

λ – geodesic convexity of the squared Sliced Wasserstein Distance

Giacomo Cozzi*, Filippo Santambrogio[†]

October 14, 2026

Abstract

The squared Wasserstein distance with respect to a fixed target measure is not a convex function along Wasserstein geodesics over $\mathbb{R}^d$ when $d \geq 2$. Still, in dimension $d = 1$, this function is 2-geodesically convex. We prove that the square of the sliced Wasserstein distance (which is a distance constructed averaging one-dimensional Wasserstein distances) is λ -geodesically convex (possibly with $\lambda \leq 0$) on sets of measures satisfying upper bounds on their density, when the target measure satisfies some given geometric conditions. We also present some consequences of this result in the study of the sliced Wasserstein flow.

Contents

1	Introduction and main results	2
2	Preliminaries	4
2.1	Notation	4
2.2	Preliminary results in optimal transport	4
2.3	Sliced-Wasserstein distance and sliced-Wasserstein flow	7
3	Lipschitz bounds on ν imply λ-geodesic convexity of $\mathcal{F}_\nu$	8
4	Lipschitz bounds on $\nu[\rho]$	9
4.1	Lipschitz bounds if the projections of ν are bounded from below	9
4.2	Remarks about measures with projections bounded from below	10
4.3	Lipschitz bounds when ν satisfies Assumption (4.6)	11
4.4	Examples of measures satisfying (4.6)	12
5	Proof of Theorem 4.4	15
6	Uniqueness of the SWF	22

*Dipartimento di Matematica “Tullio Levi-Civita”, Università degli Studi di Padova, Via Trieste, 63, 35131 Padova PD, Italy. cozzi@math.unipd.it

[†]Université Lyon 1, Ecole Centrale de Lyon, INSA Lyon, Université Jean Monnet, CNRS, ICJ, UMR 5208, Villeurbanne, France, santambrogio@math.univ-lyon1.fr

1 Introduction and main results

Optimal transport is one of the richest fields in analysis, thanks to which many remarkable results in different areas of mathematics were achieved in the last decades. The Monge problem, first formulated in his “Mémoire sur la Théorie des Déblais et des Remblais”, aimed at minimizing the average displacement needed for moving a certain distribution of mass (source) to a different prescribed configuration (target): given two probability densities $\rho, \nu \in \mathcal{P}(\mathbb{R}^d)$, a Borel map $T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is said to be a *transport map* from ρ to ν if the nonlinear constraint

$$\nu(A) = \rho(T^{-1}(A))$$

is satisfied for any Borel set A . In this case we write $T_{\#}\rho = \nu$, and we say T is the “optimal transport map” for the Monge problem with source ρ and target ν , provided

$$\int_{\mathbb{R}^d} c(T(x), x) d\rho(x) = \inf_{S_{\#}\rho = \nu} \int_{\mathbb{R}^d} c(S(x), x) d\rho(x),$$

where $c(x, y)$ is a positive cost function. When $c(x, y) = |x - y|^p$, with $p \geq 1$, the p -th root of the infimum above defines a distance between probability measures, called the *p-Wasserstein distance*, and denoted by W_p .

Wasserstein distances enjoy many properties that make them a useful tool in theoretical research, as they metrize the space of probability measures with finite p -moments and provide it of a geodesic space structure. At the same time, they are widely used in applications, being one of the most common choices as a cost for machine learning. However, computing Wasserstein distances in dimensions $d \geq 2$ can be a computationally expensive task. This is the reason why many alternatives to the Wasserstein distance have been introduced in the literature.

One of them is the following: let ρ and ν be probability measures in $\mathbb{R}^d$, both with finite second moments. For any unit direction $\vartheta \in \mathbb{S}^{d-1}$ we can define $\rho^\vartheta = (\pi^\vartheta)_{\#}\rho$ and $\nu^\vartheta = (\pi^\vartheta)_{\#}\nu$, where $\pi^\vartheta: \mathbb{R}^d \rightarrow \mathbb{R}$ is the canonical projection on the line $\vartheta\mathbb{R}$ and the symbol $\#$ indicates the push-forward of the measure. We define

$$SW_2(\rho, \nu) := \left(\oint W_2^2(\rho^\vartheta, \nu^\vartheta) d\sigma(\vartheta) \right)^{\frac{1}{2}}. \quad (1.1)$$

Here $\oint$ stands for the averaged integral, i.e., $\oint_E dm = \frac{1}{m(E)} \int_E dm$, and $d\sigma$ is the surface measure on the sphere. We reserve σ for this $(d - 1)$ -dimensional measure and write $\mathcal{H}^k$ for the k -dimensional Hausdorff measure on other integration domains. It is well known that (1.1) defines a distance over $\mathcal{P}_2(\mathbb{R}^d)$, the space of probability measures with finite second moments on $\mathbb{R}^d$, called *sliced Wasserstein distance*: this is comparable with the usual Wasserstein distance W_2 , but it presents many advantages for applications, since it relies only on one-dimensional transport constructions [KNSBR19; NDCKSS20; PS23; RM23].

Fix a target measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. The main result of the present paper concerns the convexity of the functional

$$\mathcal{F}_\nu(\rho) = \frac{1}{2} SW_2^2(\rho, \nu)$$

along Wasserstein geodesics, under certain assumptions on ν .

Here is an informal statement of the main result, which will then be presented as Theorem 4.4:

Theorem 4.4. Let $M > 0$, $K = \{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \|\rho\|_{L^\infty} \leq M\}$, and $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Assume that ν satisfies the technical assumption (4.6). For any $\rho_0, \rho_1 \in K$, let ρ_t be the constant-speed Wasserstein geodesic connecting them. Then, it holds

$$\mathcal{F}_\nu(\rho_t) \leq (1-t)\mathcal{F}_\nu(\rho_0) + t\mathcal{F}_\nu(\rho_1) - \frac{\lambda}{2}t(1-t)W_2^2(\rho_0, \rho_1),$$

with

$$\lambda = \frac{1}{d} - cM^{\frac{1}{d}},$$

where c is a constant only depending on ν and not on $\rho_0, \rho_1 \in K$.

This is in contrast with what happens for the Wasserstein distance W_2 which, rather, is in general not expected to be convex along geodesics when $d \geq 2$ [AGS05]. We stress here that, in general, λ is not expected to be positive and the convexity degenerates (i.e., λ goes to $-\infty$) as the L^∞ bound M for the set K increases, as well as when we change the constants appearing in the precise formulation of the assumptions needed on ν , namely in equation (4.6).

An interesting application of this result concerns the study of the *sliced Wasserstein flow* (SWF), which is the isotropic and continuous-in-time version of the *Iterative Distribution Transfer* (IDT), first proposed in [PKD07]. This is an algorithm based on a sequence of one-dimensional optimal matchings between projections of two given measures along different axes, chosen randomly at each time step.

More precisely: fix a time step $\tau > 0$. Starting from a probability density ρ_0 , choose randomly an orthonormal basis $\{e_1, \dots, e_d\}$, and consider the projections $(\pi_{\#}^{e_1} \rho_0, \dots, \pi_{\#}^{e_d} \rho_0) =: (\rho_0^1, \dots, \rho_0^d)$. Let $(T^1, \dots, T^d)$ be the optimal transport map (for the quadratic cost) between $(\rho_0^1, \dots, \rho_0^d)$ and $(\pi_{\#}^{e_1} \nu, \dots, \pi_{\#}^{e_d} \nu) =: (\nu_1, \dots, \nu_d)$. We now want to move each point $x \in \text{spt } \rho_0$ in the direction e_i of a displacement given by the displacement of the map T^i at the point $x \cdot e_i$, and combine these displacements for all the elements of the randomly chosen basis. To this aim, consider $v^{IDT}(x) := \sum_{i=1}^d (T^i(x \cdot e_i) - x \cdot e_i) e_i$ and define $\rho_1^\tau = (\text{id} + \tau v)_{\#} \rho_0$. Heuristically, one expects ρ_1^τ to be “closer” than ρ_0 to ν . Finally, we iterate the construction, obtaining a sequence $\{\rho_k^\tau\}_{k \geq 0}$. When $k \rightarrow +\infty$, one expects (or hopes) to recover the target probability ν .

Starting from this, Bernot proposed the sliced Wasserstein flow operating the following two modifications: first, instead of choosing a random orthonormal basis, at each iteration, we consider the collection of all projections $\{\rho_0^\vartheta\}_{\vartheta \in \mathbb{S}^{d-1}}$ and $\{\nu^\vartheta\}_{\vartheta \in \mathbb{S}^{d-1}}$. Denoting by T^ϑ the optimal transport map between ρ_0^ϑ and ν^ϑ , we define the isotropic version of the velocity field v^{IDT} . Namely, we compute

$$v[\rho_0](x) := \int_{\mathbb{S}^{d-1}} (T^\vartheta(x \cdot \vartheta) - x \cdot \vartheta) \vartheta d\sigma(\vartheta), \quad (1.2)$$

and define $\rho_1^\tau := (\text{id} + \tau v(x))_{\#} \rho_0$. We iterate the construction obtaining a sequence $\{\rho_k^\tau\}_{k \geq 0}$. The second difference from the IDT algorithm consists in considering the continuous-in-time analog of this construction, that is, we take the limit as $\tau \rightarrow 0^+$ of the aforementioned sequence. Thus, we obtain an absolutely continuous curve $(\rho_t)_{t \geq 0}$ that solves

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho v[\rho]) = 0, \\ \rho(0) = \rho_0, \end{cases} \quad (1.3)$$

This flow is expected to be an evolution which lets ρ_t approach the target ν , but it only relies on one-dimensional constructions, avoiding the growth of computational costs when the dimension d of the space is large (curse of dimensionality).

In Section 2.3 we will show that the sliced Wasserstein flow with target measure ν corresponds to the Wasserstein gradient descent for the functional $\mathcal{F}_\nu$. This fact is relevant since it allows to apply the theory of Wasserstein gradient flows to the study of the SWF. In particular, the existence of solutions to the SWF was proved by Bonnotte in his PhD thesis [Bon13] by JKO methods [San17]. In the same work he also proved bounds for the L^p norm along the flow (see Theorem 2.6 below) and addressed some questions regarding the long time asymptotics and uniqueness of the flow. Regarding the first of these two questions, we can mention [CS25], where the authors proved that, if ν is the Gaussian measure, the flow converges to ν . Other results in this directions have been obtained in [VKM25]. For what concerns uniqueness of solutions, as a byproduct of Theorem 4.4, in this paper we are able to obtain the following

Corollary 1.1. *If $\nu \in \mathcal{P}_2$ satisfies (4.6) and $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, then the sliced Wasserstein flow is well-posed in $L^\infty(\mathbb{R}^d)$, i.e., there exists a unique solution $\{\rho_t\}_{t \geq 0}$ such that $t \mapsto \|\rho_t\|_{L^\infty}$ is locally bounded.*

This is just an example of the consequences of our λ -geodesic convexity result. Among others, we cite the fact that it also implies the convergence at a rate τ in W_2 distance of the corresponding JKO scheme to the continuous solution (from the theory developed in [AGS05]) and the fact that the functional SW_2^2 can then be used in the flow interchange technique (see [MMS09]), but we do not develop these directions in details.

The paper is structured as follows: in Section 2, we report the main results needed from the theory of optimal transport, as well as some results related to the sliced Wasserstein distance and the sliced Wasserstein flow. In Section 3, we will show how to obtain the geodesic convexity of $\mathcal{F}_\nu$, once Lipschitz bounds on the velocity field v are known. In Section 4, then, we will derive these very Lipschitz bounds, first on the class of probability measures with bounds in L^p under very strong assumptions on ν . Then, we will state and comment the main result, Theorem 4.4, which requires more natural assumptions on ν but holds under L^∞ bounds: its proof will be presented in Section 5. Finally, in the last Section we will prove Corollary 1.1 and, more generally, discuss the uniqueness for the Sliced Wasserstein Flow.

2 Preliminaries

2.1 Notation

For a function $u : \mathbb{R}^d \mapsto \mathbb{R}$ smooth enough, we denote by ∇u its gradient and D^2u its Hessian matrix. Moreover, for a map $T : \mathbb{R}^d \mapsto \mathbb{R}^d$, its Jacobian matrix is denoted by DT . We denote by “id” the identity map. When referring to the target measure ν , we will denote by $N_\delta(r) = \nu^\delta((-\infty, r])$, where $\nu^\delta = \pi_\#^\delta \nu$. Similarly, for a generic distribution ρ , we use the notation $R_\delta(r) = \rho^\delta((-\infty, r])$. We will often make an abuse of notation using ρ, ν as a symbol both for the probability measures or their densities, when no confusion can arise. We will denote by $B_r^{(d)}(x)$ the d -dimensional Euclidean ball centered at x with radius r . When there is no ambiguity, we will write $B_r(x)$, and possibly just B_r when $x = 0$. Throughout the paper “spt” stands for the support of a measure, i.e., the smallest closed set on which it is concentrated. Finally, for any $p \geq 1$, we denote by p' the Hölder conjugate exponent: $1/p + 1/p' = 1$. We denote by $\omega_d = \mathcal{L}^d(B_1(0))$ the volume of the unit ball in $\mathbb{R}^d$; thus $\sigma(\mathbb{S}^{d-1}) = d\omega_d$.

2.2 Preliminary results in optimal transport

Given any open subset $\Omega \subset \mathbb{R}^d$, we denote by $\mathcal{P}(\Omega)$ the set of probability measures with support contained in Ω . For $p \geq 1$, we denote by $\mathcal{P}_p(\Omega)$ the set of probability measures on Ω with finite

moment of order p (in case Ω is unbounded), so that we have $\mathcal{P}_p(\Omega) = \{\mu \in \mathcal{P}(\Omega) : \int |x|^p d\mu < +\infty\}$.

Given $\mu, \nu \in \mathcal{P}(\Omega)$, we say that T is a transport map from μ to ν , and that ν is the push-forward of μ via T (and we write $\nu = T_{\#}\mu$), if for every Borel set $A \subset \Omega$ we have

$$\nu(A) = \mu(T^{-1}(A))$$

or, equivalently, if for any Borel map $\phi: \Omega \rightarrow \mathbb{R}$ it holds

$$\int_{\Omega} \phi(x) d\nu(x) = \int_{\Omega} \phi(T(x)) d\mu(x).$$

Definition 2.1. Let $\mu, \nu \in \mathcal{P}(\Omega)$ and let $c: \Omega \times \Omega \rightarrow \mathbb{R}$ be a continuous cost function. The map $T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is said to be an optimal transport map for the cost c if it solves the problem

$$\min_S \left\{ \int_{\Omega} c(x, S(x)) d\mu(x), \quad S \text{ Borel map such that } S_{\#}\mu = \nu \right\}. \quad (2.1)$$

The map $\varphi: \mathbb{R}^d \rightarrow \mathbb{R}$ is said to be a Kantorovich potential if it solves the problem

$$\max_{\psi} \left\{ \int_{\Omega} \psi(x) d\mu(x) + \int_{\Omega} \psi^c(x) d\nu(x) \right\}, \quad (2.2)$$

where $\psi^c(x) := \inf_{y \in \mathbb{R}^d} \{c(x, y) - \psi(y)\}$.

Let $\Pi_{\mu}^{\nu} := \{\gamma \in \mathcal{P}(\Omega \times \Omega) : \gamma_{\#}^1 = \mu, \gamma_{\#}^2 = \nu\}$, where $\gamma^i(x_1, x_2) = x_i$. We define the p -Wasserstein distance from μ to ν to be

$$W_p(\mu, \nu) = \left(\int_{\Omega \times \Omega} |x_1 - x_2|^p d\pi(x_1, x_2) \right)^{\frac{1}{p}} = \inf_{\gamma \in \Pi_{\mu}^{\nu}} \left(\int_{\Omega \times \Omega} |x_1 - x_2|^p d\gamma(x_1, x_2) \right)^{\frac{1}{p}}.$$

It is well known that W_p defines a distance over $\mathcal{P}_p(\Omega)$ for any $p \geq 1$ (see, for example, [San15, Chapter 5]).

Remark 2.1. The problems (2.1) and (2.2) do not always admit a solution. However, when they exist, the two quantities defined by (2.1) with “min” replaced by “inf”, and (2.2) with “max” replaced by “sup” coincide. Moreover, when a solution T to the Monge problem exists, then the joint density π appearing in the definition of W_p is the measure concentrated on $\text{gr}(T) \subset \Omega \times \Omega$, where $\text{gr}T$ denotes the graph of T , namely $\text{gr}T = \{(x, T(x)) \in \Omega \times \Omega : x \in \Omega\}$. As a consequence, if T is an optimal transport map and φ is the Kantorovich potential for the associated problem, it holds

$$W_p^p(\mu, \nu) = \int_{\Omega} |T(x) - x|^p d\mu = \int_{\Omega} \varphi(x) d\mu(x) + \int_{\Omega} \varphi^c(x) d\nu(x).$$

We are interested in the case $\Omega = \mathbb{R}^d$ and $c(x, y) = \frac{1}{2}|x - y|^2$. In this setting we have a relation between the optimal T and the optimal φ (which is differentiable almost everywhere with respect to the source measure, when it is absolutely continuous) which is

$$T(x) = x - \nabla \varphi(x). \quad (2.3)$$

In particular, we have $T = \nabla u$ where $u(x) = \frac{1}{2}|x|^2 - \varphi(x)$, and one can check that u is convex. This is summarized in the following celebrated theorem.

Theorem 2.2 (Brenier's theorem, see [Bre87]). Let μ, ν be two absolutely continuous probability measures in $\mathcal{P}_2(\mathbb{R}^d)$. A map $T: \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $T_{\#}\mu = \nu$ is optimal for the optimal transport problem (2.1) with the cost $c(x, y) = \frac{1}{2}|x - y|^2$ if and only if there exists a convex function $u: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $T = \nabla u$ a.e. Moreover if this is the case, then u must satisfy the Monge-Ampère equation

$$\nu(\nabla u(x)) \det D^2 u(x) = \mu(x)$$

a.e. on $\mathbb{R}^d$.

For the one-dimensional problem, we are able to characterize optimal transport maps. A detailed analysis of the one-dimensional case is available in [San15], Chapter 2.

Theorem 2.3 (Brenier's theorem in $d = 1$). Let μ, ν be two probability densities with finite second moments. Then there exists a unique non-decreasing transport map from μ to ν , which is optimal and given by

$$T(x) = F_{\nu}^{[-1]}(F_{\mu}(x)),$$

where $F_{\mu}(x) = \mu((-\infty, x])$ and $F^{[-1]}(y) = \inf\{t \in \mathbb{R}: F(t) \geq y\}$.

The metric space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a geodesic space, namely: given any two elements $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a curve $\rho_t: [0, 1] \rightarrow \mathcal{P}_2(\mathbb{R}^d)$ connecting them with the property that

$$W_2(\rho_s, \rho_t) = |t - s|W_2(\rho_0, \rho_1) \quad \text{for any } 0 < s < t < 1. \quad (2.4)$$

Definition 2.2. A curve satisfying (2.4) is called *constant speed geodesic* from ρ_0 to ρ_1 .

A functional $\mathcal{F}: \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is λ -geodesically convex if for every two points $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a constant speed geodesic $\rho(t)$ connecting $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$ such that

$$\mathcal{F}(\rho_t) \leq (1 - t)\mathcal{F}(\rho_0) + t\mathcal{F}(\rho_1) - \frac{\lambda}{2}t(1 - t)W_2^2(\rho_0, \rho_1). \quad (2.5)$$

The case in which $\lambda > 0$ is generally referred to as *strong convexity*, whereas for $\lambda = 0$ we recover the usual definition of convexity. If $t \mapsto \mathcal{F}(\rho_t)$ is twice differentiable, (2.5) is equivalent to

$$\frac{d^2}{dt^2} \mathcal{F}(\rho_t) \geq \lambda W_2^2(\rho_0, \rho_1).$$

The convexity of a functional over the space of probability measures, though not being essential for proving the well posedness of the associated gradient flow in the Wasserstein space, is the typical assumption under which uniqueness and contractivity results may be obtained (cf. [AGS05; San17]).

Throughout this paper, we will make use several times of the following two results.

Theorem 2.4. [San15, Theorem 5.27] Let $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$ with ρ_0 absolutely continuous with respect to the Lebesgue measure, and suppose there exists an optimal transport map T for the quadratic cost $c(x, y) = |x - y|^2$, i.e., a solution of (2.1). Then, the unique geodesic $\rho_t: [0, 1] \rightarrow \mathcal{P}_2(\mathbb{R}^d)$ from $\rho(0) = \rho_0$ to $\rho(1) = \rho_1$ is given by

$$\rho_t = (T_t)_{\#}\rho_0,$$

where

$$T_t(x) = (1 - t)x + tT(x)$$

for any $x \in \text{spt } \rho_0$. Then T_t is injective for $t \in [0, 1)$ and the curve $(\rho_t)_{t \geq 0}$ solves

$$\partial_t \rho_t + \nabla \cdot (\rho_t w) = 0,$$

where $w_t(x) = (T - \text{id}) \circ T_t^{-1}(x)$.

2.3 Sliced-Wasserstein distance and sliced-Wasserstein flow

Given two probability measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, we define their sliced Wasserstein distance as in (1.1). As said in the introduction, the sliced Wasserstein distance does not coincide with the classical W_2 . However, the two distances generate the same topology (see [BG21]). Moreover, whenever $\text{spt } \mu, \text{spt } \nu \subset B(0, R)$, there exist an exponent $\alpha > 0$ and constant $C(R) > 0$ such that it holds $C(R)W_2(\mu, \nu) \leq SW_2^\alpha(\mu, \nu) \leq W_2^\alpha(\mu, \nu)$ (see [Bon13, Chapter 5] and [CFMW25]). Many properties, which highlight surprising differences in the behavior of SW_2 compared to W_2 , have been studied in [PS23]. We also mention the possibility of defining a p -sliced distance SW_p , via

$$SW_p := \left(\int_{\mathbb{S}^{d-1}} W_p^p((\pi^\vartheta)_\# \mu, (\pi^\vartheta)_\# \nu) d\sigma(\vartheta) \right)^{1/p},$$

which is sometimes more convenient for sharp comparisons with the standard Wasserstein distances (see [CFMW25]). Also, a distance of the form

$$SW_{p,q}(\mu, \nu) := \left(\int_{\mathbb{S}^{d-1}} W_p^q(\mu^\vartheta, \nu^\vartheta) d\sigma(\vartheta) \right)^{\frac{1}{q}}$$

can be introduced, with a particular interest when $q = +\infty$, in which case one wants to maximal Wasserstein distance among projections.

In this Section we want also to show that the velocity field associated with the SWF corresponds to the velocity field of the gradient descent for the functional $\mathcal{F}_\nu$. Namely, we claim that $v[\rho_t](x)$ as defined in (1.2) coincides with the negative gradient of the first variation of $\mathcal{F}_\nu$:

$$-\nabla \frac{\delta \mathcal{F}_\nu}{\delta \rho}(\rho) = \int_{\mathbb{S}^{d-1}} (T^\vartheta(x \cdot \vartheta) - x \cdot \vartheta) \vartheta d\sigma(\vartheta) \text{ for any } x \in \text{spt } \rho.$$

Indeed, since the first variation of the functional $\rho \mapsto \frac{1}{2} W_2^2(\rho, \nu)$ is the Kantorovich potential for the transport problem from μ to ν (see [San15], proposition 7.17), we get

$$\frac{\delta \mathcal{F}_\nu}{\delta \rho}(\rho)(x) = \int_{\mathbb{S}^{d-1}} \frac{1}{2} \frac{\delta W_2^2(\rho^\vartheta, \nu^\vartheta)}{\delta \rho}(x) d\sigma(\vartheta) = \int_{\mathbb{S}^{d-1}} \varphi^\vartheta(x \cdot \vartheta) d\sigma(\vartheta),$$

where we denoted by φ^ϑ the Kantorovich potential between $\rho^\vartheta := (\pi^\vartheta)_\# \rho$ and $\nu^\vartheta := (\pi^\vartheta)_\# \nu$. Using also (2.3), this in turn leads to

$$-\nabla \frac{\delta \mathcal{F}_\nu}{\delta \rho}(\rho) = - \int_{\mathbb{S}^{d-1}} (\varphi^\vartheta)'(x \cdot \vartheta) \vartheta d\sigma(\vartheta) = \int_{\mathbb{S}^{d-1}} (T^\vartheta(x \cdot \vartheta) - x \cdot \vartheta) \vartheta d\sigma(\vartheta),$$

which is precisely the velocity field defined in (1.2). Therefore the SWF with target ν and source ρ_0 corresponds to a solution of (1.3)

We conclude this section recalling the main result on the Sliced Wasserstein Flow, obtained by viewing it as the gradient flow of SW_2^2 , which provides existence together with L^p estimates for the solutions:

Theorem 2.5. [Bon13, Theorem 5.5.3] *Let $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $p \in [1, +\infty]$. Given an absolutely continuous probability measure $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a continuous curve $(\rho_t)_{t \geq 0}$ solution of (1.3). Moreover, for each time $t \geq 0$, the measure ρ_t has a density in L^p and*

$$\|\rho_t\|_{L^p} \leq e^{t \sqrt{d}/p'} \|\rho_0\|_{L^p}. \quad (2.6)$$

Remark 2.6. In [Bon13], the above result is proved under the assumption that ν is strictly positive and supported in $\bar{B}(0, 1)$. However, since the estimate obtained does not depend on this condition, by approximation we can deduce the theorem in the form in which we presented it.

3 Lipschitz bounds on v imply λ -geodesic convexity of $\mathcal{F}_v$

In this section, we show how to derive the geodesic convexity of $\mathcal{F}_v$ when the velocity field v is Lipschitz continuous in space. Lipschitz bounds on v will in turn be obtained in Section 4.

Lemma 3.1. *Let $v[\rho]$ be defined by (1.2). Assume that $K \subset \mathcal{P}_2(\mathbb{R}^d)$ is a geodesically convex set and that for any $\rho \in K$*

$$\text{Lip}\left(v[\rho] + \frac{\text{id}}{d}\right) \leq L.$$

Then $\rho \mapsto \mathcal{F}_v(\rho)$ is λ -geodesically convex on K for $\lambda = \frac{1}{d} - L$.

Proof. Let ρ_0, ρ_1 be two probability densities in K , and let $(\rho_t)_{t \geq 0} \subset K$ be the Wasserstein geodesic connecting them. Let T be the optimal transport map from ρ_0 to ρ_1 and $T_t := (1-t)\text{id} + tT$. Then, by Theorem 2.4 ρ_t solves

$$\partial_t \rho_t + \text{div}(w_t \rho_t) = 0,$$

where $w_t(x) := (T - \text{id}) \circ T_t^{-1}(x)$. We have

$$\begin{aligned} \frac{d}{dt} \mathcal{F}_v(\rho_t) &= \int \frac{\delta \mathcal{F}_v}{\delta \rho}(\rho_t) \partial_t \rho_t = \int \nabla \frac{\delta \mathcal{F}_v}{\delta \rho}(\rho_t) \cdot w_t \rho_t = - \int v[\rho_t] \cdot w_t \rho_t, \\ \frac{d^2}{dt^2} \mathcal{F}_v(\rho_t) &= - \underbrace{\int \partial_t v[\rho_t] \cdot w_t \rho_t}_{:= I_1} - \underbrace{\int v[\rho_t] \cdot \partial_t(w_t \rho_t)}_{:= I_2}. \end{aligned} \quad (3.1)$$

The term I_2 can be estimated as follows; let $\bar{v}$ be a time independent vector field, with the property that

$$\left\| D\left(\bar{v}(x) + \frac{\text{id}}{d}\right) \right\|_{L^\infty} \leq L.$$

Then,

$$\begin{aligned} \frac{d}{dt} \int \bar{v} \cdot w_t \rho_t &= \frac{d}{dt} \int \bar{v} \cdot (T - \text{id})(T_t^{-1}(x)) T_{t\#} \rho_0 \\ &= \frac{d}{dt} \int \bar{v}(T_t(x)) \cdot (T(x) - x) \rho_0 = \int (T(x) - x) \cdot D\bar{v}(T_t(x)) (T(x) - x) \rho_0 \\ &= \int (T(x) - x) \cdot D\left(\bar{v} + \frac{\text{id}}{d}\right)(T_t(x)) (T(x) - x) \rho_0 - \frac{1}{d} \int |T(x) - x|^2 \rho_0 \\ &\leq \left(L - \frac{1}{d}\right) \int |T(x) - x|^2 \rho_0. \end{aligned}$$

Taking $\bar{v} = v[\rho_t]$, we obtain

$$-I_2 = - \int v[\rho_t] \cdot \partial_t(w_t \rho_t) \geq \left(\frac{1}{d} - L\right) W_2^2(\rho_0, \rho_1). \quad (3.2)$$

On the other hand, we claim that I_1 is negative. Indeed,

$$\begin{aligned}
I_1 &= \int \partial_t(v[\rho_t]) \cdot w_t \rho_t = \int \partial_t \left(\int_{\mathbb{S}^{d-1}} T_{t,\vartheta}(x \cdot \vartheta) d\sigma(\vartheta) \right) \cdot w_t \rho_t dx \\
&= \int \partial_t \left[\int_{\mathbb{S}^{d-1}} N_{\vartheta}^{-1}(R_{t,\vartheta}(x \cdot \vartheta)) d\sigma(\vartheta) \right] \cdot w_t \rho_t dx \\
&= \int \int_{\mathbb{S}^{d-1}} \frac{\partial_t \int_{y \cdot \vartheta \leq x \cdot \vartheta} \rho_t(y) dy}{\nu_{\vartheta}(T_{t,\vartheta}(x \cdot \vartheta))} \vartheta \cdot w_t \rho_t d\sigma(\vartheta) dx \\
&= - \int \int_{\mathbb{S}^{d-1}} \frac{\int_{y \cdot \vartheta \leq x \cdot \vartheta} \operatorname{div}(w_t \rho_t) dy}{\nu_{\vartheta}(T_{t,\vartheta}(x \cdot \vartheta))} \vartheta \cdot w_t \rho_t d\sigma(\vartheta) dx.
\end{aligned}$$

By Fubini, we change the order of integration and perform the integration on level sets $x \cdot \vartheta = c$:

$$\begin{aligned}
I_1 &= - \int_{\mathbb{S}^{d-1}} \int_{-\infty}^{+\infty} \frac{1}{\nu_{\vartheta}(T_{t,\vartheta}(c))} \left(\int_{y \cdot \vartheta \leq c} \operatorname{div}(w_t \rho_t) dy \right) \left(\int_{x \cdot \vartheta = c} \vartheta \cdot w_t \rho_t dx \right) dc d\sigma(\vartheta) \\
&= - \int_{\mathbb{S}^{d-1}} \int_{-\infty}^{+\infty} \frac{1}{\nu_{\vartheta}(T_{t,\vartheta}(c))} \left(\int_{x \cdot \vartheta = c} \vartheta \cdot w_t \rho_t dx \right)^2 d\sigma(\vartheta) dc \leq 0.
\end{aligned} \tag{3.3}$$

Combining (3.1), (3.2) and (3.3), we obtain

$$\frac{d^2}{dt^2} \mathcal{F}_v(\rho_t) \geq \left(\frac{1}{d} - L \right) W_2^2(\rho_0, \rho_1).$$

□

4 Lipschitz bounds on $v[\rho]$

In this Section, we present two results, both giving bounds to the Lipschitz constant of v . First, we state and prove Lemma 4.1. This result shows that $\mathcal{F}_v$ is a λ -convex functional on the set of measures with bounded L^p norm if ν has projections bounded away from zero. Then, we state Theorem 4.4, which instead proves λ -convexity on the set of measures with bounded L^∞ norm, provided that ν satisfies a certain weaker condition, involving bounds on the density of ν itself and not of its projections. Indeed, we will see that lower bounds on the projections are very demanding: if ν is bounded from below and above on a bounded domain, most of the densities of its projections will vanish on the boundary of their support, because the measure of the section tends to 0. Even more strikingly, the existence of measures with lower bounds on the densities of the projections is very restrictive, and is essentially a two- or three-dimensional affair. The precise conditions on ν will then be the object of study in the last part of this Section, where we will present some remarks and provide examples.

4.1 Lipschitz bounds if the projections of ν are bounded from below

Lemma 4.1. *Let $v[\rho]$ be defined by (1.2), $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, and $K = \{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \|\rho\|_{L^p} \leq M\}$, for some $p > d'$, $M > 0$. Assume that, for almost every $\vartheta \in \mathbb{S}^{d-1}$, the support of ν_{ϑ} is an interval I_{ϑ} , and that $\nu_{\vartheta} \geq c_0 > 0$ almost everywhere on I_{ϑ} . Then, there exists $L = L(c_0, M, p, d)$ such that, for any $\rho \in K$,*

$$\operatorname{Lip} \left(v[\rho] + \frac{\operatorname{id}}{d} \right) \leq L,$$

and $\mathcal{F}_v$ is λ -geodesically convex on K with $\lambda = \frac{1}{d} - L$.

Proof. Deriving the relation (see Theorem 2.3)

$$N_{\vartheta}(T_{\vartheta}(r)) = R_{\vartheta}(r) \quad \text{for any } r \in \text{spt } \rho^{\vartheta},$$

we deduce that

$$T'_{\vartheta}(r) = \frac{\rho^{\vartheta}(r)}{\nu^{\vartheta}(T_{\vartheta}(r))}.$$

Therefore,

$$\begin{aligned} \left| Dv[\rho](x) + \frac{\text{id}}{d} \right| &= \left| \int_{\mathbb{S}^{d-1}} T'_{\vartheta}(x \cdot \vartheta) \vartheta \otimes \vartheta \right| \leq \int_{\mathbb{S}^{d-1}} |T'_{\vartheta}(x \cdot \vartheta)| d\sigma(\vartheta) \\ &\leq \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{\nu^{\vartheta}(T_{\vartheta}(x \cdot \vartheta))} d\sigma(\vartheta) \leq \frac{1}{c_0} \int_{\mathbb{S}^{d-1}} \rho^{\vartheta}(x \cdot \vartheta) d\sigma(\vartheta) \\ &= \frac{1}{c_0} \int_{\mathbb{S}^{d-1}} \int_{e \cdot \vartheta=0} \int_0^{+\infty} \rho(x + er) r^{d-2} dr d\mathcal{H}^{d-2}(e) d\sigma(\vartheta) = \frac{(d-1)\omega_{d-1}}{d\omega_d c_0} \int \frac{\rho(y)}{|x-y|} dy, \end{aligned}$$

where we used [CS25, Lemma 5.1] to obtain the last equality. Moreover, if $\rho \in L^p$ for $p > d'$, we have

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{\rho(y)}{|x-y|} dy &\leq \int_{B_R(x)} \frac{\rho(y)}{|x-y|} dy + \int_{B_R^c(x)} \frac{\rho(y)}{|x-y|} dy \\ &\leq \left(\int_{B_R} \frac{1}{|y|^{p'}} dy \right)^{\frac{1}{p'}} \left(\int_{\mathbb{R}^d} \rho^p dx \right)^{\frac{1}{p}} + \frac{1}{R} \leq c(d, p) R^{\frac{d}{p'}-1} \|\rho\|_{L^p} + \frac{1}{R}. \end{aligned} \tag{4.1}$$

Optimizing over $R > 0$, we choose $R = \|\rho\|_{L^p}^{-\frac{p'}{d}} (c(d, p)(-1 + d/p'))^{-\frac{p'}{d}}$, thus obtaining

$$\frac{(d-1)\omega_{d-1}}{d\omega_d c_0} \int_{\mathbb{R}^d} \frac{\rho(y)}{|x-y|} dy \leq \frac{(d-1)\omega_{d-1} M^{\frac{p'}{d}}}{d\omega_d c_0} \frac{d}{d-p'} \left(\frac{c(d, p)(d-p')}{p'} \right)^{\frac{p'}{d}} =: L \tag{4.2}$$

for any $\rho \in K$. Since K is geodesically convex (see, for example [San15, Chapter 7]), we can apply Lemma 3.1 and deduce the claim. $\square$

4.2 Remarks about measures with projections bounded from below

We collect here some remarks regarding measures $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ satisfying the condition (appearing in Lemma 4.1) that $\nu^{\vartheta}(r) \geq c_0 > 0$ for a.e. $\theta \in \mathbb{S}^{d-1}$. We will also focus on the case where the ν^{ϑ} are the uniform measure on a same interval.

Remark 4.2. We do not expect that measures with the property that $\nu^{\vartheta} \geq c_0 > 0$ for a.e. $\vartheta \in \mathbb{S}^{d-1}$ exist in $\mathbb{R}^d$ for $d > 3$. We will show in this remark that this is the case when we strengthen condition (4.8) replacing it with the following property (which in turn implies (4.8)): for almost every $\vartheta \in \mathbb{S}^{d-1}$ there exists an interval I_{ϑ} such that

$$\begin{aligned} \nu^{\vartheta} &\geq c_0 \quad \text{on } I_{\vartheta} \\ \nu^{\vartheta} &= 0 \quad \text{on } \mathbb{R} \setminus I_{\vartheta} \\ C &:= \bigcap_{\vartheta \in \mathbb{S}^{d-1}} \{x \cdot \vartheta \in I_{\vartheta}\} \quad \text{is uniformly convex} \end{aligned} \tag{4.3}$$

Indeed, for any $\vartheta \in \mathbb{S}^{d-1}$, let $r_\vartheta = \sup_{x \in C} x \cdot \vartheta$ and $x_\vartheta \in \partial C$ be the boundary point such that $x_\vartheta \cdot \vartheta = r_\vartheta$. Then, consider $S_{\vartheta, \varepsilon} = \{x \in C: |x \cdot \vartheta - r_\vartheta| \leq \varepsilon\}$, which is a slice of size $\varepsilon > 0$ containing x_ϑ in its boundary. Since, by assumption, C is uniformly convex and smooth (notice that without the assumption it would just be convex), given any $\varepsilon > 0$, for any ball $B_{\sqrt{\varepsilon}}(x_\vartheta)$, it holds $S_{\vartheta, \varepsilon} \subset B_{\sqrt{\varepsilon}}(x_\vartheta)$. By (4.3), we have that $\nu(S_{\vartheta, \varepsilon}) \gtrsim \varepsilon$, and hence $\nu(B_{\sqrt{\varepsilon}}(x_\vartheta)) \gtrsim \varepsilon$. We now consider a collection of points $\{y_i\} \subset \partial C$ such that $|y_i - y_j| \geq 2\sqrt{\varepsilon}$. The number of such points is of the order $(\sqrt{\varepsilon})^{-(d-1)}$, since ∂C is a $d-1$ dimensional surface. Define $(\partial C)_\varepsilon := \{x \in \mathbb{R}^d: \text{dist}(\partial C, x) \leq \varepsilon\}$. Considering the above argument for any ball $B_{\sqrt{\varepsilon}}(y_i)$ we obtain that

$$1 = \nu(C) \geq \nu((\partial C)_\varepsilon) \geq \sum_{i=1}^{\varepsilon^{-\frac{d-1}{2}}} \nu(B_{\sqrt{\varepsilon}}(y_i)) \gtrsim \sum_{i=1}^{\varepsilon^{-\frac{d-1}{2}}} \varepsilon \gtrsim \varepsilon^{\frac{-d+3}{2}},$$

where we denoted by $\gtrsim$ inequalities holding up to multiplication by a dimensional constant. By arbitrariness of ε , we deduce that it is necessary to assume $d \leq 3$. Moreover, when $d = 3$, we also observe that ν must give positive mass to ∂C .

Remark 4.3. The measure ν with uniform projections (i.e., $\nu^\vartheta(r) = \chi_{[-1/2, 1/2]}(r)$ for any $\vartheta \in \mathbb{S}^{d-1}$) satisfies (4.3): by Remark 4.2 such a measure only exists when $d = 2, 3$, and in this case the set C mentioned there coincides with the closed ball of radius $1/2$ centered at the origin. More precisely, for $d = 3$ the measure ν is the uniform probability measure on the sphere of radius $1/2$, and for $d = 2$ it is an absolutely continuous measure with density $\nu(x) = \frac{1}{\pi \sqrt{1/4 - |x|^2}}$ for $|x| < 1/2$, and zero outside this ball. One can check that in this case we have

$$\frac{SW_2^2(\rho, \nu)}{2} = \int \frac{|x|^2}{2d} \rho(x) dx - \frac{\omega_{d-1}}{2d\omega_d} \iint \rho(x)\rho(y)|x-y| dx dy + \frac{1}{24} \quad \text{for any } \rho \in \mathcal{P}_2(\mathbb{R}^d).$$

For $\mathcal{V}(\rho) = \int V(x) d\rho(x)$ and $\mathcal{W}(\rho) = \iint W(x-y) d\rho(x) d\rho(y)$, we have that

$$\frac{\delta \mathcal{V}}{\delta \rho}(\rho) = V(x), \quad \frac{\delta \mathcal{W}}{\delta \rho}(\rho) = 2W * \rho.$$

Hence, the velocity field of the associated SWF is

$$v[\rho](x) = -\frac{x}{d} + \frac{\omega_{d-1}}{d\omega_d} \nabla W * \rho(x), \quad (4.4)$$

where

$$W(z) = |z|. \quad (4.5)$$

4.3 Lipschitz bounds when ν satisfies Assumption (4.6)

As usual, in the following result we set $\nu^\vartheta := (\pi^\vartheta)_\# \nu$ (and identify it with its density) and denote by $N_\vartheta(r)$ the cumulative distribution of ν^ϑ , i. e., $N_\vartheta(r) = \int_{-\infty}^r \nu^\vartheta(s) ds$.

Theorem 4.4. Let $M > 0$, $K = \{\rho \in \mathcal{P}_2(\mathbb{R}^d): \|\rho\|_{L^\infty} \leq M\}$, and $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Assume that there exist $0 < \alpha < (d-1)/d$, $\kappa > 0$ such that

$$\nu^\vartheta(r) \geq \kappa (\min\{N_\vartheta(r), 1 - N_\vartheta(r)\})^\alpha \quad \text{for a.e. } \vartheta \in \mathbb{S}^{d-1} \text{ and every } r \in \mathbb{R}. \quad (4.6)$$

Then, $\mathcal{F}_\nu$ is λ -geodesically convex on K with

$$\lambda = \frac{1}{d} - \frac{c(d)}{\left(1 - \frac{\alpha d}{d-1}\right) \kappa} M^{\frac{1}{d}},$$

where $c(d)$ is a constant depending only on d .

We postpone the proof of Theorem 4.4 in Section 5. Rather, here we provide some remarks on condition (4.6).

Remark 4.5. If ν satisfies (4.6), then ν must be compactly supported. Indeed, since $N'_\vartheta = \nu^\vartheta$, assumption (4.6) is equivalent to

$$N'_\vartheta(r) \geq \kappa \min\{(1 - N(r))^\alpha, N(r)^\alpha\}.$$

Moreover, since $\lim_{r \rightarrow +\infty} N(r) = 1$, for r large enough this leads to

$$N'_\vartheta(r) \geq \kappa(1 - N_\vartheta(r))^\alpha. \quad (4.7)$$

This in turn gives $(1 - N_\vartheta)^{1-\alpha} \leq -c < 0$, and it is not possible to have a nonnegative solution of this differential inequality on an unbounded domain for $\alpha < 1$.

Remark 4.6. We could relax assumption (4.6) by asking that $\nu^\vartheta(r) \geq \kappa g(\min\{N_\vartheta(r), 1 - N_\vartheta(r)\})$ where $g(t)$ satisfies the following properties:

1. $g(ct) \geq c^\alpha g(t)$, for some $\alpha \in \mathbb{R}$,
2. $\frac{1}{g\left(t^{\frac{d}{d-1}}\right)} \in L^1_{\text{loc}}$.

4.4 Examples of measures satisfying (4.6)

Example 4.7. (Projections bounded from below) Let ν be compactly supported with the property that $\text{spt } \nu^\vartheta$ is an interval I_ϑ for any $\vartheta \in \mathbb{S}^{d-1}$, and there exists $\kappa > 0$ such that

$$\nu^\vartheta(r) \geq c_0$$

for almost any $\vartheta \in \mathbb{S}^{d-1}$, $r \in I_\vartheta$. Then one trivially has

$$\nu_\vartheta(r) \geq c_0 N_\vartheta^\alpha(r), \quad (4.8)$$

since $N_\vartheta(r) \leq 1$, for any $\alpha \in [0, (d-1)/d]$.

Example 4.8. (Prescribed behavior near the boundary of their convex support) The following Lemma, whose proof is presented at the end of this list, holds true

Lemma 4.9. Let $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ be such that $\text{spt } (\nu) \subset C \subset \mathbb{R}^d$ where C is a uniformly convex, smooth set. For any $r > 0$, let $C_r := \{x \in C : \text{dist}(x, \partial C) \leq r\}$. Assume there exist $r_0, c_1, c_2 > 0$ and $\beta_1 \geq \beta_2 > -1$ with

$$0 < \frac{\frac{d-1}{2} + \beta_1}{\frac{d+1}{2} + \beta_2} < \frac{d-1}{d}, \quad (4.9)$$

such that the restriction of ν to C_{r_0} is absolutely continuous with respect to Lebesgue measure, with density satisfying

$$c_1 \text{dist}(x, \partial C)^{\beta_1} \leq \nu(x) \leq c_2 \text{dist}(x, \partial C)^{\beta_2} \quad \text{for any } x \in C_{r_0}. \quad (4.10)$$

Then, ν satisfies (4.6).

Example 4.10. (Densities bounded on their convex support) We can take $\beta_1 = \beta_2 = 0$ in Example 4.8. Therefore, assumption (4.6) is satisfied by measures with support contained in a smooth and uniformly convex domain C for which there exist $r_0 > 0$ and $0 < c_1 < c_2$ such that

$$c_1 \leq \nu(x) \leq c_2 \quad \text{for any } x \in C_{r_0}.$$

Proof of Lemma 4.9. In the following, the symbols “ $\gtrsim$ ” and “ $\lesssim$ ” will denote inequalities holding up to constants only depending on the dimension, on the set C and on r_0 .

Given any $\vartheta \in \mathbb{S}^{d-1}$, denote by $r_-^\vartheta = \inf_{t \in \mathbb{R}} \{\nu^\vartheta(t) > 0\}$, $r_+^\vartheta = \sup_{t \in \mathbb{R}} \{\nu^\vartheta(t) > 0\}$.

We distinguish three cases:

1. $r \in (r_-^\vartheta + r_0, r_+^\vartheta - r_0)$,
2. $r \in (r_-^\vartheta, r_-^\vartheta + r_0)$,
3. $r \in (r_+^\vartheta - r_0, r_+^\vartheta)$.

In each of these three domains, we will provide estimates from below for ν^ϑ , and estimates from above for N^ϑ , in the order, thus finally deducing the desired comparison.

Estimates from below for ν^ϑ :

1. Up to reducing the value of r_0 if needed, for $r \in (r_-^\vartheta + r_0, r_+^\vartheta - r_0)$, the set $S_\vartheta(r, r_0) := \{x \in C : x \cdot \vartheta = r, \frac{1}{2}r_0 < \text{dist}(x, \partial C) < r_0\}$ has $(d-1)$ -measure which is at least cr_0 . This provides, by (4.10),

$$\nu^\vartheta(r) = \int_{x \in C : x \cdot \vartheta = r} \nu(x) \geq \int_{S_\vartheta(r, r_0)} \nu(x) \gtrsim r_0^{\beta_1+1}.$$

2. For $r \in (r_-^\vartheta, r_-^\vartheta + r_0)$ one has

$$\nu^\vartheta(r) = \int_{x \in C : x \cdot \vartheta = r} \nu(x) = \int_{x \in C_{r_0} : x \cdot \vartheta = r} \nu(x) \geq c_1 \int_{x \in C_{r_0} : x \cdot \vartheta = r} \text{dist}(\partial C, x)^{\beta_1}.$$

By the smoothness of C , for any $r \in (r_-^\vartheta, r_-^\vartheta + r_0)$, there exists a $(d-1)$ dimensional ball $B_R^{(d-1)}$ with radius $|r - r_-^\vartheta|^{1/2} \lesssim R \lesssim |r - r_-^\vartheta|^{1/2}$ such that $B_R^{(d-1)} \subset C_{r_0} \cap \{x \cdot \vartheta = r\}$. On every point of this ball we have $\text{dist}(\partial C, x) \leq |r - r_-^\vartheta|$, but if we restrict to a ball of radius $R/2$ we also have $\text{dist}(\partial C, x) \gtrsim |r - r_-^\vartheta|$. We then obtain

$$\nu^\vartheta(r) \gtrsim c_1 \int_{B_{R/2}^{(d-1)}} |r - r_-^\vartheta|^\beta \gtrsim c |r - r_-^\vartheta|^{\beta_1 + \frac{d-1}{2}}. \quad (4.11)$$

3. On $(r_+^\vartheta - r_0, r_+^\vartheta)$, we can repeat the same argument and obtain

$$\nu^\vartheta(r) \gtrsim c_1 |r - r_+^\vartheta|^{\beta_1 + \frac{d-1}{2}}. \quad (4.12)$$

Estimates from above for $N_\vartheta, 1 - N_\vartheta$:

1. For $r \in (r_-^\vartheta + r_0, r_+^\vartheta - r_0)$ it trivially holds

$$N_\vartheta(r) \leq 1, \quad 1 - N_\vartheta(r) \leq 1.$$

2. We start from an estimate from above on v^ϑ . For $r \in (r_-^\vartheta, r_-^\vartheta + r_0)$, we have that

$$v^\vartheta(r) \leq c_2 \int_{x \in C: x \cdot \vartheta = r} \text{dist}(x, \partial C)^{\beta_2} dx,$$

where we used that for such an r we have $C_{r_0} \cap \{x \cdot \vartheta = r\} = C \cap \{x \cdot \vartheta = r\}$. If $\beta_2 \geq 0$, then one has that $\text{dist}(x, \partial C) \leq |r - r_-^\vartheta|$ for any $x \in C_{r_0} \cap \{x \cdot \vartheta = r\}$. Moreover, by the uniform convexity of C , for any $r \in (r_-^\vartheta, r_-^\vartheta + r_0)$, there exists a $(d-1)$ dimensional ball $B_R^{(d-1)}$ with radius $|r - r_-^\vartheta|^{1/2} \lesssim R \lesssim |r - r_-^\vartheta|^{1/2}$ such that $C \cap \{x \cdot \vartheta = r\} \subset B_R^{(d-1)}$. Thus

$$v^\vartheta(r) \leq c_2 \int_{x \in C: x \cdot \vartheta = r} |r - r_-^\vartheta|^{\beta_2} dx \lesssim c_2 \int_{B_R^{(d-1)}} |r - r_-^\vartheta|^{\beta_2} dx \lesssim c_2 |r - r_-^\vartheta|^{\beta_2 + \frac{d-1}{2}}. \quad (4.13)$$

If instead, $\beta_2 < 0$, we need a more refined argument, since upper bounds on the distance to ∂C are not anymore what we need, and we cannot restrict to a smaller ball on which we can easily estimate from below such a distance, as we did when we were bounding v^ϑ from below. We note that the function $x \mapsto \text{dist}(x, \partial C)$ is concave (see Lemma 4.12 below). This function takes value 0 on the relative boundary of $C \cap \{x : x \cdot \vartheta = r\}$. Moreover, there is a point $y_0 \in C \cap \{x : x \cdot \vartheta = r\}$ such that $\text{dist}(y_0, \partial C) \gtrsim |r - r_-^\vartheta|$, that can be taken to be the projection onto the hyperplane $\{x : x \cdot \vartheta = r\}$ of the point $x_0 \in \partial C$ realizing the infimum in $\inf_{x \in \partial C} x \cdot \vartheta$. This means, by concavity, that for every $x \in C \cap \{x : x \cdot \vartheta = r\}$ we have $\text{dist}(x, \partial C) \gtrsim |r - r_-^\vartheta| u(x)$, where the value $u(x)$ is defined as follows: let $y(x)$ be a point in the relative boundary of $C \cap \{x : x \cdot \vartheta = r\}$ such that the segment $[y_0, y(x)]$ passes through x , and define $u(x) := |x - y(x)|/|y_0 - y(x)|$.

We then obtain

$$v^\vartheta(r) \lesssim \int_{x \in C: x \cdot \vartheta = r} |r - r_-^\vartheta|^{\beta_2} u(x)^{\beta_2} dx.$$

We then notice that we have $|\nabla u| \geq 1/R$, where R is the radius of the ball $B_R^{d-1}(y_0)$ containing $C \cap \{x : x \cdot \vartheta = r\}$ and then

$$\int_{x \in C: x \cdot \vartheta = r} u(x)^{\beta_2} dx \leq R \int_{x \in C: x \cdot \vartheta = r} u(x)^{\beta_2} |\nabla u| dx = R \int_0^1 t^{\beta_2} \text{Per}(\{u > t\}) dt \lesssim R \cdot R^{d-2} \int_0^1 t^{\beta_2} dt.$$

Here we used first the coarea formula and then the fact that the sets $\{u > t\}$ are convex sets homothetical to $C \cap \{x : x \cdot \vartheta = r\}$, so that their perimeter is smaller than the one of the ball containing the latter set. Finally, we use $\beta_2 > -1$ for the finiteness of the integral, and we then obtain

$$v^\vartheta(r) \leq |r - r_-^\vartheta|^{\beta_2} R^{d-1} \simeq |r - r_-^\vartheta|^{\beta_2 + \frac{d-1}{2}}.$$

Using the previous estimates, we deduce

$$N_\vartheta(r) = \int_{r_-^\vartheta}^r v_\vartheta(s) ds \lesssim \int_0^{r-r_-^\vartheta} c_2 s^{\frac{d-1}{2} + \beta_2} ds \leq \frac{c_2}{\frac{d+1}{2} + \beta_2} |r - r_-^\vartheta|^{\frac{d+1}{2} + \beta_2}.$$

3. For $r \in (r_+^\vartheta - r_0, r_+^\vartheta)$, one similarly has

$$1 - N_\vartheta(r) \lesssim \frac{c_2}{\frac{d+1}{2} + \beta_2} |r_+^\vartheta - r|^{\frac{d+1}{2} + \beta_2}.$$

Comparing v^ϑ and N_ϑ :

1. On $(r_-^\vartheta + r_0, r_+^\vartheta - r_0)$, it holds

$$v^\vartheta(r) \gtrsim \frac{c_1}{\beta_1 + 1} r_0^{\beta_1 + 1} \gtrsim \frac{c_1}{\beta_1 + 1} r_0^{\beta_1 + 1} \min\{N_\vartheta(r), 1 - N_\vartheta(r)\}^\alpha,$$

for any $\alpha \in (0, (d-1)/d)$;

2. On $(r_-^\vartheta, r_-^\vartheta + r_0)$, denoting by κ a constant only depending on v , it holds

$$v^\vartheta(r) \gtrsim \kappa N_\vartheta(r)^{\frac{\frac{d-1}{2} + \beta_1}{\frac{d+1}{2} + \beta_2}} = \kappa \min\{N_\vartheta(r), 1 - N_\vartheta(r)\}^{\frac{\frac{d-1}{2} + \beta_1}{\frac{d+1}{2} + \beta_2}}.$$

3. On $(r_+^\vartheta - r_0, r_+^\vartheta)$, one similarly has

$$v^\vartheta(r) \gtrsim \kappa \min\{N_\vartheta(r), 1 - N_\vartheta(r)\}^{\frac{\frac{d-1}{2} + \beta_1}{\frac{d+1}{2} + \beta_2}}.$$

Imposing condition (4.9), we obtain the claim. $\square$

Remark 4.11. From the proof of Lemma 4.9 we deduce that, in the case when $\beta_1 = \beta_2 = \beta$, assumption (4.6) is satisfied as soon as $\max\{-1, -\frac{d-1}{2}\} < \beta < \frac{d-1}{2}$.

Lemma 4.12. Let K be convex, and let $R(x) = \text{dist}(x, \partial K)$. Then R is concave on K .

Proof. Given $x, y \in K$, we have $B(x, R(x)), B(y, R(y)) \subset K$. Then, by convexity of K , we have $B(tx + (1-t)y, tR(x) + (1-t)R(y)) \subset K$, which shows $R(tx + (1-t)y) \geq tR(x) + (1-t)R(y)$, hence the concavity of R . $\square$

5 Proof of Theorem 4.4

We will prove the following Theorem, which is a sufficient condition to obtain Theorem 4.4:

Proposition 5.1. Let $K = \{\rho \in \mathcal{P}_2(\mathbb{R}^d) : \|\rho\|_{L^\infty} \leq M\}$, for some $M > 0$. If v satisfies (4.6), then there exists a constant $c(d) > 0$ depending only on d , such that

$$\text{Lip}\left(v[\rho] + \frac{\text{id}}{d}\right) \leq \frac{c(d)}{(1 - \frac{\alpha d}{d-1})\kappa} M^{\frac{1}{d}} =: L \quad (5.1)$$

for any $\rho \in K$. In particular, $\mathcal{F}_v$ is λ -geodesically convex on K for $\lambda = \frac{1}{d} - L$.

Proof. Thanks to Lemma 3.1 and the fact that the set K is geodesically convex, we just need to prove that (5.1) holds true for any $\rho \in K$ and any v satisfying assumption (4.6). By definition of $v[\rho, v]$, fixing a point $x \in \text{spt } \rho$, we have (the first inequality is obtained as in the proof of Lemma 4.1)

$$\begin{aligned} \left| Dv[\rho](x) + \frac{\text{id}}{d} \right| &\leq \int_{\mathbb{S}^{d-1}} \frac{\rho^\vartheta(x \cdot \vartheta)}{v^\vartheta(T_\vartheta(x \cdot \vartheta))} d\sigma(\vartheta) \\ &\leq \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^\vartheta(x \cdot \vartheta)}{(\min\{N_\vartheta(T_\vartheta(x \cdot \vartheta)), 1 - N_\vartheta(T_\vartheta(x \cdot \vartheta))\})^\alpha} d\sigma(\vartheta) \\ &= \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^\vartheta(x \cdot \vartheta)}{(\min\{R_\vartheta(x \cdot \vartheta), 1 - R_\vartheta(x \cdot \vartheta)\})^\alpha} d\sigma(\vartheta), \end{aligned}$$

where in the last equality, we used again the relation

$$N_{\vartheta}(T_{\vartheta}(r)) = R_{\vartheta}(r) \quad \text{for any } r \in \text{spt } \rho^{\vartheta}.$$

Denoting by $\chi^+(\vartheta) = \chi_{\{R_{\vartheta}(x \cdot \vartheta) \leq 1/2\}}(\vartheta)$ and $\chi^-(\vartheta) = \chi_{\{R_{\vartheta}(x \cdot \vartheta) \geq 1/2\}}(\vartheta)$, we obtain

$$\begin{aligned} \left| Dv[\rho](x) + \frac{\text{id}}{d} \right| &\leq \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{R_{\vartheta}(x \cdot \vartheta)^{\alpha}} \chi^+(\vartheta) d\sigma(\vartheta) + \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{(1 - R_{\vartheta}(x \cdot \vartheta))^{\alpha}} \chi^-(\vartheta) d\sigma(\vartheta) \\ &\leq \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{R_{\vartheta}(x \cdot \vartheta)^{\alpha}} d\sigma(\vartheta) + \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{(1 - R_{\vartheta}(x \cdot \vartheta))^{\alpha}} d\sigma(\vartheta) \\ &= \frac{2}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\rho^{\vartheta}(x \cdot \vartheta)}{R_{\vartheta}(x \cdot \vartheta)^{\alpha}} d\sigma(\vartheta), \end{aligned} \quad (5.2)$$

where the last equality was obtained by symmetry. For any $e \in \mathbb{S}^{d-1}$, define

$$m_x(e) = \int_0^{+\infty} \rho(x + e r) r^{d-2} dr. \quad (5.3)$$

We have that, for any $\vartheta \in \mathbb{S}^{d-1}$

$$\begin{aligned} \rho^{\vartheta}(x \cdot \vartheta) &= \int_{e \cdot \vartheta = 0} m_x(e) d\mathcal{H}^{d-2}(e), \\ R_{\vartheta}(x \cdot \vartheta) &= \int_{e \cdot \vartheta \leq 0} \int_0^{+\infty} \rho(x + e r) r^{d-1} dr d\sigma(e). \end{aligned} \quad (5.4)$$

Here and below, integrals over equators $\{e \in \mathbb{S}^{d-1} : e \cdot \vartheta = 0\}$ and their subsets are taken with respect to $\mathcal{H}^{d-2}$, while integrals over hemispheres use σ . When $d = 2$, $\mathcal{H}^0$ is counting measure.

Applying (5.4) in (5.2), we obtain

$$\left| Dv[\rho](x) + \frac{\text{id}}{d} \right| \leq \frac{2}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta = 0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} \int_0^{+\infty} \rho(x + e r) r^{d-1} dr d\sigma(e) \right)^{\alpha}} d\sigma(\vartheta). \quad (5.5)$$

Consider now

$$q(M, m_x(e)) := \inf_{0 \leq \mu \leq M} \left\{ \int_0^{+\infty} \mu(r) r^{d-1} dr : \int_0^{+\infty} \mu(r) r^{d-2} dr = m_x(e) \right\}.$$

The above infimum is achieved by $\mu(r) = M \chi_{(0, R)}(r)$. Here, necessarily,

$$R = \left(\frac{(d-1)m_x(e)}{M} \right)^{\frac{1}{d-1}},$$

since

$$m_x(e) = M \int_0^{+\infty} \chi_{(0, R)}(r) r^{d-2} dr = M \int_0^R r^{d-2} dr = \frac{M}{d-1} R^{d-1}.$$

Therefore,

$$q(M, m_x(e)) = M^{-\frac{1}{d-1}} \frac{(d-1)^{\frac{d}{d-1}}}{d} m_x(e)^{\frac{d}{d-1}}. \quad (5.6)$$

Recalling (5.3), and thanks to (5.6), we are able to estimate the denominator on the right hand side of (5.5), obtaining

$$\begin{aligned} \left| Dv[\rho](x) + \frac{\text{id}}{d} \right| &\leq \frac{2d^\alpha M^{\frac{\alpha}{d-1}}}{(d-1)^{\frac{d\alpha}{d-1}} \kappa} \int_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta=0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m_x^{\frac{d}{d-1}}(e) d\sigma(e) \right)^\alpha} d\sigma(\vartheta) \\ &\leq \frac{c(d) M^{\frac{\alpha}{d-1}}}{\kappa} \int_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta=0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m_x(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\sigma(\vartheta), \end{aligned} \quad (5.7)$$

where $c(d) > 0$ denotes a constant depending only on d (the term $(d-1)^{\frac{d\alpha}{d-1}}$ can indeed be bounded uniformly for $0 \leq \alpha < (d-1)/d$).

We first present the proof in dimension $d = 2$, for the reader's convenience. Next, we will generalize the same argument to any dimension.

Proof in dimension $d = 2$

Since in this case it holds

$$\int_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta=0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m_x(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\sigma(\vartheta) = \int_{\mathbb{S}^1} \frac{m_x(\vartheta + \pi/2) + m_x(\vartheta - \pi/2)}{\left(\int_{\vartheta - \pi/2}^{\vartheta + \pi/2} m_x(e) d\sigma(e) \right)^{2\alpha}} d\sigma(\vartheta),$$

the idea is to estimate this integral in the variable ϑ writing the numerator as a derivative of the denominator. Fix an orthonormal basis (e_1, e_2) of $\mathbb{R}^2$. We parametrize $\mathbb{S}^1$ defining $\theta(e_1, e_2, t) = e_1 \cos t + e_2 \sin t$. Define $A(e_1, e_2, t) = \{e \in \mathbb{S}^{d-1} : e \cdot \theta(e_1, e_2, t) \leq 0\}$, and $A^+(e_1, e_2, t) = A(e_1, e_2, t) \cap \{e \in \mathbb{S}^{d-1} : e \cdot e_2 \geq 0\}$. Define also

$$\varphi_{e_1, e_2}(t) = \int_{A^+(e_1, e_2, t)} m_x(e) d\sigma(e).$$

The function φ is positive, decreasing for $t \in [0, \pi/2) \cup [3\pi/2, 2\pi)$ and increasing in $[\pi/2, 3\pi/2)$. Moreover it holds

$$\begin{aligned} \theta(e_1, e_2, t) &= \theta(e_1, -e_2, -t), \\ \varphi_{e_1, e_2}(t) &\leq \int_{A(e_1, e_2, t)} m_x(e) d\sigma(e), \\ |\varphi'_{e_1, e_2}(t)| + |\varphi'_{e_1, -e_2}(-t)| &= \int_{e \cdot \theta(e_1, e_2, t)=0} m_x(e) d\mathcal{H}^{d-2}(e). \end{aligned} \quad (5.8)$$

Incidentally, let us note that two of the above identities have been obtained by exploiting the symmetries in our construction when considering an orthonormal basis with a different orientation with respect to the fixed one. This trick, even if more involved, will be the same at the core of the identities in Lemma 5.2, needed for the construction in dimension $d \geq 3$.

We return to the proof in dimension $d = 2$. By our definitions and (5.8), we have

$$\begin{aligned}
& \int_{\mathbb{S}^1} \frac{\int_{e \cdot \vartheta = 0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m_x(e) d\sigma(e) \right)^{2\alpha}} d\sigma(\vartheta) = \int_0^{2\pi} \frac{\int_{e \cdot \theta(e_1, e_2, t) = 0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{A(e_1, e_2, t)} m_x(e) d\sigma(e) \right)^{2\alpha}} dt \\
&= \int_0^{2\pi} \frac{|\varphi'_{e_1, e_2}(t)|}{\left(\int_{A(e_1, e_2, t)} m_x(e) d\sigma(e) \right)^{2\alpha}} dt + \int_0^{2\pi} \frac{|\varphi'_{e_1, -e_2}(-t)|}{\left(\int_{A(e_1, -e_2, -t)} m_x(e) d\sigma(e) \right)^{2\alpha}} dt \\
&\leq \int_0^{2\pi} \frac{|\varphi'_{e_1, e_2}(t)|}{(\varphi_{e_1, e_2}(t))^{2\alpha}} dt + \int_0^{2\pi} \frac{|\varphi'_{e_1, -e_2}(-t)|}{(\varphi_{e_1, -e_2}(-t))^{2\alpha}} dt \\
&\leq \frac{1}{1-2\alpha} TV[\varphi_{e_1, e_2}^{1-2\alpha}] + \frac{1}{1-2\alpha} TV[\varphi_{e_1, -e_2}^{1-2\alpha}],
\end{aligned} \tag{5.9}$$

Where TV stands for the total variation. Using the piecewise monotone behavior of $t \mapsto \varphi_{e_1, e_2}(t)$, we have $TV[\varphi_{e_1, e_2}^{1-2\alpha}] \leq 2 \left(\max_{t \in [0, 2\pi)} \varphi_{e_1, e_2}(t) \right)^{1-2\alpha}$. Then, by [CS25, Lemma 5.1] it holds

$$\max_{t \in [0, 2\pi)} \varphi_{e_1, e_2}(t) \leq \int_{\mathbb{S}^1} \int_0^{+\infty} \rho(x + er) dr d\sigma(e) = \int_{\mathbb{R}^2} \frac{\rho(y)}{|x - y|} dy. \tag{5.10}$$

Applying estimate (4.2) with $p = \infty$ to the right hand side of (5.10), we deduce the existence of a constant c such that

$$\max_{t \in [0, 2\pi)} \varphi_{e_1, e_2}(t) \leq c M^{\frac{1}{2}}, \tag{5.11}$$

and the estimate does not depend on the chosen basis (e_1, e_2) . Combining (5.7), (5.9) and (5.11), we conclude the proof of Proposition 5.1 when $d = 2$.

Proof in dimension $d \geq 3$

Define E to be the $(2d-3)$ -dimensional set in $\mathbb{R}^{2d}$ given by couples $(e_1, e_2) \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ such that $e_1 \cdot e_2 = 0$, and endow it with the restriction of $\mathcal{H}^{2d-3}$. For any given triplet $(e_1, e_2, t) \in E \times [0, 2\pi)$, define $\theta(e_1, e_2, t) = e_1 \cos t + e_2 \sin t$. For fixed t , the fibres of $(e_1, e_2) \mapsto \theta(e_1, e_2, t)$ carry $\mathcal{H}^{d-2}$; the fibres of the full map $\theta : E \times [0, 2\pi) \rightarrow \mathbb{S}^{d-1}$ carry $\mathcal{H}^{d-1}$. All these measures are induced by the Euclidean product metric. As before, define

$$\begin{aligned}
A(e_1, e_2, t) &= \{e \in \mathbb{S}^{d-1} : e \cdot \theta(e_1, e_2, t) \leq 0\}, \\
A^+(e_1, e_2, t) &= A(e_1, e_2, t) \cap \{e \in \mathbb{S}^{d-1} : e \cdot e_2 \geq 0\}, \\
\partial A(e_1, e_2, t) &= \{e \in \mathbb{S}^{d-1} : e \cdot \theta(e_1, e_2, t) = 0\}, \\
\partial^+ A(e_1, e_2, t) &= \partial A(e_1, e_2, t) \cap \{e \in \mathbb{S}^{d-1} : e \cdot e_2 \geq 0\}.
\end{aligned}$$

Since no ambiguity can occur, we will simplify the notation, removing the dependence on e_i, t . We introduce again the function

$$\varphi_{e_1, e_2}(t) := \int_{A^+} m_x(e) d\sigma(e).$$

Notice that $\varphi_{e_1, e_2}(t)$ is a positive function, decreasing for $t \in [0, \pi/2) \cup [3\pi/2, 2\pi)$ and increasing in $[\pi/2, 3\pi/2)$. For the rest of the proof, we will make use of the following

Lemma 5.2. *Let m_x, φ, θ be defined as above, and $(e_1, e_2, t) \in E \times [0, 2\pi]$. Then,*

1. There exists a function $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ such that, for any $t \in [0, 2\pi)$, and for any $e_1, e_2 \in E$, it holds

$$\left| \int_{\partial^+ A} m(e) h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e) \right| = |\varphi'_{e_1, e_2}(t)|$$

2. Fix $t \in [0, 2\pi)$. The value of

$$\int_{\theta(e_1, e_2, t) = \vartheta, e \cdot e_2 \geq 0} h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e_1, e_2)$$

is constant with respect to $(e, \vartheta) \in E$.

3. There exists a dimensional constant $c(d)$ such that, for any two given measurable functions f, g it holds

$$\oint_{\mathbb{S}^{d-1}} \frac{\widetilde{f}(\vartheta)}{g(\vartheta)} d\sigma(\vartheta) = c(d) \int_{E \times [0, 2\pi)} \frac{f(e_1, e_2, t)}{g(\theta(e_1, e_2, t))} d\mathcal{H}^{2d-3}(e_1, e_2) dt,$$

where

$$\widetilde{f}(\vartheta) = \int_{\{\theta(e_1, e_2, t) = \vartheta\}} f(e_1, e_2, t) d\mathcal{H}^{d-1}(e_1, e_2, t).$$

We postpone the proof of Lemma 5.2. Combining the results therein contained, we can now conclude the proof of Proposition 5.1. Fix $t \in [0, 2\pi)$. Using the fact that, for any $(e_1, e_2, t) \in E \times [0, 2\pi]$ it holds

$$\varphi_{e_1, e_2}(t) \leq \int_A m_x(e) d\sigma(e),$$

together with point (1) of Lemma 5.2, one obtains

$$\int_E \frac{|\varphi'_{e_1, e_2}(t)|}{(\varphi_{e_1, e_2}(t))^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{2d-3}(e_1, e_2) \geq \left| \int_E \frac{\int_{\partial^+ A} m(e) h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e)}{\left(\int_A m(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{2d-3}(e_1, e_2) \right|.$$

Integrating both sides with respect to $t \in [0, 2\pi)$, we get

$$\int_0^{2\pi} \int_E \frac{|\varphi'_{e_1, e_2}(t)|}{(\varphi_{e_1, e_2}(t))^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{2d-3}(e_1, e_2) dt \geq \left| \int_{E \times [0, 2\pi)} \frac{\int_{\partial^+ A} m(e) h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e)}{\left(\int_A m(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{2d-3}(e_1, e_2) dt \right|.$$

Applying point (3) of Lemma 5.2 to the right hand side of the above expression, we deduce

$$\begin{aligned} & \int_0^{2\pi} \int_E \frac{|\varphi'_{e_1, e_2}(t)|}{(\varphi_{e_1, e_2}(t))^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{2d-3}(e_1, e_2) dt \\ & \geq c(d) \left| \oint_{\mathbb{S}^{d-1}} \int_{\{(e_1, e_2, t): \theta(e_1, e_2, t) = \vartheta\}} \frac{\int_{e \cdot \vartheta = 0, e \cdot e_2 \geq 0} m(e) h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\mathcal{H}^{d-1}(e_1, e_2, t) d\sigma(\vartheta) \right| \\ & = c(d) \left| \oint_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta = 0} m_x(e) \int_{\{(e_1, e_2, t): \theta(e_1, e_2, t) = \vartheta, e \cdot e_2 \geq 0\}} h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-1}(e_1, e_2, t) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m(e) d\sigma(e) \right)^{\frac{d\alpha}{d-1}}} d\sigma(\vartheta) \right|. \end{aligned}$$

Finally, by point (2) of Lemma 5.2, we infer that

$$\begin{aligned} \oint_{\mathbb{S}^{d-1}} \frac{\int_{e \cdot \vartheta=0} m_x(e) d\mathcal{H}^{d-2}(e)}{\left(\int_{e \cdot \vartheta \leq 0} m_x(e) d\sigma(e)\right)^{\frac{d\alpha}{d-1}}} d\sigma(\vartheta) &\leq c(d) \int_E \int_0^{2\pi} \frac{|\varphi'_{e_1, e_2}(t)|}{(\varphi_{e_1, e_2}(t))^{\frac{d\alpha}{d-1}}} dt d\mathcal{H}^{2d-3}(e_1, e_2) \\ &\leq \frac{c(d)}{1 - \frac{d\alpha}{d-1}} 2 \int_E \max_{t \in (0, 2\pi)} |\varphi_{e_1, e_2}(t)|^{-\frac{d\alpha}{d-1} + 1} d\mathcal{H}^{2d-3}(e_1, e_2). \end{aligned}$$

On the other hand, using again [CS25, Lemma 5.1] together with the estimate proved in (4.2) with $p = \infty$, we have

$$\max_{t \in (0, 2\pi)} \varphi_{e_1, e_2}(t) \leq \int_{\mathbb{S}^{d-1}} \int_0^{+\infty} \rho(x + er) r^{d-2} dr d\sigma(e) = \int_{\mathbb{R}^d} \frac{\rho(y)}{|x - y|} dy \leq c_d M^{1/d}.$$

Therefore, it holds

$$\int_E \max_{(0, 2\pi)} |\varphi_{e_1, e_2}(t)|^{1 - \frac{d\alpha}{d-1}} d\mathcal{H}^{2d-3}(e_1, e_2) \leq \int_E \left| c(d) M^{\frac{1}{d}} \right|^{1 - \frac{d\alpha}{d-1}} d\mathcal{H}^{2d-3}(e_1, e_2) \leq c(d) M^{\frac{1}{d} - \frac{\alpha}{d-1}}. \quad (5.12)$$

Finally, using estimate (5.12) in (5.7), we obtain

$$\left| Dv[\rho](x) + \frac{\text{id}}{d} \right| \leq \frac{c(d)}{(1 - \frac{d\alpha}{d-1})\kappa} M^{\frac{1}{d}},$$

which concludes the proof of Proposition 5.1. $\square$

Proof of Lemma 5.2. Here we prove the results contained in Lemma 5.2:

1. Fix $(e_1, e_2) \in E, t \in [0, 2\pi)$. For any $e \in \mathbb{S}^{d-1}$, define $f_{e, e_1, e_2}(t) := e \cdot \theta(e_1, e_2, t)$. We have that

$$\begin{aligned} \frac{d}{dt} \varphi_{e_1, e_2}(t) &= \frac{d}{dt} \int_{\{e \cdot e_2 \geq 0\}} m_x(e) \chi_{(-\infty, 0)} \circ f_{e, e_1, e_2}(t) d\sigma(e) \\ &= - \int_{\{e \cdot e_2 \geq 0\}} m_x(e) \delta_0(f_{e, e_1, e_2}(t)) (-e \cdot e_1 \sin t + e \cdot e_2 \cos t) d\sigma(e) \end{aligned}$$

Imposing $f_{e, e_1, e_2}(t) = 0$ (in order to have a non-zero integrand), we can write $\cos t, \sin t$ in terms of $e \cdot e_1, e \cdot e_2$, obtaining

$$|\varphi'_{e_1, e_2}(t)| = \left| \int_{e \cdot \theta(e_1, e_2, t) = 0 \cap \{e \cdot e_2 \geq 0\}} m_x(e) h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e) \right|$$

for the function $h = -e \cdot e_1 \sin t + e \cdot e_2 \cos t$.

2. Fix $t \in [0, 2\pi)$ and $(e', \vartheta') \neq (e, \vartheta)$, with $e \cdot \vartheta = e' \cdot \vartheta' = 0$. One can always find $R \in O(d)$ such that $Re = e'$ and $R\vartheta = \vartheta'$. Indeed, there always exists an isometry sending a given orthonormal basis into another. Define now $e'_1 = Re_1$ and $e'_2 = Re_2$; we deduce that

$\theta(e'_1, e'_2, t) = R\theta(e_1, e_2, t) = R\vartheta = \vartheta'$. Therefore

$$\begin{aligned}
\int_{\{(e_1, e_2): \theta(e_1, e_2, t) = \vartheta', e' \cdot e_2 \geq 0\}} h(e' \cdot e_1, e' \cdot e_2) d\mathcal{H}^{d-2}(e_1, e_2) &= \\
&= \int_{\{(e_1, e_2): \theta(e_1, e_2, t) = \vartheta', e' \cdot e_2 \geq 0\}} h(e \cdot R^{-1}e_1, e \cdot R^{-1}e_2) d\mathcal{H}^{d-2}(e_1, e_2) \\
&= \int_{\{(R^{-1}e_1, R^{-1}e_2): \theta(R^{-1}e_1, R^{-1}e_2, t) = \vartheta, e \cdot R^{-1}e_2 \geq 0\}} h(e \cdot R^{-1}e_1, e \cdot R^{-1}e_2) d\mathcal{H}^{d-2}(e_1, e_2) \\
&= \int_{\{(e_1, e_2): \theta(e_1, e_2, t) = \vartheta, e \cdot e_2 \geq 0\}} h(e \cdot e_1, e \cdot e_2) d\mathcal{H}^{d-2}(e_1, e_2).
\end{aligned}$$

We also notice that the integral in the lines above is not zero: since $e \cdot e_1 \cos t + e \cdot e_2 \sin t = 0$, recalling $h = -e \cdot e_1 \sin t + e \cdot e_2 \cos t$, it follows that

$$h(e \cdot e_1, e \cdot e_2) = \operatorname{sgn}(\cos t) \sqrt{(e \cdot e_1)^2 + (e \cdot e_2)^2}.$$

3. We will identify any point $p \in E \times [0, 2\pi)$ with a point $(e_1, e_2, t) \in \mathbb{R}^d \times \mathbb{R}^d \times [0, 2\pi)$. For any $(v, w, a) \in T_p(E \times [0, 2\pi))$ and for some $\delta > 0$, consider a curve

$$\begin{aligned}
\gamma: (-\delta, \delta) &\rightarrow \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \\
\tau &\mapsto (\gamma_1(\tau), \gamma_2(\tau), \gamma_3(\tau)),
\end{aligned}$$

with the property that $\gamma(0) = p \in E \times [0, 2\pi)$, $\gamma'_1(0) = v$, $\gamma'_2(0) = w$ and $\gamma'_3(0) = a$. Without loss of generality, we can assume that $e_1 = (1, 0, \dots, 0)$ and $e_2 = (0, 1, 0, \dots, 0)$. In order to compute $D_{(e_1, e_2, t)}\theta[(v, w, a)]$, with $(v, w, a) \in T_{(e_1, e_2, t)}(E \times [0, 2\pi))$, we write

$$D\theta_{(e_1, e_2, t)}(v, w, a) = v \cos t + w \sin t + a(-e_1 \sin t + e_2 \cos t).$$

Thus, the differential is represented in ambient coordinates by the matrix

$$D\theta = \begin{pmatrix} \cos t \operatorname{id} & \sin t \operatorname{id} & \zeta(t) \end{pmatrix},$$

where

$$\zeta(t) = (-\sin t, \cos t, 0, \dots, 0)^T.$$

Let $\eta = -e_1 \sin t + e_2 \cos t$, and complete (e_1, e_2) to an orthonormal basis $(e_1, \dots, e_d)$. A basis for $T_{(e_1, e_2, t)}M$ is given by the following $2d - 2$ orthonormal vectors: for $j = 3, \dots, d$, consider

$$\begin{aligned}
\xi_1 &:= \frac{(e_2, -e_1, 0)}{\sqrt{2}}, & \xi_2 &:= (0, 0, 1), \\
\xi_j &:= (e_j, 0, 0), & \widetilde{\xi}_j &:= (0, e_j, 0).
\end{aligned}$$

We have $D\theta[\xi_1] = \eta/\sqrt{2}$, $D\theta[\xi_2] = \eta$, $D\theta[\xi_j] = \cos t e_j$, $D\theta[\widetilde{\xi}_j] = \sin t e_j$. Therefore, with respect to the orthonormal basis $(\eta, e_3, \dots, e_d)$ which spans the tangent space to $\mathbb{S}^{d-1}$, we deduce

$$D\theta D\theta^* = \operatorname{diag}\left(\frac{3}{2}, 1, \dots, 1\right),$$

and hence $J_{d-1}\theta = \sqrt{3/2}$. By the coarea formula (see [Mag12]), we have that for any integrable measurable $\phi: E \times [0, 2\pi) \rightarrow \mathbb{R}$, it holds

$$\int_{E \times [0, 2\pi)} J_\theta(p, t) \phi(p, t) d\mathcal{H}^{2d-3}(p) dt = \int_{\mathbb{S}^{d-1}} \left(\int_{\theta^{-1}(\vartheta)} \phi(p, t) d\mathcal{H}^{d-1}(p, t) \right) d\sigma(\vartheta).$$

Taking

$$\phi(p, t) = \frac{f(p, t)}{g \circ \theta(p, t)}$$

we obtain the claim.

This concludes the proof of Lemma 5.2 □

6 Uniqueness of the SWF

In this section we furnish the proof of Corollary 1.1 using Theorem 4.4. We also provide an alternative proof under the assumption that the projections of the target measure ν are bounded from below. Finally, we show a third argument which applies if ν is the measure having uniform projections: we stress again that this can happen only if $d = 2, 3$ as a consequence of Remark 4.3.

We first recall the differentiation formulas for Wasserstein distances that will be used later. These results are discussed in [San15], Chapter 5.

Lemma 6.1. *Let $p \in [1, +\infty)$ and let $(\rho_i(t))_{i \geq 0}$, $i = 1, 2$, be locally absolutely continuous curves in time with values in $(\mathcal{P}_p(\mathbb{R}^d), W_p)$, consisting of probability measures absolutely continuous with respect to Lebesgue measure and solving*

$$\partial_t \rho_i + \nabla \cdot (\rho_i w_i) = 0.$$

Then $t \mapsto W_p^p(\rho_1(t), \rho_2(t))$ is locally absolutely continuous, and the following identities hold for almost every t .

If $\varphi = \varphi_t$ is a Kantorovich potential from $\rho_1(t)$ to $\rho_2(t)$ for the cost $\frac{1}{p}|x - y|^p$, then we have

$$\frac{1}{p} \frac{d}{dt} W_p^p(\rho_1(t), \rho_2(t)) = \int \nabla \varphi \cdot w_1 d\rho_1(t) + \int \nabla \varphi^c \cdot w_2 d\rho_2(t). \quad (6.1)$$

This can be re-written with transport maps. If $p > 1$ and $T = T_t$ is the optimal transport map from $\rho_1(t)$ to $\rho_2(t)$ for the same cost, then

$$\frac{1}{p} \frac{d}{dt} W_p^p(\rho_1(t), \rho_2(t)) = \int |x - T(x)|^{p-2} (x - T(x)) \cdot (w_1(t, x) - w_2(t, T(x))) d\rho_1(t)(x),$$

where $|z|^{p-2}z$ is understood to be zero at $z = 0$.

If $p = 1$, we choose φ to be 1-Lipschitz, so that $\varphi^c = -\varphi$; for any optimal transport map T for the cost $|x - y|$, we have

$$\frac{d}{dt} W_1(\rho_1(t), \rho_2(t)) = \int \nabla \varphi \cdot (w_1 - w_2 \circ T) d\rho_1(t).$$

We now go on with some proofs of uniqueness for the Sliced Wasserstein Flow.

Proof of Corollary 1.1. The existence of a solution is proved in [Bon13] together with local (in time) L^∞ bounds. Let $(\rho_i(t))_{i \geq 0}$ for $i = 1, 2$ be two solutions of the sliced Wasserstein flow (1.3), and let $T > 0$. Let M be a constant such that $\rho_i(t, x) \leq M$ for a.e. x and every $t \in [0, T]$ and

$i = 1, 2$. By Theorem 4.4 we know that, under assumption (4.6) $\mathcal{F}_v$ is λ -geodesically convex on $K = \{\rho : \rho \leq M\}$ for some $\lambda \in \mathbb{R}$. We then use formula (6.1) and a well-known estimate due to geodesic convexity (see [San17], Lemma 4.12), to obtain

$$\frac{d}{dt} W_2^2(\rho_1(t), \rho_2(t)) \leq -2\lambda W_2^2(\rho_1(t), \rho_2(t)),$$

. Uniqueness then follows by Gronwall inequality which proves $\rho_1(t) = \rho_2(t)$ if $\rho_1(0) = \rho_2(0)$ (see [AGS05]). $\square$

Remark 6.2. We provide here an alternative proof of the uniqueness of the SWF, which does not use Theorem 4.4, by computing the variation of the distance W_1 between two solutions. We assume that, for almost every $\vartheta \in \mathbb{S}^{d-1}$, the support of ν^ϑ is an interval I_ϑ , and that there exists $\kappa > 0$, independent of ϑ , such that $\nu^\vartheta \geq \kappa$ almost everywhere on I_ϑ . The assumption on ν implies that it is compactly supported, and this uniqueness result also concerns compactly supported solutions. Let ρ_1, ρ_2 be compactly supported solutions of (1.3). We will denote by T, φ the optimal transport map from ρ_1 to ρ_2 and the associated Kantorovich potential, respectively. Moreover, we will denote by T_i the optimal transport map from ρ_i to ν , and by T_i^ϑ the one from ρ_i^ϑ to ν^ϑ . By Lemma 6.1, for almost every t ,

$$\begin{aligned} \partial_t W_1(\rho_1(t), \rho_2(t)) &= \int \varphi(\partial_t \rho_1 - \partial_t \rho_2) = \int \nabla \varphi \cdot (v[\rho_1]\rho_1 - v[\rho_2]\rho_2) \\ &= \int \nabla \varphi \cdot (v[\rho_1] - v[\rho_2] \circ T)\rho_1, \end{aligned}$$

Then, recalling that $\text{Lip}(\varphi) \leq 1$,

$$\begin{aligned} \partial_t W_1(\rho_1(t), \rho_2(t)) &= \int \nabla \varphi \cdot (v[\rho_1] - v[\rho_2]) d\rho_1 + \int \nabla \varphi \cdot (v[\rho_2] - v[\rho_2] \circ T) d\rho_1 \\ &\leq \int \int_{\mathbb{S}^{d-1}} |T_1^\vartheta(x \cdot \vartheta) - T_2^\vartheta(x \cdot \vartheta)| d\rho_1(x) d\sigma(\vartheta) + \text{Lip}(v[\rho_2]) \int |T(x) - x| d\rho_1. \end{aligned} \quad (6.2)$$

Thanks to the assumption $\nu^\vartheta \geq \kappa$ (which in turn implies that $N_\vartheta^{[-1]}$ is κ^{-1} -Lipschitz continuous for every ϑ), we have

$$\begin{aligned} |T_{\vartheta,1}(x \cdot \vartheta) - T_{\vartheta,2}(x \cdot \vartheta)| &= |N_\vartheta^{[-1]}(R_{1,\vartheta}(x \cdot \vartheta)) - N_\vartheta^{[-1]}(R_{2,\vartheta}(x \cdot \vartheta))| \\ &\leq \frac{1}{\kappa} |R_{1,\vartheta}(x \cdot \vartheta) - R_{2,\vartheta}(x \cdot \vartheta)|. \end{aligned} \quad (6.3)$$

We recall that, for any $\mu, \gamma \in \mathcal{P}_2(\mathbb{R})$ it holds

$$W_1(\mu, \gamma) = \int_0^1 |R_\mu^{-1}(t) - R_\gamma^{-1}(t)| dt = \int_{\mathbb{R}} |R_\mu(t) - R_\gamma(t)| dt. \quad (6.4)$$

Moreover, since $\rho_1 \in L^\infty$ and it is compactly supported, we can provide an upper bound to ρ_1^ϑ . Combining (6.2), (6.3) and (6.4), together with the fact that $\rho_1^\vartheta \in L^\infty$, we obtain

$$\begin{aligned} \partial_t W_1(\rho_1(t), \rho_2(t)) &\leq \frac{1}{\kappa} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}} |R_{1,\vartheta}(r) - R_{2,\vartheta}(r)| d\rho_1^\vartheta(r) d\sigma(\vartheta) + \text{Lip}(v) W_1(\rho_1(t), \rho_2(t)) \\ &\leq \frac{C}{\kappa} W_1(\rho_1(t), \rho_2(t)) \end{aligned}$$

for some constant $C > 0$ that is fixed once $T > 0$ is. By Gronwall lemma, we deduce again the uniqueness of SWF.

Remark 6.3. As shown in [CS25, Section 4.1], if in (1.3) both the initial datum ρ_0 and the target measure ν have bounded support, then the solution ρ_t also has this property for any $t \geq 0$. Under this assumption, and if the target measure ν has uniform projections (namely, $\nu^\vartheta(r) = \chi_{[-1/2, 1/2]}(r)$ for any $\vartheta \in \mathbb{S}^{d-1}$), then the uniqueness of solutions to (1.3) can be proved following a strategy presented in [LST24]. Indeed, repeating verbatim the computations made in the proof of [LST24, Theorem 4.1], for any ρ_1, ρ_2 compactly supported solutions of the SWF with target measure ν , and for $p \geq 1$, the following estimate holds. Here we fix a finite time interval $[0, T]$ and choose $R > 0$ so that the supports of both solutions lie in B_R throughout this interval; Lip_{B_R} denotes the Lipschitz constant on this ball.

$$\frac{1}{p} \frac{d}{dt} W_p^p(\rho_1(t), \rho_2(t)) \leq \underbrace{\text{Lip}_{B_R}(v[\rho_1]) \int |T_t(x) - x|^p d\rho_1}_{:=I_1} + \underbrace{W_p(\rho_1, \rho_2)^{p-1} \left(\int |v[\rho_1] - v[\rho_2]|^p (T_t(x)) d\rho_1 \right)^{1/p}}_{:=I_2}.$$

Then, for I_1 , recalling $W(z) = |z|$ and setting $a_d = \omega_{d-1}/(d\omega_d)$, we have

$$\begin{aligned} \text{Lip}_{B_R}(v[\rho_1(t)]) &\leq \frac{1}{d} + a_d \|D^2 W * \rho_1(t)\|_{L^\infty(B_R)} \\ &\leq \frac{1}{d} + a_d \|D^2 W\|_{L^{p_0}(B_{2R})} \|\rho_1(t)\|_{L^{p'_0}}. \end{aligned}$$

On the other hand, for I_2 , following again the argument presented in [LST24], we have that for any $q \geq 1$ it holds

$$I_2 \leq a_d W_p^p(\rho_1, \rho_2) \|D^2 W\|_{L^{p'_0}(B_{2R})} \|\rho_2(t)\|_{L^q}^{\frac{1}{p}}.$$

Choose $1 \leq p < p_0 < d$ and set $q = p_0/(p_0 - p)$, so that $pq' = p_0$. We deduce

$$\begin{aligned} \frac{1}{p} \frac{d}{dt} W_p^p(\rho_1(t), \rho_2(t)) &\leq I_1 + I_2 \\ &\leq W_p^p(\rho_1(t), \rho_2(t)) \left[\frac{1}{d} + a_d \|D^2 W\|_{L^{p_0}(B_{2R})} \left(\|\rho_1(t)\|_{L^{p'_0}} + \|\rho_2(t)\|_{L^q}^{1/p} \right) \right]. \end{aligned}$$

By (4.5) we observe that $|D^2 W| \lesssim |z|^{-1} \in L_{\text{loc}}^{p_0}$ for any $p_0 < d$, hence $p'_0 > d'$. Applying again the Gronwall lemma, and thanks to Theorem 2.6, we can conclude that $\rho_1 = \rho_2$.

Acknowledgements

This work was supported by the European Union via the ERC AdG 101054420 EYAWKAJKOS project and by the ARQUS alliance, through the support of a collaborative project between the teams in Lyon and Padua.

AI statement

In this work, LLM models have been used to compute some exact integral values throughout the proof, and to proofread the final manuscript.

References

- [AGS05] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005.
- [BG21] Erhan Bayraktar and Gaoyue Guo. “Strong equivalence between metrics of Wasserstein type”. In: *Electron. Commun. Probab.* 26 (2021). doi: [10.1214/21-ECP383](https://doi.org/10.1214/21-ECP383).
- [Bon13] Nicolas Bonnotte. “Unidimensional and evolution methods for optimal transportation”. PhD thesis. Université Paris Sud-Paris XI; Scuola normale superiore (Pise, Italie), 2013.
- [Bre87] Yann Brenier. “Décomposition polaire et réarrangement monotone des champs de vecteurs”. In: *CR Acad. Sci. Paris Sér. I Math.* 305 (1987), pp. 805–808.
- [CFMW25] Guillaume Carlier, Alessio Figalli, Quentin Mérigot, and Yi Wang. “Sharp comparisons between sliced and standard 1-Wasserstein distances”. In: *arXiv preprint arXiv:2510.16465* (2025).
- [CS25] Giacomo Cozzi and Filippo Santambrogio. “Long-time asymptotics of the sliced-Wasserstein flow”. In: *SIAM Journal on Imaging Sciences* 18.1 (2025), pp. 1–19.
- [KNSBR19] Soheil Kolouri, Kimia Nadjahi, Umut Simsekli, Roland Badeau, and Gustavo Rohde. “Generalized sliced wasserstein distances”. In: *Advances in neural information processing systems* 32 (2019).
- [LST24] Frédéric Lagoutière, Filippo Santambrogio, and Sébastien Tran Tien. “Vanishing viscosity limit for aggregation-diffusion equations”. In: *Journal de l’École polytechnique—Mathématiques* 11 (2024), pp. 1123–1179.
- [Mag12] Francesco Maggi. *Sets of finite perimeter and geometric variational problems: an introduction to Geometric Measure Theory*. 135. Cambridge University Press, 2012.
- [MMS09] Daniel Matthes, Robert J. McCann, and Giuseppe Savaré. “A family of nonlinear fourth order equations of gradient flow type”. In: *Communications in Partial Differential Equations* 34.11 (2009), pp. 1352–1397. doi: [10.1080/03605300903296256](https://doi.org/10.1080/03605300903296256).
- [NDCKSS20] Kimia Nadjahi, Alain Durmus, Lénaïc Chizat, Soheil Kolouri, Shahin Shahrampour, and Umut Simsekli. “Statistical and Topological Properties of Sliced Probability Divergences”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 20802–20812.
- [PS23] Sangmin Park and Dejan Slepčev. “Geometry and analytic properties of the sliced Wasserstein space”. In: *arXiv preprint arXiv:2311.05134* (2023).
- [PKD07] François Pitié, Anil C Kokaram, and Rozenn Dahyot. “Automated colour grading using colour distribution transfer”. In: *Computer Vision and Image Understanding* 107.1-2 (2007), pp. 123–137.
- [RM23] Raif M Rustamov and Subhabrata Majumdar. “Intrinsic sliced Wasserstein distances for comparing collections of probability distributions on manifolds and graphs”. In: *International Conference on Machine Learning*. PMLR. 2023, pp. 29388–29415.
- [San15] Filippo Santambrogio. *Optimal transport for applied mathematicians*. Vol. 55. 58-63. Springer, 2015.

- [San17] Filippo Santambrogio. “{Euclidean, metric, and Wasserstein} gradient flows: an overview”. In: *Bulletin of Mathematical Sciences* 7 (2017), pp. 87–154.
- [VKM25] Christophe Vauthier, Anna Korba, and Quentin Mérigot. “Towards understanding gradient dynamics of the sliced-wasserstein distance via critical point analysis”. In: *arXiv preprint arXiv:2502.06525* (2025).