} a_{i_1,\ldots,i_D}\frac{1}{D}\sum_{d=1}^D x_d^{\sum_{\delta=1}^D i_\delta}\\ &= \frac{1}{D}\sum_{d=1}^D\sum_{I\geqslant 0} x_d^I \sum_{\substack{i_1,\ldots,i_D\geqslant 0\\ \sum_{\delta=1}^D i_\delta = I}} a_{i_1,\ldots,i_D}\\ &= \frac{1}{D}\sum_{d=1}^D \xi_\dagger(x_d).\end{aligned}$$

□

Remark 7.4. As a consequence of the proof of (7.2) we have just given and the second part of Proposition 7.1, if the power series of ξ only has terms of even degree, then (7.2) holds for $x\in\mathbb{R}^D$.

7.2. **Positivity principle for multiple species.** Roughly speaking, the positivity principle is the statement that if $\sigma,\tau\in\mathbb{R}^{D\times N}$ are independent random variables drawn from the Gibbs measure, then almost surely in the limit $N\to+\infty$ the overlaps $\frac{\sigma_d\cdot\tau_d}{N}$ are all nonnegative. When $D=1$, it is well known that the positivity principle is verified as soon as the Gibbs measure satisfies the so-called Ghirlanda–Guerra identities [20, Theorem 3.4]. When $D>1$ and P_N is the uniform probability measure on a product of D spheres, it was shown that the statement still holds true [3, Section 3.2]. In this section, for the convenience of the reader and the sake of completeness, we briefly explain how the arguments given to justify [3, Lemma 3.3] can be adapted into a proof of the positivity principle for models with $D>1$ and

$P_N = P_1^{\otimes N}$. The proofs are unchanged, and we simply replace the uniform probability measure on a product of D spheres by $P_1^{\otimes N}$.

Given $\sigma, \tau \in \mathbb{R}^{D\times N}$, we let $R(\sigma,\tau) = (\frac{\sigma_d \cdot \tau_d}{N})_{1\leqslant d \leqslant D}$. Given $w \in [0,1]^D$, we define

$$R^w(\sigma,\tau) = \frac{w\cdot R(\sigma,\tau)}{D} = \frac{1}{D}\sum_{d=1}^{D} w_d \frac{\sigma_d \cdot \tau_d}{N}.$$

We let $\mathscr{W} = (w_q)_{q\geqslant 1}$ denote a countable dense subset of $[0,1]^D$. We assume that $\mathscr{W}$ contains $e_1 \dots, e_D$ and does not contain the null vector. We fix $p \geqslant 1$, for every $q \geqslant 1$, we consider the centered Gaussian Process $H^{\text{pert}}_{N,p,q}$ with covariance

$$\mathbb{E} H^{\text{pert}}_{N,p,q}(\sigma) H^{\text{pert}}_{N,p,q}(\tau) = \frac{N}{4^{p+q}} \left(R^{w_q}(\sigma,\tau)\right)^p. \tag{7.3}$$

Such a process can be explicitly defined by setting

$$H^{\text{pert}}_{N,p,q}(\sigma) = \frac{2^{-(p+q)}}{N^{\frac{p-1}{2}} D^{\frac{p}{2}}} \sum_{d_1,\dots,d_p=1}^{D} \sum_{i_1,\dots,i_p=1}^{N} J_{(d_1,\dots,d_p),(i_1,\dots,i_p)} \prod_{k=1}^{p} \sqrt{w_{d_k}} \sigma_{d_k,i_k} \tau_{d_k,i_k},$$

Where the coefficients $J_{(d_1,\dots,d_p),(i_1,\dots,i_p)}$ are independent standard normal random variables. We fix $(u_{p,q})_{p,q\geqslant 1}$ a sequence of numbers in $[1,2]$ and we let

$$H^{\text{pert}}_N(\sigma) = \sum_{p,q\geqslant 1} u_{p,q} H^{\text{pert}}_{N,p,q}(\sigma). \tag{7.4}$$

Soon we are going to choose random coefficients $(u_{p,q})_{p,q\geqslant 1}$ that are independent and uniformly distributed in $[1,2]$ and independent of every other source of randomness. We denote by $\mathbb{E}_u$ the expectation with respect to this law. We let $c_N = N^{-\omega}$ with $0 < \omega < 1/2$, and for every nonrandom function $H : \mathbb{R}^{D\times N} \to \mathbb{R}$, we consider

$$\widetilde{H}_N(\sigma) = H(\sigma) + c_N H^{\text{pert}}_N(\sigma).$$

First, note that by introducing the perturbation $c_N H^{\text{pert}}_N$ we do not change the limiting value of the free energy. Indeed, by applying Jensen's inequality twice, we obtain

$$0 \leqslant \left(-\frac{1}{N}\mathbb{E}\log\int e^{H(\sigma)}\mathrm{d}P_N(\sigma)\right) - \left(-\frac{1}{N}\mathbb{E}\log\int e^{\widetilde{H}(\sigma)}\mathrm{d}P_N(\sigma)\right) \leqslant \frac{c_N^2}{2}. \tag{7.5}$$

We let $\widetilde{G}_N$ denote the Gibbs measure associated to $\widetilde{H}_N$, that is

$$d\widetilde{G}_N(\sigma) \propto e^{\widetilde{H}_N(\sigma)} dP_N(\sigma). \tag{7.6}$$

We denote by $\mathbb{E}_{H^{\text{pert}}_N}$ the expectation with respect to the randomness of the Gaussian process H^{pert}_N. We denote by $\widetilde{G}_N^{\otimes k}(A)$ the probability of the event A under the probability measure $\widetilde{G}_N^{\otimes k}$. As usual, let σ, τ denote two random variables sampled independently with law $\widetilde{G}_N$. The following lemma is [3, Lemma 3.3].

Lemma 7.5 ([3]). *For every $\varepsilon > 0$,*

$$\lim_{N\to+\infty}\sup_{H(\sigma)}\mathbb{E}_u\mathbb{E}_{H_N^{pert}}\widetilde{G}_N^{\otimes 2}\left(\exists d,\ \frac{\sigma_d\cdot\tau_d}{N}\leqslant-\varepsilon\right)=0,$$

where the sup is taken over all nonrandom functions $H:\mathbb{R}^{D\times N}\to\mathbb{R}$ satisfying $\int\exp|H(\sigma)|\mathrm{d}P_N(\sigma)<+\infty$ and $\mathbb{E}_u$ denote the expectation with respect to the parameters $(u_{p,q})_{p,q\geqslant 1}$ in (7.4).

Proof. It suffices to show that for every $d\in\{1,\ldots,D\}$,

$$\lim_{N\to+\infty}\sup_{H(\sigma)}\mathbb{E}_u\mathbb{E}_{H_N^{\mathrm{pert}}}\widetilde{G}_N^{\otimes 2}\left(\frac{\sigma_d\cdot\tau_d}{N}\leqslant-\varepsilon\right)=0.$$

Let $(\sigma^\ell)_{\ell\geqslant 1}$ be independent and identically distributed random variables with law $\widetilde{G}_N$. We let $R_{\ell,\ell'}=R(\sigma^\ell,\sigma^{\ell'})$, and for every $d\in\{1,\ldots,D\}$, $R^d_{\ell,\ell'}=\frac{\sigma_d^\ell\cdot\sigma_d^{\ell'}}{N}$. The result will follow from the fact that the array $(R^d_{\ell,\ell'})_{\ell,\ell'\geqslant 1}$ satisfies the Ghirlanda-Guerra identities in the limit $N\to+\infty$. More precisely, if we define for any $n\geqslant 1$, $R(n)=(R_{\ell,\ell'})_{1\leqslant\ell,\ell'\leqslant n}$ and for every $d\in\{1,\ldots,D\}$, for every real valued bound measurable function $g=g(R(n))$ and every continuous function, $h:\mathbb{R}\to\mathbb{R}$,

$$\Delta^d(g,n,h)=\Bigg|\mathbb{E}\widetilde{G}_N\left(gh(R^d_{1,n+1})\right)-\frac{1}{n}\mathbb{E}\widetilde{G}_N\left(gh(R^d_{1,2})\right)-\frac{1}{n}\sum_{l=2}^{n}\mathbb{E}\widetilde{G}_N\left(gh(R^d_{1,\ell})\right)\Bigg|.$$

Then, according to [3, Theorem A.3], we have

$$\lim_{N\to+\infty}\sup_{H}\mathbb{E}_u\Delta^d(g,n,h)=0.$$

We can then proceed as in [20, Theorem 3.4] to deduce the desired result. □

7.3. **The interpolation.** Given $\varsigma\in\mathbb{R}^N$, we define

$$H_N^{\mathrm{sym}}(\varsigma)=H_N(\varsigma,\ldots,\varsigma).$$

We let $H_N^{\mathrm{sym},1},\ldots,H_N^{\mathrm{sym},D}$ be D independent copies of the Gaussian process H_N^{sym} and we assume that $(H_N^{\mathrm{sym},1},\ldots,H_N^{\mathrm{sym},D})$ is independent of H_N. For every $\sigma=(\sigma_1,\ldots,\sigma_D)\in\mathbb{R}^{D\times N}$, we define

$$K_N(\sigma)=\frac{1}{\sqrt{D}}\sum_{d=1}^{D}H_N^{\mathrm{sym},d}(\sigma_d).\tag{7.7}$$

The Gaussian process K_N satisfies

$$\mathbb{E}\left[K_N(\sigma)K_N(\tau)\right]=N\Xi\left(R(\sigma,\tau)\right),$$

where $\Xi(x)=\frac{1}{D}\sum_{d=1}^{D}\xi_\dagger(x_d)$. Note that, $\Xi_\dagger=\xi_\dagger$, also note that according to Proposition 7.2, the function Ξ is convex on $\mathbb{R}^D_+$. In particular, the family of Gaussian processes $(K_N)_{N\geqslant 1}$ is covered by Theorem 1.3. Thus, the free energy of the Gaussian process K_N converges when $N \to +\infty$ and the limit free energy can be expressed using a variational formula. We let $\overline{F}_N$ denote the free energy of H_N as in (1.2), and we let $\overline{G}_N$ denote the free energy of K_N.

Theorem 7.6. *Assume that ξ satisfies (7.2), we have for every $t \geqslant 0$,*

$$\limsup_{N\to+\infty} \overline{F}_N(t) \leqslant \lim_{N\to+\infty} \overline{G}_N(t). \tag{7.8}$$

Remark 7.7. According to Proposition 7.2, as soon as ξ is permutation-invariant, ξ satisfies (7.2). So we can deduce Theorem 1.4 from Theorem 7.6 by applying Theorem 1.3 to the Hamiltonian K_N and observing that,

$$\left(\xi_\dagger\left(\frac{\cdot}{D}\right)\right)^*(\lambda) = \Xi^*(\lambda,\dots,\lambda).$$

Remark 7.8. Note that in fact, as stated, neither P_1 nor ξ are required to be permutation-invariant for Theorem 7.6 to hold. But if P_1 or ξ is not permutation-invariant, there is no reason to believe that (7.8) is a tight upper bound, and in fact the proof of Theorem 7.10 given below breaks down in this case.

Proof. Let H_N^{pert} be the Gaussian process defined in (7.4), we take it independent of H_N and K_N. Recall that the covariance of H_N^{pert} depends on some parameter u and we denote $\mathbb{E}_u$ the expectation with respect to this parameter. Without loss of generality, throughout the proof we assume that $t = 1/2$. For every $\lambda \in [0,1]$, define

$$\begin{aligned} H_{N,\lambda}(\sigma) = &\left(\sqrt{1-\lambda}H_N(\sigma) - (1-\lambda)\frac{N}{2}\xi\left(R(\sigma,\sigma)\right)\right) \\ &+ \left(\sqrt{\lambda}K_N(\sigma) - \lambda\frac{N}{2}\Xi\left(R(\sigma,\sigma)\right)\right) \\ &+ c_N H_N^{\mathrm{pert}}(\sigma). \end{aligned}$$

Where $c_N = N^{-\omega}$ with $0 < \omega < 1/2$. We let $\langle\cdot\rangle_\lambda$ denote the Gibbs measure associated to $H_{N,\lambda}(\sigma)$ and we also define the associated free energy

$$\varphi_N(\lambda) = -\frac{1}{N}\mathbb{E}\log\int \exp\left(H_{N,\lambda}(\sigma)\right)\mathrm{d}P_N(\sigma)$$

Here, as previously, we use the symbol $\mathbb{E}$ to denote expectation with respect to every source of randomness involved. In particular, $\mathbb{E}$ integrates out the randomness of u and $\varphi_N(\lambda)$ is a nonrandom quantity. Note that crucially, thanks to (7.5), the nature of the term $c_N H_N^{\mathrm{pert}}$ is indeed perturbative as it does not influence the value of the limit free energy of $\langle\cdot\rangle_\lambda$. Inequality (7.8) can be rewritten as

$$\limsup_{N\to+\infty}\varphi_N(0) \leqslant \limsup_{N\to+\infty}\varphi_N(1).$$

We will prove this inequality by observing that φ_N is almost a nondecreasing function of λ. For every $\lambda \in (0,1)$, φ_N is differentiable at λ and we have

$$\begin{aligned}\varphi_N'(\lambda) &= -\frac{1}{N}\mathbb{E}\left\langle \frac{d}{d\lambda}H_{N,\lambda}(\sigma)\right\rangle_\lambda \\ &= -\frac{1}{N}\left(\mathbb{E}\left\langle -\frac{1}{2\sqrt{1-\lambda}}H_N(\sigma) + \frac{1}{2\sqrt{\lambda}}K_N(\sigma) + \frac{N}{2}(\xi-\Xi)\left(R(\sigma,\sigma)\right)\right\rangle_\lambda\right),\end{aligned}$$

Let us now treat the first term in the above display using Gaussian integration by parts [20, Lemma 1.1]. For every $\sigma, \tau \in \mathbb{R}^D$, we have

$$\begin{aligned}&\mathbb{E}\left[\left(-\frac{1}{2\sqrt{1-\lambda}}H_N(\sigma) + \frac{1}{2\sqrt{\lambda}}K_N(\sigma)\right)H_{N,\lambda}(\tau)\right] \\ &= \mathbb{E}\left[-\frac{1}{2\sqrt{1-\lambda}}H_N(\sigma)\sqrt{1-\lambda}H_N(\tau)\right] + \mathbb{E}\left[\frac{1}{2\sqrt{\lambda}}K_N(\sigma)\sqrt{\lambda}K_N(\tau)\right] \\ &+ \mathbb{E}[c_N H_N^{\mathrm{pert}}(\sigma)]\mathbb{E}\left[-\frac{1}{2\sqrt{1-\lambda}}H_N(\sigma) + \frac{1}{2\sqrt{\lambda}}K_N(\sigma)\right] \\ &+ ((1-\lambda)\xi + \lambda\Xi)\,\mathbb{E}\left[-\frac{1}{2\sqrt{1-\lambda}}H_N(\sigma) + \frac{1}{2\sqrt{\lambda}}K_N(\sigma)\right] \\ &= -\frac{1}{2}N\xi\left(R(\sigma,\tau)\right) + \frac{1}{2}N\Xi\left(R(\sigma,\tau)\right) + 0 + 0 \\ &= \frac{N}{2}(\Xi-\xi)\left(R(\sigma,\tau)\right).\end{aligned}$$

Thus, the Gaussian integration by parts formula yields

$$\varphi_N'(\lambda) = \frac{1}{2}\mathbb{E}\left\langle(\Xi-\xi)\left(R(\sigma,\tau)\right)\right\rangle.$$

From Remark 7.4, if the power series that defines the function ξ does not include any term of odd degree, then $\Xi-\xi$ is nonnegative on $\mathbb{R}^D$ which implies that $\varphi_N'(\lambda) \geqslant 0$ and we do not need the perturbation $c_N H_N^{\mathrm{pert}}(\sigma)$ in this case. However, if some terms of odd degree are present, then according to Proposition 7.2 and Remark 7.3, we can only claim that $\Xi-\xi$ is nonnegative on $\mathbb{R}_+^D$. Since the overlaps $\frac{\sigma_d\cdot\tau_d}{N}$ can take values in $[-1,1]$, the previous display only implies

$$\varphi_N'(\lambda) \geqslant \delta(\varepsilon) - \frac{\|\Xi-\xi\|_\infty}{2}\mathbb{E}\left\langle \mathbf{1}_{\exists d,\ \frac{\sigma_d\cdot\tau_d}{N}\leqslant-\varepsilon}\right\rangle_\lambda,$$

where $\delta(\varepsilon) = \min\left\{\frac{(\Xi-\xi)(R)}{2}\,\middle|\,\forall d,\ R_d \in (-\varepsilon,1]\right\}$. Since $\Xi-\xi$ is locally Lipschitz and nonnegative on $\mathbb{R}_+^D$, there exists $C>0$ such that $\delta(\varepsilon) \geqslant -C\varepsilon$. Therefore,

$$\varphi_N(1) \geqslant \varphi_N(0) - C\varepsilon - \frac{\|\Xi-\xi\|_\infty}{2}\int_0^1 \mathbb{E}\left\langle \mathbf{1}_{\exists d,\ \frac{\sigma_d\cdot\tau_d}{N}\leqslant-\varepsilon}\right\rangle_\lambda \mathrm{d}\lambda.$$

To control the remaining term, observe that

$$\mathbb{E}\left\langle \mathbf{1}_{\exists d,\ \frac{\sigma_d\cdot\tau_d}{N}\leqslant-\varepsilon}\right\rangle_\lambda \leqslant \sup_H \mathbb{E}_u\mathbb{E}_{H_N^{\mathrm{pert}}}\widetilde{G}_N^{\otimes 2}\left(\exists d,\ \frac{\sigma_d\cdot\tau_d}{N}\leqslant -\varepsilon\right),$$

where, as in Lemma 7.5, the supremum is taken over all nonrandom Hamiltonians H, $\mathbb{E}_{H_N^{\mathrm{pert}}}$ denotes the expectation with respect to H_N^{pert} and $\widetilde{G}_N$ is the Gibbs measure of $H + c_N H_N^{\mathrm{pert}}$. In particular, the right-hand side in the previous display is independent of λ and according to Lemma 7.5, it vanishes in the limit $N \to +\infty$. Letting $N \to +\infty$, we obtain

$$\limsup_{N\to+\infty} \varphi_N(1) \geqslant \limsup_{N\to+\infty} \varphi_N(0) - C\varepsilon - 0.$$

Since $\varepsilon > 0$ is arbitrary, we can conclude by letting $\varepsilon \to 0$. □

7.4. **Connections with the Hamilton-Jacobi approach.** When ξ is nonconvex, the Parisi formula completely breaks down. To the best of our knowledge, until recently, it seems that there was no clear conjecture on what the limit of the free energy should be in this case. In [16, Conjecture 2.6], it is proposed that results such as Theorem 1.6 should generalize to nonconvex models. It was later shown in [17, 15] that the lim inf of $\overline{F}_N(t)$ as $N \to +\infty$ is lower bounded in terms of the viscosity solution of a Hamilton-Jacobi equation. This lower bound holds, regardless of convexity and permutation invariance.

Recall that $\xi_\dagger(\lambda) = \xi(\lambda, \ldots, \lambda)$ and $\psi^\dagger(p) = \psi(p, \ldots, p)$, also recall that

$$\Xi(x) = \frac{1}{D}\sum_{d=1}^{D} \xi(x_d, \ldots, x_d).$$

If we combine the lower bound of [17, 15] with Theorem 7.6 and 5.9, we obtain the following proposition.

Proposition 7.9. *Assume that ξ is permutation-invariant and only depends on the diagonal coefficients of its argument, assume that P_1 is permutation-invariant. Then, even when ξ is nonconvex on S^D_+, we have*

$$f(t,0) \leqslant \liminf_{N\to+\infty} \overline{F}_N(t) \leqslant \limsup_{N\to+\infty} \overline{F}_N(t) \leqslant g(t,0), \tag{7.9}$$

where f and g are the viscosity solutions of the following equations,

$$\begin{cases} \partial_t f - \int \xi(\nabla f) = 0 \text{ on } (0,\infty) \times (\mathcal{Q}(\mathbb{R}^D_+) \cap L^2) \\ f(0,\cdot) = \psi \text{ on } \mathcal{Q}(\mathbb{R}^D_+) \cap L^2, \end{cases} \tag{7.10}$$

$$\begin{cases} \partial_t g - \int \xi_\dagger\left(\frac{\nabla g}{D}\right) = 0 \text{ on } (0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+) \cap L^2) \\ g(0,\cdot) = \psi^\dagger \text{ on } \mathcal{Q}(\mathbb{R}_+) \cap L^2. \end{cases} \tag{7.11}$$

The results of Section 5.2 can be regarded as a proof of the fact $f(t,0) = g(t,0)$ when ξ is convex. To prove this statement we only relied on the convexity of $\xi_\dagger$ and the fact that f is not simply a solution in the viscosity sense but also a Gateaux differentiable solution. To the best of our knowledge, there is no known proof of the fact that f is Gateaux differentiable when ξ is nonconvex. Indeed, in the convex case, proving that f is Gateaux differentiable amounts to observing that it is semi-convex thanks to the Hopf-Lax representation and semi-concave thanks to the fact it can be written as the limit of the enriched free energy (see [7, Proposition 8.6]). Since the Hopf-Lax representation for viscosity solutions is not available when ξ is nonconvex, some innovative new ideas seem to be needed to prove (or disprove) the Gateaux differentiability of f in the nonconvex case. Nevertheless, for nonconvex models, we can formulate the following hypothesis.

$$(\mathrm{H}_{\xi,P_1}) \qquad \begin{array}{c}\text{The viscosity solution of (7.10) is Gateaux} \\ \text{differentiable on } (0,\infty)\times\left(\mathcal{Q}_\uparrow(\mathbb{R}^D_+)\cap L^\infty\right).\end{array}$$

In physics terms, (H_{ξ,P_1}) can be regarded as the statement that no first order phase transition occurs in the mean-field spin glass model with covariance function ξ and reference measure P_1. According to [7, Propositions 7.2 & 8.6], when ξ is convex, (H_{ξ,P_1}) holds. When ξ is nonconvex, assuming (H_{ξ,P_1}) holds, we obtain the following representation for the limit free energy.

Theorem 7.10. *Assume that ξ is permutation-invariant and only depends on the diagonal coefficients of its argument, assume that P_1 is permutation-invariant. If (H_{ξ,P_1}) is true, then for every $t>0$, the free energy $\overline{F}_N(t)$ converges as $N\to+\infty$ and*

$$\lim_{N\to+\infty}\overline{F}_N(t) = g(t,0), \tag{7.12}$$

where g is the viscosity solution of (7.11). Furthermore, we have $g(t,0)=f(t,0)$, where f is the viscosity solution of (7.10) and

$$g(t,0)=\sup_{p\in\mathcal{Q}(\mathbb{R}_+)\cap L^\infty}\left\{\psi(pid_D)-t\int_0^1\Xi^*\left(\frac{p(u)id_D}{t}\right)\mathrm{d}u\right\}.$$

Remark 7.11. In fact, even if (H_{ξ,P_1}) only holds on a region of the form $[0,t_c)\times(\mathcal{Q}_\uparrow(\mathbb{R}^D_+)\cap L^\infty)$ for some $t_c>0$ then (7.12) still holds for $t\leqslant t_c$.

Remark 7.12. It seems plausible that there exists $t_c>0$ such that the weaker version of (H_{ξ,P_1}) mentioned in Remark 7.11 holds. Indeed, ψ is Gateaux differentiable on $\mathcal{Q}(\mathbb{R}^D_+)\cap L^2$ and its gradient is a Lipschitz function [7, Corollary 5.2]. Therefore, the characteristic curves

$$q\mapsto q-t\nabla\xi(\nabla\psi(q)),$$

are injective when $t\geqslant 0$ is small enough ($t<1/\|\nabla\xi(\nabla\psi)\|_{\mathrm{Lip}}$). Usually, the viscosity solution remains differentiable as long as the characteristic curves are injective.

When $D=1$, the Parisi formula is a differentiable function of $\beta=\sqrt{2t}$ [19]. Therefore, the right-hand side of (7.12) is differentiable on $(0,+\infty)$ as a function of t. So at least (H_{ξ,P_1}) is not in direct contradiction with (7.12).

Conditionally on (H_{ξ,P_1}), Theorem 7.10 confirms [16, Conjecture 2.6] for permutation-invariant models. Note that the argument given below will not show that $(t,q) \mapsto \lim_{N\to+\infty} \overline{F}_N(t,q)$ is the viscosity solution of

$$\begin{cases} \partial_t f - \int \xi(\nabla f) = 0 \text{ on } (0,\infty) \times (\mathcal{Q}(\mathbb{R}^D_+) \cap L^2) \\ f(0,\cdot) = \psi \text{ on } \mathcal{Q}(\mathbb{R}^D_+) \cap L^2. \end{cases}$$

In fact, even under (H_{ξ,P_1}), it does not follow from the proof below that $\overline{F}_N(t,q)$ converges as $N \to +\infty$ when q is not of the form $(p,\ldots,p)$.

It is worth mentioning that Theorem 7.10 still holds if (H_{ξ,P_1}) is replaced by the assumption that the enriched free energy converges as $N \to +\infty$ and the limit is Gateaux differentiable on $(0,\infty) \times \left(\mathcal{Q}_\uparrow(\mathbb{R}^D_+) \cap L^\infty\right)$. To prove this version of Theorem 7.10, one can adapt the argument given below, replacing f the viscosity solution of (7.10) by the limit free energy and appealing to [7, Proposition 7.2]. Note that when choosing to state Theorem 7.10 this way, the convergence of the free energy as $N \to +\infty$ is a hypothesis, not a conclusion.

Proof of Theorem 7.10. Let f and g denote the viscosity solutions of (7.10) and (7.11). Given (7.9), it suffices to verify that $f(t,0) = g(t,0)$ to prove the desired result. As previously, for every $p \in \mathcal{Q}(\mathbb{R}_+) \cap L^1$, we let

$$f^\dagger(t,p) = f(t,(p,\ldots,p)).$$

Step 1. We show that for every $s \in \mathscr{S}_D$,

$$f(t,(p_1,\ldots,p_D)) = f(t,(p_{s(1)},\ldots,p_{s(D)})).$$

Given $s \in \mathscr{S}_D$, we let

$$\begin{aligned} f^s(t,(p_1,\ldots,p_D)) &= f(t,(p_{s(1)},\ldots,p_{s(D)})) \\ \psi^s(p_1,\ldots,p_D) &= \psi(p_{s(1)},\ldots,p_{s(D)}). \end{aligned}$$

Since ψ depends only on P_1 and not on ξ, according to Proposition 5.1, we have $\psi^s = \psi$. Using the permutation invariance of ξ, it can be checked that f^s is a viscosity solution of (7.10). By uniqueness of the viscosity solution of (7.10), we obtain $f^s = f$.

Step 2. We show that $f^\dagger$ is a strong solution of

$$\begin{cases} \partial_t f^\dagger - \int \Xi_\dagger\left(\frac{\nabla f^\dagger}{D}\right) = 0 \text{ on } (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty) \\ f^\dagger(0,\cdot) = \psi^\dagger \text{ on } \mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty. \end{cases}$$

Using Step 1, we can proceed as in the proof of Proposition 5.7 to discover that for every $(t,p) \in (0,+\infty) \times (\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty)$, $f^\dagger(t,\cdot)$ is Gateaux differentiable at p and

$$\nabla f(t,(p,\ldots,p)) = \left(\frac{\nabla f^\dagger(t,p)}{D},\ldots,\frac{\nabla f^\dagger(t,p)}{D}\right). \tag{7.13}$$

It then follows from (H_{ξ,P_1}) that

$$\partial_t f^\dagger(t,p) - \int \xi_\dagger\left(\frac{\nabla f^\dagger(t,p)}{D}\right) = \partial_t f(t,(p,\dots,p)) - \int \xi(\nabla f(t,(p,\dots,p))) = 0.$$

Finally, since $\xi_\dagger = \Xi_\dagger$ we conclude Step 2.

Step 3. We show that $f(t,0) = g(t,0)$.

According to Proposition 7.2, Ξ is convex on $\mathbb{R}_+^D$ and Ξ is the covariance function of the process defined by (7.7). In Theorem 5.9 we have shown that, when the nonlinearity is convex, any strong solution of a Hamilton-Jacobi equation on $\mathcal{Q}_\uparrow(\mathbb{R}_+) \cap L^\infty$ is also a viscosity solution on $\mathcal{Q}(\mathbb{R}_+) \cap L^2$. Therefore, using Step 2 and applying Theorem 5.9 to $\Xi_\dagger$ and $f^\dagger$, we discover that $f^\dagger$ the viscosity solution of

$$\begin{cases} \partial_t f^\dagger - \int \Xi_\dagger\left(\frac{\nabla f^\dagger}{D}\right) = 0 \text{ on } (0,+\infty) \times (\mathcal{Q}(\mathbb{R}_+) \cap L^2) \\ f^\dagger(0,\cdot) = \psi^\dagger \text{ on } \mathcal{Q}(\mathbb{R}_+) \cap L^2. \end{cases}$$

Once again, since $\Xi_\dagger = \xi_\dagger$, $f^\dagger$ is in fact the viscosity solution of (7.11). Finally, by uniqueness of the viscosity solution, it follows that $f^\dagger = g$ and choosing $p = 0$, we deduce $f(t,0) = g(t,0)$. Thus $F_N(t)$ converges as $N \to +\infty$ and

$$\lim_{N\to+\infty} \overline{F}_N(t) = f(t,0) = g(t,0).$$

Step 4. We show that

$$g(t,0) = \sup_{p \in \mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ \psi(p\,\mathrm{id}_D) - t\int_0^1 \Xi^*\left(\frac{p(u)\mathrm{id}_D}{t}\right) \mathrm{d}u \right\}.$$

According to Proposition 7.2, $\xi_\dagger$ is convex on $\mathbb{R}_+$. Therefore, it follows from the one dimensional Hopf-Lax representation [9, Proposition A.3] that,

$$g(t,0) = \sup_{p \in \mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ \psi^\dagger(p) - t\int_0^1 \left(\xi_\dagger\left(\frac{\cdot}{D}\right)\right)^* \left(\frac{p(u)}{t}\right) \mathrm{d}u \right\}.$$

In addition, Ξ is convex on $\mathbb{R}_+^D$ and we have

$$\left(\xi_\dagger\left(\frac{\cdot}{D}\right)\right)^*(\lambda) = \Xi^*(\lambda,\dots,\lambda).$$

Thus,

$$g(t,0) = \sup_{p \in \mathcal{Q}(\mathbb{R}_+)\cap L^\infty} \left\{ \psi(p\,\mathrm{id}_D) - t\int_0^1 \Xi^*\left(\frac{p(u)\mathrm{id}_D}{t}\right) \mathrm{d}u \right\}.$$

□

References

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

[2] Adriano Barra, Giuseppe Genovese, Francesco Guerra, and Daniele Tantari. How glassy are neural networks? *J. Stat. Mech.*, 2012(07):P07009, 2012.

[3] Erik Bates and Youngtak Sohn. Free energy in multi-species mixed p-spin spherical models. *Electron. J. Probab.*, 27:1–75, 2022.

[4] Erik Bates and Youngtak Sohn. Parisi formula for balanced Potts spin glass. Preprint, arXiv:2310.06745.

[5] Hong-Bin Chen. On Parisi measures of Potts spin glasses with correction. Preprint, arXiv:2311.11699.

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

[7] Hong-Bin Chen and Jean-Christophe Mourrat. On the free energy of vector spin glasses with non-convex interactions. Preprint, arXiv:2311.08980.

[8] Hong-Bin Chen and Jiaming Xia. Fenchel-Moreau identities on convex cones. Preprint, arXiv:2011.06979.

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

[10] Hong-Bin Chen and Jiaming Xia. Hamilton-Jacobi equations with monotone nonlinearities on convex cones. Preprint, arXiv:2206.12537.

[11] David Elderfield and David Sherrington. The curious case of the Potts spin glass. *J. Phys. C: Solid State Phys.*, 16(15):497–503, 1983.

[12] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

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

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

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

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

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

[18] Ben O'Neill. The double-constant matrix, centering matrix and equicorrelation matrix: Theory and applications. Preprint, arXiv:2109.05814.

[19] Dmitry Panchenko. On differentiability of the Parisi formula. *Elect. Comm. in Probab.*, 13:241–247, 2008.

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

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

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

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

[24] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43(23):1754–1756, 1979.

[25] Ralph Tyrrell Rockafellar. *Convex analysis*, volume 36. Princeton university press, 1970.

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