Independence of synthetic Curvature Dimension conditions on transport distance exponent[*]

Afiny Akdemir[†], Fabio Cavalletti[‡], Andrew Colinet[†], Robert McCann[†], Flavia Santarcangelo[‡]

July 22, 2020

Abstract

The celebrated Lott-Sturm-Villani theory of metric measure spaces furnishes synthetic notions of a Ricci curvature lower bound K joint with an upper bound N on the dimension. Their condition, called the Curvature-Dimension condition and denoted by $\mathsf{CD}(K, N)$, is formulated in terms of a modified displacement convexity of an entropy functional along W_2-Wasserstein geodesics. We show that the choice of the squared-distance function as transport cost does not influence the theory. By denoting with $\mathsf{CD}_p(K, N)$ the analogous condition but with the cost as the pth power of the distance, we show that $\mathsf{CD}_p(K, N)$ are all equivalent conditions for any $p > 1$ — at least in spaces whose geodesics do not branch.

Following [13], we show that the trait d'union between all the seemingly unrelated $\mathsf{CD}_p(K, N)$ conditions is the needle decomposition or localization technique associated to the L^1-optimal transport problem. We also establish the local-to-global property of $\mathsf{CD}_p(K, N)$ spaces.

Contents

[†]Department of Mathematics, University of Toronto, Toronto Ontario, Canada M5S 2E4 `afiny@math.toronto.edu`, `andrew.colinet@mail.utoronto.ca`, `mccann@math.toronto.edu`

[‡]Mathematics Area, SISSA, Trieste (Italy) `cavallet@sissa.it`, `fsantarc@sissa.it`

1

1 Introduction

The theory of optimal transport has been successfully used to study geometric and analytic properties of possibly singular spaces. In their seminal works, Lott–Villani [29] and Sturm [40, 41] have established for metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$ consisting of a complete separable metric space (X, d) endowed with a Radon reference measure $\mathfrak{m}$, a synthetic condition which encodes, in a generalized sense, a combined lower bound $K \in \mathbb{R}$ on the Ricci curvature and upper bound $N \in [1, \infty)$ on the dimension. Their condition is called the Curvature-Dimension condition $\mathsf{CD}(K, N)$; a general account on its history, huge developments and impacts goes far beyond the scope of this introduction.

For our purposes, the cornerstone of the Curvature-Dimension condition is the equivalence between a lower bound on the Ricci curvature combined with an upper bound on the dimension and a certain convexity property of an entropy functional along W_2-Wasserstein geodesics in the setting of weighted manifolds [18] [38]. The strength of the optimal transport approach permitted Lott–Villani and Sturm to obtain a stable notion of convergence, with stability intended with respect to a suitable distance over the class of metric measure spaces. We refer to Section 2.2 for precise definitions.

As the $\mathsf{CD}(K, N)$ condition for smooth manifolds is equivalent to a joint lower bound on the Ricci curvature and an upper bound on the dimension, it is natural to consider whether the squared-distance cost function plays a special role in the theory. Among the possible transport cost functions, the power distance costs, namely d^p with $p > 1$, are related to the geometry of the underlying space. The power distance costs have already appeared in the literature in the definition of the p-Wasserstein distance W_p that turns the space of probability measures with finite p^{th}-moments into a complete and separable metric space $(\mathcal{P}_p(X), W_p)$. Another natural setting for such spaces can also be seen in the case of doubly-degenerate diffusion dyanamics [35] [1]. Accordingly, the modified displacement convexity of the entropy functional can be considered with respect to W_p-geodesics – and this in turn furnishes a straightforward and legitimate extension of the definition of $\mathsf{CD}(K, N)$ condition proposed by Kell [27] and denoted by $\mathsf{CD}_p(K, N)$. The notation $\mathsf{CD}(K, N)$ will be reserved for the classical case $p = 2$. While Kell established the equivalence of all $\mathsf{CD}_p(K, N)$ in the smooth setting via the use of Ricci curvature, no previous results are known in the context of nonsmooth metric measure spaces.

Our approach to obtaining this equivalence in the nonsmooth setting utilizes two closely related L^1 optimal transportation curvature dimension conditions introduced by Cavalletti and Milman [13], which we denote by $\mathsf{CD}^1(K, N)$ and $\mathsf{CD}^1_{Lip}(K, N)$. The $\mathsf{CD}^1(K, N)$ condition has been successfully

used in [13] to establish the local-to-global property of $\mathsf{CD}(K, N)$ spaces. Cavalletti and Milman's formulation is partially based on the needle or localization paradigm introduced by Klartag [28] in the smooth setting, which was later generalized to the metric setting by Cavalletti and Mondino [14].

Cavalletti and Milman established the local-to-global property by demonstrating the equivalence of the local version of $\mathsf{CD}(K, N)$ condition (namely $\mathsf{CD}_{loc}(K, N)$, Definition 2.5) to the $\mathsf{CD}^1(K, N)$ condition. In particular, the trait d'union between all of the curvature-dimension conditions is in the behaviour of the gradient flow lines of signed-distance functions, also known as transport rays.

In this paper we will use this point of view to link two different curvature dimension conditions: we will demonstrate the equivalence of $\mathsf{CD}_p(K, N)$ and $\mathsf{CD}_q(K, N)$ for a general metric measure space $(X, \mathsf{d}, \mathfrak{m})$, for $p, q > 1$ and $K, N \in \mathbb{R}$ with $N > 1$, under the requirement that $(X, \mathsf{d}, \mathfrak{m})$ is either non-branching or at least satisfies appropriate versions of the essential non-branching condition of Definition 2.1. More specifically, we obtain the following results:

Theorem 1.1 (Equivalence of CD_p on $p > 1$). *Let $(X, \mathsf{d}, \mathfrak{m})$ be such that $\mathfrak{m}(X) = 1$. Assume $(X, \mathsf{d}, \mathfrak{m})$ is p-essentially non-branching and verifies $\mathsf{CD}_p(K, N)$ for some $p > 1$. If $(X, \mathsf{d}, \mathfrak{m})$ is also q-essentially non-branching for some $q > 1$, then it verifies $\mathsf{CD}_q(K, N)$.*

Recently, Cavalletti, Gigli, and Santarcangelo [11] have characterized $\mathsf{CD}^1_{Lip}(K, N)$ in terms of a modified displacement convexity of an entropy functional along a certain family of W_1 Wasserstein geodesics. Hence, Theorem 1.1 completes the picture by showing that for any $p \geq 1$, all of the $\mathsf{CD}_p(K, N)$ conditions, when expressed in terms of displacement convexity, are equivalent, provided the space X satisfies the appropriate essentially non-branching condition.

Since we employ the strategy used in [13] to distance costs with powers other than $p = 2$, we also establish the local-to-global property for $\mathsf{CD}_p(K, N)$ spaces.

Theorem 1.2 (Local-to-Global). *Fix any $p > 1$ and $K, N \in \mathbb{R}$ with $N > 1$. Let $(X, \mathsf{d}, \mathfrak{m})$ be a p-essentially non-branching metric measure space verifying $\mathsf{CD}_{p,loc}(K, N)$ from Definition 2.5 and such that (X, d) is a length space with $\mathrm{spt}(\mathfrak{m}) = X$ and $\mathfrak{m}(X) = 1$. Then $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_p(K, N)$.*

In Theorem 1.1 and Theorem 1.2 we are assuming $\mathfrak{m}(X) = 1$. This assumption is also used in [13] but we believe that it is most likely a purely technical assumption. At the moment, the main obstacle to the case of a general Radon measure $\mathfrak{m}$ is the lack of a canonical disintegration theorem once a "measurable" partition is given. For some preliminary results in this direction we refer to [15].

Another motivation to studying distance costs with powers other than $p = 2$ comes from the recent works of McCann [31] and Mondino-Suhr [32], where the authors analyze the relation between optimal transportation and timelike Ricci curvature bounds in the smooth Lorentzian setting. Analogously to the Riemannian setting, timelike Ricci curvature lower bounds can be equivalently characterised in terms of convexity properties of the Bolzmann-Shannon entropy functional along ℓ_p-geodesics of probability measures, where ℓ_p denotes the causal transport distance with exponent $p \in (0, 1]$. This point of view has been pushed forward in [16] and [30] where the authors proposed a synthetic formulation of the Strong Energy condition, denoted by $\mathsf{TCD}_p(K, N)$, which is valid for non-smooth Lorentzian spaces. Unlike the Riemmannian case, the Lorentzian setting does not have a distinguished p; and one of the next steps of the theory will be to address whether $TCD_p(K, N)$ depends on p or not.

1.1 Structure of the paper

We start this note by recalling basic definitions of Optimal Transport as well as the notions of synthetic lower curvature bounds as introduced by Lott-Sturm-Villani in Section 2.

Section 3 is devoted to a careful analysis of Kantorovich potentials and their evolution via the Hopf-Lax semigroup with a general exponent $p > 1$. In particular, we will obtain second and third order information on the time behaviour of φ_t leading to the fundamental Theorem 3.19 where a new third order inequality is obtained that plays a crucial role in the rest of the paper.

In Section 4 we show that a local version of $\mathsf{CD}_p(K, N)$ implies $\mathsf{CD}^1(K, N)$ in the version reported in Theorem 4.6. Finally, in Section 5 we obtain a complete equivalence of all $\mathsf{CD}_p(K, N)$ (Theorem 1.1) and each of them also enjoys the local-to-global property (Theorem 1.2).

1.2 Brief Overview

Throughout this overview we will be working on a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ satisfying suitable hypotheses. We will also be considering the transport of a measures μ_0 to μ_1 where both measures are absolutely continuous with respect to $\mathfrak{m}$. We denote the interpolant measure by μ_t and we write ρ_t for their densities with respect to $\mathfrak{m}$.

In Section 3 the goal is to obtain information about the time derivative of the t-propagated s-Kantorovich potential Φ_s^t as defined in Section 3.6. This quantity is crucial for the Jacobian factor that appears when comparing interpolant measures, μ_t, between measures μ_0 and μ_1 along a transport geodesic at two times. To achieve this goal, Section 3.1 - 3.3 are dedicated to a detailed study of the regularity properties of the Hopf-Lax transform. In particular we establish second order regularity for the Hopf-Lax transform of a Kantorovich potential as well as a few identities related to the positional information stored in a Kantorovich potential. From here, Section 3.4 demonstrates, through a delicate argument, third order temporal regularity of time propagated Kantorovich potentials along transport geodesics.

In Section 4, we remind the reader of the standard definitions of L^1-optimal transport. We show in Section 4.1 that the non-branched transport set partitions a space into transport rays. This partition allows us to decompose measures into a collection of one-dimensional conditional measures concentrated on transport rays. This disintegration also gives the advantage of passing curvature information from the total space down to the L^1-transport rays no matter from which $\mathsf{CD}_p(K, N)$ we are starting, as we show in Section 4.2. This is highlighted in Theorem 4.6 where we demonstrate that any p-essentially non-branching metric measure space verifying $\mathsf{CD}_p(K, N)$ also verifies $\mathsf{CD}_{Lip}^1(K, N)$. This will be useful in Section 5 when we want to compare the restriction of measure to a Kantorovich geodesic at two different times. To propagate a measure from one time to another we will use the time propagated Kantorovich potential from Section 3.

In Section 5, the goal is to transfer the curvature properties along transport geodesics back to the total space through q-Wasserstein geodesics and hence proving that an enhanced version of $\mathsf{CD}_{Lip}^1(K, N)$ implies $\mathsf{CD}_q(K, N)$. This will be done by proving, in the terminology of [13], an "LY"-decomposition for the densities ρ_t of the q-Wasserstein geodesic μ_t (see Theorem 5.12). More precisely, this "LY"-decomposition provides a factorization of the ratio ρ_t/ρ_s into two factors: the first one — denoted by L — is a concave function taking into account only the one dimensional distortion due to the volume stretching in the direction of the geodesic. The second factor is denoted by Y and contains the volume distortion in the transversal directions.

To achieve this goal we first use the Disintegration theorem from Section 4 to represent $\mathfrak{m}$ as an average of measures that live on L^1-transport geodesics for the signed distance to any given level set of a p-Kantorovich potential. In this disintegration of $\mathfrak{m}$ we follow the evolution of a specific collection of Kantorovich geodesics. More specifically, we fix $a \in \mathbb{R}$ and $s \in (0, 1)$, and consider q-Kantorovich geodesics γ which satisfy $\varphi_s(\gamma_s) = a$, where φ_s is the evolved Kantorovich potential for

the q-Wasserstein geodesic. We denote such geodesics by $G_{a,s}$ and we disintegrate $\mathfrak{m}$ over $\{\gamma_t \colon \gamma \in G_{a,s}\}_{t\in[0,1]}$ to obtain

$$\mathfrak{m}\llcorner_{\mathrm{e}_{[0,1]}(G_{a,s})} = \int_{[0,1]} \mathfrak{m}_t^{a,s}\, \mathcal{L}^1(dt),$$

Then we compare this to a disintegration of $\mathfrak{m}$ over $\{\varphi_s^{-1}(a)\}_{a\in\mathbb{R}}$ on the time t evaluation of a sufficiently large set of Kantorovich geodesics denoted by G. Specifically, we obtain

$$\mathfrak{m}\llcorner_{\mathrm{e}_t(G)} = \int_{\varphi_s(\mathrm{e}_s(G))} \mathfrak{m}_{a,s}^t \mathcal{L}^1(da)$$

This leads to two measures, $\mathfrak{m}_t^{a,s}$ and $\mathfrak{m}_{a,s}^t$, that live on $\mathrm{e}_t(G_{a,s})$. In Section 5.3 we compare these two disintegrations to deduce that $\mathfrak{m}_t^{a,s}$ and $\mathfrak{m}_{a,s}^t$ differ only by $\partial_t \Phi_s^t$. This information is used in Section 5.4 to deduce the Jacobian factor between $\rho_t(\gamma_t)$ and $\rho_s(\gamma_s)$. This formula for the Jacobian factor allows us to conclude the desired "LY" decomposition. Once the "LY" decomposition is at our disposal, we can invoke [13] to conclude that the space satisfies $\mathsf{CD}_q(K, N)$.

2 Prerequisites

2.1 Geodesics and Measures

Let (X, d) be a complete and separable metric space. A map $\gamma : [0,1] \to X$ satisfying $\mathsf{d}(\gamma_t, \gamma_s) = |t - s|\mathsf{d}(\gamma_0, \gamma_1)$ for all $s, t \in [0,1]$ is called a geodesic connecting γ_0 to γ_1. We regard $\mathrm{Geo}(X)$ as a subset of all Lipschitz curves $\mathrm{Lip}([0,1], X)$ endowed with the uniform topology.

We say the metric space (X, d) is a geodesic metric space if for each $x, y \in X$ there is a geodesic connecting x and y.

For any $t \in [0,1]$, we denote the continuous evaluation map $\mathrm{e}_t : \mathrm{Geo}(X) \to X$ as $\mathrm{e}_t(\gamma) = \gamma_t$. We will also adopt the following abbreviations: given $I \subset [0,1]$ and $G \subset \mathrm{Geo}(X)$

$$\mathrm{e}_t(G) = G(t) = \{\gamma_t \; ; \; \gamma \in G\}, \quad \mathrm{e}_I(G) := \cup_{t\in I}\mathrm{e}_t(G).$$

The space of all Borel probability measures over X is denoted by $\mathcal{P}(X)$, and $\mathcal{P}_p(X)$ is the subspace of $\mathcal{P}(X)$ consisting of measures with finite p^{th}-moment. Given a non-negative Radon measure $\mathfrak{m}$, we call the space $(X, \mathsf{d}, \mathfrak{m})$ a metric measure space, and $\mathcal{P}_p(X, \mathsf{d}, \mathfrak{m})$ will denote the subspace of $\mathcal{P}_p(X)$ consisting of probability measures that are absolutely continuous with respect to $\mathfrak{m}$. Unless otherwise noted, we assume $\mathfrak{m}(X) = 1$ to permit disintegration of $\mathfrak{m}$ into conditional measures as needed. For any $p \geq 1$, the L^p-Wasserstein distance W_p is defined for any $\mu_0, \mu_1 \in \mathcal{P}(X)$ as

$$W_p^p(\mu_0, \mu_1) := \inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{X \times X} \mathsf{d}^p(x, y)\, \pi(dx, dy), \tag{2.1}$$

where $\Pi(\mu_0, \mu_1)$ is the set of $\pi \in \mathcal{P}(X \times X)$ with $(P_1)_\sharp \pi = \mu_0$ and $(P_2)_\sharp \pi = \mu_1$.

It is known that the infimum in (2.1) is always attained for any $\mu_0, \mu_1 \in \mathcal{P}(X)$; the set of optimal plans will be denoted by $\mathrm{Opt}_p(\mu_0, \mu_1)$.

As (X, d) is a complete and separable metric space, so is $(\mathcal{P}_p(X), W_p)$. It is also known that (X, d) is geodesic if and only if $(\mathcal{P}_p(X), W_p)$ is geodesic. Moreover, if (X, d) is a geodesic space, then the following two statements are equivalent (see for instance [3, Theorem 3.10]):

- $[0,1] \ni t \mapsto \mu_t \in \mathcal{P}_p(X)$ is a W_p-geodesic;

- there exists $\nu \in \mathcal{P}(\mathrm{Geo}(X))$ such that $(e_0, e_1)_\sharp \nu \in \mathrm{Opt}_p(\mu_0, \mu_1)$ and $\mu_t = (e_t)_\sharp \nu$.

The set of $\nu \in \mathcal{P}(\mathrm{Geo}(X))$ verifying the last point are called dynamical optimal plans and are denoted by $\mathrm{OptGeo}_p(\mu_0, \mu_1)$. Notice that if $\nu \in \mathrm{OptGeo}_p(\mu_0, \mu_1)$, then also $(e_t, e_s)_\sharp \nu$ is p-optimal between its marginals.

Definition 2.1 (p-Essentially Non-Branching). A subset $G \subset \mathrm{Geo}(X)$ of geodesics is called non-branching if for any $\gamma^1, \gamma^2 \in G$ the following holds:

$$\gamma_0^1 = \gamma_0^2, \ \gamma_{\bar{t}}^1 = \gamma_{\bar{t}}^2, \ \bar{t} \in (0,1) \quad \Longrightarrow \quad \gamma_s^1 = \gamma_s^2, \quad \forall s \in [0,1].$$

The space (X, d) is called *non-branching* if $\mathrm{Geo}(X)$ is non-branching; $(X, \mathsf{d}, \mathfrak{m})$ is called *p-essentially non-branching* if for all $\mu_0, \mu_1 \in \mathcal{P}_p(X, \mathsf{d}, m)$, any $\nu \in \mathrm{OptGeo}_p(\mu_0, \mu_1)$ is concentrated on a Borel non-branching set $G \subset \mathrm{Geo}(X)$, in agreement with the terminology of [37] when $p = 2$.

We remark that examples of branched spaces which are essentially non-branching may be found in Ohta [34].

2.2 Curvature-Dimension conditions

We recall the definition of volume distortion coefficients.

Definition 2.2 ($\sigma_{K,\mathcal{N}}$-coefficients). Given $K \in \mathbb{R}$ and $\mathcal{N} \in (0, \infty]$, define:

$$D_{K,\mathcal{N}} := \begin{cases} \frac{\pi}{\sqrt{K/\mathcal{N}}} & K > 0, \ \mathcal{N} < \infty, \\ +\infty & \text{otherwise.} \end{cases}$$

In addition, given $t \in [0,1]$ and $0 \le \theta < D_{K,\mathcal{N}}$, define:

$$\sigma_{K,\mathcal{N}}^{(t)}(\theta) := \frac{\sin(t\theta\sqrt{\frac{K}{\mathcal{N}}})}{\sin(\theta\sqrt{\frac{K}{\mathcal{N}}})} = \begin{cases} \frac{\sin(t\theta\sqrt{\frac{K}{\mathcal{N}}})}{\sin(\theta\sqrt{\frac{K}{\mathcal{N}}})} & K > 0, \ \mathcal{N} < \infty \\ t & K = 0 \text{ or } \mathcal{N} = \infty \\ \frac{\sinh(t\theta\sqrt{\frac{-K}{\mathcal{N}}})}{\sinh(\theta\sqrt{\frac{-K}{\mathcal{N}}})} & K < 0, \ \mathcal{N} < \infty \end{cases}$$

and set $\sigma_{K,\mathcal{N}}^{(t)}(0) = t$ and $\sigma_{K,\mathcal{N}}^{(t)}(\theta) = +\infty$ for $\theta \ge D_{K,\mathcal{N}}$.

Definition 2.3 ($\tau_{K,N}$-coefficients). Given $K \in \mathbb{R}$ and $N = \mathcal{N} + 1 \in (1, \infty]$, define:

$$\tau_{K,N}^{(t)}(\theta) := t^{\frac{1}{N}} \sigma_{K,N-1}^{(t)}(\theta)^{1-\frac{1}{N}}.$$

When $N = 1$, set $\tau_{K,1}^{(t)}(\theta) = t$ if $K \le 0$ and $\tau_{K,1}^{(t)}(\theta) = +\infty$ if $K > 0$.

We will use the following definition introduced in [41] for the case $p = 2$. Recall that given $N \in [1, \infty)$, the N-Rényi relative-entropy functional $\mathcal{E}_N : \mathcal{P}(X) \to [0, \infty]$ is defined as:

$$\mathcal{E}_N(\mu) := \int \rho^{1-\frac{1}{N}} d\mathfrak{m},$$

where $\mu = \rho\mathfrak{m} + \mu^{\mathrm{sing}}$ is the Lebesgue decomposition of μ with $\mu^{\mathrm{sing}} \perp \mathfrak{m}$. It is known [41] that $\mathcal{E}_N$ is upper semi-continuous with respect to the weak topology on $\mathcal{P}(X)$.

Definition 2.4 ($\mathsf{CD}_p(K,N)$)**.** Given $K,N \in \mathbb{R}$ with $N \geq 1$, $(X,\mathsf{d},\mathfrak{m})$ is said to satisfy $\mathsf{CD}_p(K,N)$ if for all $\mu_0,\mu_1 \in \mathcal{P}_p(X,\mathsf{d},\mathfrak{m})$, there exists $\nu \in \mathrm{OptGeo}_p(\mu_0,\mu_1)$ so that for all $t \in [0,1]$, $\mu_t := (\mathrm{e}_t)_\sharp \nu \ll \mathfrak{m}$, and for all $N' \geq N$:

$$\mathcal{E}_{N'}(\mu_t) \geq \int_{X \times X} \left(\tau_{K,N'}^{(1-t)}(\mathsf{d}(x_0,x_1))\rho_0^{-1/N'}(x_0) + \tau_{K,N'}^{(t)}(\mathsf{d}(x_0,x_1))\rho_1^{-1/N'}(x_1) \right) \pi(dx_0, dx_1), \qquad (2.2)$$

where $\pi = (\mathrm{e}_0,\mathrm{e}_1)_\sharp(\nu)$ and $\mu_i = \rho_i \mathfrak{m}$, $i = 0,1$.

When we omit the subscript p from $\mathsf{CD}_p(K,N)$, we tacitly mean the classical $p = 2$, as introduced independently by Lott-Villani in [29] and Sturm in [40, 41].

As a natural curvature notion, $\mathsf{CD}_p(K,N)$ has a local version that is denoted by $\mathsf{CD}_{p,loc}(K,N)$.

Definition 2.5 ($\mathsf{CD}_{p,loc}(K,N)$)**.** Given $K,N \in \mathbb{R}$ with $N \geq 1$, $(X,\mathsf{d},\mathfrak{m})$ is said to satisfy $\mathsf{CD}_{p,loc}(K,N)$ if for any $o \in \mathrm{spt}(\mathfrak{m})$, there exists a neighborhood $X_o \subset X$ of o, so that for all $\mu_0,\mu_1 \in \mathcal{P}_p(X,\mathsf{d},\mathfrak{m})$ supported in X_o, there exists $\nu \in \mathrm{OptGeo}_p(\mu_0,\mu_1)$ so that for all $t \in [0,1]$, $\mu_t := (\mathrm{e}_t)_\sharp \nu \ll \mathfrak{m}$, and for all $N' \geq N$, (2.2) holds.

Note that $(\mathrm{e}_t)_\sharp \nu$ from the definition of $\mathsf{CD}_{loc}^p(K,N)$ is not required to be supported in X_o for intermediate times $t \in (0,1)$.

Requiring the $\mathsf{CD}(K,N)$ condition to hold whenever μ_1 degenerates to δ_o, a delta-measure at $o \in \mathrm{spt}(\mathfrak{m})$, goes by the name of Measure Contraction Property and is denoted by $\mathsf{MCP}(K,N)$. This property was introduced independently by Ohta in [33] and Sturm in [41]. Since $\mathrm{OptGeo}_p(\mu_0,\delta_o)$ does not depend on p, whenever $p > 1$, the superscript will be omitted. We now record the version of the definition of $\mathsf{MCP}(K,N)$ found in [33].

Definition 2.6 ($\mathsf{MCP}(K,N)$)**.** The space $(X,\mathsf{d},\mathfrak{m})$ is said to satisfy $\mathsf{MCP}(K,N)$ if for any $o \in \mathrm{spt}(\mathfrak{m})$ and $\mu_0 \in \mathcal{P}_2(X,\mathsf{d},\mathfrak{m})$ of the form $\mu_0 = \frac{1}{\mathfrak{m}(A)}\mathfrak{m}\llcorner A$ for some Borel set $A \subset X$ with $0 < \mathfrak{m}(A) < \infty$ (and with $A \subset B(o,\pi\sqrt{(N-1)/K})$ if $K > 0$), there exists $\nu \in \mathrm{OptGeo}_2(\mu_0,\delta_o)$ such that:

$$\frac{d}{d\mathfrak{m}}\left[(\mathrm{e}_t)_\sharp \left(\tau_{K,N}^{(1-t)}(\mathsf{d}(\gamma_0,\gamma_1))^N \nu(d\gamma)\right)\right] \leq \frac{1}{\mathfrak{m}(A)} \quad \forall t \in [0,1]. \qquad (2.3)$$

As one would expect, $\mathsf{CD}_p(K,N)$ implies $\mathsf{MCP}(K,N)$ (see [13, Lemma 6.11] for the case $p = 2$; the proof works the same for any $p > 1$), without any type of essential non-branching. When coupled with the p-essentially non-branching condition, MCP yields nice properties for W_p-geodesics. A weaker contraction property [12] of $(X,\mathsf{d},\mathfrak{m})$ is called *qualitative non-degeneracy*, which asserts for each ball $B_R(x_0)$, there is a ratio $f(t) \in (0,1]$ with $\limsup_{t \to 0} f(t) > 1/2$ which bounds the decrease in measure whenever any Borel set $A \subset B_R(x_0)$ is contracted a fraction t of the distance towards any $x \in B_R(x_0)$:

$$\mathfrak{m}(\mathrm{e}_t(G)) \geq f(t)\mathfrak{m}(\mathrm{e}_0((G)) \qquad (2.4)$$

for $G = (\mathrm{e}_0 \times \mathrm{e}_1)^{-1}(A \times \{x\})$. Thus MCP permits one to invoke the following:

Theorem 2.7 (Optimal dynamic transport is unique iff the space is essentially non-branching [26])**.** *Let $(X,\mathsf{d},\mathfrak{m})$ be a metric measure space with $\mathfrak{m}$ qualitatively non-degenerate. Then the following properties are equivalent:*

i) $(X,\mathsf{d},\mathfrak{m})$ is p-essentially non-branching;

ii) for every $\mu_0, \mu_1 \in \mathcal{P}_p(X)$ with $\mu_0 \ll \mathfrak{m}$ there is a unique $\nu \in \mathrm{OptGeo}_p(\mu_0, \mu_1)$. Moreover, the p-optimal coupling $(e_0, e_1)_\sharp \nu$ is induced by a transport map and each interpolant $\mu_t = (e_t)_\sharp \nu$, where $t \in [0, 1)$, is absolutely continuous with respect to $\mathfrak{m}$.

Remark 2.8. It is also worth recalling that the local version of $\mathsf{CD}(K, N)$, denoted by $\mathsf{CD}_{loc}(K, N)$, is known to imply $\mathsf{MCP}(K, N)$ provided that (X, d) is a non-branching length space, see [17]. Since any $\mathsf{CD}_{p,loc}(K, N)$ gives the same information when considered for Wasserstein geodesics arriving at a Dirac mass, we can conclude that the same argument of [17] shows $\mathsf{CD}_{loc,p}(K, N)$ implies $\mathsf{MCP}(K, N)$, provided (X, d) is a non-branching length space.

Moreover, it has already been observed and used in the literature that the non-branching assumption can be weakened to essentially non-branching when $p = 2$: the non-branching property in [17] was used to obtain a partition of X formed of all geodesics arriving at the same point $o \in X$ and subsequently to ensure uniqueness of a dynamical optimal plan connecting μ_0 to μ_1 with $\mu_0 \ll \mathfrak{m}$. Both properties can be deduced from p-essentially non-branching together with Theorem 2.7; for more details see Section 5.1. Hence, we will tacitly use that for each $p > 1$, a metric measure space satisfying $\mathsf{CD}_{p,loc}(K, N)$ and which is a p-essentially non-branching length space also verifies $\mathsf{MCP}(K, N)$.

We conclude this subsection with the $\mathsf{CD}^1(K, N)$ condition introduced in [13]. Notice that this definition uses the additional assumption that $\mathfrak{m}(X) = 1$. We will also need to recall some classical terminology from the distance cost optimal transport theory that we briefly recall.

To any 1-Lipschitz function $u : X \to \mathbb{R}$ there is a naturally associated d-cyclically monotone set

$$\Gamma_u := \{(x, y) \in X \times X : u(x) - u(y) = \mathsf{d}(x, y)\}, \tag{2.5}$$

which we call the *transport ordering*. We write $x \geq_u y$ if and only if $(x, y) \in \Gamma_u$; the 1-Lipschitz condition on u implies $\geq_u$ is a partial-ordering. The *transport relation* R_u and the *transport set* $\mathcal{T}_u$ are defined as:

$$R_u := \Gamma_u \cup \Gamma_u^{-1}, \quad \mathcal{T}_u := P_1(R_u \setminus \{x = y\}), \tag{2.6}$$

where $\{x = y\}$ denotes the diagonal $\{(x, y) \in X^2 : x = y\}$, P_i the projection onto the i-th component and $\Gamma_u^{-1} = \{(x, y) \in X \times X : (y, x) \in \Gamma_u\}$. Since u is 1-Lipschitz, Γ_u, Γ_u^{-1} and R_u are closed sets, and so are $\Gamma_u(x)$ and $R_u(x)$ (recall that $\Gamma_u(x) = \{y \in X ; (x, y) \in \Gamma_u\}$ and similarly for $R_u(x)$). Consequently $\mathcal{T}_u$ is a projection of a Borel set and hence analytic; it follows that it is universally measurable, and in particular, $\mathfrak{m}$-measurable [39].

Following [13, Definition 7.7], a maximal chain R in $(X, \mathsf{d}, \leq_u)$ is called a *transport ray* if it is isometric to a closed interval I in $(\mathbb{R}, |\cdot|)$ of positive (possibly infinite) length.

Definition 2.9. ($\mathsf{CD}_u^1(K, N)$ when $\mathrm{spt}(\mathfrak{m}) = X$) Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space such that $\mathrm{spt}(\mathfrak{m}) = X$ and $\mathfrak{m}(X) = 1$. Let us consider $K, N \in \mathbb{R}, N > 1$ and let $u : (X, \mathsf{d}) \to \mathbb{R}$ be a 1-Lipschitz function. We say that $(X, \mathsf{d}, \mathfrak{m})$ satisfies the CD_u^1 condition if there exists a family $\{X_\alpha\}_{\alpha \in Q} \subset X$ such that:

(1) There exists a disintegration of $\mathfrak{m} \llcorner_{\mathcal{T}_u}$ on $\{X_\alpha\}_{\alpha \in Q}$:

$$\mathfrak{m} \llcorner_{\mathcal{T}_u} = \int_Q \mathfrak{m}_\alpha \, \mathfrak{q}(d\alpha), \quad \text{where } \mathfrak{m}_\alpha(X_\alpha) = 1, \text{ for } \mathfrak{q}\text{-a.e. } \alpha \in Q.$$

(2) For $\mathfrak{q}$-a.e. $\alpha \in Q$, X_α is a *transport ray* for Γ_u.

8

(3) For $\mathfrak{q}$-a.e. $\alpha \in Q$, $\mathfrak{m}_\alpha$ is supported on X_α.

(4) For $\mathfrak{q}$-a.e. $\alpha \in Q$, the metric measure space $(X_\alpha, \mathsf{d}, \mathfrak{m}_\alpha)$ satisfies $\mathsf{CD}(K, N)$.

Remark 2.10 (The assumption $\mathfrak{m}(X) = 1$). For an overview (and a self-contained proof) of the Disintegration Theorem we refer to [8, 23] (see also [13]). It is worth mentioning here that the assumption $\mathfrak{m}(X) = 1$ is most probably purely technical. In the framework of general Radon measure, the Disintegration Theorem does not furnish a unique family of conditional measures and one has to consider an additional normalization function; for additional details we refer to [15] where a localization of synthetic lower Ricci curvature bounds has been obtained also for general Radon measure.

Let us recall that it is well known that the last condition of Definition 2.9 is equivalent to asking $\mathfrak{m}_\alpha \sim h_\alpha \mathcal{L}^1 {\llcorner_{[0,|X_\alpha|]}}$ where $|X_\alpha| = \ell(X_\alpha)$ denotes the length of the transport ray X_α and $\sim$ means up to isometry of the space, and the density h_α has to satisfy the power-concavity inequality

$$\left(h_\alpha^{1/(N-1)}\right)'' + \frac{K}{N-1} h_\alpha^{1/(N-1)} \leq 0,$$

in the distributional sense.

Finally, we will say that the metric measure space $(X, \mathsf{d}, \mathfrak{m})$ satisfies $\mathsf{CD}_{Lip}^1(K, N)$ if $(\mathrm{spt}(\mathfrak{m}), \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_u^1(K, N)$ for all 1-Lipschitz functions $u : (\mathrm{spt}(\mathfrak{m}), \mathsf{d}) \to \mathbb{R}$, and satisfies $\mathsf{CD}^1(K, N)$ if $(\mathrm{spt}(\mathfrak{m}), \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_u^1(K, N)$ whenever u is a signed distance function defined as follows: given a continuous function $f : (X, \mathsf{d}) \to \mathbb{R}$ such that $\{f = 0\} \neq \emptyset$, the function

$$d_f : X \to \mathbb{R}, \qquad d_f(x) := \mathrm{dist}(x, \{f = 0\})\mathrm{sgn}(f), \tag{2.7}$$

is called the signed distance function (from the zero-level set of f). Notice that d_f is 1-Lipschitz on $\{f \geq 0\}$ and $\{f \leq 0\}$. If (X, d) is a length space, then d_f is 1-Lipschitz on the entire X.

2.3 Derivatives

In order to carry out a third order analysis of Kantorovich potentials, we will frequently use incremental ratios over arbitrary subsets of $\mathbb{R}$. We will use the following notation: for a function $g : A \to \mathbb{R}$ on a subset $A \subset \mathbb{R}$, denote its upper and lower derivatives at a point $t_0 \in A$ which is an accumulation point of A by:

$$\frac{\overline{d}}{dt}g(t_0) = \limsup_{A \ni t \to t_0} \frac{g(t) - g(t_0)}{t - t_0} \ , \quad \underline{\frac{d}{dt}}g(t_0) = \liminf_{A \ni t \to t_0} \frac{g(t) - g(t_0)}{t - t_0}.$$

We will say that g is differentiable at t_0 iff $\frac{d}{dt}g(t_0) := \frac{\overline{d}}{dt}g(t_0) = \underline{\frac{d}{dt}}g(t_0) \in \mathbb{R}$. This is a slightly more general definition of differentiability than the traditional one which requires that t_0 is an interior point of A.

Remark 2.11. Note that there are only a countable number of isolated points in A, so a.e. point in A is an accumulation point. In addition, it is clear that if $t_0 \in B \subset A$ is an accumulation point of B and g is differentiable at t_0, then $g|_B$ is also differentiable at t_0 with the same derivative. In particular, if g is a.e. differentiable on A then $g|_B$ is also a.e. differentiable on B and the derivatives coincide.

Remark 2.12. Denote by $A_1 \subset A$ the subset of density one points of A (which are in particular accumulation points of A). By Lebesgue's Density Theorem $\mathcal{L}^1(A \setminus A_1) = 0$, where we denote by $\mathcal{L}^1$

the Lebesgue measure on $\mathbb{R}$ throughout this work. If $g : A \to \mathbb{R}$ is Lipschitz, consider any Lipschitz extension $\hat{g} : \mathbb{R} \to \mathbb{R}$ of g. Then it is easy to check that for $t_0 \in A_1$, g is differentiable in the above sense at t_0 if and only if $\hat{g}$ is differentiable at t_0 in the usual sense, in which case the derivatives coincide. In particular, as $\hat{g}$ is a.e. differentiable on $\mathbb{R}$, it follows that g is a.e. differentiable on A_1 and hence on A, and it holds that $\frac{d}{dt}g = \frac{d}{dt}\hat{g}$ a.e. on A.

If $f : I \to \mathbb{R}$ is a convex function on an open interval $I \subset \mathbb{R}$, it is a well-known fact that the left and right derivatives $f'^{,-}$ and $f'^{,+}$ exist at every point in I and that f is locally Lipschitz. In particular, f is differentiable at a given point if and only if the left and right derivatives coincide there. Denoting by $D \subset I$ the differentiability points of f in I, it is also well-known that $I \setminus D$ is at most countable. Consequently, any point in D is an accumulation point, and we may consider the differentiability in D of $f' : D \to \mathbb{R}$ as defined above.

We will recall the following classical one-dimensional result about twice differentiability a.e. of convex functions on $\mathbb{R}^n$. The result extends to locally semi-convex and semi-concave functions as well; recall that a function $f : I \to \mathbb{R}$ is called semi-convex (semi-concave) if there exists $C \in \mathbb{R}$ so that $I \ni x \mapsto f(x) + Cx^2$ is convex (concave).

Lemma 2.13 (Second Order Differentiability of Convex Function). *Let $f : I \to \mathbb{R}$ be a convex function on an open interval $I \subset \mathbb{R}$, and let $\tau_0 \in I$ and $\Delta \in \mathbb{R}$. Then the following statements are equivalent:*

1. *f is differentiable at τ_0, and if $D \subset I$ denotes the subset of differentiability points of f in I, then $f' : D \to \mathbb{R}$ is differentiable at τ_0 with:*

$$(f')'(\tau_0) := \lim_{D \ni \tau \to \tau_0} \frac{f'(\tau) - f'(\tau_0)}{\tau - \tau_0} = \Delta.$$

2. *The right derivative $f'^{,+} : I \to \mathbb{R}$ is differentiable at τ_0 with $(f'^{,+})'(\tau_0) = \Delta$.*

3. *The left derivative $f'^{,-} : I \to \mathbb{R}$ is differentiable at τ_0 with $(f'^{,-})'(\tau_0) = \Delta$.*

4. *f is differentiable at τ_0 and has the following second order expansion there:*

$$f(\tau_0 + \varepsilon) = f(\tau_0) + f'(\tau_0)\varepsilon + \Delta\frac{\varepsilon^2}{2} + o(\varepsilon^2) \ \text{as} \ \varepsilon \to 0.$$

In this case, f is said to have a second Peano derivative at τ_0.

For a locally semi-convex or semi-concave function f, we will say that f is twice differentiable at τ_0 if any (all) of the above equivalent conditions hold for some $\Delta \in \mathbb{R}$, and write $(\frac{d}{d\tau})^2|_{\tau=\tau_0} f(\tau) = \Delta$.

Finally, we will recall the following slightly different version of the second order differential.

Definition 2.14 (Upper and lower second Peano derivatives). Given an open interval $I \subset \mathbb{R}$ and a function $f : I \to \mathbb{R}$ which is differentiable at $\tau_0 \in I$, we define its upper and lower second Peano derivatives at τ_0, denoted $\overline{\mathcal{P}}_2 f(\tau_0)$ and $\underline{\mathcal{P}}_2 f(\tau_0)$ respectively, by:

$$\overline{\mathcal{P}}_2 f(\tau_0) := \limsup_{\varepsilon \to 0} \frac{h(\varepsilon)}{\varepsilon^2} \geq \liminf_{\varepsilon \to 0} \frac{h(\varepsilon)}{\varepsilon^2} =: \underline{\mathcal{P}}_2 f(\tau_0), \tag{2.8}$$

where:

$$h(\varepsilon) := 2(f(\tau_0 + \varepsilon) - f(\tau_0) - \varepsilon f'(\tau_0)). \tag{2.9}$$

We say that f has a second Peano derivative at τ_0 iff $\overline{\mathcal{P}}_2 f(\tau_0) = \underline{\mathcal{P}}_2 f(\tau_0) \in \mathbb{R}$.

Lemma 2.15. *Given an open interval $I \subset \mathbb{R}$ and a locally absolutely continuous function $f : I \to \mathbb{R}$ which is differentiable at $\tau_0 \in I$, we have:*

$$\frac{\underline{d}}{dt} f'(\tau_0) \leq \underline{\mathcal{P}}_2 f(\tau_0) \leq \overline{\mathcal{P}}_2 f(\tau_0) \leq \frac{\overline{d}}{dt} f'(\tau_0).$$

2.4 Notation

Given a subset $D \subset X \times \mathbb{R}$, we denote its sections by:

$$D(t) := \{x \in X \ ; \ (x,t) \in D\} \ , \ D(x) := \{t \in \mathbb{R} \ ; \ (x,t) \in D\} \,.$$

Given a subset $G \subset \text{Geo}(X)$, we denote by $\mathring{G} := \left\{\gamma|_{(0,1)} \ ; \ \gamma \in G\right\}$ the corresponding open-ended geodesics on $(0,1)$. For a subset of (closed or open) geodesics $\tilde{G}$, we denote:

$$Im(\tilde{G}) := \left\{(x,t) \in X \times \mathbb{R} \ ; \ \exists \gamma \in \tilde{G} \ , \ t \in \text{Dom}(\gamma) \ , \ x = \gamma_t\right\}. \tag{2.10}$$

3 Hopf-Lax transform with exponent p

In this section we review the basic properties of the Hopf-Lax transform in a metric measure space setting with a general exponent $p > 1$. Some of following properties are well-known for the case $p = 2$, hence we omit the proofs for general p whenever they follow the same line of reasoning as the corresponding proofs for $p = 2$. The main references for most of the definitions and proofs will be [5, 6, 13]; further developments related to ours may also be found in [2] [25] [24] [7] and their references.

As motivation for the needed properties of the metric measure space Hopf-Lax transform we remind the reader of the relationship between the Hopf-Lax transform and the Eulerian view of optimal transport. We also provide a comparison between the results found in this paper to familiar results from Euclidean spaces.

We illustrate the main relationship for the case $(\mathbb{R}^n, d)$ with d as the Euclidean distance, and the cost function $c(x,y) = \frac{d(x,y)^p}{p}$ where $p > 1$. Recall that in the Eulerian view of optimal transport, the Wasserstein distance can be interpreted as the minimizing energy to the problem

$$\begin{cases} \rho_t + \nabla \cdot (\rho v) = 0 & \text{in } \mathbb{R}^n \times (0,1) \\ \rho(\cdot, 0) = \rho_0 & \text{in } \mathbb{R}^n \\ \rho(\cdot, 1) = \rho_1 & \text{in } \mathbb{R}^n \end{cases} \tag{3.1}$$

where ρ, v are the distribution of mass and the velocity at position x at time t respectively [4, Theorem 8.3.1]. By choosing $v = DH(\nabla\varphi)$, where in our case $H(w) = |w|^{p'}/p'$, and φ is a solution to the Hamilton-Jacobi equation

$$\begin{cases} \partial_t \varphi + H(\nabla\varphi) = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ \varphi(x, 0) = \varphi_0(x) & \text{for } x \in \mathbb{R}^n, \end{cases} \tag{3.2}$$

where φ_0 is a Kantorovich potential for the optimal transport problem and p' is the real number satisfying $\frac{1}{p} + \frac{1}{p'} = 1$. That is, p' is the Hölder dual of p. The method of characteristics gives a solution

to the Hamilton-Jacobi equation for a convex Hamiltonian H [19]. Furthermore, this solution can be expressed by the Hopf-Lax formula

$$\varphi(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ \varphi_0(y) + tL\left(\frac{x-y}{t}\right) \right\},$$

where the Lagrangian L is defined by

$$L(z) = \inf_{w \in \mathbb{R}^n} \left\{ z \cdot w - H(w) \right\}.$$

In our case, the Lagrangian is explicitly computed as $L(v) = \frac{|v|^p}{p}$, hence

$$\varphi(x,t) = \inf_{y \in \mathbb{R}^n} \left\{ \varphi_0(y) + \frac{|x-y|^p}{pt^{p-1}} \right\}. \tag{3.3}$$

Finally, in the context of smooth manifolds, we can compute the spatial gradient as

$$\nabla \varphi(x) = \frac{|x-y|^{p-2}(x-y)}{t^{p-1}},$$

where y is chosen to be a minimizer in the Hopf-Lax infimum (3.3). Hence,

$$\frac{|\nabla \varphi(x,t)|^{p'}}{p'} = \frac{(p-1)|x-y|^p}{pt^p}. \tag{3.4}$$

Note that, due to (3.4), (3.2) and (3.3) can be compared to conclusion 3 of Theorem 3.3 and Corollary 3.11 respectively. In particular, the expression in (3.4) depends on x only through its distance to the minimizing value y. This should be compared to Definition 3.2. With the above in mind, we now present the details of the nonsmooth case.

In the following sections, we will only consider the cost function $c = \mathsf{d}^p/p$ on $X \times X$.

Definition (*c*-Concavity, Kantorovich Potential)**.** The *c*-transform of a function $\psi : X \to \mathbb{R} \cup \{\pm\infty\}$ is defined as the following (upper semi-continuous) function:

$$\psi^c(x) = \inf_{y \in X} \frac{\mathsf{d}(x,y)^p}{p} - \psi(y).$$

A function $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$ is called *c*-concave if $\varphi = \psi^c$ for some ψ as above. It is well known that φ is *c*-concave iff $(\varphi^c)^c = \varphi$. A *c*-concave function $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ which is not identically equal to $-\infty$ is also known as a Kantorovich (or *p*-Kantorovich) potential, and this is how we will refer to such functions in this work. In that case, $\varphi^c : X \to \mathbb{R} \cup \{-\infty\}$ is also a Kantorovich potential, called the dual or conjugate potential.

In these sections, we only assume that (X, d) is a **proper geodesic metric space**. (Here proper refers to the requirement that closed balls are compact).

3.1 General definitions

Definition 3.1 (Hopf-Lax transform)**.** Let $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be not identically $+\infty$ and $t > 0$, $p > 1$. The Hopf-Lax transform $Q_t f : X \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$Q_t f(x) := \inf_{y \in X} \frac{\mathsf{d}(x, y)^p}{p t^{p-1}} + f(y). \tag{3.5}$$

If $Q_t f(\bar{x}) \in \mathbb{R}$ for some $\bar{x} \in X$ and $t > 0$, then $Q_s f(x) \in \mathbb{R}$ for all $x \in X$ and $0 < s \leq t$. Hence defining

$$t_*(f) := \sup\{t > 0 : Q_t f \not\equiv -\infty\},$$

where we set $t_*(f) = 0$ if the supremum is over an empty set, it holds that $Q_t f(x) \in \mathbb{R}$ for every $x \in X, t \in (0, t_*(f))$. Moreover, we set $Q_0 f := f$. The definition of $Q_t f$ can be extended to negative times $t < 0$ by setting

$$Q_t f(x) = -Q_{-t}(-f)(x) = \sup_{y \in X} -\frac{\mathsf{d}(x, y)^p}{p(-t)^{p-1}} + f(y), \quad t < 0. \tag{3.6}$$

If (X, d) is a length space (and in particular, if it is geodesic), the Hopf-Lax transform is in fact a semi-group on $[0, \infty)$:

$$Q_{s+t} f = Q_s \circ Q_t f \quad \forall t, s \geq 0.$$

Being the infimum of continuous functions in (t, x), the map $(0, \infty) \times X \ni (t, x) \mapsto Q_t f(x)$ is upper semi-continuous. Moreover, by definition $[0, \infty) \ni t \mapsto Q_t f(x)$ is monotone non-increasing; hence, it is continuous from the left.

We define the *distance progressed* as the length of the geodesic segment in X along which information propagates from the initial values to (t, x); this geodesic plays the role of a characteristic curve. Since we are modeling optimal transport, shocks do not form before unit time has elapsed [42].

Definition 3.2. (Distance progressed $D_f^\pm$)**.** Given $f : X \to \mathbb{R} \cup \{+\infty\}$ not identically $+\infty$, we define

$$D_f^+(x, t) := \sup \limsup_{n \to +\infty} \mathsf{d}(x, y_n) \geq \inf \liminf_{n \to +\infty} \mathsf{d}(x, y_n) =: D_f^-(x, t)$$

where the supremum and the infimum are taken on the set of minimizing sequences $\{y_n\}_{n \in \mathbb{N}}$ in the definition of Hopf-Lax transform. Using a diagonal argument, it is possible to show that the supremum and infimum are attained, though they may differ in the presence of shocks.

For $p = 2$, the following properties were established in [5, Chapter 3]. For a proof adopted to a similar framework we refer to [13, Section 3.2].

Theorem 3.3 (Hopf-Lax solution to metric space Hamilton-Jacobi equations)**.** *For any metric space* (X, d) *the following properties hold:*

1. *Both functions $D_f^\pm(x, t)$ are locally finite on $X \times (0, t_*(f))$ and $(x, t) \mapsto Q_t f(x)$ is locally Lipschitz there.*

2. *The map $(x, t) \mapsto D_f^+(x, t)$ $\left((x, t) \mapsto D_f^-(x, t)\right)$ is upper (lower) semi-continuous on $X \times (0, t_*(f))$.*

13

3. For every $x \in X$,

$$\partial_t^\pm Q_t f(x) = -\frac{(p-1)D_f^\pm(x,t)^p}{pt^p}, \qquad \forall\, t \in (0, t_*(f)),$$

where ∂_t^- and ∂_t^+ denote the left and right partial derivatives respectively. In particular, the map $(0, t_*(f)) \ni t \mapsto Q_t f(x)$ is locally Lipschitz and locally semi-concave. Moreover, it is differentiable at $t \in (0, t_*(f))$ if and only if $D_f^+(x,t) = D_f^-(x,t)$.

Proof. For the readers' convenience we will only address *3*. The claim can be found [4, Remark 3.1.7] and the proof for $p = 2$ is given in [4, Theorem 3.1.4].

Fix $t_0 < t_1 \in (0, t_*(f))$. By Lemma 3.4, there exists $x_{t_1} \in \operatorname{argmin}\left\{\frac{\mathsf{d}(x,y)^p}{pt_1^{p-1}} + f(y)\right\}$, for which $\mathsf{d}(x, x_{t_1}) = D_f^+(x, t_1)$. In particular, it holds:

$$
\begin{aligned}
Q_{t_0} f(x) - Q_{t_1} f(x) &\leq \frac{\mathsf{d}(x, x_{t_1})^p}{pt_0^{p-1}} - \frac{\mathsf{d}(x, x_{t_1})^p}{pt_1^{p-1}} \\
&= \frac{D_f^+(x, t_1)^p}{p} \cdot \left(\frac{t_1^{p-1} - t_0^{p-1}}{t_0^{p-1} \cdot t_1^{p-1}}\right).
\end{aligned}
$$

Applying again Lemma 3.4, there exists $x_{t_0} \in \operatorname{Argmin}\left\{\frac{\mathsf{d}(x,y)^p}{pt_0^{p-1}} + f(y)\right\}$ for which $\mathsf{d}(x, x_{t_0}) = D_f^+(x, t_0)$. Arguing as before, we get:

$$
\begin{aligned}
Q_{t_0} f(x) - Q_{t_1} f(x) &\geq \frac{\mathsf{d}(x, x_{t_0})^p}{pt_0^{p-1}} - \frac{\mathsf{d}(x, x_{t_0})^p}{pt_1^{p-1}} \\
&= \frac{D_f^+(x, t_0)^p}{p} \cdot \left(\frac{t_1^{p-1} - t_0^{p-1}}{t_0^{p-1} \cdot t_1^{p-1}}\right).
\end{aligned}
$$

Dividing by $t_1 - t_0 > 0$, we obtain:

$$\frac{D_f^+(x, t_0)^p}{p} \cdot \left(\frac{t_1^{p-1} - t_0^{p-1}}{(t_1 - t_0) \cdot t_0^{p-1} \cdot t_1^{p-1}}\right) \leq \frac{Q_{t_0} f(x) - Q_{t_1} f(x)}{t_1 - t_0} \leq \frac{D_f^+(x, t_1)^p}{p} \cdot \left(\frac{t_1^{p-1} - t_0^{p-1}}{(t_1 - t_0) \cdot t_0^{p-1} \cdot t_1^{p-1}}\right)$$

Sending t_1 to t_0 from the right we obtain:

$$\partial_t^+ Q_t f(x) = -\frac{(p-1)D_f^+(x,t)^p}{pt^p}, \qquad \forall\, t \in (0, t_*(f)),$$

The same holds with the minus sign. $\qquad\square$

The next property will be used throughout the paper; we include a proof for the readers' convenience.

Lemma 3.4 (Hopf-Lax attainment). *Let X be a proper metric space, $f : X \to \mathbb{R}$ a lower semi-continuous function, and $t_*(f) > 0$. For fixed $x \in X$ and $t \in (0, t_*(f))$, there exist $y_t^\pm \in X$ so that*

$$Q_t f(x) = \frac{\mathsf{d}(x, y_t^\pm)^p}{pt^{p-1}} + f(y_t^\pm). \tag{3.7}$$

Moreover, the following holds: $\mathsf{d}(x, y_t^\pm) = D_f^\pm(x, t)$.

14

Proof. Let $\{y_t^{\pm,n}\}$ be a minimizing sequence such that

$$Q_t f(x) = \lim_{n\to\infty} \frac{\mathsf{d}(x, y_t^{\pm,n})^p}{pt^{p-1}} + f(y_t^{\pm,n}) \text{ and } D_f^{\pm}(x,t) = \lim_{n\to\infty} \mathsf{d}(x, y_t^{\pm,n})$$

By local finiteness of $D_f^{\pm}$, it follows that $D_f^{\pm}(x,t) < R$ for some $R < \infty$. The properness of the space X guarantees that the closed geodesic ball $B_R(x)$ is compact, hence $\{y_t^{\pm,n}\}$ admits a subsequence converging to $\{y_t^{\pm}\}$. Using the lower semi-continuity of f, we get:

$$Q_t f(x) = \inf_{y\in X} \frac{\mathsf{d}(x,y)^p}{pt^{p-1}} + f(y) = \min_{y\in B_R(x)} \frac{\mathsf{d}(x,y)^p}{pt^{p-1}} + f(y) = \frac{\mathsf{d}(x, y_t^{\pm})^p}{pt^{p-1}} + f(y_t^{\pm}).$$

Hence, the claim holds true. $\qquad\square$

Lemma 3.5 (Time monotonicity of distance progressed)**.** *Let X be a proper metric space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous function. Then, for every $x \in X$, both functions $(0, t^*(f)) \ni t \mapsto D_f^{\pm}(x,t)$ are monotone non-decreasing and coincide except where they have jump discontinuities.*

Proof. Since trivially $D_f^{-} \le D_f^{+}$, it is sufficient to prove that

$$D_f^{+}(x,s) \le D_f^{-}(x,t), \qquad 0 < s < t < t^*(f)$$

in order to conclude. By Lemma 3.4, there exist y_s^{+}, y_t^{-} such that

$$\frac{\mathsf{d}(x, y_s^{+})^p}{ps^{p-1}} + f(y_s^{+}) = Q_s(f)(x) \le \frac{\mathsf{d}(x, y_t^{-})^p}{ps^{p-1}} + f(y_t^{-}),$$

$$\frac{\mathsf{d}(x, y_t^{-})^p}{pt^{p-1}} + f(y_t^{-}) = Q_t(f)(x) \le \frac{\mathsf{d}(x, y_s^{+})^p}{pt^{p-1}} + f(y_s^{+}).$$

Summing the two, we get

$$\mathsf{d}(x, y_s^{+})^p \cdot \left(\frac{1}{s^{p-1}} - \frac{1}{t^{p-1}} \right) \le \mathsf{d}(x, y_t^{-})^p \cdot \left(\frac{1}{s^{p-1}} - \frac{1}{t^{p-1}} \right).$$

Since the Lemma 3.4 also guarantees that $\mathsf{d}(x, y_t^{-}) = D_f^{-}(x,t)$ and $\mathsf{d}(x, y_s^{+}) = D_f^{+}(x,s)$, the claim follows. $\qquad\square$

3.2 Intermediate-time Kantorovich potentials

Definition 3.6. (Interpolating Intermediate-Time Kantorovich Potentials). Given a Kantorovich potential $\varphi : X \to \mathbb{R}$, the interpolating p-Kantorovich potential at time $t \in [0,1]$, denoted by $\varphi_t : X \to \mathbb{R}$, is defined for all $t \in [0,1]$ by:

$$\varphi_t(x) := Q_{-t}(\varphi) = -Q_t(-\varphi). \tag{3.8}$$

Note that $\varphi_0 = \varphi$, $\varphi_1 = -\varphi^c$, and:

$$-\varphi_t(x) = \inf_{y\in X} \frac{\mathsf{d}^p(x,y)}{pt^{p-1}} - \varphi(y) \quad \forall t \in (0,1].$$

15

Applying the previous general properties of the Hopf-Lax semi-group we directly obtain that

1. $(x, t) \mapsto \varphi_t(x)$ is lower semi-continuous on $X \times (0, 1]$ and continuous on $X \times (0, 1)$.

2. For every $x \in X$, $[0, 1] \ni t \mapsto \varphi_t(x)$ is monotone non-decreasing and continuous on $(0, 1]$.

We also recall the following terminology: given a Kantorovich potential $\varphi : X \to \mathbb{R}$, $\gamma \in \mathrm{Geo}(X)$ is called a (φ, p)-*Kantorovich geodesic* if

$$\varphi(\gamma_0) + \varphi^c(\gamma_1) = \frac{\mathsf{d}(\gamma_0, \gamma_1)^p}{p} = \frac{\ell(\gamma)^p}{p}. \tag{3.9}$$

The set of all Kantorovich geodesics will be denoted with G_φ; the upper semi-continuity of φ and φ^c implies that G_φ is a closed subset of $\mathrm{Geo}(X)$. Using the modified triangular inequality

$$\mathsf{d}(x, y)^p \leq \frac{\mathsf{d}(x, z)^p}{t^{p-1}} + \frac{\mathsf{d}(z, y)^p}{(1 - t)^{p-1}}, \tag{3.10}$$

valid for every choice of $x, y, z \in X$, we may conclude that along (φ, p)-Kantorovich geodesics, φ_t is affine in time, and it verifies the following nice expression:

$$\varphi_t(\gamma_t) = (1 - t)\frac{\mathsf{d}(\gamma_0, \gamma_1)^p}{p} - \varphi^c(\gamma_1). \tag{3.11}$$

This result easily implies the following corollary.

Corollary 3.7. *Let γ be a (φ, p)-Kantorovich geodesic. Then, for any $s, r \in (0, 1)$, we have:*

$$\varphi_s(\gamma_s) - \varphi_r(\gamma_r) = (r - s)\frac{\mathsf{d}(\gamma_0, \gamma_1)^p}{p}. \tag{3.12}$$

Lemma 3.8. *Let x, y, z be points in X and let $t \in (0, 1)$. If*

$$\frac{\mathsf{d}(x, y)^p}{pt^{p-1}} - \varphi(y) = \varphi^c(z) - \frac{\mathsf{d}(x, z)^p}{p(1 - t)^{p-1}}, \tag{3.13}$$

then x is a t-intermediate point between y and z with

$$\mathsf{d}(y, z) = \frac{\mathsf{d}(x, y)}{t} = \frac{\mathsf{d}(x, z)}{1 - t}. \tag{3.14}$$

Moreover there exists a (φ, p)-Kantorovich geodesic $\gamma : [0, 1] \to X$ with $\gamma_0 = y$, $\gamma_t = x$, $\gamma_1 = z$.

Proof. By definition of the c-transform, from the assumption (3.13) it follows that

$$\frac{\mathsf{d}(x, y)^p}{pt^{p-1}} + \frac{\mathsf{d}(x, z)^p}{p(1 - t)^{p-1}} = \varphi(y) + \varphi^c(z) \leq \frac{\mathsf{d}(y, z)^p}{p}.$$

Hence, the equality holds since the reverse inequality is trivially satisfied by (3.10). In particular, requiring the equality in the Hölder inequality implies that

$$\frac{\mathsf{d}(x, z)^p}{(1 - t)^p} = \mathsf{d}(y, z)^p = \frac{\mathsf{d}(x, y)^p}{t^p}. \tag{3.15}$$

16

So the concatenation $\gamma : [0, 1] \to X$ of any constant speed geodesic $\gamma^1 : [0, t] \to X$ between x and y with any constant speed geodesic $\gamma^2 : [t, 1] \to X$ between y and z so that $\gamma_0 = y$, $\gamma_t = x$, $\gamma_1 = z$ must be a constant speed geodesic itself by the triangle inequality. In particular also

$$\varphi(y) + \varphi^c(z) \leq \frac{\mathsf{d}(y, z)^p}{p}.$$

must hold as equality, impling γ to be a (φ, p)-Kantorovich geodesic. $\qquad \square$

In what follows, forward and backward evolution via the Hopf-Lax semi-group will permit us to obtain regularity properties and key estimates on the intermediate-time Kantorovich potential. However, it is immediate to show by inspecting the definitions that we always have

$$Q_{-s} \circ Q_s f \leq f \text{ on } X \quad \forall s > 0;$$

note that for $f = -\varphi$ where φ is a Kantorovich potential, we do have equality for $s = 1$, and in fact for all $s \in [0, 1]$; for $f = Q_t(-\varphi)$, $t \in (0, 1)$ and $s = 1 - t$, we can only assert an *inequality*

$$(\varphi^c)_{1-t} = Q_{-(1-t)} \circ Q_1(-\varphi) \leq Q_t(-\varphi) = -\varphi_t \text{ on } X, \tag{3.16}$$

and equality need not hold at every point of X.

Definition (Time-Reversed Interpolating Potential)**.** Given a Kantorovich potential $\varphi : X \to \mathbb{R}$, define the time-reversed interpolating Kantorovich potential at time $t \in [0, 1]$, $\bar{\varphi}_t : X \to \mathbb{R}$, as:

$$\bar{\varphi}_t := -(\varphi^c)_{1-t} = Q_{1-t}(-\varphi^c) = -Q_{-(1-t)} \circ Q_{1-t}(-\varphi_t).$$

Note that $\bar{\varphi}_0 = \varphi$, $\bar{\varphi}_1 = -\varphi^c$, and:

$$\bar{\varphi}_t(x) = \inf_{y \in X} \frac{\mathsf{d}^p(x, y)}{p(1-t)^{p-1}} - \varphi^c(y) \quad \forall t \in [0, 1).$$

Note that, since any Kantorovich potential φ is upper semi-continuous, Lemma 3.4 applies to $f = -\varphi$.

Lemma 3.9 (Relating forward to reverse evolution of potentials)**.** *The following properties hold true:*

1. *$\varphi_0 = \bar{\varphi}_0 = \varphi$ and $\varphi_1 = \bar{\varphi}_1 = -\varphi^c$;*

2. *For all $t \in [0, 1]$, $\varphi_t \leq \bar{\varphi}_t$;*

3. *For any $t \in (0, 1)$, $\varphi_t(x) = \bar{\varphi}_t(x)$ if and only if $x \in \mathsf{e}_t(G_\varphi)$.*

Proof. Point *1.* is a trivial consequence of the definitions. Also *2.* is straightforward, since

$$\bar{\varphi}_t := Q_{1-t}(-\varphi^c) = -Q_{-(1-t)} \circ Q_{1-t}(-\varphi_t) \geq \varphi_t.$$

To demonstrate *3.*, let us consider a point $x = \gamma_t$ with $\gamma \in G_\varphi$ and use the following notation $\ell(\gamma) = \mathsf{d}(\gamma_0, \gamma_1)$ for length. Applying Corollary 3.7 to γ with $s = 0$ and $r = t$ we get

$$\varphi(\gamma_0) - \varphi_t(\gamma_t) = t \frac{\ell(\gamma)^p}{p},$$

17

while applying the same result to $\gamma^c \in G_{\varphi^c}$, the time reversed curve, with $s = 1$, and $r = (1 - t)$ we obtain

$$-\varphi(\gamma_0) - \varphi^c_{1-t}(\gamma_t) = (\varphi^c)_1(\gamma^c_1) - (\varphi^c)_{1-t}(\gamma^c_{1-t})$$
$$= -t\frac{\ell(\gamma^c)^p}{p} = -t\frac{\ell(\gamma)^p}{p}.$$

Summing the two identities, it follows that $\varphi_t(\gamma_t) = -(\varphi^c)_{1-t}(\gamma_t) = \bar{\varphi}_t(\gamma_t)$.

For the other implication, let us assume that for some $x \in X$, $t \in (0,1)$ $\varphi_t(x) = -(\varphi^c)_{1-t}(x)$. Applying Lemma 3.4 to the lower semi-continuous functions $-\varphi$ and $-\varphi^c$, it turns out that there exist $y_t, z_t \in X$ such that

$$-\varphi_t(x) = Q_t(-\varphi)(x) = \frac{\mathsf{d}(x, y_t)^p}{pt^{p-1}} - \varphi(y_t),$$

$$\varphi_t(x) = Q_{1-t}(-\varphi^c)(x) = \frac{\mathsf{d}(x, z_t)^p}{pt^{p-1}} - \varphi^c(z_t).$$

Summing the two equations, we get that

$$\frac{\mathsf{d}(x, y_t)^p}{pt^{p-1}} - \varphi(y_t) = \varphi^c(z_t) - \frac{\mathsf{d}(x, z_t)^p}{p(1-t)^{p-1}},$$

so we are in position to apply Lemma 3.8, obtaining the claim. $\qquad\square$

Motivated by Lemma 3.9 we will also consider the following set

$$D(\mathring{G}_\varphi) = \{(x, t) \in X \times (0,1) \; ; \; \varphi_t(x) = \bar{\varphi}_t(x)\}, \tag{3.17}$$

which is a closed subset of $X \times (0,1)$.

3.3 First and Second Order inequalities

Let us now introduce the speed along which each characteristic is traversed; since the particles move freely, this coincides with the total length of the characteristic, which is why the same functions are called length functions ℓ_t in [13]. To emphasize the dynamic point of view, we shall also refer to $(p-1)\ell_t^p/p = (\ell_t^{p-1})^{p'}/p'$ as the *energy*, though it is really the energy per unit mass transported.

Definition 3.10 (Speed functions $\ell_t^\pm, \bar{\ell}_t^\pm$). Given a Kantorovich potential $\varphi : X \to \mathbb{R}$, define the speed functions $\ell_t^\pm, \bar{\ell}_t^\pm$ as follows:

$$\ell_t^\pm(x) := \frac{D_{-\varphi}^\pm(x, t)}{t}, \quad \bar{\ell}_t^\pm(x) := \frac{D_{-\varphi^c}^\pm(x, 1-t)}{1-t}, \quad (x, t) \in X \times (0,1).$$

Let us mention that we will shortly see that if $x = \gamma_t$ with $\gamma \in G_\varphi$ and $t \in (0,1)$, then:

$$\ell_t^+(x) = \ell_t^-(x) = \bar{\ell}_t^+(x) = \bar{\ell}_t^-(x) = \ell(\gamma).$$

In particular, all (φ, p)-Kantorovich geodesics having x as their t-mid-point have necessarily the same length. For $\tilde{\ell} \in \{\ell, \bar{\ell}\}$, we define the set:

$$D_{\tilde{\ell}} := \{(x, t) \in X \times (0,1) : \tilde{\ell}_t^+(x) = \tilde{\ell}_t^-(x)\}. \tag{3.18}$$

On $D_{\tilde{\ell}}$ we set $\tilde{\ell}_t(x) := \tilde{\ell}_t^-(x) = \tilde{\ell}_t^+(x)$. Recalling that $\varphi_t = -Q_t(-\varphi)$ and $\bar{\varphi}_t = Q_{1-t}(-\varphi^c)$, we can apply Theorem 3.3 to deduce the following:

18

Corollary 3.11 (Time semi-continuity of speeds). *Let $\varphi : X \to \mathbb{R}$ denote a Kantorovich potential. Then:*

1. *Choosing $\tilde{\ell} \in \{\ell, \bar{\ell}\}$ and $\tilde{\varphi} \in \{\varphi, \bar{\varphi}\}$ correspondingly, $\tilde{\ell}_t^{\pm}(x)$ are locally finite on $X \times (0,1)$, and $(x,t) \mapsto \tilde{\varphi}_t(x)$ is locally Lipschitz there.*

2. *For $\tilde{\ell} \in \{\ell, \bar{\ell}\}$ the map $(x,t) \mapsto \tilde{\ell}_t^+(x)$ $((x,t) \mapsto \tilde{\ell}_t^-(x))$ is upper (lower) semi-continuous on $X \times (0,1)$. In particular, $D_{\tilde{\ell}} \subset X \times (0,1)$ is Borel and $(x,t) \mapsto \tilde{\ell}_t(x)$ is continuous on $D_{\tilde{\ell}}$.*

3. *For every $x \in X$ we have:*

$$\partial_t^{\pm} \varphi_t(x) = \frac{(p-1)\ell_t^{\pm}(x)^p}{p}, \quad \partial_t^{\pm} \bar{\varphi}_t(x) = \frac{(p-1)\bar{\ell}_t^{\pm}(x)^p}{p} \quad \forall t \in (0,1).$$

   In particular, for $\tilde{\ell} \in \{\ell, \bar{\ell}\}$ and the corresponding $\tilde{\varphi} \in \{\varphi, \bar{\varphi}\}$, the map $(0,1) \ni t \mapsto \tilde{\varphi}_t(x)$ is locally Lipschitz, and it is differentiable at $t \in (0,1)$ iff $t \in D_{\tilde{\ell}}(x)$, the set on which both maps $(0,1) \ni t \mapsto \tilde{\ell}_t^{\pm}(x)$ coincide. $D_{\tilde{\ell}}(x)$ is precisely the set of continuity points of both maps, and thus coincides with $(0,1)$ with at most countably exceptions.

All four maps $(0,1) \ni t \mapsto t\ell_t^{\pm}(x)$ and $(0,1) \ni t \mapsto (t-1)\bar{\ell}_t^{\pm}(x)$ are monotone non-decreasing; in particular, both $D_\ell(x) \ni t \mapsto \ell_t^p(x)$ and $D_{\bar{\ell}}(x) \ni t \mapsto \bar{\ell}_t^p(x)$ are differentiable a.e.. From monotonicity it is straightforward to deduce

$$\underline{\partial}_t \ell_t(x) \geq -\frac{1}{t}\ell_t(x) \quad \forall t \in D_\ell(x),$$

as well as a similar estimate for $\bar{\ell}_t$. In particular, the following estimates holds (see [13, Corollary 3.10]).

Corollary 3.12 (Energies are locally Lipschitz in time). *The following estimates hold for every $x \in X$:*

$$\underline{\partial}_t \frac{\ell_t^p(x)}{p} \geq -\frac{1}{t}\ell_t^p(x), \quad \forall t \in D_\ell(x). \tag{3.19}$$

$$\bar{\partial}_t \frac{\bar{\ell}_t^p(x)}{p} \leq \frac{1}{1-t}\bar{\ell}_t^p(x), \quad \forall t \in D_{\bar{\ell}}(x). \tag{3.20}$$

The first and the last points of the next Theorem can be compared with [13, Theorem 2.13] in the case $p = 2$.

Theorem 3.13 (Time-derivatives of energies bound second time-derivatives of potentials). *Let $\varphi : X \to \mathbb{R}$ be a Kantorovich potential. Then the following holds true:*

1. *For all $x \in e_t(G_\varphi)$ with $t \in (0,1)$, we have:*

$$\ell_t^+(x) = \ell_t^-(x) = \bar{\ell}_t^+(x) = \bar{\ell}_t^-(x) = \ell(\gamma).$$

2. *For all $x \in X$, $\mathring{G}_\varphi(x) \ni t \mapsto \ell_t(x) = \bar{\ell}_t(x)$ is locally Lipschitz and, provided $\ell(\gamma) > 0$, the following estimate holds true*

$$\frac{1-s}{1-t} \leq \frac{\ell_t(x)}{\ell_s(x)} \leq \frac{s}{t}, \quad 0 < t \leq s < 1.$$

3. For all $(x,t) \in D(\mathring{G}_\varphi) \subset D_\ell \cap D_{\bar\ell}$ we have that the following estimate holds true for the upper and lower second derivatives $z \in \{\underline{\mathcal{P}}_2\bar\varphi_t(x), \overline{\mathcal{P}}_2\varphi_t(x)\}$ in time of (2.8):

$$-\frac{p-1}{t}\ell_t^p(x) \le \underline{\partial}_t \frac{(p-1)\ell_t^p(x)}{p} \le \underline{\mathcal{P}}_2\varphi_t(x) \le z \le \overline{\mathcal{P}}_2\bar\varphi_t(x) \le \bar\partial_t \frac{(p-1)\bar\ell_t^p(x)}{p} \le \frac{p-1}{1-t}\ell_t^p(x).$$

Proof. Let $(x,t) \in D(\mathring{G}_\varphi)$ (recall (3.17)). An application of Lemma 3.4 implies that there exist $y^\pm, z^\pm \in X$ such that

$$-\varphi_t(x) = \frac{\mathsf{d}(x,y_\pm)^p}{pt^{p-1}} - \varphi(y_\pm),$$

$$-\bar\varphi_t(x) = -\frac{\mathsf{d}(x,z_\pm)^p}{pt^{p-1}} + \varphi^c(z_\pm).$$

Since $\varphi_t(x) = \bar\varphi_t(x)$ by Lemma 3.9, we can equate the two expressions, obtaining that the assumption (3.13) in the Lemma 3.8 is satisfied. Hence x is the t-midpoint of a geodesic connecting $y_\pm$ and $z_\pm$ for all four possibilities. The same lemma guarantees that

$$\frac{\mathsf{d}(x,y^\pm)}{t} = \frac{\mathsf{d}(x,z^\pm)}{1-t}$$

and thus $\ell_t^\pm(x) = \bar\ell_t^\pm(x)$. Recall now that if $x = \gamma_t$ for some $\gamma \in G_\varphi$ then Corollary 3.7 implies that

$$Q_t(-\varphi)(x) = -\varphi_t(x) = \frac{\mathsf{d}(x,\gamma_0)^p}{pt^{p-1}} - \varphi(\gamma_0)$$

and thus the sequence $\{y_n\}$ with $y_n \equiv \gamma_0$ is in the class of admissible sequences for the infimum and supremum in the definition of $D_{-\varphi}^\pm(x,t)$. Hence

$$t\ell_t^-(x) = D_{-\varphi}^-(x,t) \le \mathsf{d}(x,\gamma_0) = t\ell(\gamma) \le D_{-\varphi}^+(x,t) = t\ell_t^+(x),$$

and *1.* follows.

In order to prove *2.*, we use that, by the discussion following Corollary 3.13, for $x \in X$ the maps $t \mapsto t\ell_t^\pm(x)$ are monotone non-decreasing and the maps $t \mapsto (1-t)\bar\ell_t^\pm$ are monotone non-increasing combined with the previous conclusion of this theorem to obtain that for $x \in X$ and $t, s \in \mathring{G}_\varphi(x)$ with $t < s$:

$$t\ell_t(x) \le s\ell_s(x), \quad (1-s)\ell_s(x) \le (1-t)\ell_t(x).$$

For $\gamma \in G_\varphi$ with $\ell(\gamma) > 0$ we conclude the desired statement by rearranging. This allows us to conclude that $\ell_\cdot(x)$ is locally Lipschitz.

To obtain *3.*, as in (2.9) let us define $\tilde h = h, \bar h$ as

$$\tilde h(\varepsilon) := 2(\tilde\varphi_{t_0+\varepsilon}(x) - \tilde\varphi_{t_0}(x) - \varepsilon\partial_t\tilde\varphi_{t_0}(x)).$$

Recall that, by Lemma 3.9, for all $t \in [0,1]$ it holds $\varphi_t \le \bar\varphi_t$ with the equality satisfied in the case $x \in \mathsf{e}_t(G_\varphi)$. Moreover, since $\mathring{G}_\varphi(x) \subset D_\ell(x) \cap D_{\bar\ell}(x)$, the maps $t \mapsto \tilde\varphi_t(x)$ are differentiable at $t_0 \in \mathring{G}_\varphi(x)$ and $(p-1)\ell_{t_0}^p(x)/p = \partial_t|_{t=t_0}\varphi_t(x) = \partial_t|_{t=t_0}\bar\varphi_t(x) = (p-1)\bar\ell_{t_0}^p(x)/p$. These facts imply that $h \le \tilde h$ on $(-t_0, 1-t_0)$. Dividing by ε^2 and taking subsequential limits, we obtain

$$\underline{\mathcal{P}}_2\varphi_t(x) \le \underline{\mathcal{P}}_2\bar\varphi_t(x), \quad \overline{\mathcal{P}}_2\varphi_t(x) \le \overline{\mathcal{P}}_2\bar\varphi_t(x).$$

Combining these inequalities with those of Lemma 2.15, (3.19) and (3.20) we get the claim. $\qquad\square$

We conclude with the following result; for its proof we refer to [13, Corollary 3.13].

Corollary 3.14. *For all $x \in X$, for a.e. $t \in \mathring{G}_\varphi(x)$, $\partial_t \ell_t^p(x)$ and $\partial_t \overline{\ell}_t^p(x)$ exist, coincide, and satisfy:*

$$
-\frac{\ell_t^p(x)}{t} \le \partial_t \frac{\ell_t^p(x)}{p} = \partial_t \frac{\ell_t^p(x) \mid_{\mathring{G}_\varphi(x)}}{p}
$$
$$
= \partial_t \frac{\overline{\ell}_t^p(x) \mid_{\mathring{G}_\varphi(x)}}{p} = \partial_t \frac{\overline{\ell}_t^p(x)}{p} \le \frac{\overline{\ell}_t^p(x)}{1-t}. \tag{3.21}
$$

Remark 3.15. Recall that we already proved that $\partial_\tau^\pm|_{\tau=s} \varphi_\tau(x) = (p-1)\ell_s^\pm(x)^p/p$ and $\ell_s^\pm(\gamma_s) = \ell$ for all $s \in (0,1)$.

3.4 Third order inequality

Just as the solution to a Hamilton-Jacobi equation with Hamiltonian $H(w) = |w|^{p'}/p'$ behaves affinely in time on its characteristics, (3.11) similarly shows that the t interpolant φ_t of a Kantorovich potential becomes an affine function of time t along a φ-Kantorovich geodesic γ_t. The goal of this and the next sections is to show that $\partial_t^2 \varphi_t$ is non-decreasing along such curves and provide a positive lower bound (3.31)–(3.32) for the slope of $z(t) := [\partial_t^2 \varphi_t](\gamma_t)$ — at least under certain regularity hypotheses which can be subsequently verified for a large enough family of φ-Kantorovich geodesics that serve our purposes. For $p = p' = 2$, such estimates were discovered in [13], but their proof does not generalize to our case. However, Cavalletti and Milman [13] also provided a heuristic argument in the smooth setting which can be adapted to $p \ne 2$ as follows.

Start from the Hamilton-Jacobi equation

$$
\partial_t \varphi_t = H(\nabla \varphi_t)
$$

satisfied by the time t interpolant φ_t of a Kantorovich potential φ on a Riemannian manifold. Differentiating in t gives

$$
\partial_t^2 \varphi_t = DH|_{\nabla \varphi_t}(\nabla \partial_t \varphi_t). \tag{3.22}
$$

Setting $z(t) = \left[\partial_t^2 \varphi_t\right](\gamma_t)$ where γ_t is the time t evaluation of a φ-Kantorovich geodesic, we observe using $\gamma'(t) = -DH(\nabla \varphi_t)$ that

$$
z'(t) = \partial_t^3 \varphi_t(\gamma_t) - \left\langle \nabla \partial_t^2 \varphi_t(\gamma_t), DH(\nabla \varphi_t(\gamma(t))) \right\rangle.
$$

On the other hand

$$
\partial_t^3 \varphi_t = D^2 H|_{\nabla \varphi_t}(\nabla \partial_t \varphi_t, \nabla \partial_t \varphi_t) + DH|_{\nabla \varphi_t}(\nabla \partial_t^2 \varphi_t).
$$

Inserting this into the previous equation yields

$$
z'(t) = D^2 H|_{\nabla \varphi_t}(\nabla \partial_t \varphi_t, \nabla \partial_t \varphi_t)
$$
$$
= |\nabla \varphi_t(\gamma(t))|^{p'-2}|\nabla \partial_t \varphi_t(\gamma(t))|^2 + (p'-2)|\nabla \varphi_t(\gamma(t))|^{p'-4} \left\langle \nabla \varphi_t(\gamma(t)), \nabla \partial_t \varphi_t(\gamma(t)) \right\rangle^2.
$$

Convexity of $H(w) = |w|^{p'}/p'$ shows that $z(t)$ is non-decreasing (hence confirming differentiability a.e.) and allows its derivative to be estimated from below in terms of $|\nabla \varphi_t(\gamma(t))|$ and $|\nabla \partial_t \varphi_t(\gamma(t))|$

— both of which exist a.e. since φ_t is locally semiconvex in the halfspace $t > 0$. The Cauchy-Schwarz inequality gives

$$z'(t) \geq (p'-1)|\nabla\varphi_t(\gamma(t))|^{p'-4}\langle\nabla\varphi_t(\gamma(t)), \nabla\partial_t\varphi_t(\gamma(t))\rangle^2$$
$$= \frac{1}{p-1}\frac{z(t)^2}{\ell_t^p},$$

where $\ell_t = |DH(\nabla\varphi_t)|$ and $(p-1)(p'-1) = 1$, and (3.22) has been used to identify $z(t) = |\nabla\varphi_t(\gamma_t)|^{p'-2}\langle\nabla\varphi_t(\gamma_t), \nabla\partial_t\varphi_t(\gamma_t)\rangle$. At least heuristically, this establishes (3.32).

In order to obtain rigorous estimates on third order variations of Kantorovich potentials, we introduce the quantities $\tilde{r} \in \{r, \bar{r}\}$ which measure the time partial of energies along a fixed φ-Kantorovich geodesic (which plays the role of a characteristic in the nonsmooth setting); for every $s \in (0,1)$ set

$$\tilde{r}_+^\gamma(s) = \tilde{r}_+(s) := \overline{\partial}_\tau|_{\tau=s}\frac{(p-1)}{p}\tilde{\ell}_\tau^p(\gamma_s) = (p-1)\tilde{\ell}^{p-1}\overline{\partial}_\tau|_{\tau=s}\tilde{\ell}_\tau(\gamma_s),$$

$$\tilde{r}_-^\gamma(s) = \tilde{r}_-(s) := \underline{\partial}_\tau|_{\tau=s}\frac{(p-1)}{p}\tilde{\ell}_\tau^p(\gamma_s) = (p-1)\tilde{\ell}^{p-1}\underline{\partial}_\tau|_{\tau=s}\tilde{\ell}_\tau(\gamma_s).$$

By definition, $\tilde{r}_-(s) \leq \tilde{r}_+(s)$; moreover, equality holds $\tilde{r}_-(s) = \tilde{r}_+(s) = \tilde{r}$ if and only if the map $\tau \mapsto (p-1)\tilde{\ell}_\tau^p(\gamma_s)/p$ is differentiable at $\tau = s$ with derivative $\tilde{r}$.

We also define upper and lower second order Peano derivatives in time (Definition 2.14) $\tilde{q}_\pm \in \{q_\pm, \bar{q}_\pm\}$ of the (forward and backward) interpolated Kantorovich potentials respectively, evaluated along the same characteristic, as follows:

$$\tilde{q}_+(s) := \overline{\mathcal{P}}_2\tilde{\varphi}_s(x)|_{x=\gamma_s} = \limsup_{\varepsilon\to 0}\frac{\tilde{h}(s,\varepsilon)}{\varepsilon^2},$$

$$\tilde{q}_-(s) := \underline{\mathcal{P}}_2\tilde{\varphi}_s(x)|_{x=\gamma_s} = \liminf_{\varepsilon\to 0}\frac{\tilde{h}(s,\varepsilon)}{\varepsilon^2},$$

where $\tilde{h}(s,\varepsilon)$ is defined analogously to (2.9). By definition, $\tilde{q}_-(s) = \tilde{q}_+(s) = \tilde{q}$ hold if and only if the map $\tau \mapsto \tilde{\varphi}_\tau(\gamma_s)$ has second-order Peano derivative at $\tau = s$ given by $\tilde{q}$. We summarize the relation between $\tilde{q}_\pm$ and $\tilde{r}_\pm$ implied by Lemma 2.13 and Lemma 2.15 in the following :

Corollary 3.16 (First differentiability of energy is equivalent to second differentiability of potential). *The following statements are equivalent for a given $s \in (0,1)$:*

1. *$\tilde{r}_-(s) = \tilde{r}_+(s) = \tilde{r} \in \mathbb{R}$, i.e. the map $D_{\tilde{\ell}}(\gamma_s) \ni \tau \mapsto (p-1)\tilde{\ell}_\tau^p(\gamma_s)/p$ is differentiable at $\tau = s$ with derivative $\tilde{r}$.*

2. *$\tilde{q}_-(s) = \tilde{q}_+(s) = \tilde{q} \in \mathbb{R}$, i.e. the map $(0,1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_s)$ has second order Peano derivative at $\tau = s$ given by $\tilde{q}$.*

If one of the two conditions above is satisfied, the map $(0,1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_s)$ is twice differentiable at $\tau = s$, and we have :

$$\partial_\tau^2|_{\tau=s}\tilde{\varphi}_\tau(\gamma_s) = \partial_\tau|_{\tau=s}\frac{(p-1)\tilde{\ell}_\tau^p(\gamma_s)}{p} = (p-1)\tilde{\ell}^{p-1}\cdot\partial_\tau|_{\tau=s}\ell_\tau(\gamma_s) = \tilde{r} = \tilde{q}.$$

We are now in a position to obtain lower bounds on the incremental ratio of $\tilde{q}$. This provides the required third-order information concerning φ_t even when the upper and lower derivatives in question do not agree. For the geometric interpretation of the following discretized differential inequalities, we refer to the discussion of the case $p = 2$ in [13, Section 5.1].

Theorem 3.17 (Third-order difference quotient bounds on potential along its characteristics). *For all $0 < s < t < 1$ and both possibilities for $\pm$, we have*

$$\frac{q_+(t) - q_-(s)}{t - s} \geq \frac{s}{t} \frac{r_\pm(s)^2}{(p-1)\ell^p}, \tag{3.23}$$

$$\frac{\bar{q}_+(t) - \bar{q}_-(s)}{t - s} \geq \frac{1-t}{1-s} \frac{\bar{r}_\pm(t)^2}{(p-1)\ell^p}. \tag{3.24}$$

The proof of the analogous estimate for $p = 2$ ([13, Theorem 5.2]) does not work in the general case $p > 1$.

Proof. By definition of the Hopf-Lax transform and by Lemma 3.4, we have that

$$-\varphi_{s+\varepsilon}(\gamma_s) = Q_{s+\varepsilon}(-\varphi)(\gamma_s) = \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_s)^p}{p(s+\varepsilon)^{p-1}} - \varphi(y_\varepsilon^\pm),$$

with $\mathsf{d}(y_\varepsilon^\pm, \gamma_s) = D_{-\varphi}^\pm(\gamma_s, s+\varepsilon) = (s+\varepsilon)\ell_{s+\varepsilon}^\pm(\gamma_s) =: D_{s+\varepsilon}^\pm$. Moreover, the following inequality trivially holds:

$$-\varphi_{t+\varepsilon}(\gamma_t) \leq \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_t)^p}{p(t+\varepsilon)^{p-1}} - \varphi(y_\varepsilon^\pm).$$

Subtracting the two expressions above, we obtain:

$$\varphi_{t+\varepsilon}(\gamma_t) - \varphi_{s+\varepsilon}(\gamma_s) \geq -\frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_t)^p}{p(t+\varepsilon)^{p-1}} + \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_s)^p}{p(s+\varepsilon)^{p-1}},$$

hence recalling (2.9)

$$\frac{1}{2}(h(t,\varepsilon) - h(s,\varepsilon)) \geq -\varphi_t(\gamma_t) + \varphi_s(\gamma_s) - \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_t)^p}{p(t+\varepsilon)^{p-1}} + \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_s)^p}{p(s+\varepsilon)^{p-1}},$$

$$= (t-s)\frac{\ell^p}{p} - \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_t)^p}{p(t+\varepsilon)^{p-1}} + \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_s)^p}{p(s+\varepsilon)^{p-1}},$$

$$= (t-s)\frac{\ell^p}{p} - \frac{\mathsf{d}(y_\varepsilon^\pm, \gamma_t)^p}{p(t+\varepsilon)^{p-1}} + \frac{(s+\varepsilon)(\ell_{s+\varepsilon}^\pm(\gamma_s))^p}{p}. \tag{3.25}$$

We need now an estimate from below of the second term. In order to do that, let us observe that

$$\mathsf{d}(y_\varepsilon^\pm, \gamma_t) \leq \mathsf{d}(y_\varepsilon^\pm, \gamma_s) + \mathsf{d}(\gamma_s, \gamma_t) = D_{s+\varepsilon}^\pm + D_t - D_s,$$

where we put $D_r = r\ell = \mathsf{d}(\gamma_r, \gamma_0)$, for $r = s, t$. In particular,

$$D_{s+\varepsilon}^\pm + D_t - D_s = (s+\varepsilon)(\ell_{s+\varepsilon}^\pm(\gamma_s) - \ell_s(\gamma_s)) + (t+\varepsilon)\ell$$

$$= (t+\varepsilon)\left[\frac{s+\varepsilon}{t+\varepsilon}\ell_{s+\varepsilon}^\pm(\gamma_s) + \left(1 - \frac{s+\varepsilon}{t+\varepsilon}\right)\ell\right].$$

23

Thus, substituting this expression in (3.25), we get

$$\frac{h(t,\varepsilon) - h(s,\varepsilon)}{2\varepsilon^2} \geq \frac{t+\varepsilon}{p\varepsilon^2}\left[\frac{t-s}{t+\varepsilon}\ell^p + \frac{s+\varepsilon}{t+\varepsilon}(\ell^{\pm}_{s+\varepsilon}(\gamma_s))^p - \left(\frac{s+\varepsilon}{t+\varepsilon}\ell^{\pm}_{s+\varepsilon}(\gamma_s) + \frac{t-s}{t+\varepsilon}\ell\right)^p\right]. \tag{3.26}$$

In other words, denoting with $f(x) := x^p$ and defining for every $\lambda \in [0,1]$ the functions

$$s_{x,y}(\lambda) = \lambda f(x) + (1-\lambda)f(y), \qquad g_{x,y}(\lambda) = f(\lambda x + (1-\lambda)y),$$

we want to estimate from below the quantity $s_{x,y}(\lambda) - g_{x,y}(\lambda)$ for the following choices of λ, x, y:

$$\lambda = \frac{s+\varepsilon}{t+\varepsilon}, \qquad x = \ell^{\pm}_{s+\varepsilon}(\gamma_s), \qquad y = \ell = \ell_s(\gamma_s). \tag{3.27}$$

Appling the following inequality $s_{x,y}(\lambda) - g_{x,y}(\lambda) \geq \min_{[y,x]} f'' \cdot \frac{\lambda(1-\lambda)}{2}(x-y)^2$, for all $\lambda \in [0,1]$, with the choices of x, y, λ given by (3.27), we get

$$\frac{h(t,\varepsilon) - h(s,\varepsilon)}{2\varepsilon^2} \geq \frac{t+\varepsilon}{p\varepsilon^2}\left[\min_{z\in[\ell_s(\gamma_s),\ell^{\pm}_{s+\varepsilon}(\gamma_s)]} z^{p-2} \cdot \frac{p(p-1)}{2} \cdot \frac{t-s}{t+\varepsilon} \cdot \frac{s+\varepsilon}{t+\varepsilon} \cdot (\ell^{\pm}_{s+\varepsilon}(\gamma_s) - \ell_s(\gamma_s))^2\right]. \tag{3.28}$$

Taking appropriate subsequential limits as $\varepsilon \to 0$, we obtain

$$\frac{q^+(t) - q^-(s)}{2(t-s)} \geq \frac{s}{t}\frac{(p-1)}{2}\ell^{p-2}(\partial_\tau|_{\tau=s}\ell^{\pm}_\tau(\gamma_s))^2.$$

In particular, it turns out that

$$\frac{q^+(t) - q^-(s)}{t-s} \geq \frac{s}{t}\frac{r_{\pm}(s)^2}{(p-1)\ell^p}.$$

Next, we will deduce inequality (3.24) from (3.23) by simply using the duality between φ and φ^c. Indeed, since by definition it holds that $\bar{\varphi}_t = -\varphi^c_{1-t}$, we deduce that :

$$\bar{h}^\varphi_\gamma(r,\varepsilon) = -h^{\varphi^c}_{\gamma^c}(1-r,-\varepsilon).$$

Moreover, it holds

$$\frac{(p-1)(\ell^{\varphi^c,\pm}_{1-r-\varepsilon}(\gamma^c_{1-r}))^p}{p} = -\partial^{\mp}_r\varphi^c_{1-r-\varepsilon}(\gamma^c_{1-r}) = \partial^{\mp}_r\varphi_{r+\varepsilon}(\gamma_r) = \frac{(p-1)(\bar{\ell}^{\varphi,\pm}_{r+\varepsilon}(\gamma_r))^p}{p};$$

hence, choosing as φ, γ, ε, s, t respectively φ^c, γ_c, $-\varepsilon$, $1-t$, $1-s$ we get the second claim. $\square$

3.5 Consequences

We start by noticing an immediate consequence of Theorem 3.17:

Corollary 3.18. *For both $\tilde{q} = q, \bar{q}$, the functions $t \mapsto \tilde{q}_{\pm}(t)$ are monotone non-decreasing on $(0,1)$.*

We now put together previous regularity results on time behaviour of Kantorovich potential together with Theorem 3.17 in order to have a clear statement on the third order variation of Kantorovich potentials.

Theorem 3.19 (A priori third-order bounds for potential along its characteristics). *Assume that for* ***a.e.*** $t \in (0,1)$:

$$(0,1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_t) \quad \text{is twice differentiable at } \tau = t \quad \text{for both} \quad \tilde{\varphi} = \varphi, \bar{\varphi}, \tag{3.29}$$

in any of the equivalent senses of Corollary 3.16 and that moreover:

$$\partial_\tau^2|_{\tau=t}\varphi_\tau(\gamma_t) = \partial_\tau^2|_{\tau=t}\bar{\varphi}_\tau(\gamma_t) \quad \text{for a.e. } t \in (0,1).$$

If there exists a continuous function z for which

$$\partial_\tau^2|_{\tau=t}\varphi_\tau(\gamma_t) = \partial_\tau^2|_{\tau=t}\bar{\varphi}_\tau(\gamma_t) = z(t) \quad \text{for a.e. } t \in (0,1),$$

*then (3.29) holds for **all** $t \in (0,1)$ and for all $t \in (0,1)$*

$$\partial_\tau^2|_{\tau=t}\varphi_\tau(\gamma_t) = \partial_\tau^2|_{\tau=t}\bar{\varphi}_\tau(\gamma_t) = \partial_\tau|_{\tau=t}\frac{(p-1)\ell_\tau^p(\gamma_t)}{p} = \partial_\tau|_{\tau=t}\frac{(p-1)\bar{\ell}_\tau^p(\gamma_t)}{p} = z(t). \tag{3.30}$$

Finally, the following third order information on $\varphi_t(x)$ at $x = \gamma_t$ holds true:

$$\frac{z(t) - z(s)}{t-s} \geq \sqrt{\frac{s}{t}\frac{1-t}{1-s}}\frac{|z(s)||z(t)|}{(p-1)\ell^p}, \quad \forall\, 0 < s < t < 1. \tag{3.31}$$

In particular, for any point $t \in (0,1)$ where $z(t)$ is differentiable we have

$$z'(t) \geq \frac{z(t)^2}{(p-1)\ell^p}. \tag{3.32}$$

Proof. By Corollary 3.16, it follows that $\tilde{q}_-(t) = \tilde{q}_+(t) = z(t)$ for a.e. $t \in (0,1)$. More precisely, the same holds true for every $t \in (0,1)$ by the monotonicity of $\tilde{q}_\pm$ and the continuity of z; thus, (3.30) is satisfied. Moreover, Corollary 3.16 also implies that $\tilde{r}_-(t) = \tilde{r}_+(t) = z(t)$ for both $\tilde{r} = r, \bar{r}$ and for all $t \in (0,1)$. Taking the geometric mean of (3.23) and (3.24), we get (3.31). Finally, passing to the limit as $s \to t$ in (3.31), we obtain (3.32). $\qquad\square$

The assumptions of Theorem 3.19 will hold true for a.e. $t \in (0,1)$ only for a certain family of Kantorovich geodesics. Nonetheless, this family shall be sufficient for our purposes.

Finally, inequality (3.32) will be crucial to deduce concavity of certain one-dimensional factors. We include here a result that will be used later. For its proof we refer to [13, Lemma 5.7].

Lemma 3.20 (Concavity restatement). *Assume that for some locally absolutely continuous function z on $(0,1)$ we have:*

$$\partial_\tau|_{\tau=t}\frac{(p-1)\ell_\tau^p(\gamma_t)}{p} = z(t) \quad \text{for a.e. } t \in (0,1).$$

Then for any fixed $r_0 \in (0,1)$, the function:

$$L(r) = \exp\left(-\frac{1}{\ell^p(p-1)}\int_{r_0}^r \partial_\tau|_{\tau=t}\frac{(p-1)\ell_\tau^p(\gamma_t)}{p}\, dt\right) = \exp\left(-\frac{1}{\ell^p(p-1)}\int_{r_0}^r z(t)\, dt\right)$$

is concave on $(0,1)$.

3.6 Time propagation of Intermediate Kantorovich potentials

Finally we recall the definition of time-propagated intermediate Kantorovich potentials as introduced in [13].

Definition 3.21. Given a Kantorovich potential $\varphi : X \to \mathbb{R}$ and $s, t \in (0, 1)$, define the t-propagated s-Kantorovich potential Φ_s^t on the domain $D_\ell(t)$ where forward speed is well-defined and its time-reversed version $\bar{\Phi}_s^t$ on the domain $D_{\bar{\ell}}(t)$ from (3.18), by:

$$\Phi_s^t := \varphi_t + (t - s)\frac{\ell_t^p}{p} \text{ on } D_\ell(t), \qquad \bar{\Phi}_s^t := \bar{\varphi}_t + (t - s)\frac{\bar{\ell}_t^p}{p} \text{ on } D_{\bar{\ell}}(t).$$

Using Theorem 3.13, it follows that for all $s, t \in (0, 1)$:

$$\Phi_s^t = \bar{\Phi}_s^t = \varphi_s \circ e_s \circ (e_t|_{G_\varphi}^{-1}), \qquad \text{on } e_t(G_\varphi). \tag{3.33}$$

Indeed, for any $\gamma \in G_\varphi$ it holds

$$\Phi_s^t(\gamma_t) = \varphi_t(\gamma_t) + (t - s)\frac{\ell_t(\gamma_t)^p}{p} = \varphi_t(\gamma_t) + (t - s)\frac{\ell(\gamma)^p}{p} = \varphi_s(\gamma_s).$$

Consequently, on $e_t(G_\varphi)$, $\Phi_s^t = \bar{\Phi}_s^t$ is identified as the push-forward of φ_s via $e_t \circ e_s^{-1}$, i.e. its propagation along G_φ from time s to time t.

Proposition 3.22 (Linear expansion of energy in time generates propagation of potential)**.** *For any $s \in (0, 1)$, the following properties hold:*

1. *The maps $(x, t) \mapsto \Phi_s^t(x)$ and $(x, t) \mapsto \bar{\Phi}_s^t(x)$ are continuous on D_ℓ and on $D_{\bar{\ell}}$ respectively;*

2. *For each $x \in X$, denoting $\tilde{\Phi} \in \{\Phi, \bar{\Phi}\}$ and the corresponding $\tilde{\ell} \in \{\ell, \bar{\ell}\}$, the map $D_{\tilde{\ell}}(x) \ni t \mapsto \tilde{\Phi}_s^t(x)$ is differentiable at t if and only if $D_{\tilde{\ell}}(x) \ni t \mapsto \tilde{\ell}_t^p(x)$ is differentiable at t or if $t = s \in D_{\tilde{\ell}}(x)$. In particular, $t \mapsto \tilde{\Phi}_s^t(x)$ is a.e. differentiable. At any point of differentiability:*

$$\partial_t \tilde{\Phi}_s^t(x) = \tilde{\ell}_t^p(x) + (t - s)\frac{\partial_t \tilde{\ell}_t^p(x)}{p}$$

   In particular, if $s \in D_{\tilde{\ell}}(x)$ then $\partial_t|_{t=s}\tilde{\Phi}_s^t(x)$ exists and is given by $\tilde{\ell}_t^p(x)$.

3. *For each $x \in X$, the map $G_\varphi \ni t \mapsto \Phi_s^t(x) = \bar{\Phi}_s^t(x)$ is locally Lipschitz;*

4. *For all $t \in (0, 1)$:*

$$\begin{cases} \underline{\partial}_t \Phi_s^t(x) \geq \dfrac{s}{t}\ell_t^p(x), & t \geq s \\[2mm] \overline{\partial}_t \Phi_s^t(x) \leq \dfrac{s}{t}\ell_t^p(x), & t \leq s \end{cases} \quad \forall x \in D_\ell(t); \qquad \begin{cases} \overline{\partial}_t \bar{\Phi}_s^t(x) \leq \dfrac{1-s}{1-t}\bar{\ell}_t^p(x), & t \geq s \\[2mm] \underline{\partial}_t \bar{\Phi}_s^t(x) \geq \dfrac{1-s}{1-t}\bar{\ell}_t^p(x), & t \leq s \end{cases} \quad \forall x \in D_{\bar{\ell}}(t).$$

$$\tag{3.34}$$

Proof. By lower semi-continuity and Corollary 3.7, 1) and 2) follow trivially. By Corollary 3.11 and Theorem 3.13, 3) holds true. To see 4), observe that for every $x \in D_{\tilde{\ell}}(t)$,

$$\underline{\partial}_t \tilde{\Phi}_s^t(x) = \tilde{\ell}_t^p(x) + (t-s)\underline{\partial}_t \frac{\tilde{\ell}_t^p(x)}{p}, \quad t \geq s$$

$$\underline{\partial}_t \tilde{\Phi}_s^t(x) = \tilde{\ell}_t^p(x) + (t-s)\overline{\partial}_t \frac{\tilde{\ell}_t^p(x)}{p}, \quad t \leq s$$

with analogous identities holding for $\overline{\partial}_t \tilde{\Phi}_s^t(x)$. Using estimates (3.19) and (3.20) of Corollary 3.12, the claim follows. $\qquad \square$

4 Curvature-Dimension conditions: from $p > 1$ to $p = 1$

We will now focus on the main goal of this paper: to show that for essentially non-branching spaces, the synthetic ($p = 2$) curvature-dimension condition can be equivalently formulated in terms of entropic convexity conditions along p-Wasserstein geodesics for any other $p > 1$. Our approach is to show that for essentially non-branching spaces, the $\mathsf{CD}_p(K, N)$ for $p > 1$ is equivalent to $\mathsf{CD}^1(K, N)$, which is an appropriate concavity statement about the factor measures which arise whenever $\mathfrak{m}$ is disintegrated along the needles of the signed distance to the zero level-set of an arbitrary continuous function.

The first implication that we will address is the following one: if $(X, \mathsf{d}, \mathfrak{m})$ is a p-essentially non-branching metric measure space verifying $\mathsf{CD}_p(K, N)$ then it satisfies $\mathsf{CD}^1(K, N)$ (actually the stronger $\mathsf{CD}^1_{Lip}(K, N)$).

Before we begin the proof, we recall the concepts used in L^1 optimal transport theory. For simplicity, we will illustrate the case $X = \mathbb{R}^n$ paired with the Euclidean metric, and a restriction of Lebesgue as the ambient measure.

To a 1-Lipschitz function, $u : \mathbb{R}^n \to \mathbb{R}$, we associate a transport ordering, Γ_u, defined as in (2.5). Membership of (x, y) in this set should be understood as "y travels to x along a transport ray determined by u". In particular, it is helpful to consider $u(x) = |x|$ in which case the transport rays are polar rays emanating from 0, and $(x, y) \in \Gamma_u$ means that x and y lie on the same polar ray with x being larger in norm than y.

It is desirable to associate the points that travel along a geodesic with the geodesic itself. Towards this goal, consider a symmetric relation R_u composed of Γ_u together with its inverse relation, and denote the projection of R_u onto its first component by $\mathcal{T}_u$ (see (2.6)). We refer to R_u as the transport relation and $\mathcal{T}_u$ as the transport set. Even though R_u is a symmetric relation over $\mathcal{T}_u$, it is not transitive. This obstruction to transitivity is called branching, where two distinct points, z and w, travel to or from a point x but no transport ray of u transports z to w or vice versa. More specifically we have the forward and backward branching points A_+, A_- as defined in (4.2) and (4.3). To overcome this difficulty we simply remove the offending points and consider the resulting equivalence relation. That is, we consider $\mathcal{T}_u^b := \mathcal{T}_u \setminus (A_+ \cup A_-)$ and $R_u^b = R_u \cap (\mathcal{T}_u^b \times \mathcal{T}_u^b)$. As seen in Theorem 4.3, this procedure removes only a negligible set of points.

We then use this equivalence relation to break $\mathfrak{m}$, restricted to the branched transport set, into measures supported on each of the transport rays determined by u. We do this by applying the Disintegration Theorem. Using our example of $u(x) = |x|$ over $(\mathbb{R}^n, |\cdot|, \frac{1}{\omega_n} \mathcal{L}^n\big|_{B_1(0)})$ we arrive at

$$\mathcal{L}^n {\llcorner}_{B_1(0)}(\mathrm{d}x) = \int_{S^{n-1}} |x|^{n-1} \mathcal{H}^1 {\llcorner}_{[0,\alpha]} \mathcal{H}^{n-1}(\mathrm{d}\alpha) \tag{4.1}$$

where $\mathcal{H}^k$ is k-dimensional Hausdorff measure and $\frac{\omega_n}{n} = \mathcal{L}^n(B_1(0))$. Hence, in this case, disintegration gives polar integration. We observe that (4.1) can be compared to (4.4) where $Q = S^{n-1}$ and polar rays form the set of non-branched transport geodesics.

Finally, we remind the reader about $\mathsf{CD}(K, N)$. The $\mathsf{CD}(K, N)$ condition represents, in a very generalized sense, a Ricci curvature bound from below by $K \in \mathbb{R}$ and a dimension bound from above by $N \in (1, \infty)$. In particular, over an interval in $\mathbb{R}$ and a measure $\mathfrak{m} = h\mathcal{L}^1 \big|_{[0,L]}$, where $h > 0$ on $(0, L)$, the $\mathsf{CD}(K, N)$ condition reduces to

$$\left(h^{\frac{1}{N-1}}\right)'' + \frac{K}{N-1} h^{\frac{1}{N-1}} \leq 0.$$

This condition is equivalent to (K, N) convexity of $-\log(h)$ [22]. That is, $-\log(h)$ satisfies

$$(-\log h)'' \geq \frac{1}{N-1}\left((-\log h)'\right)^2 + K.$$

Hence, in one-dimensional space, the $\mathsf{CD}(K, N)$ condition amounts to a concavity condition on the density of the reference measure with respect to Lebesgue. In particular, using the example of $\mathfrak{m} = nr^{n-1}\mathcal{L}^1 \llcorner_{(0,1)}$ over a polar ray of length one, the $\mathsf{CD}(K, N)$ conditions becomes

$$Kr^2 \leq (n-1)\left(1 - \frac{n-1}{N-1}\right) \quad \text{for } 0 \leq r \leq 1.$$

This is satisfied for all $K \leq 0$ and $N \geq n$ which is consistent with the curvature and dimensionality of $\mathbb{R}^n$.

4.1 $\,\,L^1$ optimal transport

We recall a standard fact about 1-Lipschitz functions and their associated transport set.

To any 1-Lipschitz function $u : X \to \mathbb{R}$ there is a naturally associated d-cyclically monotone set Γ_u defined in (2.5) that we call the transport ordering; we write $x \geq_u y$ if and only if $(x, y) \in \Gamma_u$ and we recall that $\geq_u$ is a partial-ordering. The *transport relation R_u* and the *transport set $\mathcal{T}_u$* are defined in (2.6).

The transport "flavor" of the previous definitions can be seen in the next property that is immediate to verify: for any $\gamma \in \mathrm{Geo}(X)$ such that $(\gamma_0, \gamma_1) \in \Gamma_u$, then

$$(\gamma_s, \gamma_t) \in \Gamma_u, \quad \forall\, 0 \leq s \leq t \leq 1.$$

Finally, recall the definition of the *forward and backward branching points* of $\mathcal{T}_u$ that was introduced in [9]:

$$A_+ := \{x \in \mathcal{T}_u : \exists\, z, w \in \Gamma_u(x), (z, w) \notin R_u\}, \tag{4.2}$$

$$A_- := \{x \in \mathcal{T}_u : \exists\, z, w \in \Gamma_u(x)^{-1}, (z, w) \notin R_u\}. \tag{4.3}$$

Once branching points are removed, we obtain the *non-branched transport set* and the *non-branched transport relation*,

$$\mathcal{T}_u^b := \mathcal{T}_u \setminus (A_+ \cup A_-), \quad R_u^b := R_u \cap (\mathcal{T}_u^b \times \mathcal{T}_u^b).$$

The following was obtained in [9] and highlights the motivation for removing branching points.

Proposition 4.1 (Transport relation is an equivalence relation on the non-branched transport set)**.** *The non-branched transport relation $R_u^b \subset X \times X$ is an equivalence relation on $\mathcal{T}_u^b$.*

Noticing that once we fix $x \in \mathcal{T}_u^b$, for any choice of $z, w \in R_u(x)$, there exists $\gamma \in \text{Geo}(X)$ such that

$$\{x, z, w\} \subset \{\gamma_s : s \in [0, 1]\},$$

it is not hard to deduce that each equivalence class is a geodesic. These equivalence classes are sometimes called transport rays [20] or needles [28].

It is a classical procedure then to construct an $\mathfrak{m}$-measurable quotient map $\mathfrak{Q}$ for the equivalence relation R_u^b over $\mathcal{T}_u^b$; in particular, there will be an $\mathfrak{m}$-measurable quotient set $Q \subset \mathcal{T}_u^b$ which is the image of $\mathfrak{Q}$. The Disintegration Theorem then implies the following disintegration formula:

$$\mathfrak{m}{\llcorner}_{\mathcal{T}_u^b} = \int_Q \mathfrak{m}_\alpha \mathfrak{q}(d\alpha), \tag{4.4}$$

where $\mathfrak{q} = \mathfrak{Q}_\sharp \mathfrak{m}{\llcorner}_{\mathcal{T}_u^b}$, and for $\mathfrak{q}$-a.e. $\alpha \in Q$ we have $\mathfrak{m}_\alpha \in \mathcal{P}(X)$, $\mathfrak{m}_\alpha(X \setminus X_\alpha) = 0$, where we have used the notation X_α to denote the equivalence class of the element $\alpha \in Q$ (indeed $X_\alpha = R_u^b(\alpha)$).

Remark 4.2. For a brief account on the Disintegration Theorem, we refer to [8] and references therein (see also [13]). It is worth mentioning here that the map $Q \ni \alpha \mapsto \mathfrak{m}_\alpha \in \mathcal{P}(X)$ is essentially unique (meaning that any two maps for which (4.4) holds true have to coincide $\mathfrak{q}$-a.e.) thanks to the assumption $\mathfrak{m}(X) = 1$, while $\mathfrak{m}_\alpha(X \setminus X_\alpha) = 0$ (also called strongly consistence of the disintegration) is a consequence of the existence an $\mathfrak{m}$-measurable quotient map $\mathfrak{Q}$.

Again in [9], it was proved also that assuming the $\mathsf{RCD}(K, N)$ condition (which enhances $\mathsf{CD}(K, N)$ with an infinitesmal Hilbertianity assumption), the measure of the set of branching points is zero. As already observed several times in the literature, the $p = 2$ proof only requires all optimal plans to be maps, and so the same argument works for $CD_p(K, N)$ and any $p > 1$:

Theorem 4.3 (Negligibility of forward and backward branching points)**.** *Let $(X, \mathsf{d}, \mathfrak{m})$ be a m.m.s. such that for any $\mu_0, \mu_1 \in \mathcal{P}_p(X)$ with $\mu_0 \ll \mathfrak{m}$ any optimal transference plan for W_p is concentrated on the graph of a function. Then*
$$\mathfrak{m}(A_+) = \mathfrak{m}(A_-) = 0.$$

From Theorem 2.7, the p-essentially non-branching hypothesis implies that for every $\mu_0, \mu_1 \in \mathcal{P}_p(X)$ with $\mu_0 \ll m$ there exists a unique p-optimal plan and it is induced by a map. Hence, the assumptions of Theorem 4.3 are satisfied, and therefore

$$\mathfrak{m}(A_+) = \mathfrak{m}(A_-) = 0. \tag{4.5}$$

Putting together (4.4) and (4.5) we obtain:

$$\mathfrak{m}{\llcorner}_{\mathcal{T}_u} = \int_Q \mathfrak{m}_\alpha \mathfrak{q}(d\alpha). \tag{4.6}$$

In what follows we will prove that $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_u^1(K, N)$.

4.2 Curvature estimates

Recalling Definition 2.9, one will observe that to prove $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_u^1(K, N)$ it suffices to show that, for $\mathfrak{q}$-a.e. $\alpha \in Q$, the one dimensional metric measure space $(X_\alpha, \mathsf{d}, \mathfrak{m}_\alpha)$ is a $\mathsf{CD}(K, N)$ space, i.e. if X_α is isometric to $[0, L_\alpha]$ where L_α is the length of X_α then,

$$\mathfrak{m}_\alpha = h_\alpha \mathcal{L}^1 {\llcorner}_{[0, L_\alpha]}, \qquad \left(h_\alpha^{\frac{1}{N-1}}\right)'' + \frac{K}{N-1} h_\alpha^{\frac{1}{N-1}} \leq 0,$$

where the inequality has to be understood in the distributional sense. Notice indeed that, by construction, the transport rays X_α are the maximal totally-ordered subsets of $\mathcal{T}_u^b \subset X$ under the partial-ordering $\leq_u$ given by Γ_u.

First we recall a result relating d^p-cyclically monotone sets to d-cyclically monotone set, presented in [10] for $p = 2$.

Lemma 4.4 (Certain d-cyclically monotone sets are also d^p-cyclical monotone). *Let $p > 1$ be any real number and let $\Delta \subset \Gamma_u$ be any set such that*

$$(x_0, y_0), (x_1, y_1) \in \Delta \implies (u(y_1) - u(y_0)) \cdot (u(x_1) - u(x_0)) \geq 0.$$

Then Δ is d^p-cyclically monotone.

Proof. By hypothesis the set

$$\Lambda := \{(u(x), u(y)) : (x, y) \in \Delta\} \subset \mathbb{R}^2$$

is monotone in the Euclidean sense. Since $\Lambda \subset \mathbb{R}^2$, it is a standard fact that it is also c-cyclically monotone, for any cost $c(x, y) = \vartheta(|x - y|)$ with $\vartheta : [0, +\infty) \to [0, +\infty)$ convex and such that $\vartheta(0) = 0$. Hence, in particular, Λ is $|\cdot|^p$-cyclically monotone.

Fix now $\{(x_i, y_i)\}_{i=1}^n \subset \Delta$. Using that u is 1-Lipschitz and $\Delta \subset \Gamma$, it turns out that

$$\sum_{i=1}^n \mathsf{d}^p(x_i, y_i) = \sum_{i=1}^n |u(x_i) - u(y_i)|^p$$
$$\leq \sum_{i=1}^n |u(x_i) - u(y_{i+1})|^p \leq \sum_{i=1}^n \mathsf{d}^p(x_i, y_{i+1}).$$

Hence the claim. □

Example 4.5 (Dimensional count in the smooth case). If d is the geodesic distance on an n-dimensional Riemannian manifold X (or Euclidean space), then — away from the cut locus — any d-cyclically monotone subset Δ is contained in a $n + 1$ dimensional subset of X^2, the extra dimension being due to the degeneracy of d along the direction of transport [36]. On the other hand, if the left projection $P_1(\Delta) \subset \{\tilde{u} = 0\}$ for some C^1 function $\tilde{u}$ whose derivative is non-vanishing on its zero set, we expect the dimension of Δ to be reduced to n, which coincides with the dimensional bound on a d^p-cyclically monotone set for $p > 1$. This example helps motivate both the previous lemma and the construction to follow.

Similarly, in the nonsmooth setting, fixing $\delta \in \mathbb{R}$ and considering pairs $\Delta \subset \Gamma_u$ of partners $(x, y) \in \Delta$ whose lower endpoint lies on a fixed level set $u(y) = \delta$, it follows that Δ is d^p-cyclically monotone for all $p > 1$. Equivalently, for each $C \subset \mathcal{T}_u^b$ and $\delta \in \mathbb{R}$, the set $\Delta := (C \times \{u = \delta\}) \cap \Gamma_u$ is d^p-cyclically monotone. Setting

$$C_\delta = P_1((C \times \{u = \delta\}) \cap \Gamma_u),$$

we see that if $\mathfrak{m}(C_\delta) > 0$, then by Theorem 2.7, there exists a unique $\nu \in \mathrm{OptGeo}_p(\mu_0, \mu_1)$ such that

$$(\mathrm{e}_0)_\sharp \nu = \mathfrak{m}(C_\delta)^{-1} \mathfrak{m} \llcorner C_\delta, \qquad (\mathrm{e}_0, \mathrm{e}_1)_\sharp \nu (C \times \{u = \delta\} \cap \Gamma_u) = 1,$$

and whose push-forwards by e_t verify the entropic concavity statement (2.2) for all $t \in [0, 1]$. Letting C and δ vary, it is a standard procedure, see for example [9], to deduce that:

- for $\mathfrak{q}$-a.e. $\alpha \in Q$, the conditional probabilities $\mathfrak{m}_\alpha$ are absolutely continuous w.r.t. $\mathcal{L}^1 \llcorner X_\alpha$;

- if $\mathfrak{m}_\alpha = h_\alpha \mathcal{L}^1 \llcorner X_\alpha$, then $h_\alpha > 0$ in the relative interior of X_α and is locally Lipschitz.
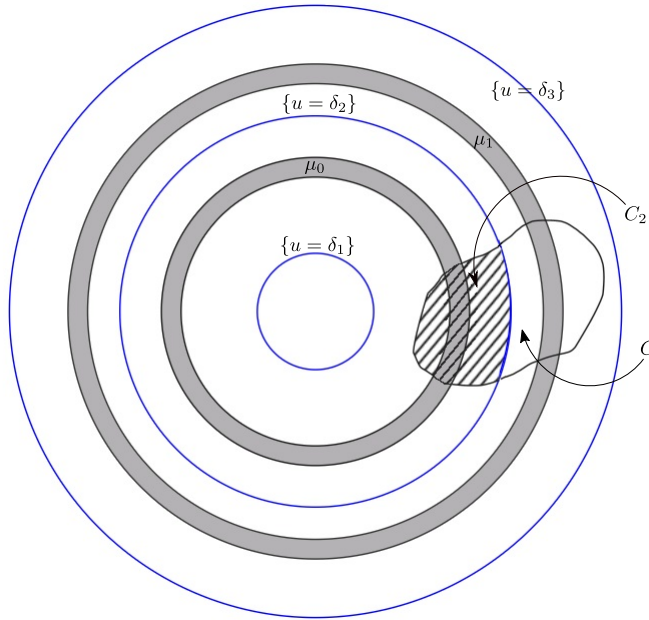

Figure 1: (The sets C_δ) Transporting the sets μ_0 to μ_1 along radial transport geodesics determined by a radial 1-Lipschitz function u associated to the radial Kantorovich potential φ. If we assume that u behaves like the euclidean norm, then we see that $C_{\delta_1} = \varnothing, C_{\delta_2} = C_2, C_{\delta_3} = C$.

The next step is to prove the $\mathsf{CD}(K, N)$ inequality for $\mathfrak{q}$-a.e. one-dimensional density h_α. This follows repeating verbatim the proof of [14, Theorem 4.2] where the same implication was proved assuming $\mathsf{CD}_{2,loc}(K, N)$ and 2-essentially non-branching. The main ingredient being Lemma 4.4 for $p = 2$, the argument carries over for any $p > 1$.

Putting together what has been discussed so far, we see that we have obtained the following:

Theorem 4.6 (Non-branching $\mathsf{CD}_{p,loc}$ spaces are CD^1_{Lip})**.** *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a p-essentially non-branching m.m.s. satisfying the* $\mathsf{CD}_{p,loc}(K, N)$ *condition for some* $p \in (1, +\infty)$, $K \in \mathbb{R}$, *and* $N \in [1, \infty)$ *and* $\mathfrak{m}(X) = 1$.

Then, for any fixed 1-Lipschitz function $u : X \to \mathbb{R}$, *the transport relation* R^b_u *induces on the transport set a disintegration of* $\mathfrak{m}\llcorner_{\mathcal{T}_u}$ *into conditional measures,* $\mathfrak{m}_\alpha$, *that for* $\mathfrak{q}$-*a.e.* α *satisfy* $\mathfrak{m}_\alpha = h_\alpha \mathcal{L}^1 \llcorner_{X_\alpha}$ *and:*

$$h_\alpha((1-s)t_0 + st_1)^{1/(N-1)} \geq \sigma^{(1-s)}_{K,N-1}(t_1 - t_0)h_\alpha(t_0)^{1/(N-1)} + \sigma^{(s)}_{K,N-1}(t_1 - t_0)h_\alpha(t_1)^{1/(N-1)},$$

for all $s \in [0,1]$ *and for* $t_0, t_1 \in [0, L_\alpha]$ *with* $t_0 < t_1$, *where we have identified the transport ray* X_α *with the real interval* $[0, L_\alpha]$ *having the same length.*

Notice that the $\mathfrak{q}$-measurability of the disintegration, ensured by the Disintegration Theorem, implies joint measurability of the map $(\alpha, t) \to h_\alpha(t)$.

Remark 4.7 (Enhancing CD^1_{Lip}). It is worth underlining that the conclusion of Theorem 4.6 is actually stronger than claiming that $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}^1_{Lip}(K, N)$. Notice, indeed, that while $\mathsf{CD}^1_{Lip}(K, N)$ asks for *a* disintegration of $\mathfrak{m}\llcorner_{\mathcal{T}_u}$ (no partition required, see Definition 2.9) where each conditional measure is concentrated along a maximal transport ray and verifies $\mathsf{CD}(K, N)$, Theorem 4.6 shows that we have a *partition* of the transport set made of maximal transport rays and the associated *essentially unique* disintegration verifies $\mathsf{CD}(K, N)$ (recall Remark 4.2). In what follows we will show that this property is enough to prove that $(X, \mathsf{d}, \mathfrak{m})$ also verifies $\mathsf{CD}_q(K, N)$ for any $q > 1$, provided it is also q-essentially non-branching.

To complete the picture we mention that in [13, Proposition 8.13] it is shown that $\mathsf{CD}^1_{Lip}(K, N)$ coupled with essentially non-branching (hence $p = 2$) implies that the disintegration of $\mathfrak{m}\llcorner_{\mathcal{T}^b_u}$ coming from the partition induced by the transport relation R^b_u indeed verifies all the conditions required by $\mathsf{CD}^1_{Lip}(K, N)$. We refer to [13, Proposition 8.13] for additional details.

Remark 4.8 (Strategy of proof). Here we briefly comment on the technique used in [14] to prove Theorem 4.6. The idea is to first establish the existence of a disintegration of $\mathfrak{m}$ into a collection of conditional measures, $\{\mathfrak{m}_\alpha\}_{\alpha \in Q}$, that are supported along transport rays determined by an arbitrary 1-Lipschitz function u as in (4.6). In particular, we may express $\mathfrak{m}_\alpha$, the portion of measure of $\mathfrak{m}$ that lives on the transport geodesic of parameter α, as

$$\mathfrak{m}_\alpha = g(\alpha, \cdot)_\sharp \left(h_\alpha(t) \mathcal{L}^1 (dt) \right)$$

where $g : Q \times [0, 1] \to X$ is such that for each α we have that $\mathrm{dom}(g(\alpha, \cdot))$ is convex and $h_\alpha : \mathrm{dom}(g(\alpha, \cdot)) \to [0, \infty)$. Next, the quotient set Q, which labels the various transport rays, is covered by a countable disjoint collection of sets $\{Q_i\}_{i \in I}$ where each Q_i is contained in a rational level set of u. Finally, along each Q_i we consider the transport of one uniform measure to another, of possibly differing size, along the transport rays of u. More specifically, our countable decomposition is constructed to provide for each i a uniform subinterval

$$(a_0, a_1) \subset \mathrm{dom}(g(\alpha, \cdot)) \quad \text{for all } \alpha \in Q_i$$

as well as real numbers $A_0, A_1 \in (a_0, a_1)$ and $L_0, L_1 \in (0, \infty)$ such that

$$A_0 + L_0 < A_1 \text{ and } A_1 + L_1 < a_1.$$

32

This allows us to consider the measures

$$\mu_0 = \int_{Q_i} g(\alpha,\cdot)_\sharp \left(\frac{1}{L_0} \mathcal{L}^1_{[A_0,A_0+L_0]}(dt) \right) \mathfrak{q}(d\alpha), \qquad \mu_1 = \int_{Q_i} g(\alpha,\cdot)_\sharp \left(\frac{1}{L_1} \mathcal{L}^1_{[A_1,A_1+L_1]}(dt) \right) \mathfrak{q}(d\alpha).$$

Transporting these measures allows us to deduce concavity information for the density h_α of $\mathfrak{q}$-a.e. $\mathfrak{m}_\alpha$ from the entropic concavity (2.2) asserted by $\mathsf{CD}_p(K,N)$.

5   Curvature-Dimension conditions: from $p=1$ to $q>1$

Before tackling Theorems 1.1 and 1.2 we explore an example which illustrates some of the strategies and notations used.

Example 5.1 (Radial transport). Let $X = \mathbb{R}^n$, d be Euclidean distance, and set $\mathfrak{m} = \mathcal{L}^n$. Let $\mu_0(dx) = \frac{1}{\omega_n |x|^{n-1}} \mathcal{L}^n \llcorner_{A_{1,2}}(dx)$ and $\mu_1(dx) = \frac{1}{\omega_n |x|^{n-1}} \mathcal{L}^n \llcorner_{A_{3,4}}(dx)$ where for $0 < s < r < \infty$, $A_{s,r}$ is defined as the spherical shell

$$A_{s,r} = B_r(0) \setminus B_s(0).$$

We use the cost $c(x,y) = \frac{\mathsf{d}(x,y)^q}{q}$ where $1 < q < \infty$. For this transport problem, the optimal map is $T(x) = (|x|+2)\frac{x}{|x|}$, the Kantorovich potential is $\varphi(x) = -2^{q-1}|x|$, and its interpolated potentials are

$$\varphi_t(x) = \begin{cases} \frac{-|x|^q}{qt^{q-1}}, & \text{if } |x| \le 2t, \\ -2^{q-1}\left[|x| - \frac{2t}{q'}\right], & \text{if } 2t < |x|, \end{cases}$$

where q' is the Hölder dual to q. It is possible to show that the set G_φ of (φ,q)-Kantorovich geodesics (3.9) consist of all segments of length two pointed away from the origin. Notice that not all such geodesics are involved in the transport of μ_0 to μ_1: indeed only those starting in the source $A_{1,2}$ (and therefore ending in the target $A_{3,4}$) are. In particular, only the subset of geodesics starting at a point in $A_{1,2}$ will have mass passing along them at all times $t \in (0,1)$. This restriction should be compared to condition 3 from Definition 5.4. In particular, we use G to denote a good subset of G_φ of full measure which meet the stipulations of Definition 5.4.

Since we wish to apply the Disintegration Theorem, we have to associate the geodesics of G_φ with the transport rays of a 1-Lipschitz function. We do so by choosing our 1-Lipschitz function to be the signed distance to a level set of φ. In our example we can use the norm since φ is a monotone radial function. However, in the general case, we must use the signed distance $d_{a,s} := d_{\varphi_s - a}$ with respect to the a level set of φ. Note that the ordinary distance function was not used so that we could refer to the level sets of $d_{a,s}$ uniquely. This idea is the basis of the discussion in subsection 5.1. In both cases we see that we are working with a subset, G, of the transport set according to the 1-Lipschitz function we chose. This should be compared to Lemma 5.5.

Finally, we demonstrate how the change of variables formula from Theorem 5.10 applies to our example. For $0 < t < 1$ and $\gamma \in G$, the interpolating maps, measures, and densities are given by:

$$T_t(x) = (|x|+2t)\frac{x}{|x|},$$

$$\mu_t(dx) = \frac{1}{\omega_n |x|^{n-1}} \mathcal{L}^n \llcorner_{A_{1+2t,2+2t}}(dx),$$

$$\text{and } \rho_t(\gamma_t) = \frac{1}{\omega_n \left(|\gamma_0| + 2t\right)^{n-1}}.$$

Hence, for $s, t \in (0, 1)$ we have

$$\frac{\rho_t(\gamma_t)}{\rho_s(\gamma_s)} = \left(\frac{1 + \ell + 2s}{1 + \ell + 2t}\right)^{n-1} \tag{5.1}$$

if $|\gamma_0| = 1 + \ell$. For fixed $s \in (0, 1)$, we note that if $a = -2^{q-1}\left[1 + \ell + \frac{2s}{q}\right]$ for $0 \le \ell \le 1$ and $\gamma \in G$ is a geodesic such that $|\gamma_0| = 1 + \ell$ then

$$\varphi_s(\gamma_s) = a.$$

In particular, using this notation we can write $G_{a,s} = \{\gamma \in G : \varphi_s(\gamma_s) = a\}$. Hence,

$$e_s(G_{a,s}) = \partial B_{1+\ell+2s}(0)$$

$$e_{[0,1]}(G_{a,s}) = A_{1+\ell, 3+\ell}$$

where $-\ell = 2^{1-q}a + 1 + \frac{2s}{q}$. Using the Disintegration Theorem, as in (4.1), for any $1 \le \ell \le 1$, we obtain

$$\mathcal{L}^n \llcorner A_{1+\ell, 3+\ell}(dx) = \int_{\partial B_1(0)} |x|^{n-1} \mathcal{H}^1 \llcorner_{\{r\alpha | 1 + \ell \le r \le 3 + \ell\}} \mathcal{H}^{n-1}(d\alpha)$$

where $\mathcal{H}^k$ denotes k-dimensional Hausdorff measure. Notice that we can rewrite this as

$$\mathcal{L}^n \llcorner A_{1+2\ell, 3+2\ell} = \int_{\partial B_{1+\ell+2s}(0)} g^{a,s}(\alpha, \cdot)_\sharp \left(2\left(\frac{1 + \ell + 2t}{1 + \ell + 2s}\right)^{n-1} \chi_{[0,1]}(t)dt\right) \mathcal{H}^{n-1}(d\alpha)$$

$$= \int_0^1 g^{a,s}(\cdot, t)_\sharp \left(2\left(\frac{1 + \ell + 2t}{1 + \ell + 2s}\right)^{n-1} \chi_{[0,1]}(t)d\mathcal{H}^{n-1}\right) \mathcal{L}^1(dt) \tag{5.2}$$

where $g^{a,s} : e_s(G_{a,s}) \times [0, 1] \to X$ and $g^{a,s}(\alpha, \cdot) = e_s \llcorner_{G_{a,s}}^{-1}(\alpha)$. Hence, $h_\alpha^{a,s}(t) = \left(\frac{1+\ell+2t}{1+\ell+2s}\right)^{n-1}$ where we have normalized this function so that $h_\alpha^a(s) = 1$. Next notice that

$$\Phi_s^t(x) = -2^{q-1}\left[|x| - 2t + \frac{2s}{q}\right]. \tag{5.3}$$

We may also compute that

$$\partial_\tau\Big|_{\tau=t} \Phi_s^\tau(x) = 2^q.$$

This allows us to show that

$$\frac{\partial_\tau\big|_{\tau=t} \Phi_s^\tau(\gamma_t)}{\ell^q(\gamma)} \cdot \frac{1}{h_{\gamma_s}^{\varphi_s(\gamma_s), s}(t)} = \left(\frac{1 + \ell + 2s}{1 + \ell + 2t}\right)^{n-1}$$

if $|\gamma_0| = 1 + \ell$ which, of course, matches (5.1) and verifies Theorem 5.10. Note that in general one will not have such explicit information. As such, an expression like (5.2) will be not at disposal; hence, it is necessary to deduce information by comparing the disintegration described in (5.2) with another one. Observe that the measure being pushed forward in (5.2) lives on $e_t(G_{a,s})$ and was obtained from a disintegration with respect to a time varying partition of $e_t(G_{a,s})$. For the second disintegration we instead focus on varying the level set values a to form a partition of $e_t(G)$. This description should be compared with subsection 5.2 and the comparison done in subsection 5.3.

34

Let $(X, \mathsf{d}, \mathfrak{m})$ be a p-essentially non-branching metric measure space satisfying $\mathsf{CD}_p(K, N)$ and, consequently from Theorem 4.6, also the enhanced $\mathsf{CD}^1_{Lip}(K, N)$ described in Remark 4.7. This will be needed to close the argument: the enhanced $\mathsf{CD}^1_{Lip}(K, N)$ will give a "canonical" way of disintegrating the measure $\mathfrak{m}$ that will be crucial in the implementation of the strategy outlined in the last few lines of Example 5.1.

Given any $q > 1$, we will prove that $(X, \mathsf{d}, \mathfrak{m})$ also verifies $\mathsf{CD}_q(K, N)$, provided the space is q-essentially non-branching as well. Recall that without loss of generality we can assume $\mathrm{spt}(\mathfrak{m}) = X$ and we have the standing assumption that $\mathfrak{m}(X) = 1$.

Fix $\mu_0, \mu_1 \in \mathcal{P}_q(X, \mathsf{d}, \mathfrak{m})$. From the curvature assumption it follows that (X, d) is a geodesic space, hence, from Section 2.1, $(\mathcal{P}_q(X), W_q)$ is a geodesic space as well; therefore the set of q-optimal dynamical plan $\mathrm{OptGeo}_q(\mu_0, \mu_1)$ is non-empty.

Recall moreover that $\mathsf{CD}_p(K, N)$ implies qualitative non-degeneracy (2.4) by [26], hence Theorem 2.7 yields a unique $\nu \in \mathrm{OptGeo}_q(\mu_0, \mu_1)$ and

$$[0, 1] \ni t \mapsto \mu_t := (\mathrm{e}_t)_\sharp \nu = \rho_t \mathfrak{m}.$$

Finally, let $\varphi : X \to \mathbb{R}$ be a Kantorovich potential for the Optimal transport problem from μ_0 to μ_1 associated to the cost $c := \mathsf{d}^q/q$. Recall that $G_\varphi \subset \mathrm{Geo}(X)$ denotes the set of (φ, q)-Kantorovich geodesics, i.e. all the geodesics γ for which

$$\varphi(\gamma_0) + \varphi^c(\gamma_1) = \frac{\mathsf{d}^q(\gamma_0, \gamma_1)}{q}.$$

We denote with G^0_φ the set of null (φ, q)-Kantorovich geodesics defined as follows:

$$G^0_\varphi := \{\gamma \in G_\varphi : \ell(\gamma) = 0\},$$

and its complement in G_φ by G^+_φ.

Using [13, Proposition 9.1], the $\mathsf{MCP}(K, N)$ condition implies some non-trivial regularity properties on the time behaviour of the density ρ_t: indeed the implication $(1) \Rightarrow (4)$ of [13, Proposition 9.1] gives a Lipschitz-type bound whenever μ_1 reduces to a Dirac mass δ_o for some $o \in X$ (notice that from [13, Remark 9.4] this implication does not require any type of essential non-branching property). Then the case of a general μ_1 can be obtained via approximation: using the q-essential non-branching property in its equivalent formulation given by Theorem 2.7, one can repeat the arguments of [13, Proposition 9.1] in the implications $(4) \Rightarrow (2)$ and $(2) \Rightarrow (3)$ where the main points were uniqueness of optimal dynamical plans and upper semi-continuity of entropies, which are both still valid in our framework. We summarize this discussion in the next statement:

Corollary 5.2 (Logarithmic finite difference bounds for interpolating densities along characteristics). *Let $(X, \mathsf{d}, \mathfrak{m})$ be a q-essentially non-branching m.m.s. verifying $\mathsf{MCP}(K, N)$. Then for all $\mu_0, \mu_1 \in \mathcal{P}_q(X)$ with $\mu_0 \ll \mathfrak{m}$ there exists a unique $\nu \in \mathrm{OptGeo}_q(\mu_0, \mu_1)$ and a map $S : X \to \mathrm{Geo}(X)$ such that $\nu = S_\sharp \mu_0$.*

Moreover $\mu_t = (\mathrm{e}_t)_\sharp \nu \ll \mathfrak{m}$ for $t \in [0, 1)$ and there exist versions of the densities $\rho_t = \frac{d\mu_t}{d\mathfrak{m}}$, such that for ν-a.e. $\gamma \in \mathrm{Geo}(X)$, for all $0 \leq s \leq t < 1$, it holds

$$\rho_s(\gamma_s) > 0, \quad \left(\tau^{(\frac{s}{t})}_{K,N}(\mathsf{d}(\gamma_0, \gamma_t))\right)^N \leq \frac{\rho_t(\gamma_t)}{\rho_s(\gamma_s)} \leq \left(\tau^{(\frac{1-t}{1-s})}_{K,N}(\mathsf{d}(\gamma_s, \gamma_1))\right)^{-N}. \tag{5.4}$$

In particular, for ν-a.e. γ, the map $t \mapsto \rho_t(\gamma_t)$ is locally Lipschitz on $(0, 1)$ and upper semi-continuous at $t = 0$.

35

A further consequence of Corollary 5.2 can be obtained considering the regularity property of the map $t \mapsto \mathfrak{m}(e_t(G))$, for some compact subset G of φ-Kantorovich geodesics (see for instance [13, Proposition 9.6]).

Proposition 5.3 (Near continuity of the evolution of $\operatorname{spt}\mu_t$). *Let $(X, \mathsf{d}, \mathfrak{m})$ be a q-essentially non-branching m.m.s. verifying $\mathsf{MCP}(K, N)$. For $\mu_0, \mu_1 \in \mathcal{P}_q(X)$ with $\mu_0 \ll \mathfrak{m}$, let ν denote the unique element of $\operatorname{OptGeo}_q(\mu_0, \mu_1)$.*

Then for any compact set $G \subset \operatorname{Geo}(X)$ with $\nu(G) > 0$, such that (5.4) holds true for all $\gamma \in G$ and $0 \leq s \leq t < 1$, it holds for any $t \in (0,1)$:

$$\lim_{\varepsilon \to 0+} \frac{\mathcal{L}^1\big(G(x) \cap (t - \varepsilon, t + \varepsilon)\big)}{2\varepsilon} = 1 \quad in \ L^1(e_t(G), \mathfrak{m}),$$

where $G(x) = \bigcup_{\gamma \in G} \gamma^{-1}(x)$.

Finally, we conclude this first part by recalling the definition of a special class of Kantorovich geodesics.

Definition 5.4 (Good collections of geodesics). Given $\mu_0, \mu_1 \in \mathcal{P}_q(X)$ with $\mu_0 \ll \mathfrak{m}$, we say that $G \subset G_\varphi^+$ is a *good* subset of geodesics if the following properties hold true:

1. G is compact;

2. there exists a constant $c > 0$ such that for every $\gamma \in G$: $c \leq \ell(\gamma) \leq 1/c$;

3. for every $\gamma \in G$, $\rho_t(\gamma_t) > 0$ for all $t \in [0,1]$ and the map $(0,1) \ni t \mapsto \rho_t(\gamma_t)$ is continuous;

4. the claim of Proposition 5.3 holds true for G;

5. The map $e_t|_G : G \to X$ is injective.

From now on we will assume $G \subset G_\varphi^+$ to be a good subset. In particular all the results contained in Sections 5.1, 5.2, and 5.3 will be obtained tacitly assuming any optimal dynamical plan to be concentrated on a good subset of geodesics.

We will dispose of this assumption in Section 5.4 via an approximation argument. Notice indeed that under q-essentially non-branching and $\mathsf{MCP}(K, N)$ for any $\nu \in \operatorname{OptGeo}_q(\mu_0, \mu_1)$ with $\mu_0 \ll \mathfrak{m}$, and any $\varepsilon > 0$ there exists a *good* compact subset $G^\varepsilon \subset G_\varphi^+$ such that $\nu(G^\varepsilon) \geq \nu(G_\varphi^+) - \varepsilon$ for any $\varepsilon > 0$. Without loss of generality, we can also assume that G^ε increases along any given sequence of ε decreasing to 0.

In what follows we will use a suitable collection of L^1-optimal transport problems to decompose the Jacobian of the evolution of the W_q-geodesic $t \to \mu_t$ and to obtain key estimates on both components: our interest will be focused on finding a codimension-1 Jacobian orthogonal to the evolution and a one-dimensional counterpart. For both of these factors, curvature estimates will be obtained via L^1-optimal transport techniques, in particular Theorem 4.6, by comparing two families of conditional measures: one coming from the aforementioned L^1-optimal transport problem and the other one from the q-Kantorovich potential.

The decomposition technique will be very similar to the one developed in [13]; we will not repeat all the proofs but just list the main differences and include additional details where needed.

5.1 L^1 **Partition**

For $s \in [0,1]$ and $a \in \mathbb{R}$, we define the set of geodesics $G_{a,s} \subset G_{\varphi}$ as follows:

$$G_{a,s} = \{\gamma \in G : \varphi_s(\gamma_s) = a\}.$$

Let us observe that since G is compact and $\mathrm{e}_s : G \to X$ is continuous, $\mathrm{e}_s(G)$ is still compact. Moreover, for $s \in (0,1)$, $\varphi_s : X \to \mathbb{R}$ is continuous and hence $G_{a,s}$ is compact as well.
Let us fix $a \in \varphi_s(\mathrm{e}_s(G))$. The aim of the next subsection will be to analyze the structure of the evolution of the set $G_{a,s}$, i.e. $\mathrm{e}_{[0,1]}(G_{a,s})$.

From now on we will denote the signed-distance function from a level set a of φ_s with $d_{a,s} := d_{\varphi_s - a}$ (recall the notation of (2.7)). Since $d_{a,s}$ is a 1-Lipschitz function, we can associate to it all the sets introduced in Section 4.1, including the transport ordering $\Gamma_{d_{a,s}} = \leq_{d_{a,s}}$, relation $R_{d_{a,s}} = \Gamma_{d_{a,s}} \cup \Gamma_{d_{a,s}}^{-1}$ and set $\mathcal{T}_{d_{a,s}} \subset P_1(R_{d_{a,s}})$.

Lemma 5.5. *Let (X, d) be a geodesic space. Once $s \in [0,1]$ and $a \in \varphi_s(\mathrm{e}_s(G))$ are fixed, then for each $\gamma \in G_{a,s}$ and for every $0 \leq r \leq t \leq 1$, $(\gamma_r, \gamma_t) \in \Gamma_{d_{a,s}}$. In particular,*

$$\mathrm{e}_{[0,1]}(G_{a,s}) \subset \mathcal{T}_{d_{a,s}}.$$

The proof goes along the same lines of [13, Lemma 10.3] which we have included for the reader's convenience.

Proof. Let us fix $\gamma \in G_{a,s}$. By Corollary 3.7 and Lemma 3.9 (2), we have that if $s \in [0,1)$ then for any $x \in \{\varphi_s = a\}$, it holds

$$\frac{\mathsf{d}^p(\gamma_s, \gamma_1)}{p(1-s)^{p-1}} = \varphi_s(\gamma_s) + \varphi^c(\gamma_1) = \varphi_s(x) + \varphi^c(\gamma_1) \leq \bar{\varphi}_s(x) + \varphi^c(\gamma_1) \leq \frac{\mathsf{d}^p(x, \gamma_1)}{p(1-s)^{p-1}}.$$

Hence $\mathsf{d}(\gamma_s, \gamma_1) \leq \mathsf{d}(x, \gamma_1)$. In the same way, if $s \in (0,1]$, then for any $y \in \{\varphi_s = a\}$ we have that

$$\frac{\mathsf{d}^p(\gamma_s, \gamma_0)}{p s^{p-1}} = \varphi(\gamma_0) - \varphi_s(\gamma_s) = \varphi(\gamma_0) - \varphi_s(y) \leq \frac{\mathsf{d}^p(y, \gamma_0)}{p s^{p-1}}.$$

So $\mathsf{d}(\gamma_s, \gamma_0) \leq \mathsf{d}(y, \gamma_0)$, which is also trivially satisfied in the case $s = 0$. Thus, for any $x, y \in \{\varphi_s = a\}$ we have

$$\mathsf{d}(\gamma_0, \gamma_1) \leq \mathsf{d}(\gamma_0, x) + \mathsf{d}(y, \gamma_1).$$

Taking the infimum over x and y we get that

$$\mathsf{d}(\gamma_0, \gamma_1) \leq d_{a,s}(\gamma_0) - d_{a,s}(\gamma_1),$$

where the sign of $\mathsf{d}_{a,s}$ was determined by the fact that $s \mapsto \varphi_s(\gamma_s)$ is decreasing. More precisely, the latter relation turns out to hold as an equality by 1-Lipschitz regularity of $d_{a,s}$, thus $(\gamma_0, \gamma_1) \in \Gamma_{\mathsf{d}_{a,s}}$. This implies that for every $0 \leq r \leq t \leq 1$, $(\gamma_r, \gamma_t) \in \Gamma_{d_{a,s}}$. $\qquad\square$

By Theorem 4.6 we have that, choosing $u = d_{a,s}$, the following disintegration formula holds

$$\mathfrak{m} \llcorner \mathcal{T}_{d_{a,s}} = \int_Q \hat{\mathfrak{m}}_\alpha^{a,s} \hat{\mathfrak{q}}^{a,s}(d\alpha), \tag{5.5}$$

where Q is a section of the partition of $\mathcal{T}_{d_{a,s}}^b$ given by the equivalence classes $\{R_{d_{a,s}}^b(\alpha)\}_{\alpha \in Q}$, and for $\hat{q}^{a,s}$-a.e. $\alpha \in Q$, $\hat{\mathfrak{m}}_\alpha^{a,s}$ is a probability measure supported on the transport ray $X_\alpha = R_{d_{a,s}}(\alpha)$ and $(X_\alpha, \mathsf{d}, \hat{\mathfrak{m}}_\alpha^{a,s})$ verifies $\mathsf{CD}(K,N)$. By Lemma 5.5, it follows that:

$$\mathfrak{m} \llcorner_{\mathsf{e}_{[0,1]}(G_{a,s})} = \int_Q \hat{\mathfrak{m}}_\alpha^{a,s} \llcorner_{\mathsf{e}_{[0,1]}(G_{a,s})} \hat{\mathfrak{q}}^{a,s}(d\alpha).$$

From the very definition of $G_{a,s}$ and the p-essentially non-branching property, in the previous disintegration formula the quotient set Q can be naturally identified with $\mathsf{e}_s(G_{a,s})$; moreover, we can consider the Borel parametrization

$$g^{a,s} : \mathsf{e}_s(G_{a,s}) \times [0,1] \to X, \quad g^{a,s}(\alpha, \cdot) = (\mathsf{e}_{s \llcorner G_{a,s}})^{-1}(\alpha),$$

yielding the following disintegration formula:

$$\mathfrak{m} \llcorner_{\mathsf{e}_{[0,1]}(G_{a,s})} = \int_{\mathsf{e}_s(G_{a,s})} g^{a,s}(\alpha, \cdot)_\sharp \left(h_\alpha^{a,s} \cdot \mathcal{L}^1 \llcorner_{[0,1]} \right) \mathfrak{q}^{a,s}(d\alpha), \tag{5.6}$$

where $\mathfrak{q}^{a,s}$ is a Borel measure concentrated on $\mathsf{e}_s(G_{a,s})$, and for $\mathfrak{q}^{a,s}$-a.e. $\alpha \in \mathsf{e}_s(G_{a,s})$, $h_\alpha^{a,s}$ is a $\mathsf{CD}(\ell_s(\alpha)^2 K, N)$ density on $[0,1]$. Notice that the factor $\ell_s(\alpha)^2 = \mathcal{H}^1(X_\alpha)^2$ is due to the reparametrization of the transport ray on $[0,1]$.

This permits, invoking Fubini's theorem, to reverse the order of integration so to have:

$$\mathfrak{m} \llcorner_{\mathsf{e}_{[0,1]}(G_{a,s})} = \int_{[0,1]} g^{a,s}(\cdot, t)_\sharp (h_\cdot^{a,s}(t) \cdot \mathfrak{q}^{a,s}) \mathcal{L}^1(dt) = \int_{[0,1]} \mathfrak{m}_t^{a,s} \mathcal{L}^1(dt), \tag{5.7}$$

where we defined

$$\mathfrak{m}_t^{a,s} := g^{a,s}(\cdot, t)_\sharp (h_\cdot^{a,s}(t) \cdot \mathfrak{q}^{a,s}).$$

Finally, the previous disintegration formula does not change if we multiply and divide conditional measures by $h_\alpha^{a,s}(s)$; therefore, changing $\mathfrak{q}^{a,s}$, we can assume $h_\alpha^{a,s}(s) = 1$, yielding $\mathfrak{m}_s^{a,s} = \mathfrak{q}^{a,s}$ and

$$\mathfrak{m}_t^{a,s} := g^{a,s}(\cdot, t)_\sharp (h_\cdot^{a,s}(t) \cdot \mathfrak{m}_s^{a,s}). \tag{5.8}$$

Moreover (see [13, Proposition 10.7]), for any $s \in (0,1)$ and $a \in \varphi_s(\mathsf{e}_s(G))$, the map

$$(0,1) \ni t \mapsto \mathfrak{m}_t^{a,s}$$

is continuous in the weak topology and if $\mathfrak{m}(\mathsf{e}_{[0,1]}(G_{a,s})) > 0$, then $\mathfrak{m}_t^{a,s}(\mathsf{e}_t(G_{a,s})) > 0$, for all $t \in (0,1)$. Finally,

$$\forall t \in [0,1] \quad \mathfrak{m}_t^{a,s}(\mathsf{e}_t(G_{a,s})) = \|\mathfrak{m}_t^{a,s}\| \le C \, \mathfrak{m}(\mathsf{e}_{[0,1]}(G_{a,s})),$$

for some $C > 0$ depending only on K, N and $\{\ell(\gamma) : \gamma \in G_{a,s}\}$.

5.2 L^q partition

We will now consider a decomposition of $\mathfrak{m}$ into conditional measures induced by Kantorovich potentials.

Hence for any $s, t \in (0,1)$, let us consider $a \in \Phi_s^t(\mathsf{e}_t(G)) = \varphi_s(\mathsf{e}_s(G))$. With such a choice of a, the compact set $\mathsf{e}_t(G)$ admits a partition given by $\mathsf{e}_t(G) \cap \{\Phi_s^t = a\}_{a \in \mathbb{R}}$.

Continuity of Φ_s^t makes it possible to apply the Disintegration Theorem. Since $\mathfrak{m}[e_t(G)] < \infty$, there exists an essentially unique disintegration of $\mathfrak{m}\llcorner_{e_t(G)}$ strongly consistent with respect to the quotient map Φ_s^t:

$$\mathfrak{m}\llcorner_{e_t(G)} = \int_{\varphi_s(e_s(G))} \hat{\mathfrak{m}}_{a,s}^t \mathfrak{q}_s^t(da) \tag{5.9}$$

where $\mathfrak{q}_s^t = (\Phi_s^t)_\sharp \mathfrak{m}\llcorner_{e_t(G)}$ and $\hat{\mathfrak{m}}_{a,s}^t$ is a probability measure concentrated on the set $e_t(G) \cap \{\Phi_s^t = a\} = e_t(G_{a,s})$.

Notice that, as one would expect, being the image of a time propagation of an intermediate Kantorovich potential, the quotient set $\varphi_s(e_s(G))$ does not depend on t.

The next follows with no modification from [13, Proposition 10.8].

Proposition 5.6. *The following properties hold true:*

- *For any $s, t, \tau \in (0, 1)$, the quotient measures $\mathfrak{q}_s^t$ and $\mathfrak{q}_s^\tau$ are mutually absolutely continuous;*

- *For any $s, t \in (0, 1)$, the quotient measure $\mathfrak{q}_s^t$ is absolutely continuous with respect to Lebesgue measure $\mathcal{L}^1$ on $\mathbb{R}$.*

Employing what we obtained so far, we can rewrite (5.9) in the following way:

$$\mathfrak{m}\llcorner_{e_t(G)} = \int_{\varphi_s(e_s(G))} \mathfrak{m}_{a,s}^t \mathcal{L}^1(da), \tag{5.10}$$

where $\mathfrak{m}_{a,s}^t := (d\mathfrak{q}_s^t/d\mathcal{L}^1) \cdot \hat{\mathfrak{m}}_{a,s}^t$ is concentrated on $e_t(G_{a,s})$ for $\mathcal{L}^1$-a.e. $a \in \varphi_s(e_s(G))$.

Over the set $e_t(G)$ we also have the measure μ_t; as it can be lifted to the set $\mathrm{Geo}(X)$, it makes sense to notice that the family of sets $\{G_{a,s}\}_{a \in \mathbb{R}}$ provides a partition of G. Hence an application of the Disintegration Theorem guarantees the existence of an essentially unique disintegration of ν strongly consistent with respect $\varphi_s \circ e_s$:

$$\nu = \int_{\varphi_s(e_s(G))} \nu_{a,s} \mathfrak{q}_s^\nu(da) \tag{5.11}$$

where the probability measure $\nu_{a,s}$ is concentrated on $G_{a,s}$ for $\mathfrak{q}_s^\nu$-a.e. $a \in \varphi_s(e_s(G))$. In particular, $q_s^\nu(\varphi_s(e_s(G))) = ||\nu|| = 1$.

Multiplying (5.10) by ρ_t and applying $(e_t)_\sharp$ to (5.11) produces the same measure μ_t: this permits to deduce what follows. For all the missing details we refer to [13, Corollary 10.10].

Corollary 5.7. *We have the following*

1. *For any $s \in (0, 1)$, the quotient measure $\mathfrak{q}_s^\nu$ is mutually absolutely continuous with respect to $\mathfrak{q}_s^s$. In particular, it is absolutely continuous with respect to $\mathcal{L}^1$.*

2. *For any $s, t \in (0, 1)$ and $\mathcal{L}^1$-a.e. $a \in \varphi_s(e_s(G))$:*

$$\rho_t \cdot \mathfrak{m}_{a,s}^t = q_s^\nu(a) \cdot (e_t)_\sharp \nu_{a,s},$$

   where $q_s^\nu := d\mathfrak{q}_s^\nu/d\mathcal{L}^1$. In particular, $\mathfrak{m}_{a,s}^t$ and $(e_t)_\sharp \nu_{a,s}$ are mutually absolutely continuous for $\mathfrak{q}_s^\nu$-a.e. $a \in \varphi_s(e_s(G))$.

3. *For any $s \in (0, 1)$ and $\mathfrak{q}_s^\nu$-a.e. $a \in \varphi_s(e_s(G))$, the maps*

$$[0, 1] \ni t \mapsto \rho_t \cdot \mathfrak{m}_{a,s}^t, \quad [0, 1] \ni t \mapsto (e_t)_\sharp \nu_{a,s}$$

   coincide for $\mathcal{L}^1$-a.e. $t \in [0, 1]$ up to a positive multiplicative constant $C_{a,s}$ depending only on a, s.

5.3 Comparison between conditional measures

We will now link the seemingly unrelated disintegrations (5.7) and (5.10).

Observe that $\mathfrak{m}_{a,s}^t$ and $\mathfrak{m}_t^{a,s}$ are concentrated on $\mathrm{e}_t(G_{a,s})$, for each $t \in (0,1)$ for $\mathcal{L}^1$-a.e. $a \in \varphi_s(\mathrm{e}_s(G))$ and for each $a \in \varphi_s(\mathrm{e}_s(G))$ and all $t \in (0,1)$, respectively.

The common feature of the two families of conditional measures $\mathfrak{m}_{a,s}^t$ and $\mathfrak{m}_t^{a,s}$ is that they are both coming from a disintegration formula with quotient measure the Lebesgue measure. We can exploit this property in the next lemma.

Lemma 5.8. *For every $s, t \in (0,1)$ and $a \in \varphi_s(\mathrm{e}_s(G))$, the limit*

$$\mathfrak{m}_t^{a,s} = \lim_{\epsilon \to 0} \frac{1}{2\varepsilon} \mathfrak{m} \llcorner_{\mathrm{e}_{[t-\varepsilon, t+\varepsilon]}(G_{a,s})}$$

holds true in the weak topology.

Proof. Since $(0,1) \ni t \mapsto \mathfrak{m}_t^{a,s}$ is continuous in the weak topology, and so together with (5.7), we see that for any $f \in C_b(X)$:

$$\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_X f(z) \mathfrak{m} \llcorner_{\mathrm{e}_{[t-\varepsilon, t+\varepsilon]}(G_{a,s})}(dz) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} \left(\int_X f(z) \mathfrak{m}_\tau^{a,s}(dz) \right) \mathcal{L}^1(d\tau) = \int_X f(z) \mathfrak{m}_t^{a,s}(dz),$$

thereby concluding the proof. $\qquad\square$

We are now in position to compare $\mathfrak{m}_{a,s}^t$ and $\mathfrak{m}_t^{a,s}$ by comparing $\mathfrak{m}$ in a neighborhood of $\mathrm{e}_t(G_{a,s})$ obtained varying t and then varying a. We refer to [13, Theorem 11.3] for all the details in the case $q = 2$ and simply note that the argument works for any $q > 1$; (the main ingredients needed for the proof are the disintegration formulas (5.7), (5.10) and temporal regularity of Φ_s^t obtained in Section 3).

Theorem 5.9 (Relating factorization by potential values and by φ-Kantorovich geodesics via Fubini). *For any $s \in (0,1)$,*

$$\mathfrak{m}_s^{a,s} = \ell_s^p \cdot \mathfrak{m}_{a,s}^s, \quad \text{for} \quad \mathcal{L}^1\text{-a.e. } a \in \varphi_s(\mathrm{e}_s(G)).$$

Moreover, for any $s \in (0,1)$ and $\mathcal{L}^1$-a.e. $t \in (0,1)$ including at $t = s$, $\partial_t \Phi_s^t(x)$ exists and is positive, and for $\mathcal{L}^1$-a.e. $a \in \varphi_s(\mathrm{e}_s(G))$ and $\mathfrak{m}_{a,s}^t$-a.e. x we have:

$$\mathfrak{m}_t^{a,s} = \partial_t \Phi_s^t \cdot \mathfrak{m}_{a,s}^t, \quad \text{for} \quad \mathcal{L}^1\text{-a.e.} \quad a \in \varphi_s(\mathrm{e}_s(G)). \tag{5.12}$$

5.4 Change of variable formula

Building on Theorem 5.9, we are now in position to write the Jacobian associated to the evolution of μ_t as the product of two factors.

All the results obtained until now will be used to prove the following:

Theorem 5.10 (Change of variables formula). *Let $(X, \mathsf{d}, \mathfrak{m})$ be a p-essentially non branching m.m.s. satisfying $\mathsf{CD}_p(K, N)$ and assume it is also q-essentially non branching.*

Let us consider $\mu_0, \mu_1 \in \mathcal{P}_q(X, \mathsf{d}, \mathfrak{m})$ and let ν denote the unique element of $\mathrm{OptGeo}_q(\mu_0, \mu_1)$. Setting $\mu_t = (\mathrm{e}_t)_\sharp \nu \ll \mathfrak{m}$, we will consider the densities $\rho_t := d\mu_t/d\mathfrak{m}$, $t \in [0,1]$, given by Corollary 5.2.

Then for any $s \in (0,1)$, for $\mathcal{L}^1$-a.e. $t \in (0,1)$ and ν-a.e. $\gamma \in G_\varphi^+$, $\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t)$ exists and the following formula holds:

$$\frac{\rho_t(\gamma_t)}{\rho_s(\gamma_s)} = \frac{\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t)}{\ell^p(\gamma)} \cdot \frac{1}{h_{\gamma_s}^{\varphi_s(\gamma_s),s}(t)}. \tag{5.13}$$

Here $h_{\gamma_s}^{\varphi_s(\gamma_s),s}$ is the $\mathsf{CD}(\ell(\gamma)^2 K, N)$ density on $[0,1]$ from (5.6), renormalized in such a way $h_{\gamma_s}^{\varphi_s(\gamma_s),s}(s) = 1$. Finally, for all $\gamma \in G_\varphi^0$, it holds:

$$\rho_t(\gamma_t) = \rho_s(\gamma_s), \quad \forall t, s \in [0,1]. \tag{5.14}$$

Proof. By [13, Lemma 6.11] and the discussion below Definition 2.6, $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{MCP}(K, N)$ and Corollary 5.2 guarantees the existence of versions of the densities satisfying (5.4). For any $\varepsilon > 0$, there exists a *good* compact subset $G^\varepsilon \subset G_\varphi^+$ such that $\nu(G^\varepsilon) \geq \nu(G_\varphi^+) - \varepsilon$ and such that G_ε increases along a sequence of ε decreasing to 0. Fixing $\varepsilon > 0$ on this sequence and the good subset G^ε, let us set

$$\nu^\varepsilon = \frac{1}{\nu(G^\varepsilon)}\nu \llcorner G^\varepsilon, \quad \mu_t^\varepsilon := (\mathrm{e}_t)_\sharp \nu^\varepsilon \ll \mathfrak{m}.$$

In particular we have that $\mu_t^\varepsilon = \frac{1}{\nu(G^\varepsilon)}\mu \llcorner \mathrm{e}_t(G_\varepsilon)$, for all $t \in [0,1]$ and therefore:

$$\mu_t^\varepsilon = \rho_t^\varepsilon \mathfrak{m}, \quad \rho_t^\varepsilon := \frac{1}{\nu(G^\varepsilon)}\rho_t|_{\mathrm{e}_t(G^\varepsilon)}, \ \forall t \in [0,1].$$

As we proved in Corollary 5.7, for each $s \in (0,1)$ and $\mathfrak{q}_s^{\varepsilon,s}$-a.e. $a \in \varphi_s(\mathrm{e}_s(G^\varepsilon))$, the map $[0,1] \ni t \mapsto \rho_t \cdot \mathfrak{m}_{a,s}^{\varepsilon,t}$ coincides for $\mathcal{L}^1$-a.e. $t \in [0,1]$ with the geodesic $t \mapsto (\mathrm{e}_t)_\sharp \nu_{a,s}^\varepsilon$ up to a constant $C_{a,s}^\varepsilon > 0$. Hence, for such s and a, for $\mathcal{L}^1$ a.e $t \in [0,1]$, we have that for any Borel set $H \subset G^\varepsilon$ the quantity

$$\int_{\mathrm{e}_t(H)} \rho_t^\varepsilon(x)\mathfrak{m}_{a,s}^{\varepsilon,t}(dx) = C_{a,s}^\varepsilon \int_{\mathrm{e}_t(H)} (\mathrm{e}_t)_\sharp \nu_{a,s}^\varepsilon(dx) = C_{a,s}^\varepsilon \nu_{a,s}^\varepsilon(H) \tag{5.15}$$

is constant in t, where in the last equality we used the injectivity of the map $\mathrm{e}_t : G^\varepsilon \to X$. By Theorem 5.9, for $\mathcal{L}^1$-a.e. $t \in (0,1)$ and $\mathcal{L}^1$-a.e. $a \in \varphi_s(G_s^\varepsilon)$, $\partial_t\Phi_s^t(x)$ exists and is positive for $\mathfrak{m}_{a,s}^{\varepsilon,t}$-a.e. x; moreover (5.12) holds. Thus, for all a, s and t for which the previous condition and (5.15) hold, we have

$$C_{a,s}^\varepsilon \nu_{a,s}^\varepsilon(H) = \int_{\mathrm{e}_t(H)} \rho_t^\varepsilon(x)\mathfrak{m}_{a,s}^{\varepsilon,t}(dx) = \int_{\mathrm{e}_t(H)} \rho_t^\varepsilon(x)(\partial_t\Phi_s^t(x))^{-1}\mathfrak{m}_t^{\varepsilon,a,s}(dx) \tag{5.16}$$

$$= \int_{\mathrm{e}_s(H)} \rho_t^\varepsilon(g^{a,s}(\alpha,t))(\partial_\tau|_{\tau=t}\Phi_s^\tau(g^{a,s}(\alpha,t)))^{-1}h_\alpha^{a,s}(t)\mathfrak{m}_s^{\varepsilon,a,s}(d\alpha)$$

$$= \int_{\mathrm{e}_s(H)} \rho_t^\varepsilon(g^{a,s}(\alpha,t))(\partial_\tau|_{\tau=t}\Phi_s^\tau(g^{a,s}(\alpha,t)))^{-1}h_\alpha^{a,s}(t)\ell_s^p(\alpha)\mathfrak{m}_{a,s}^{\varepsilon,s}(d\alpha)$$

where the two last equalities follow from (5.8) and Theorem 5.9, respectively.

Since the left-hand side of (5.16) does not depend on t, it follows that for all $s \in (0,1)$ and for $q_s^{\varepsilon,s}$-a.e. $a \in \varphi_s(\mathrm{e}_s(G^\varepsilon))$, there exists a subset $T \subset (0,1)$ of full $\mathcal{L}^1$ measure such that for all $H \subset G_{a,s}^\varepsilon$ the map

$$T \ni t \mapsto \int_{\mathrm{e}_s(H)} \rho_t^\varepsilon(g^{a,s}(\alpha,t))(\partial_\tau|_{\tau=t}\Phi_s^\tau(g^{a,s}(\alpha,t)))^{-1}h_\alpha^{a,s}(t)\ell_s^p(\alpha)\mathfrak{m}_{a,s}^{\varepsilon,s}(d\alpha),$$

is constant. In particular, since any Borel subset of $e_s(G_{a,s})$ can be written in the form $e_s(H)$, we have that for $t, t' \in T$

$$\rho_{t'}^\varepsilon(\gamma_{t'})(\partial_\tau|_{\tau=t'}\Phi_s^\tau(\gamma_{t'}))^{-1}h_{\gamma_s}^{a,s}(t') = \rho_t^\varepsilon(\gamma_t)(\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t))^{-1}h_{\gamma_s}^{a,s}(t), \tag{5.17}$$

for $\mathfrak{m}_{a,s}^{\varepsilon,s}$-a.e. $\alpha \in e_s(G_{a,s}^\varepsilon)$ where $\gamma = e_s^{-1}(\alpha) = g^{a,s}(\alpha, \cdot) \in G_{a,s}^\varepsilon$, with the exceptional set depending on t, t'. Recall that, by Corollary 5.7, given $t' \in T$, $\partial_\tau|_{\tau=t'}\Phi_s^\tau(\gamma_{t'}^\alpha)$ exists for $\mathfrak{m}_{a,s}^{\varepsilon,s}$-a.e. $\alpha \in e_s(G_{a,s}^\varepsilon)$. Thus, in particular, the equality (5.17) holds for a countable sequence of $\{t'\} \subset T$ dense in $(0,1)$. Using the normalization $h_{\gamma_s}^{a,s}(s) = 1$, the continuity of $h_{\gamma_s}^{a,s}(\cdot)$, $\rho_\cdot^\varepsilon(\gamma_\cdot)$ and the fact that

$$\lim_{T \ni t' \to s} \partial_\tau|_{\tau=t'}\Phi_s^\tau(\gamma_{t'}^\alpha) = \ell_s(\gamma_s^\alpha)^p = \ell(\gamma^\alpha)^p,$$

it is possible to pass to the limit for $t' \to s$ in (5.17)

$$\rho_s^\varepsilon(\gamma_s)\ell(\gamma)^{-p} = \rho_t^\varepsilon(\gamma_t)(\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t))^{-1}h_{\gamma_s}^{a,s}(t), \tag{5.18}$$

for $\mathfrak{m}_{a,s}^{\varepsilon,s}$-a.e. $\alpha \in e_s(G_{a,s}^\varepsilon)$, with $\gamma = e_s^{-1}(\alpha) \in G_{a,s}^\varepsilon$.

By corollary 5.7, the measures $\mathfrak{m}_{a,s}^{\varepsilon,s}$ and $(e_s)_\sharp \nu_{a,s}^\varepsilon$ are mutually absolutely continuous for $\mathfrak{q}_s^{\varepsilon,s}$-a.e. $a \in \varphi_s(e_s(G^\varepsilon))$. In particular, this implies that for all $s \in (0,1)$, for $\mathfrak{q}_s^{\varepsilon,s}$-a.e. $a \in \varphi_s(e_s(G^\varepsilon))$ and $\mathcal{L}^1$-a.e. $t \in (0,1)$, the equality (5.17) holds for $\nu_{a,s}$-a.e. γ. By Corollary 5.7, it follows that the measures $\mathfrak{q}_s^{\varepsilon,s}$ and $\mathfrak{q}_s^{\varepsilon,\nu}$ are mutually absolutely continuous; thus, by the disintegration formula (5.11), it follows that for all $s \in (0,1)$ and $\mathcal{L}^1$-a.e. $t \in (0,1)$:

$$\rho_s^\varepsilon(\gamma_s)\ell(\gamma)^{-p} = \rho_t^\varepsilon(\gamma_t)(\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t))^{-1}h_{\gamma_s}^{\varphi_s(\gamma_s),s}(t),$$

for ν-a.e. $\gamma \in G^\varepsilon$. Passing to the limit as $\varepsilon \to 0$ along the chosen sequence, it turns out that all $s \in (0,1)$, $\mathcal{L}^1$-a.e. $t \in (0,1)$ and ν-a.e. $\gamma \in G_\varphi^+$ satisfy

$$\rho_s(\gamma_s)\ell(\gamma)^{-p} = \rho_t(\gamma_t)(\partial_\tau|_{\tau=t}\Phi_s^\tau(\gamma_t))^{-1}h_{\gamma_s}^{\varphi_s(\gamma_s),s}(t).$$

By Fubini 's Theorem, for ν-a.e. $\gamma \in G_\varphi^+$, we have that (5.13) holding for $\mathcal{L}^1$-a.e. $s, t \in (0,1)$. $\qquad\square$

Remark 5.11. All of the results of this section also hold for $\bar{\Phi}_s^t$ in place of Φ_s^t. Indeed, recall that for all $x \in X$, $\Phi_s^t(x) = \bar{\Phi}_s^t(x)$ for $t \in \mathring{G}_\varphi(x)$, and that by Proposition 3.22, $\partial_t \Phi_s^t(x) = \partial_t \bar{\Phi}_s^t(x)$ for a.e. $t \in \mathring{G}_\varphi(x)$. As these were the only two properties used in the above derivation the assertion follows.

By Proposition 3.22, we know that the differentiability points of $\tau \mapsto \tilde{\Phi}_s^\tau(x)$ and $\tau \mapsto \tilde{\ell}_\tau^p(x)$ coincide for all $\tau \neq s$ and at these points

$$\partial_\tau \tilde{\Phi}_s^\tau(x) = \tilde{\ell}_\tau^p(x) + (\tau - s)\partial_\tau \frac{\tilde{\ell}_\tau^p(x)}{p}.$$

Hence by Remark 5.11, we deduce that for ν-a.e. geodesic $\gamma \in G_\varphi^+$ and for a.e. $t \in (0,1)$ both quantities

$$\partial_\tau|_{\tau=t}\ell_\tau^p(\gamma_t) = \partial_\tau|_{\tau=t}\bar{\ell}_\tau^p(\gamma_t)$$

exist and coincide. We can therefore rewrite the change of variable formula in the following way: for ν-a.e. geodesic $\gamma \in G_\varphi^+$

$$\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{h_{\gamma_s}^{\varphi_s(\gamma_s),s}(t)}{1 + (t-s)\frac{\partial_\tau|_{\tau=t}\ell_\tau^p(\gamma_t)}{p\ell(\gamma)^p}} = \frac{h_{\gamma_s}^{\varphi_s(\gamma_s),s}(t)}{1 + (t-s)\partial_\tau|_{\tau=t}\log\bar{\ell}_\tau(\gamma_t)}, \quad \text{for a.e. } t, s \in (0,1). \tag{5.19}$$

For sake of brevity, once the geodesic γ is fixed, we will use the following notation: $\rho(t) = \rho_t(\gamma_t)$, $h_s(t) := h_{\gamma_s}^{\varphi_s(\gamma_s)}(t)$ and $K_0 = K \cdot \ell(\gamma)^2$. We recall that, by Corollary 5.2 and (5.6), given by Theorem 4.6, the following properties hold true for ν-a.e $\gamma \in G_\varphi^+$:

(A) $(0,1) \ni t \mapsto \rho(t)$ is locally Lipschitz and strictly positive.

(B) For all $s \in (0,1)$, h_s is a $\mathsf{CD}(K_0, N)$ density on $[0,1]$ satisfying $h_s(s)=1$.

Fix now a geodesic $\gamma \in G_\varphi^+$ satisfying the change of variable formula (5.21), (A), (B) above.

The formula (5.19) implies that there exists a set $I \subset (0,1)$ of full measure such that for all $s \in I$ the functions

$$t \mapsto \partial_\tau|_{\tau=t} \frac{\tilde{\ell}_\tau^p / p(\gamma_t)}{\tilde{\ell}(\gamma)^p}, \quad t \mapsto z_s(t) := \frac{\frac{\rho(t)}{\rho(s)} h_s(t) - 1}{t - s}$$

coincide a.e. on $(0,1)$ for both $\tilde{\ell} \in \{\ell, \bar{\ell}\}$, with z_s defined on $(0,1) \setminus \{s\}$. Hence, by continuity, the functions $\{z_s\}_{s \in I}$ must all coincide, where defined, with a unique function $t \mapsto z(t)$ defined on $(0,1)$ such that

$$z(t) = \frac{\partial}{\partial \tau}\bigg|_{\tau=t} \log \ell_\tau(\gamma_t) = \frac{\partial}{\partial \tau}\bigg|_{\tau=t} \log \bar{\ell}_\tau(\gamma_t), \quad \text{for a.e. } t \in (0,1). \tag{5.20}$$

Since $\mathsf{CD}(K, N)$ densities are locally Lipschitz in the interior of the domain where they are defined, we see that z is locally Lipschitz in $(0,1)$ from (5.21). Combining (5.20) with the third order information provided by Theorem 3.19 (up to constant factors) yields:

(C) $(0,1) \ni t \mapsto z(t)$ is locally Lipschitz. Moreover, for any $\delta \in (0,1/2)$ there exists $C_\delta > 0$ so that:

$$\frac{z(t) - z(s)}{t - s} \geq (1 - C_\delta(t - s))|z(s)||z(t)|, \quad \forall 0 < \delta \leq s < t \leq 1 - \delta < 1.$$

In particular, $z'(t) \geq z^2(t)$ for a.e. $t \in (0,1)$.

To summarize, the change of variable formula can be rewritten in the following form:

$$\frac{\rho(s)}{\rho(t)} = \frac{h_s(t)}{1 + (t - s)z(t)}, \quad \text{for all } t, s \in (0,1), \tag{5.21}$$

where $z(t)$ coincides for all $t \in (0,1)$ with the second Peano derivative of $\tau \mapsto \varphi_\tau(\gamma_t)$ and of $\tau \mapsto \bar{\varphi}_\tau(\gamma_t)$ at $\tau = t$. These second Peano derivatives exist for all $t \in (0,1)$ and are a continuous function. We are therefore in position to obtain the aforementioned factorization of the "Jacobian". It has been already proved in [13] (see Theorem 12.3) that properties (A), (B), (C) together with the change of variable formula (5.21) are enough to obtain a factorization of the real function $1/\rho(t)$ into a product $L(t)Y(t)$, in which the first factor $L(t)$ is concave due to dilational and dimensional effects (analogous to the Brunn-Minkowski inequality on $(\mathbb{R}^n, |\cdot|, \mathcal{L}^n)$), while the latter term $Y(t)$ captures the effects of the curvature of $(X, \mathsf{d}, \mathfrak{m})$. In the smooth case $\rho(t)^{-1/n}$ would be interpreted as the mean-free path between particles during transport.

Theorem 5.12 (Isolating curvature effects in the volume distortion along the direction transported [13, Theorem 12.3]). *If the change of variable formula* (5.21) *holds and the properties (A), (B), (C) are satisfied, then*

$$\frac{1}{\rho_t(\gamma_t)} = L(t)Y(t) \quad \forall t \in (0,1),$$

where L is concave and Y is a $\mathsf{CD}(K_0, N)$ density on $(0,1)$.

5.5 Main Theorems

Finally, putting together the result proved so far in Section 4 and Section 5 we close the circle by proving:

Theorem 5.13 (Non-branching CD_p spaces are CD^1_{Lip} hence CD_q). *Let $(X, \mathsf{d}, \mathfrak{m})$ be a p-essentially non-branching m.m.s. verifying $\mathsf{CD}_p(K, N)$ for some $p > 1$. If $(X, \mathsf{d}, \mathfrak{m})$ is also q-essentially non-branching for some $q > 1$, then it verifies $\mathsf{CD}_q(K, N)$.*

Proof. Consider $\mu_0, \mu_1 \in \mathcal{P}_q(X, \mathsf{d}, \mathfrak{m})$. Recall that $\mathsf{CD}_p(K, N)$ implies (X, d) to be a geodesic space, hence the same is true for $(\mathcal{P}_q(X), W_q)$. Moreover, it implies (X, d) is $\mathsf{MCP}(K, N)$, hence qualitatively non-degenerate. Since $(X, \mathsf{d}, \mathfrak{m})$ is assumed to be q-essentially non-branching, Theorem 2.7 yields a unique $\nu \in \mathrm{OptGeo}_q(\mu_0, \mu_1)$ and

$$[0, 1] \ni t \mapsto \mu_t := (\mathrm{e}_t)_\sharp \nu \ll \mathfrak{m}.$$

Let $\rho_t := d\mu_t / d\mathfrak{m}$ be the versions of the densities guaranteed by Corollary 5.2.

Finally let $\varphi : X \to \mathbb{R}$ be a Kantorovich potential for the optimal transport problem from μ_0 to μ_1, with cost $c := \mathsf{d}^q / q$. Recall that $G_\varphi \subset \mathrm{Geo}(X)$ denote the set of (φ, q)-Kantorovich geodesics, i.e. all the geodesics γ for which

$$\varphi(\gamma_0) + \varphi^c(\gamma_1) = \frac{\mathsf{d}^q(\gamma_0, \gamma_1)}{q}.$$

As already observed, ν will be concentrated on $G_\varphi = G_\varphi^+ \cup G_\varphi^0$, where G_φ^+ and G_φ^0 denote the subsets of positive and zero length (φ, q)-Kantorovich geodesics respectively.

By the change of variables formula obtained in Theorem 5.10 (which relies on the $\mathsf{CD}^1_{Lip}(K, N)$ conclusion of Theorem 4.6), for ν-a.e. geodesic $\gamma \in G_\varphi^+$:

$$\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{h^{\varphi_s(\gamma_s), s}_{\gamma_s}(t)}{1 + (t - s)\frac{\partial_\tau|_{\tau = t}\ell^p_\tau(\gamma_t)}{p\ell(\gamma)^p}} = \frac{h^{\varphi_s(\gamma_s), s}_{\gamma_s}(t)}{1 + (t - s)\frac{\partial_\tau|_{\tau = t}\bar{\ell}^p_\tau(\gamma_t)}{p\ell(\gamma)^p}}, \quad \text{for a.e. } t, s \in (0, 1) \tag{5.22}$$

where for all $s \in (0, 1), h_s = h^{\varphi_s(\gamma_s), s}_{\gamma_s}$ is a $\mathsf{CD}(K_0, N)$ density, with $K_0 = \ell(\gamma)^2 K$ and $h_s(s) = 1$. Since Corollary 5.2 implies the Lipschitz regularity of $t \mapsto \rho_t(\gamma_t)$, assumptions (A) and (B) of the Theorem 5.12 are satisfied. Moreover, the third order information on the Kantorovich potential φ guarantees also the validity of the assumption (C) of the Theorem 5.12. Hence for ν-a.e. $\gamma \in G_\varphi^+$, it holds

$$\frac{1}{\rho_t(\gamma_t)} = L(t)Y(t), \ \forall t \in (0, 1)$$

where L is a concave function and Y is a $\mathsf{CD}(K_0, N)$ density on $(0, 1)$.

It is now a standard application of Hölder's inequality that gives us the validity of the $\mathsf{CD}_q(K, N)$ inequality along the W_q-geodesic μ_t: fix $t_0, t_1 \in (0, 1)$ and set $t_\alpha = \alpha t_1 + (1 - \alpha)t_0$, where $\alpha \in [0, 1]$. Using that $\sigma^{(\alpha)}_{K_0, N}(\theta) = \sigma^{(\alpha)}_{K, N}(\theta \ell(\gamma))$, it holds true:

$$
\begin{aligned}
\rho_{t_\alpha}^{-\frac{1}{N}}(\gamma_{t_\alpha}) &= L^{\frac{1}{N}}(t_\alpha)Y^{\frac{1}{N}}(t_\alpha) \\
&\geq \left(\alpha L(t_1) + (1 - \alpha)L(t_0)\right)^{\frac{1}{N}} \cdot \left(\sigma^{(\alpha)}_{K_0, N-1}(|t_1 - t_0|)Y^{\frac{1}{N-1}}(t_1) + \sigma^{(1-\alpha)}_{K_0, N-1}(|t_1 - t_0|)Y^{\frac{1}{N-1}}(t_0)\right)^{\frac{N-1}{N}} \\
&\geq \alpha^{\frac{1}{N}}\sigma^{(\alpha)}_{K_0, N-1}(|t_1 - t_0|)^{\frac{N-1}{N}}Y^{\frac{1}{N}}(t_1)L^{\frac{1}{N}}(t_1) + (1 - \alpha)^{\frac{1}{N}}\sigma^{(1-\alpha)}_{K_0, N-1}(|t_1 - t_0|)^{\frac{N-1}{N}}Y^{\frac{1}{N}}(t_0)L^{\frac{1}{N}}(t_0) \\
&= \alpha^{\frac{1}{N}}\sigma^{(\alpha)}_{K, N-1}(|t_1 - t_0|\ell(\gamma))^{\frac{N-1}{N}}\rho_{t_1}^{-\frac{1}{N}}(\gamma_{t_1}) + (1 - \alpha)^{\frac{1}{N}}\sigma^{(1-\alpha)}_{K, N-1}(|t_1 - t_0|\ell(\gamma))^{\frac{N-1}{N}}\rho_{t_0}^{-\frac{1}{N}}(\gamma_{t_0}) \\
&= \tau^{(\alpha)}_{K, N}(\mathsf{d}(\gamma_{t_0}, \gamma_{t_1}))\rho_{t_1}^{-\frac{1}{N}}(\gamma_{t_1}) + \tau^{(1-\alpha)}_{K, N}(\mathsf{d}(\gamma_{t_0}, \gamma_{t_1}))\rho_{t_0}^{-\frac{1}{N}}(\gamma_{t_0}).
\end{aligned} \tag{5.23}
$$

Recall that, by Corollary 5.2, the function $t \mapsto \rho_t(\gamma_t)$ is upper semi-continuous at the endpoints; so, it follows that for ν-a.e. $\gamma \in G_\varphi^+$ the inequality (5.23) holds true for all $t_0, t_1 \in [0,1]$. In particular, setting $t_0 = 0, t_1 = 1$, we have that for all $\alpha \in [0,1]$:

$$\rho_\alpha^{-\frac{1}{N}}(\gamma_\alpha) \geq \tau_{K,N}^{(\alpha)}(\mathsf{d}(\gamma_0,\gamma_1))\rho_1^{-\frac{1}{N}}(\gamma_1) + \tau_{K,N}^{(1-\alpha)}(\mathsf{d}(\gamma_0,\gamma_1))\rho_0^{-\frac{1}{N}}(\gamma_0); \tag{5.24}$$

the latter inequality being satisfied for ν-a.e.$\gamma \in G_\varphi^+$. We now claim that (5.24) is also satisfied for every $\gamma \in G_\varphi^0$, confirming in this way the validity of the $\mathsf{CD}(K,N)$ condition. Indeed, in this case the map $\alpha \mapsto \rho_\alpha(\gamma_\alpha)$ turns out to be constant by the Theorem 5.10 and then (5.24) is trivially satisfied as an equality, since $\tau_{K,N}^{(\alpha)}(0) = \alpha$, for every $\alpha \in [0,1]$. Thus, the claim. $\qquad\square$

Corollary 5.14 (Local-to-Global). *Fix any $p > 1$ and $K, N \in \mathbb{R}$ with $N > 1$. Let $(X, \mathsf{d}, \mathfrak{m})$ be a p-essentially non-branching metric measure space verifying $\mathsf{CD}_{p,loc}(K,N)$ and such that (X, d) is a length space with $\mathrm{spt}(\mathfrak{m}) = X$. Then $(X, \mathsf{d}, \mathfrak{m})$ verifies $\mathsf{CD}_p(K,N)$.*

References

[1] M. AGUEH: *Rates of decay to equilibria for p-Laplacian type equations.* Nonlinear Anal., **68** (2008), no. 7, 1909–1927.

[2] L. AMBROSIO AND J. FENG: *On a class of first order Hamilton-Jacobi equations in metric spaces.* In J. Differential Equations **256**, (2014), 2194–2245.

[3] L. AMBROSIO AND N. GIGLI: *A user's guide to optimal transport.* Modelling and Optimisation of Flows on Networks, Piccoli, B., Rascle, M. (eds), volume 2062 of *Lecture Notes in Math.*, pages 1–155. Springer, Heidelberg, 2013.

[4] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ: *Gradient flows in metric spaces and in the space of probability measure.* Second edition. Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[5] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ: *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below,* Invent. Math., **195**, 2, (2014), 289–391.

[6] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ: *Metric measure spaces with Riemannian Ricci curvature bounded from below,* Duke Math. J., **163**, (2014), 1405–1490.

[7] U. BESSI: *Hamilton-Jacobi in metric spaces with a homological terms.* Preprint at `arXiv:2001.11823`

[8] S. BIANCHINI AND L. CARAVENNA: *On the extremality, uniqueness and optimality of transference plans,* Bull. Inst. Math. Acad. Sin.(N.S.), **4**, (2009), 353–454.

[9] F. CAVALLETTI: *Monge problem in metric measure spaces with Riemannian curvature-dimension condition*, Nonlinear Anal. **99**, (2014), 136–151.

[10] F. CAVALLETTI: *Decomposition of geodesics in the Wasserstein space and the globalization property,* Geom. Funct. Anal., **24**, (2014), 493 – 551.

[11] F. CAVALLETTI, N. GIGLI AND F. SANTARCANGELO: *Displacement convexity of Entropy and the distance cost Optimal Transportation,* arXiv:2005.00243

[12] F. CAVALLETTI AND M. HUESMANN: *Existence and uniqueness of optimal transport maps,* Ann. I. H. Poincaré AN, **32**, (2015), 1367–1377.

[13] F. CAVALLETTI AND E. MILMAN: *The Globalization Theorem for the Curvature Dimension Condition,* preprint arXiv:1612.07623.

[14] F. CAVALLETTI AND A. MONDINO: *Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds,* Invent. Math., **208**, (2017), 803–849.

[15] F. CAVALLETTI AND A. MONDINO: *New formulas for the Laplacian of distance functions and applications,* Analysis & PDE, arXiv:1803.09687.

[16] F. CAVALLETTI AND A. MONDINO: *Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications,* arXiv:2004.08934.

[17] F. Cavalletti and K.-T. Sturm: *Local curvature-dimension condition implies measure-contraction property*, J. Funct. Anal., **262**, (2012), 5110–5127.

[18] D. Cordero-Erausquin, R.J. McCann and M. Schmuckenschläger: *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb.* Invent. Math. **146**, (2001), 219–257.

[19] L. C. Evans: *Partial Differential Equations Second Edition* volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2010

[20] L. C. Evans and W. Gangbo: *Differential equations methods for the Monge-Kantorovich mass transfer problem*, Mem. Amer. Math. Soc., **137**, (1999), 1–66.

[21] L. C. Evans and R. F. Gariepy: *Measure Theory and Fine Properties of Functions* volume 5 of *Studies in Advanced Mathematics*

[22] M. Erbar, K. Kuwada and K.T. Sturm: *On the Equivalence of the Entropic Curvature-Dimension Condition and Bochner's Inequality on Metric Measure Space*, Invent. Math., **201**, (2015), no. 3, 993–1071.

[23] D.H. Fremlin:, *Measure Theory*, volume 4. Torres Fremlin, (2002).

[24] W. Gangbo and A. Święch: *Metric viscosity solutions of Hamilton-Jacobi equations depending on local slopes.* Calc. Var. Partial Differential Equations, **54**, (2015), 1183–1218.

[25] N. Gozlan, C. Roberto and P.M. Samson: *Hamilton Jacobi equations on metric spaces and transport entropy inequalities.* Rev. Mat. Iberoam., **30**, (2014), 133–163.

[26] M. Kell: *Transport maps, non-branching sets of geodesics and measure rigidity*, Adv. Math., **320**, (2017), 520–573.

[27] M. Kell: *On interpolation and curvature via Wasserstein geodesics*, Adv. Calc. Var., **10**(2), (2017),125–167.

[28] B. Klartag: *Needle decomposition in Riemannian geometry*, Mem. Amer. Math. Soc., **249**, (2017), no. 1180, v + 77 pp.

[29] J.Lott and C.Villani: *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), **169**, (2009), 903–991.

[30] R.J. McCann: in preparation.

[31] R.J. McCann: *Displacement convexity of Boltzmann's entropy characterizes the strong energy condition from general relativity,* To appear in Camb. Math. J., Preprint at `arXiv1808.1536v2`.

[32] A. Mondino and S. Suhr: *An optimal transport formulation of the Einstein equations of general relativity,* Preprint at `arXiv:1810.13309v2`.

[33] S.-I. Ohta: *On the measure contraction property of metric measure spaces*, Comment. Math. Helv., **82**, (2007), 805–828.

[34] S.-I. Ohta: *Examples of spaces with branching geodesics satisfying the curvature-dimension condition*, Bull. Lond. Math. Soc. **46** (2014), no. 1, 19-25.

[35] F. Otto: *Doubly degenerate diffusion equations as steepest descent* Unpublished manuscript, circa 1998.

[36] B. Pass: *On the local structure of optimal measures in the multi-marginal optimal transportation problem*, Calc. Var. Partial Differential Equations **43**, (2012), 529–536.

[37] T. Rajala and K.T. Sturm: *Non-branching geodesics and optimal maps in strong* $\mathsf{CD}(K,\infty)$*-spaces*, Calc. Var. Partial Differential Equations, **50**, (2014), 831–846.

[38] M.-K. von Renesse and K.T. Sturm: Transport inequalities, gradient estimates, entropy and Ricci curvature. *Comm. Pure Appl. Math.* **58**, (2005), 923–940.

[39] S.M. Srivastava: *A course on Borel sets*, Graduate Texts in Mathematics, Springer 1998.

[40] K.T. Sturm: *On the geometry of metric measure spaces. I*, Acta Math. **196**, (2006), 65–131.

[41] K.T. Sturm: *On the geometry of metric measure spaces. II*, Acta Math. **196**, (2006), 133–177.

[42] C. Villani: *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[43] C. Villani: *Optimal transport - old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009.