The globalization theorem for the Curvature-Dimension condition

Fabio Cavalletti¹ · Emanuel Milman²

Received: 21 February 2020 / Accepted: 24 February 2021 / Published online: 2 August 2021 © The Author(s) 2021

Abstract The Lott–Sturm–Villani Curvature-Dimension condition provides a synthetic notion for a metric-measure space to have Ricci-curvature bounded from below and dimension bounded from above. We prove that it is enough to verify this condition locally: an essentially non-branching metric-measure space \((X, d, m)\) (so that \((\text{supp}(m), d)\) is a length-space and \(m(X) < \infty\)) verifying the local Curvature-Dimension condition \(\text{CD}_{\text{loc}}(K, N)\) with parameters \(K \in \mathbb{R}\) and \(N \in (1, \infty)\), also verifies the global Curvature-Dimension condition \(\text{CD}(K, N)\). In other words, the Curvature-Dimension condition enjoys the globalization (or local-to-global) property, answering a question which had remained open since the beginning of the theory. For the proof, we establish an equivalence between \(L^1\)- and \(L^2\)-optimal-transport–based interpolation. The challenge is not merely a technical one, and several new conceptual ingredients which are of independent interest are developed: an explicit change-of-variables formula for densities of Wasserstein geodesics

E. Milman: The research leading to these results is part of a project that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 637851).

¹ SISSA, 34136 Trieste, Italy
² Mathematics Department, Technion - I.I.T., 32000 Haifa, Israel

E. Milman: The research leading to these results is part of a project that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 637851).
depending on a second-order temporal derivative of associated Kantorovich potentials; a surprising third-order theory for the latter Kantorovich potentials, which holds in complete generality on any proper geodesic space; and a certain rigidity property of the change-of-variables formula, allowing us to bootstrap the a-priori available regularity. As a consequence, numerous variants of the Curvature-Dimension condition proposed by various authors throughout the years are shown to, in fact, all be equivalent in the above setting, thereby unifying the theory.

Contents

1 Introduction ....................................... 3
  1.1 Disentangling volume-distortion coefficients .......... 6
  1.2 Comparing $L^2$- and $L^1$-Optimal-Transport and main result .... 7
  1.3 Main new ingredients of proof ........................... 10
Part I Temporal theory of Optimal-Transport ..................... 15
  2 Preliminaries ....................................... 15
    2.1 Geodesics ...................................... 15
    2.2 Derivatives ..................................... 16
  3 Temporal theory of intermediate-time Kantorovich potentials: first and second order 18
    3.1 Hopf-Lax semi-group ............................... 19
    3.2 Distance functions ................................. 21
    3.3 Intermediate-time duality and time-reversed potential ....... 23
    3.4 Length functions $\ell_t^{\pm}$ and $\bar{\ell}_t^{\pm}$ .......... 26
    3.5 Null-geodesics ................................... 33
  4 Temporal theory of intermediate-time Kantorovich potentials: time-propagation ... 34
    4.1 Monotonicity .................................... 36
    4.2 Properties of $\Phi_1$ .................................. 38
  5 Temporal theory of intermediate-time Kantorovich potentials: third order 40
    5.1 Formal argument .................................. 41
    5.2 Notation ...................................... 42
    5.3 Main inequality ................................... 43
    5.4 Consequences ................................... 45
Part II Disintegration theory of Optimal-Transport ..................... 48
  6 Preliminaries ....................................... 48
    6.1 Geometry of Optimal-Transport on metric measure spaces ... 48
    6.2 Curvature-Dimension conditions .......................... 50
    6.3 Disintegration theorem ............................... 55
  7 Theory of $L^1$-Optimal-Transport ............................ 58
    7.1 Preliminaries .................................... 58
    7.2 Maximalty of transport rays on non-branched transport-set .... 61
  8 The $CD^1$ condition ................................ 65
    8.1 Definitions of $CD^1$ and $MCP^1$ .......................... 66
    8.2 $MCP^1$ implies $MCP_\epsilon$ .......................... 68
    8.3 On essentially non-branching spaces ...................... 72
  9 Temporal-regularity under $MCP$ ................................ 75
 10 Two families of conditional measures .......................... 84
    10.1 $L^1$ partition ..................................... 86
    10.2 $L^2$ partition ..................................... 93
11 Comparison between conditional measures ........................................... 97
   11.1 Equivalence of conditional measures ........................................... 98
   11.2 Change-of-variables formula .................................................. 103
Part III Putting it all together ..................................................... 107
12 Combining change-of-variables formula with Kantorovich 3rd order information .... 107
   12.1 Change-of-variables rigidity .................................................. 107
   12.2 Formal argument ...................................................................... 110
   12.3 Rigorous argument ............................................................... 112
13 Final results ................................................................................... 119
   13.1 Proof of the Main Theorem 1.1 .................................................. 119
   13.2 \( \text{RCD}(K, N) \) spaces ......................................................... 123
   13.3 Concluding remarks ............................................................... 124
A Appendix: One dimensional \( \text{CD}(K, N) \) densities .................... 126
   A.1 Differential characterization ...................................................... 127
   A.2 A-priori estimates ................................................................. 130
   A.3 Logarithmic convolutions ....................................................... 131
References ......................................................................................... 134

1 Introduction

The Curvature-Dimension condition \( \text{CD}(K, N) \) was first introduced in the 1980’s by Bakry and Émery [15, 16] in the context of diffusion generators, having in mind primarily the setting of weighted Riemannian manifolds, namely smooth Riemannian manifolds endowed with a smooth density with respect to the Riemannian volume. The \( \text{CD}(K, N) \) condition serves as a generalization of the classical condition in the non-weighted Riemannian setting of having Ricci curvature bounded below by \( K \in \mathbb{R} \) and dimension bounded above by \( N \in [1, \infty) \) (see e.g. [56, 60] for further possible extensions). Numerous consequences of this condition have been obtained over the past decades, extending results from the classical non-weighted setting and at times establishing new ones directly in the weighted one. These include diameter bounds, volume comparison theorems, heat-kernel and spectral estimates, Harnack inequalities, topological implications, Brunn–Minkowski-type inequalities, and isoperimetric, functional and concentration inequalities—see e.g. [17, 48, 77] and the references therein.

Being a differential and Hilbertian condition, it was for many years unclear how to extend the Bakry–Émery definition beyond the smooth Riemannian setting, as interest in (measured) Gromov-Hausdorff limits of Riemannian manifolds and other non-Hilbertian singular spaces steadily grew. In parallel, and apparently unrelatedly, the theory of Optimal-Transport was being developed in increasing generality following the influential work of Brenier [21] (see e.g. [2, 36, 53, 65, 75–77]). Given two probability measures \( \mu_0, \mu_1 \) on a common geodesic space \( (X, d) \) and a prescribed cost of transporting a single mass from point \( x \) to \( y \), the Monge-Kantorovich idea is to optimally couple \( \mu_0 \) and \( \mu_1 \) by minimizing the total transportation cost, and as a byproduct obtain
a Wasserstein geodesic \([0, 1] \ni t \mapsto \mu_t\) connecting \(\mu_0\) and \(\mu_1\) in the space of probability measures \(\mathcal{P}(X)\). This gives rise to the notion of displacement convexity of a given functional on \(\mathcal{P}(X)\) along Wasserstein geodesics, introduced and studied by McCann [52]. Following the works of Cordero-Erausquin–McCann–Schmuckenschläger [33], Otto–Villani [62] and von Renesse–Sturm [70], it was realized that the \(\text{CD}(K, \infty)\) condition in the smooth setting may be equivalently formulated synthetically as a certain convexity property of an entropy functional along \(W_2\) Wasserstein geodesics (associated to \(L^2\)-Optimal-Transport, when the transport-cost is given by the squared-distance function).

This idea culminated in the seminal works of Lott–Villani [51] and Sturm [73,74], where a synthetic definition of \(\text{CD}(K, N)\) was proposed on a general (complete, separable) metric space \((X, d)\) endowed with a (locally-finite Borel) reference measure \(m\) (“metric-measure space”, or m.m.s.); it was moreover shown that the latter definition coincides with the Bakry–Émery one in the smooth Riemannian setting (and in particular in the classical non-weighted one), that it is stable under measured Gromov-Hausdorff convergence of m.m.s.’s, and that it implies various geometric and analytic inequalities relating metric and measure, in complete analogy with the smooth setting. It was subsequently also shown [58,64] that Finsler manifolds and Alexandrov spaces satisfy the Curvature-Dimension condition. Thus emerged an overwhelmingly convincing notion of Ricci curvature lower bound \(K\) and dimension upper bound \(N\) for a general (geodesic) m.m.s. \((X, d, m)\), leading to a rich and fruitful theory exploring the geometry of m.m.s.’s by means of Optimal-Transport.

One of the most important and longstanding open problems in the Lott–Sturm–Villani theory (see [73,74] and [77, pp. 888, 907]) is whether the Curvature-Dimension condition on a general geodesic m.m.s. (say, having full-support \(\text{supp}(m) = X\)) enjoys the globalization (or local-to-global) property: if the \(\text{CD}(K, N)\) condition is known to hold on a neighborhood \(X_o\) of any given point \(o \in X\) (a property henceforth denoted by \(\text{CD}_{loc}(K, N)\)), does it also necessarily hold on the entire space? Clearly this is indeed the case in the smooth setting, as both curvature and dimension may be computed locally (by equivalence with the differential \(\text{CD}\) definition). However, for reasons which we will expand on shortly, this is not at all clear and in some cases is actually false on general m.m.s.’s. An affirmative answer to this question would immensely facilitate the verification of the \(\text{CD}\) condition, which at present requires testing all possible \(W_2\)-geodesics on \(X\), instead of locally on each \(X_o\). The analogous question for sectional curvature on Alexandrov spaces (where the dimension \(N\) is absent) does indeed have an affirmative answer, as shown by Topogonov, and in full generality, by Perelman (see [22]).

Several partial answers to the local-to-global problem have already been obtained in the literature. A geodesic space \((X, d)\) is called non-branching if geodesics are forbidden to branch at an interior-point into two separate
geodesics. On a non-branching geodesic m.m.s. \((X, d, m)\) having full support, it was shown by Sturm in [73, Theorem 4.17] that the local-to-global property is satisfied when \(N = \infty\) (assuming that the space of probability measures with finite \(m\)-relative entropy is geodesically convex; see also [77, Theorem 30.42] where the same globalization result was proved under a different condition involving the existence of a full-measure totally-convex subset of \(X\) of finite-dimensional points). Still for non-branching geodesic m.m.s.’s having full support, a positive answer was also obtained by Villani in [77, Theorem 30.37] for the case \(K = 0\) and \(N \in [1, \infty)\).

We stress that in these results, the restriction to non-branching spaces is not merely a technical assumption—an example of a heavily-branching m.m.s. verifying \(\text{CD}_{loc}(0, 4)\) which does not verify \(\text{CD}(K, N)\) for any fixed \(K \in \mathbb{R}\) and \(N \in [1, \infty)\) was constructed by Rajala in [67]. Consequently, a natural assumption is to require that \((X, d)\) be non-branching, or more generally, to require that the \(L^2\)-Optimal-Transport on \((X, d, m)\) be concentrated (i.e. up to a null-set) on a non-branching subset of geodesics, an assumption introduced by Rajala and Sturm in [68] under the name essentially non-branching (see Sect. 6 for precise definitions). For instance, it is known [68] that measured Gromov-Hausdorff limits of Riemannian manifolds satisfying \(\text{CD}(K, \infty)\), and more generally, \(\text{RCD}(K, \infty)\) spaces, always satisfy the essentially non-branching assumption (see Sect. 13).

In this work, we provide an affirmative answer to the globalization problem in the remaining range of parameters: for \(N \in (1, \infty)\) and \(K \in \mathbb{R}\), the \(\text{CD}(K, N)\) condition verifies the local-to-global property on an essentially non-branching geodesic m.m.s. \((X, d, m)\) having finite total-measure and full support. The exclusion of the case \(N = 1\) is to avoid unnecessary pathologies, and is not essential. Our assumption that \(m\) has finite total-measure (or equivalently, by scaling, that it is a probability measure) is most probably technical, but we did not verify it can be removed so as to avoid overloading the paper even further. This result is new even under the additional assumption that the space is infinitesimally Hilbertian (see [40])—we will say that such spaces verify \(\text{RCD}(K, \infty)\)—in which case the assumption of being (globally) essentially non-branching is in fact superfluous.

To better explain the difference between the previously known cases when \(\frac{K}{N} = 0\) and the conceptual challenge which the newly treated case \(\frac{K}{N} \neq 0\) poses, as well as to sketch our solution and its main new ingredients, which we believe are of independent interest, we provide some additional details below and refer to Sect. 6 for precise definitions.
1.1 Disentangling volume-distortion coefficients

Roughly speaking, the $\text{CD}(K, N)$ condition prescribes a synthetic second-order bound on how an infinitesimal volume changes when it is moved along a $W^2$-geodesic: the volume distortion (or transport Jacobian) $J$ along the geodesic should satisfy the following interpolation inequality for $t_0 = 0$ and $t_1 = 1$:

$$J^{\frac{1}{N}}(\alpha t_1 + (1 - \alpha)t_0) \geq \tau^{(\alpha)}_{K, N}(|t_1 - t_0| \theta) J^{\frac{1}{N}}(t_1) + \tau^{(1-\alpha)}_{K, N}(|t_1 - t_0| \theta) J^{\frac{1}{N}}(t_0) \quad \forall \alpha \in [0, 1],$$  

(1.1)

where $\tau^{(t)}_{K, N}(\theta)$ is an explicit coefficient depending on the curvature $K \in \mathbb{R}$, dimension $N \in [1, \infty]$, the interpolating time parameter $t \in [0, 1]$ and the total length of the geodesic $\theta \in [0, \infty)$ (with an appropriate interpretation of (1.1) when $N = \infty$). When $N < \infty$, the latter coefficient is obtained by geometrically averaging two different volume distortion coefficients:

$$\tau^{(t)}_{K, N}(\theta) := t^{\frac{1}{N}} \sigma^{(t)}_{K, N-1}(\theta)^{\frac{N-1}{N}},$$  

(1.2)

where the $\sigma^{(t)}_{K, N-1}(\theta)$ term encodes an $(N-1)$-dimensional evolution orthogonal to the transport and thus affected by the curvature, and the linear term $t$ represents a one dimensional evolution tangential to the transport and thus independent of any curvature information. As with the Jacobi equation in the usual Riemannian setting, the function $[0, 1] \ni t \mapsto \tau^{(t)}_{K, N}(\theta)$ does not satisfy a second-order differential characterization such as (1.3). If it did, it would be possible to express the interpolation inequality (1.1) on $[t_0, t_1] \subset [0, 1]$ as a second-order differential inequality for $J^{\frac{1}{N}}$ on $[t_0, t_1]$ (see Lemmas A.5 and A.6), and so if (1.1) were known to hold for all $\{[t_i^0, t_i^1]\}_{i=1}^k$ so that $\cup_{i=1}^k (t_i^0, t_i^1) = (0, 1)$, it would follow that (1.1) also holds for $[t_0, t_1] = [0, 1]$. However, a counterexample to the latter implication was constructed by Deng and Sturm in [34], thereby showing that:
the local-to-global property for \( \frac{K}{N} \neq 0 \), if true, cannot be obtained by a one-dimensional bootstrap argument on a single \( W_2 \)-geodesic as above, and must follow from a deeper reason involving a family of \( W_2 \)-geodesics simultaneously. (1.4)

On the other hand, the above argument does work if we were to replace \( \tau \) by the slightly smaller \( \sigma \) coefficients. This motivated Bacher and Sturm in [14] to define for \( K \in \mathbb{R} \) and \( N \in (1, \infty) \) the slightly weaker “reduced” Curvature-Dimension condition, denoted by \( \text{CD}^*(K, N) \), where the distortion coefficients \( \tau_{K,N}(t) \) are indeed replaced by \( \sigma_{K,N}(t) \). Using the above gluing argument (after resolving numerous technicalities), the local-to-global property for \( \text{CD}^*(K, N) \) was established in [14] on non-branching spaces (see also the work of Erbar–Kuwada–Sturm [35, Corollary 3.13, Theorem 3.14 and Remark 3.26] for an extension to the essentially non-branching setting, cf. [29,68]). Let us also mention here the work of Ambrosio–Mondino–Savaré [10], who independently of a similar result in [35], established the local-to-global property for \( \text{RCD}^*(K, N) \) proper spaces, \( K \in \mathbb{R} \) and \( N \in [1, \infty) \), without a-priori assuming any non-branching assumptions (but a-posteriori, such spaces must be essentially non-branching by [68]).

Without requiring any non-branching assumptions, the \( \text{CD}^*(K, N) \) condition was shown in [14] to imply the same geometric and analytic inequalities as the \( \text{CD}(K, N) \) condition, but with slightly worse constants (typically missing the sharp constant by a factor of \( \frac{N-1}{N} \)), suggesting that the latter is still the “right” notion of Curvature-Dimension. We conclude that the local-to-global challenge is to properly disentangle between the orthogonal and tangential components of the volume distortion \( J \) before attempting to individually integrate them as above. This also highlights the geometric nature of the globalization problem, and demonstrates that it is not merely a technical challenge.

1.2 Comparing \( L^2 \)- and \( L^1 \)-Optimal-Transport and main result

There have been a couple of prior attempts to disentangle the volume distortion into its orthogonal and tangential components, by comparing between \( W_2 \) and \( W_1 \) Wasserstein geodesics (associated to \( L^2 \)- and \( L^1 \)-Optimal-Transport, respectively). In [30], this strategy was implicitly employed by Cavalletti and Sturm to show that \( \text{CD}_{loc}(K, N) \) implies the measure-contraction property \( \text{MCP}(K, N) \), which in a sense is a particular case of \( \text{CD}(K, N) \) when one end of the \( W_2 \)-geodesic is a Dirac delta at a point \( o \in X \) (see [57,74]). In that case, all of the transport-geodesics have \( o \) as a common end point, so by considering a disintegration of \( m \) on the family of spheres centered at \( o \), and restricting the \( W_2 \)-geodesic to these spheres, the desired disentanglement was obtained. In the subsequent work [24], Cavalletti generalized this approach to a particular
family of $W_2$-geodesics, having the property that for a.e. transport-geodesic $\gamma$, its length $\ell(\gamma)$ is a function of $\varphi(\gamma_0)$, where $\varphi$ is a Kantorovich potential associated to the corresponding $L^2$-Optimal-Transport problem. Here the disintegration was with respect to the individual level sets of $\varphi$, and again the restriction of the $W_2$-geodesic enjoying the latter property to these level sets (formally of co-dimension one) induced a $W_1$-geodesic, enabling disentanglement.

Another application of $L^1$-Optimal-Transport, seemingly unrelated to disentanglement of $W_2$-geodesics, appeared in the recent breakthrough work of Klartag [47] on localization in the smooth Riemannian setting. The localization paradigm, developed by Payne–Weinberger [63], Gromov–Milman [44] and Kannan–Lovász–Simonovits [46], is a powerful tool to reduce various analytic and geometric inequalities on the space $(\mathbb{R}^n, d, m)$ to appropriate one-dimensional counterparts. The original approach by these authors was based on a bisection method, and thus inherently confined to $\mathbb{R}^n$. In [47], Klartag extended the localization paradigm to the weighted Riemannian setting, by disintegrating the reference measure $m$ on $L^1$-Optimal-Transport geodesics (or “rays”) associated to the inequality under study (cf. Feldman–McCann [38]), and proving that the resulting conditional one-dimensional measures inherit the Curvature-Dimension properties of the underlying manifold.

Klartag’s idea is quite robust, and permitted Cavalletti and Mondino in [27] to avoid the smooth techniques used in [47] and to extend the localization paradigm to the framework of essentially non-branching geodesic m.m.s.’s $(X, d, m)$ of full-support verifying $CD_{loc}(K, N), N \in (1, \infty)$. By a careful study of the structure of $W_1$-geodesics, Cavalletti and Mondino were able to transfer the Curvature-Dimension information encoded in the $W_2$-geodesics to the individual rays along which a given $W_1$-geodesic evolves, thereby proving that on such spaces,

$$\text{the conditional one-dimensional measures obtained by disintegration of } m \text{ on } L^1\text{-Optimal-Transport rays satisfy } CD(K, N). \quad (1.5)$$

Note that the densities of one-dimensional $CD(K, N)$ spaces are characterized via the $\sigma$ (as opposed to $\tau$) volume-distortion coefficients (see the “Appendix”), so by applying the gluing argument described in the previous subsection, only local $CD_{loc}(K, N)$ information was required in [27] to obtain global control over the entire one-dimensional transport ray.

This allowed Cavalletti and Mondino (see [27,28]) to obtain a series of sharp geometric and analytic inequalities for $CD_{loc}(K, N)$ spaces as above, in particular extending from the smooth Riemannian setting the sharp Lévy-Gromov [42] and Milman [55] isoperimetric inequalities, as well as the sharp Brunn-Minkowski inequality of Cordero-Erausquin–McCann–Schmuckenschläger [33] and Sturm [74], all in global form (see also Ohta [59]).
We would like to address at this point a certain general belief shared by some in the Optimal-Transport community, stating that the property $BM(K, N)$ of satisfying the Brunn-Minkowski inequality (with sharp coefficients correctly depending on $K, N$), should be morally equivalent to the $CD(K, N)$ condition. Rigorously establishing such an equivalence would immediately yield the local-to-global property of $CD(K, N)$, by the Cavalletti–Mondino localization proof that $CD_{loc}(K, N) \Rightarrow BM(K, N)$. However, we were unsuccessful in establishing the missing implication $BM(K, N) \Rightarrow CD(K, N)$, and in fact a careful attempt in this direction seems to lead back to the circle of ideas we were ultimately able to successfully develop in this work.

Instead of starting our investigation from $BM(K, N)$, our strategy is to directly start from a suitable modification of the property (1.5), which we dub $CD^1(K, N)$, when (1.5) is required to hold for transport rays associated to (signed) distance functions from level sets of continuous functions. A stronger condition, when (1.5) is required to hold for transport rays associated to all 1-Lipschitz functions, is denoted by $CD^1_{Lip}(K, N)$—see Sect. 8 for precise definitions. The main result of this work consists of showing that $CD_{loc}^1(K, N) \Rightarrow CD(K, N)$, by means of transferring the one-dimensional $CD(K, N)$ information encoded in a family of suitably constructed $L^1$-Optimal-Transport rays, onto a given $W_2$-geodesic, thereby obtaining the correct disentanglement between tangential and orthogonal distortions. This goes in exactly the opposite direction to the one studied by Cavalletti and Mondino in [27], and completes the cycle

$$CD_{loc}(K, N) \Rightarrow CD^1_{Lip}(K, N) \Rightarrow CD^1(K, N) \Rightarrow CD(K, N).$$

To the best of our knowledge, this decisive feature of our work—deducing $CD(K, N)$ for a given $W_2$-geodesic by considering the $CD_{loc}^1(K, N)$ information encoded in family (in accordance with (1.4)) of different associated $W_2$-geodesics (manifesting itself in the $CD^1(K, N)$ information along a family of different $L^1$-Optimal-Transport rays)—has not been previously explored.

**Main Theorem 1.1** Let $(X, d, m)$ be an essentially non-branching m.m.s. with $m(X) < \infty$, and let $K \in \mathbb{R}$ and $N \in (1, \infty)$. Then the following statements are equivalent:

1. $(X, d, m)$ verifies $CD(K, N)$.
2. $(X, d, m)$ verifies $CD^*(K, N)$.
3. $(X, d, m)$ verifies $CD^1_{Lip}(K, N)$.
4. $(X, d, m)$ verifies $CD^1(K, N)$.

If in addition $(\text{supp}(m), d)$ is a length-space, the above statements are equivalent to

5. $(X, d, m)$ verifies $CD_{loc}(K, N)$.
To this list one can also add the entropic Curvature-Dimension condition $\text{CD}^e(K, N)$ of Erbar–Kuwada–Sturm [35], which is known to be equivalent to $\text{CD}^*(K, N)$ for essentially non-branching spaces. In other words, all synthetic definitions of Curvature-Dimension are equivalent for essentially non-branching m.m.s.’s, and in particular, the local-to-global property holds for such spaces (recall that this is known to be false on m.m.s.’s where branching is allowed by [67]). The equivalence with $\text{CD}_{loc}(K, N)$ is clearly false without some global assumption ultimately ensuring that $(\text{supp}(m), d)$ is a geodesic-space, see Remark 13.4.

As already mentioned, and being slightly imprecise (see Sect. 13 for precise statements), the implications $\text{CD}(K, N) \Rightarrow \text{CD}^*(K, N) \Rightarrow \text{CD}_{loc}(K, N)$ follow from the work of Bacher and Sturm [14], and the implication $\text{CD}_{loc}(K, N) \Rightarrow \text{CD}_{Lip}^1(K, N)$ follows by adapting to the present framework what was already proved by Cavalletti and Mondino in [27] (after taking care of the important maximality requirement of transport-rays, see Theorem 7.10). So almost all of our effort goes into proving that $\text{CD}^1(K, N) \Rightarrow \text{CD}(K, N)$.

For a smooth weighted Riemannian manifold $(M, d, m)$, it is an easy exercise to show the latter implication using the Bakry–Émery differential characterization of $\text{CD}(K, N)$—simply use an appropriate umbilic hypersurface $H$ passing through a given point $p \in M$ and perpendicular to a given direction $\xi \in T_p M$, and apply the $\text{CD}^1(K, N)$ definition to the distance function from $H$. Of course, this provides no insight towards how to proceed in the m.m.s. setting, so it is natural to try and obtain an alternative synthetic proof, still in the smooth setting. While this is possible, it already poses a much greater challenge, which in some sense provided the required insight leading to the strategy we ultimately employ in this work.

### 1.3 Main new ingredients of proof

To achieve the right disentanglement, we are required to develop several new ingredients beyond the present state-of-the-art, which, being conceptual in nature, are in our opinion of independent interest.

1. The first is a change-of-variables formula for the density of an $L^2$-Optimal-Transport geodesic in $X$ (see Theorem 11.4), which depends on a second-order derivative of associated interpolating Kantorovich potentials.

Let $\text{Geo}(X)$ denote the collection of constant speed geodesics on $X$ parametrized on the interval $[0, 1]$, and let $e_t : \text{Geo}(X) \ni \gamma \mapsto \gamma_t \in X$ denote the evaluation map at time $t \in [0, 1]$. Given two Borel probability measures $\mu_0, \mu_1 \in \mathcal{P}(X)$ with finite second moments, any $W_2$-geodesic $[0, 1] \ni t \mapsto \mu_t \in \mathcal{P}(X)$ can be lifted to an optimal dynamical plan $\nu \in \mathcal{P}(\text{Geo}(X))$, so that $(e_t)_\ast \nu = \mu_t$ for all $t \in [0, 1]$. Let $\varphi$ denote a
The globalization theorem

Kantorovich potential associated to the $L^2$-transport problem between $\mu_0$ and $\mu_1$. Given $s, t \in (0, 1)$, we introduce the time-propagated intermediate Kantorovich potential $\Phi_s^t$ by pushing forward $\phi_s$ via $e_t \circ e_s^{-1}$, where $\{\phi_t\}_{t \in [0, 1]}$ is the family of interpolating Kantorovich potentials obtained via the Hopf–Lax semi-group applied to $\phi$. While $e_t^{-1}$ may be multi-valued, Theorem 3.11 ensures that $\Phi_s^t = \phi_s \circ e_s \circ e_t^{-1}$ is well-defined on $e_t(G_\phi)$, the set of $t$-mid-points of transport geodesics.

Theorem 11.4 states that if $(X, d, m)$ is an essentially non-branching m.m.s. verifying $CD^1(K, N)$ ($m(X) < \infty$ and $N \in (1, \infty)$), and if $\mu_0, \mu_1 \ll m$, then for $\nu$-a.e. transport-geodesic $\gamma \in Geo(X)$ of positive length:

$$\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{\ell^2(\gamma)}{\partial_t |_{\tau=t} \Phi_s^\tau(\gamma_\tau)} \cdot h_s^\gamma(t) \quad \text{for a.e. } t, s \in (0, 1), \quad (1.6)$$

where $\rho_t$ are appropriate versions of the densities $d\mu_t/dm$, and for every $s \in (0, 1)$, $h_s^\gamma$ is a $CD(\ell(\gamma))^2(K, N)$ density on $[0, 1]$ so that $h_s^\gamma(s) = 1$. In particular, for a.e. $t, s \in (0, 1)$, $\partial_t |_{\tau=t} \Phi_s^\tau(\gamma_\tau)$ exists and is positive. Here $h_s^\gamma$ is obtained from the $CD^1(K, N)$ condition applied to the transport-ray associated to the (signed) distance function from the level set $\{\phi_s = \phi_s(\gamma_s)\}$.

Theorem 11.4 constitutes the culmination of Part II of this work, which is mostly dedicated to introducing the $CD^1(K, N)$ condition and rigorously establishing the change-of-variables formula (1.6). Note that we refrain from making any assumptions on (the challenging) spatial regularity of $\Phi_s^t$ when $t \neq s$, so we are precluded from invoking the coarea formula in our derivation. Our main tool for deriving (1.6) is a comparison between two disintegrations of appropriate measures, one encoding $W^2$ information and another encoding $W^1$ information—see Sect. 11 for a heuristic derivation.

(2) To obtain disentanglement of the “Jacobian” $t \mapsto 1/\rho_t(\gamma_t)$ into its orthogonal and tangential components, we need to understand the first-order variation of the change-of-variables formula (1.6) at $t = s$, i.e. the second-order variation of $t \mapsto \Phi_s^t$ at $t = s$, which amounts to a third-order variation of $t \mapsto \phi_t$. Our second main new ingredient in this work is a surprising third-order bound on the variation of $t \mapsto \phi_t$ along the Hopf–Lax semi-group (Theorem 5.5), which holds in complete generality on any proper geodesic space.

To this end, we develop in Part I of this work a first, second, and finally third order temporal theory of intermediate Kantorovich potentials in a purely metric setting $(X, d)$, without specifying any reference measure $m$ and without assuming any non-branching assumptions. This part, which
may be read independently of the other components of this work, is presented first (in Sects. 2–5), since its results are constantly used throughout the rest of this work.

Our starting point here is the pioneering work by Ambrosio–Gigli–Savaré [5], [6, Section 3], who already investigated in a very general (extended) metric space setting the first and second order temporal behaviour of the Hopf-Lax semi-group $Q_t$ applied to a general function $f : X \to \mathbb{R} \cup \{+\infty\}$. However, the essential point we observe in our treatment is that when $f$ is itself a Kantorovich potential $\varphi$, characterized by the property that $\varphi = Q_1(-\varphi^c)$ and $\varphi^c = Q_1(-\varphi)$, much more may be said regarding the behaviour of $t \mapsto \varphi_t := -Q_t(-\varphi)$, even in first and second order. This is due to the fact that if we reverse time and define $\bar{\varphi}_t := Q_{1-t}(-\varphi^c)$, then we obtain two-sided control over $\varphi_t$ on the set $\{\varphi_t = \bar{\varphi}_t\}$, which turns out to coincide with the set $e_t(G_{\varphi})$. So for instance, two apparently novel observations which we constantly use throughout this work are that for all $t \in (0, 1)$, $\ell_t^2/2 := \partial_t \varphi_t$ exists on $e_t(G_{\varphi})$, and that transport geodesics having a given $x \in X$ as their $t$-midpoint all have the same length $\ell_t(x)$. In Sect. 3, we establish Lipschitz regularity properties of $t \mapsto \ell_t^2(x)$ for all $x \in X$, as well as upper and lower derivative estimates, both pointwise and a.e., for appropriate times $t$. These are then transferred in Sect. 4 to corresponding estimates for the function $\Phi_t^\prime$.

Part I culminates in Sect. 5, whose goal is to prove a quantitative version of the following (somewhat oversimplified) statement, which crucially provides second order information on $\ell_t$, or equivalently, third order information on $\varphi_t$, along $\gamma_t$:

$$\text{If } \frac{1}{\ell(\gamma)^2} \partial_t |_{\tau=t} \ell_{\tau}^2(\gamma_{\tau}) \text{ exists a.e. in } t \in (0, 1) \text{ and coincides with an absolutely continuous function } z, \text{ then } z'(t) \geq z(t)^2 \text{ for a.e. } t \in (0, 1).$$

(1.7)

Equivalently, this amounts to the statement that:

$$(0, 1) \ni r \mapsto L(r) := \exp\left(-\frac{1}{\ell(\gamma)^2} \int_{\tau=0}^{r} \partial_t |_{\tau=t} \ell_{\tau}^2(\gamma_{\tau}) dt\right) \text{ is concave},$$

(1.8)

since (formally)

$$\frac{L''}{L} = (\log L)'' + ((\log L)')^2 = -z' + z^2 \leq 0.$$

It turns out that $L(t)$ precisely corresponds to the tangential component of $1/\rho_t(\gamma_t)$, and its concavity ensures that it is synthetically controlled by
the linear term appearing in the definition of \( \tau_{K,N}(\theta) \) in (1.2). The novel observation that it is possible to extract in a general metric setting third order information from the Hopf-Lax semi-group, which formally solves the first-order Hamilton-Jacobi equation, is in our opinion one of the most surprising parts of this work. Even in the smooth Riemannian setting, we were not able to find a synthetic proof which is easier than the one in the general metric setting; a formal differential proof of (1.7) assuming both temporal and (more challenging) spatial higher-order regularity of \( \varphi_t \) is provided in Sect. 5.1, but the latter seems to wrongly suggest that it would not be possible to extend (1.7) beyond a Hilbertian setting. Our proof in the general metric setting (Theorem 5.2) is based on a careful comparison of second order expansions of \( \varepsilon \mapsto \varphi_{\tau+\varepsilon}(\gamma_\tau) \) at \( \tau = t,s \), and subtle differences between the usual second derivative and the second Peano derivative (see Sect. 2) come into play.

(3) Our third main new ingredient, described in Part III, is a certain rigidity property of the change-of-variables formula (1.6), which allows us to bootstrap the a-priori available temporal regularity, and which in combination with the first and second ingredients, enables us to achieve disentanglement.

Indeed, the definition of \( \Phi'_s \) may be naturally extended to an appropriate domain beyond \( e_t(G_\varphi) \) as follows, allowing to easily (formally) calculate its partial derivative:

\[
\Phi'_s = \varphi_t + (t-s)\frac{\ell_t^2}{2}, \quad \partial_t \Phi'_s = \ell_t^2 + (t-s)\partial_t \frac{\ell_t^2}{2}.
\]

Evaluating at \( x = \gamma_t \) and plugging this into the change-of-variables formula (1.6), it follows that for \( \nu \)-a.e. geodesic \( \gamma \):

\[
\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{h_t^\gamma(t)}{1 + (t-s)\frac{\partial_t |_{\tau=t} = \ell_t^2/2(\gamma_t)}{\ell^2(\gamma)}} \quad \text{for a.e. } t, s \in (0,1). \tag{1.9}
\]

Thanks to the idea of considering together both initial-point \( s \) and endpoint \( t \), the latter formula takes on a very rigid structure: note that on the left-hand-side the \( s \) and \( t \) variables are separated, and the denominator on the right-hand-side depends linearly on \( s \). Consequently, we can easily bootstrap the a-priori available regularity in \( s \) and \( t \) of all terms involved. It follows that \( \frac{1}{\ell^2(\gamma)} \partial_t |_{\tau=t} = \ell_t^2/2(\gamma_t) \) must coincide for a.e. \( t \in (0,1) \) with a locally-Lipschitz function \( z(t) \), so that (1.7) applies. In addition, by redefining \( \{ h_s^\gamma \} \) for \( s \) in a null subset of \( (0,1) \), we can guarantee that \((0,1) \ni s \mapsto h_s^\gamma(t) \) is locally Lipschitz (for any given \( t \in (0,1) \)), even though there is a-priori no relation between the different densities \( \{ h_s^\gamma \} \) for \( s \in (0,1) \).
At this point, if \( \rho_t(\gamma_t) \) and \( z(t) \) were known to be \( C^2 \) smooth, and equality were to hold in (1.9) for all \( s, t \in (0, 1) \), we could then define

\[
Y(r) := \exp \left( \int_{r_0}^r \partial_t|_{t=s} \log h_s^r(t) \, ds \right),
\]

and as \( \partial_t|_{t=s} \log(1 + (t - s)z(t)) = z(s) \), it would follow, recalling the definition (1.8) of \( L \), that

\[
\frac{\rho_{r_0}(\gamma_{r_0})}{\rho_r(\gamma_r)} = L(r)Y(r) \quad \forall r \in (0, 1).
\]

Using the fact that all \( \{h_s^r\}_{s \in (0, 1)} \) are CD\((\ell(\gamma)^2K, N)\) densities to control \( \partial_t^2|_{t=r} \log h_r(t) \), and surprisingly, also the concavity of \( L \) (again!) to control the mixed partial derivatives \( \partial_s \partial_t|_{s=s=r} \log h_s^r(t) \), a formal computation described in Sect. 12.2 then verifies that \( Y \) is a CD\((\ell(\gamma)^2K, N)\) density itself. A rigorous justification without all of the above non-realistic assumptions turns out to be extremely tedious, due to the difficulty in applying an approximation argument while preserving the rigidity of the equation—this is worked out in Sect. 12 and the “Appendix”.

After taking care of all these details, we finally obtain the desired disentanglement (1.11) of the Jacobian: \( L \) is concave and so controlled synthetically by a linear distortion coefficient, whereas \( Y \) is a CD\((\ell(\gamma)^2K, N)\) density and so (by definition) \( Y^{1/(N-1)} \) is controlled synthetically by the \( \sigma_{K,N-1}^{(t)}(\ell(\gamma)) \) coefficient. A standard application of Hölder’s inequality then verifies that \( J^{1/N}(r) = \rho_r(\gamma_r)^{-1/N} \) is controlled by the \( \tau_{K,N}^{(t)}(\ell(\gamma)) \) distortion coefficient, i.e. satisfies (1.1)—in fact for all \( t_0, t_1 \in [0, 1] \)—thereby establishing CD\((K, N)\), see Theorem 13.2.

The definition (1.10) of \( Y \) finally sheds light on the crucial role which the parameter \( s \in (0, 1) \) plays in our strategy—its role is to vary between the different \( W_2 \)-geodesics from which the CD\(_{loc}(K, N)\) information is extracted into the CD\(_1(K, N)\) information on the disintegration into transport-rays from the (signed) distance functions from level sets \( \{\varphi_s = \varphi_s(\gamma_s)\} \), thereby coming full circle with the observation of (1.4).

Besides establishing the local-to-global property of CD\((K, N)\) and the equivalence of its various variants (in our setting), we emphasize that as a by product of our proof, we obtain a remarkable new self-improvement property of CD\((K, N)\): the \( \tau_{K,N} \)-concavity (1.1) of the transport Jacobian \( J_t(\gamma_t) \) along all \( W_2 \)-geodesics implies the (a-priori) stronger “L-Y” decomposition \( J_t(\gamma_t) = L_\gamma(t)Y_\gamma(t) \), where \( L_\gamma \) is concave and \( Y_\gamma \) is a CD\((\ell(\gamma)^2K, N)\) density on \( (0, 1) \). As already mentioned above, this self-improvement is false for
a single $W_2$-geodesic. We believe that the stronger “L-Y” information will prove to be of further use in the study of $\text{CD}(K,N)$ essentially non-branching spaces.

We refer to Sect. 13 for the final details and for additional immediate corollaries of the Main Theorem 1.1 pertaining to $\text{RCD}(K,N)$ and strong $\text{CD}(K,N)$ spaces. We also provide there several concluding remarks and suggestions for further investigation.

Part I Temporal theory of Optimal-Transport

2 Preliminaries

2.1 Geodesics

A metric space $(X,d)$ is called a length space if for all $x, y \in X$, $d(x, y) = \inf \ell(\sigma)$, where the infimum is over all (continuous) curves $\sigma : I \to X$ connecting $x$ and $y$, and $\ell(\sigma) := \sup \sum_{i=1}^{k} d(\sigma(t_{i-1}), \sigma(t_{i}))$ denotes the curve’s length, where the latter supremum is over all $k \in \mathbb{N}$ and $t_0 \leq \cdots \leq t_k$ in the interval $I \subset \mathbb{R}$. A curve $\gamma$ is called a geodesic if $\ell(\gamma|_{[t_0,t_1]}) = d(\gamma(t_0), \gamma(t_1))$ for all $[t_0, t_1] \subset I$. If $\ell(\gamma) = 0$ we will say that $\gamma$ is a null geodesic. The metric space is called a geodesic space if for all $x, y \in X$ there exists a geodesic in $X$ connecting $x$ and $y$. We denote by $\text{Geo}(X)$ the set of all closed directed constant-speed geodesics parametrized on the interval $[0, 1]$:

$$\text{Geo}(X) := \left\{ \gamma : [0, 1] \to X ; \quad d(\gamma(s), \gamma(t)) = |s - t|d(\gamma(0), \gamma(1)) \quad \forall s, t \in [0, 1] \right\}.$$  

We regard $\text{Geo}(X)$ as a subset of all Lipschitz maps $\text{Lip}([0, 1], X)$ endowed with the uniform topology. We will frequently use $\gamma_t := \gamma(t)$.

The metric space is called proper if every closed ball (of finite radius) is compact. It follows from the metric version of the Hopf-Rinow Theorem (e.g. [22, Theorem 2.5.28]) that for complete length spaces, local compactness is equivalent to properness, and that complete proper length spaces are in fact geodesic.

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

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

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

$$D(\tilde{G}) := \left\{ (x, t) \in X \times \mathbb{R} ; \quad \exists \gamma \in \tilde{G} , \ t \in \text{Dom}(\gamma) , \ x = \gamma_t \right\} .$$
We denote by $e_t : \text{Geo}(X) \ni \gamma \mapsto \gamma_t \in X$ the (continuous) evaluation map at $t \in [0, 1]$, and abbreviate given $I \subset [0, 1]$ as follows:

$$e_t(\tilde{G}) = \tilde{G}(t) := D(\tilde{G})(t) = \left\{ \gamma_t : \gamma \in \tilde{G} \right\},$$

$$e_I(\tilde{G}) := \bigcup_{t \in I} e_t(\tilde{G}).$$

$$\tilde{G}(x) := D(\tilde{G})(x) = \left\{ t \in [0, 1] ; \exists \gamma \in \tilde{G}, t \in \text{Dom}(\gamma) , \gamma_t = x \right\}.$$

### 2.2 Derivatives

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

$$\frac{d}{dt}^+ g(t_0) = \lim \sup_{A \ni t \to t_0} \frac{g(t) - g(t_0)}{t - t_0}, \quad \frac{d}{dt}^- g(t_0) = \lim \inf_{A \ni t \to t_0} \frac{g(t) - g(t_0)}{t - t_0}.$$

We will say that $g$ is differentiable at $t_0$ iff $\frac{d}{dt} g(t_0) := \frac{d}{dt}^+ g(t_0) = \frac{d}{dt}^- g(t_0) < \infty$.

This is a slightly more general definition of differentiability than the traditional one which requires that $t_0$ be an interior point of $A$.

**Remark 2.1** 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 2.2** 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 locally Lipschitz, consider any locally 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 $A$.

Let $f : I \to \mathbb{R}$ denote a convex function on an open interval $I \subset \mathbb{R}$. It is well-known that the left and right derivatives $f^{l,-}$ and $f^{r,+}$ exist at every point in $I$ and that $f$ is locally Lipschitz there; in particular, $f$ is differentiable at a given point iff 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
require the following elementary one-dimensional version (probably due to Jessen) of the well-known Aleksandrov’s theorem about twice differentiability a.e. of convex functions on $\mathbb{R}^n$ (see [45, Theorem 5.2.1] or [20, Section 2.6], and [71, p. 31] for historical comments). Clearly, all of these results extend 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 2.3** (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 $\tau_0$, and if $D \subset I$ denotes the subset of differentiability points of $f$ in $I$, then $f' : D \to \mathbb{R}$ is differentiable at $\tau_0$ 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 $\tau_0$ with $(f',+)(\tau_0) = \Delta$.

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

4. $f$ is differentiable at $\tau_0$ and has the following second order expansion there:

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

   In this case, $f$ is said to have a second Peano derivative at $\tau_0$.

We remark that even for a differentiable function $f$, while the implication (1) $\Rightarrow$ (4) follows by Taylor’s theorem (existence of the second derivative at a point implies existence of the second Peano derivative there), the converse implication is in general false (see e.g. [61] for a nice discussion). For a locally semi-convex or semi-concave function $f$, we will say that $f$ is twice differentiable at $\tau_0$ if any (all) of the above equivalent conditions hold for some $\Delta \in \mathbb{R}$, and write $(\frac{d^2}{d\tau^2})|_{\tau=\tau_0} f(\tau) = \Delta$.

Finally, we will require the following slightly more refined notation.

**Definition** 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 $\tau_0$, denoted $\overline{P}_2 f(\tau_0)$ and $\underline{P}_2 f(\tau_0)$ respectively, by

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

$\square$ Springer
where

\[ h(\varepsilon) := 2(f(\tau_0 + \varepsilon) - f(\tau_0) - \varepsilon f'(\tau_0)). \]

Clearly \( f \) has a second Peano derivative at \( \tau_0 \) iff \( \overline{P}_2 f(\tau_0) = \underline{P}_2 f(\tau_0) < \infty \).

The following is a type of Stolz–Cesàro lemma:

**Lemma 2.4** 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

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

**Proof** By local absolute continuity, \( f \) is differentiable a.e. in \( I \) and we have for small enough \( |\varepsilon| \):

\[
\frac{1}{2} h(\varepsilon) = f(\tau_0 + \varepsilon) - f(\tau_0) - \varepsilon f'(\tau_0) = \int_{0}^{\varepsilon} (f'(\tau_0 + \delta) - f'(\tau_0)) d\delta,
\]

and hence

\[
\frac{h(\varepsilon)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \int_{0}^{\varepsilon} 2\delta \frac{f'(\tau_0 + \delta) - f'(\tau_0)}{\delta} d\delta.
\]

Taking appropriate subsequential limits as \( \varepsilon \to 0 \), the asserted inequalities readily follow. \( \square \)

### 3 Temporal theory of intermediate-time Kantorovich potentials: first and second order

In the next sections, we will only consider the quadratic cost function \( c = \frac{d^2}{2} \) on \( X \times X \).

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

\[
\psi^c(x) = \inf_{y \in X} \frac{d(x, y)^2}{2} - \psi(y).
\]

A function \( \varphi : X \to \mathbb{R} \cup \{-\infty\} \) is called \( c \)-concave if \( \varphi = \psi^c \) for some \( \psi \) as above. It is well known [76, Exercise 2.35] that \( \varphi \) is \( c \)-concave iff \((\varphi^c)^c = \varphi \).

In the context of Optimal-Transport with respect to the quadratic cost \( c \), a \( c \)-concave function \( \varphi : X \to \mathbb{R} \cup \{-\infty\} \) which is not identically equal to \(-\infty\)
is also known as a 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.

There is a natural way to interpolate between a Kantorovich potential and its dual by means of the Hopf-Lax semi-group, resulting in intermediate-time Kantorovich potentials \( \{\varphi_t\}_{t \in (0,1)} \). The goal of the next three sections is to provide first, second and third order information on the time-behavior \( t \mapsto \varphi_t(x) \) at intermediate times \( t \in (0, 1) \). In these sections, we only assume that \((X, d)\) is a \textit{proper geodesic metric space}.

In this section, we focus on first and second order information. The main new result is Theorem 3.11.

3.1 Hopf-Lax semi-group

We begin with several well-known definitions which we slightly modify and specialize to our setting.

**Definition (Hopf-Lax Transform)** Given \( f : X \to \mathbb{R} \cup \{\pm\infty\} \) which is not identically \( +\infty \) and \( t > 0 \), define the Hopf-Lax transform \( Q_t f : X \to \mathbb{R} \cup \{-\infty\} \) by

\[
Q_t f(x) := \inf_{y \in X} \frac{d(x, y)^2}{2t} + f(y). \tag{3.1}
\]

Clearly either \( Q_t f \equiv -\infty \) or \( Q_t f(x) \) is finite for all \( x \in X \) (as our metric \( d \) is finite). Consequently, we denote:

\[
t_\ast(f) := \sup\{t > 0 : Q_t f \neq -\infty\}
\]

setting \( t_\ast(f) = 0 \) if the supremum is over an empty set. Finally, we set \( Q_0 f := f \).

It is not hard to check (see e.g. [49, Theorem 2.5 (i)]) that when \((X, d)\) is a length space (and in particular geodesic), the Hopf-Lax transform is in fact a semi-group on \([0, \infty)\):

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

**Remark 3.1** It is also possible to extend the definition of \( Q_t f \) to negative times \( t < 0 \) by setting

\[
Q_t f(x) := -Q_{-t}(-f)(x) = \sup_{y \in X} \frac{d(x, y)^2}{2t} + f(y), \quad t < 0.
\]
This is called the backwards Hopf-Lax semi-group on $(-\infty, 0]$. However, $(\mathbb{R}, +) \ni t \mapsto (Q_t, \circ)$ is in general not an abelian group homomorphism, not even for $t \in [0, 1]$ when applied to a Kantorovich potential $\varphi$ (characterized by $Q_1 \circ Q_1(-\varphi) = -\varphi$)—see Sect. 3.3. This will be a rather significant nuisance we will need to cope with in this work.

Clearly $(0, \infty) \times X \ni (t, x) \mapsto Q_t f(x)$ is upper semi-continuous as the infimum of continuous functions in $(t, x)$, and by definition $(0, \infty) \ni t \mapsto Q_t f(x)$ is monotone non-increasing for each $x \in X$. Consequently, $(0, \infty) \ni t \mapsto Q_t f(x)$ must be continuous from the left.

It may also be shown (see [5, Lemma 3.1.2]) that $X \times (0, t^*_f) \ni (x, t) \mapsto Q_t f(x)$ is continuous (and in fact locally Lipschitz, see Theorem 3.4 below). Together with the left-continuity, we deduce that for every $x \in X$, $(0, t^*_f) \ni t \mapsto Q_t f(x)$ is continuous.

Note that by definition $f^c = Q_1(-f)$, and that a Kantorovich pair of conjugate potentials $\varphi, \varphi^c : X \to \mathbb{R} \cup \{-\infty\}$ are characterized by not being identically equal to $-\infty$ and satisfying:

$$\varphi = Q_1(-\varphi^c), \quad \varphi^c = Q_1(-\varphi).$$

In particular, $t^*_f(\varphi), t^*_f(\varphi^c) \geq 1$, and we a-posteriori deduce that $\varphi, \varphi^c$ are both finite on the entire space $X$ (we have used above the fact that the metric $d$ is finite, which differs from other more general treatments).

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

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

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

$$-\varphi_t(x) = \inf_{y \in X} \frac{d^2(x, y)}{2t} - \varphi(y) \quad \forall t \in (0, 1].$$

Applying the above mentioned general properties of the Hopf-Lax semi-group to $\varphi_t$, it will be useful to record:

**Lemma 3.2** (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 $(0, 1]$. 

\[ Springer\]
The globalization theorem

**Definition (Kantorovich Geodesic)** Given a Kantorovich potential \( \varphi : X \to \mathbb{R} \), a geodesic \( \gamma \in \text{Geo}(X) \) is called a \( \varphi \)-Kantorovich (or optimal) geodesic if

\[
\varphi(\gamma_0) + \varphi^c(\gamma_1) = \frac{d(\gamma_0, \gamma_1)^2}{2} = \frac{\ell(\gamma)^2}{2}.
\]

We denote all \( \varphi \)-Kantorovich geodesics by \( G_\varphi \). Note that \( \gamma \in G_\varphi \) iff \( \gamma^c \in G_{\varphi^c} \), where \( \gamma^c(t) := \gamma(1 - t) \) is the time-reversed geodesic. By upper semi-continuity of \( \varphi \) and \( \varphi^c \), it follows that \( G_\varphi \) is a closed subset of \( \text{Geo}(X) \).

The following is not hard to check (see e.g. [24, Corollary 2.16]):

**Lemma 3.3** Let \( \gamma \) be a \( \varphi \)-Kantorovich geodesic. Then

\[
\varphi_s(\gamma_s) - \varphi_r(\gamma_r) = \frac{d(\gamma_s, \gamma_r)^2}{2(r - s)} = \frac{(r - s) \ell(\gamma)^2}{2} \quad \forall s, r \in [0, 1].
\]

### 3.2 Distance functions

The following important definition was given by Ambrosio–Gigli–Savaré [5, 6]:

**Definition (Distance functions \( D_f^\pm \))** Given \( f : X \to \mathbb{R} \cup \{+\infty\} \) which is not identically \( +\infty \), denote

\[
D_f^+(x, t) := \sup_n \limsup_{n \to \infty} d(x, y_n) \geq \inf_n \liminf_{n \to \infty} d(x, y_n) =: D_f^-(x, t),
\]

where the supremum and infimum above run over the set of minimizing sequences \( \{y_n\} \) in the definition of the Hopf-Lax transform (3.1). A simple diagonal argument shows that the (outer) supremum and infimum above are in fact attained.

The following properties were established in [5], [6, Chapter 3]:

**Theorem 3.4** (Ambrosio–Gigli–Savaré) For any metric space \((X, d)\) (not necessarily proper, complete nor geodesic):

1. Both functions \( D_f^\pm(x, t) \) are locally finite on \( X \times (0, t_\ast(f)) \), and \((x, t) \mapsto Q_t \) is locally Lipschitz there.
2. \((x, t) \mapsto D_f^\pm(x, t)\) is upper \((D_f^+(x, t))\) / lower \((D_f^-(x, t))\) semi-continuous on \( X \times (0, t_\ast(f)) \).
3. For every \( x \in X \), both functions \((0, t_\ast(f)) \ni t \mapsto D_f^\pm(x, t)\) are monotone non-decreasing and coincide except where they have (at most countably many) jump discontinuities.
(4) For every \( x \in X, \partial_t^\pm Q_t f(x) = -\frac{(D_t^\pm f(x,t))^2}{2t^2} \) for all \( t \in (0, t_*(f)) \), where \( \partial_t^- \) and \( \partial_t^+ \) 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, and differentiable at \( t \in (0, t_*(f)) \) iff \( D_t^+ (x, t) = D_t^- (x, t) \).

It may be instructive to recall the proof of property (3) above, which is related to some ensuing properties, so for completeness, we present it below.

**Lemma 3.5** Given a proper metric space \( X \), a lower semi-continuous \( f : X \to \mathbb{R}, x \in X \) and \( t \in (0, t_*(f)) \), there exist \( y_t^\pm \in X \) so that

\[
Q_t f(x) = \frac{d(x, y_t^\pm)^2}{2t} + f(y_t^\pm) \quad \text{and} \quad d(x, y_t^\pm) = D_t^\pm f(x, t).
\]

Recall that \(-\varphi\) is indeed lower semi-continuous for any Kantorovich potential \( \varphi \).

**Proof of Lemma 3.5** Let \( \{y_t^{\pm,n}\} \) denote a minimizing sequence so that

\[
Q_t f(x) = \lim_{n \to \infty} \frac{d(x, y_t^{\pm,n})^2}{2t} + f(y_t^{\pm,n}) \quad \text{and} \quad D_t^\pm f(x, t) = \lim_{n \to \infty} d(x, y_t^{\pm,n}).
\]

By property (1) we know that \( D_t^\pm f(x, t) < R < \infty \), and the properness implies that the closed geodesic ball \( B_R(x) \) is compact. Consequently \( \{y_t^{\pm,n}\} \) has a converging subsequence to \( y_t^\pm \), and the lower semi-continuity of \( f \) implies that:

\[
Q_t f(x) = \inf_{y \in X} \frac{d(x, y)^2}{2t} + f(y) = \min_{y \in B_R(x)} \frac{d(x, y)^2}{2t} + f(y) = \frac{d(x, y_t^\pm)^2}{2t} + f(y_t^\pm),
\]

as asserted. \( \square \)

**Proof of (3) for proper \( X \) and lower semi-continuous \( f \).** The assertion will follow immediately after establishing

\[
D_t^+(x, s) \leq D_t^-(x, t) \quad \forall 0 < s < t < t_*(f),
\]

since trivially \( D_t^- \leq D_t^+ \) and since a monotone function can only have a countable number of jump discontinuities. By Lemma 3.5, there exist \( y_s^+ \) and \( y_s^- \) so that

\[
Q_s f(x) = \inf_{y \in X} \frac{d(x, y)^2}{2s} + f(y) = \frac{d(x, y_s^+)^2}{2s} + f(y_s^+) \quad \text{and}
\]

\[ \square \] Springer
\[ d(x, y_s^+) = D_f^+(x, s), \]

and:

\[ Q_tf(x) = \inf_{y \in X} \frac{d(x, y)^2}{2t} + f(y) = \frac{d(x, y_t^-)^2}{2t} + f(y_t^-) \]

and

\[ d(x, y_t^-) = D_f^-(x, t). \]

It follows that

\[ \frac{d(x, y_s^+)^2}{2s} + f(y_s^+) \leq \frac{d(x, y_t^-)^2}{2s} + f(y_t^-), \]

\[ \frac{d(x, y_t^-)^2}{2t} + f(y_t^-) \leq \frac{d(x, y_s^+)^2}{2t} + f(y_s^+). \]

Summing these two inequalities and rearranging terms, one deduces

\[ D_f^+(x, s) \left( \frac{1}{s} - \frac{1}{t} \right) \leq D_f^-(x, t) \left( \frac{1}{s} - \frac{1}{t} \right), \]

as required. \( \Box \)

### 3.3 Intermediate-time duality and time-reversed potential

It is immediate to show by inspecting the definitions that we always have (e.g. [77, Theorem 7.34 (iii)] or [3, Proposition 2.17 (ii)]):

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

this is an inherent group-structure incompatibility of the Hopf-Lax forward and backward semi-groups. Note that for \( f = -\varphi \) where \( \varphi \) is a Kantorovich potential, we do have equality for \( s = 1 \), and in fact for all \( s \in [0, 1] \). However, for \( f = Q_t(-\varphi), t \in (0, 1) \) and \( s = 1 - t \), we can only assert an inequality above ([77, Theorem 7.36], [3, Corollary 2.23 (i)]):

\[ (\varphi^c)_{1-t} = Q_{-(1-t)} \circ Q_1(-\varphi) \leq Q_t(-\varphi) = -\varphi_t \text{ on } X, \quad (3.2) \]

and equality may not hold at every point of \( X \) (cf. [77, Remark 7.37]). Nevertheless, in our setting, the subset where equality is attained may be characterized as in the next proposition. We first introduce the following very convenient:
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

\[
\bar{\varphi}_t(x) = \inf_{y \in X} \frac{d^2(x, y)}{2(1-t)} - \varphi^c(y) \quad \forall t \in [0, 1).
\]

Proposition 3.6 (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 \( x \in e_t(G\varphi) \). In other words:

\[
D(\hat{G}\varphi) = \{(x, t) \in X \times (0, 1) : \varphi_t(x) = \bar{\varphi}_t(x)\}.
\] (3.3)

(1) is immediate by \( c \)-concavity, and (2) is a reformulation of (3.2), so the only assertion requiring proof is (3). The if direction is well-known (e.g. [77, Theorem 7.36], [3, Corollary 2.23 (ii)]), but the other direction appears to be new. It is based on the following simple lemma, which we will use again later on:

Lemma 3.7 Assume that for some \( x, y, z \in X \) and \( t \in (0, 1) \):

\[
\frac{d(x, y)^2}{2t} - \varphi(y) = \varphi^c(z) - \frac{d(x, z)^2}{2(1-t)}.
\]

Then \( x \) is a \( t \)-intermediate point between \( y \) and \( z \):

\[
d(y, z) = \frac{d(x, y)}{t} = \frac{d(x, z)}{1-t},
\] (3.4)

and there exists a \( \varphi \)-Kantorovich geodesic \( \gamma : [0, 1] \to X \) with \( \gamma(0) = y, \gamma(t) = x \) and \( \gamma(1) = z \).

Proof Using that

\[
\varphi(y) + \varphi^c(z) \leq \frac{d(y, z)^2}{2},
\] (3.5)

our assumption yields

\[
\frac{d(x, y)^2}{2t} + \frac{d(x, z)^2}{2(1-t)} \leq \frac{d(y, z)^2}{2}.
\]
On the other hand, the reverse inequality is always valid by the triangle and Cauchy–Schwarz inequalities:

\[
\frac{d(y, z)^2}{2} \leq \frac{(d(x, y) + d(x, z))^2}{2} \leq \frac{d(x, y)^2}{2t} + \frac{d(x, z)^2}{2(1 - t)}.
\]

It follows that we must have equality everywhere above, and (3.4) amounts to the equality case in the Cauchy–Schwarz inequality. Consequently, the concatenation \( \gamma : [0, 1] \to X \) of any constant speed geodesic \( \gamma_1 : [0, t] \to X \) between \( y \) and \( x \), with any constant speed geodesic \( \gamma_2 : [t, 1] \to X \) between \( x \) and \( z \), so that \( \gamma(0) = y, \gamma(t) = x \) and \( \gamma(1) = z \), must be a constant speed geodesic itself (by the triangle inequality). Lastly, the equality in (3.5) implies that \( \gamma \in G_{\varphi} \), thereby concluding the proof.

**Proof of Proposition 3.6 (3)** We begin with the known direction. Let \( x = \gamma_t \) with \( \gamma \in G_{\varphi} \). Apply Lemma 3.3 to \( \gamma \) with \( s = 0 \) and \( r = t \):

\[
\varphi(\gamma_0) - \varphi_t(\gamma_t) = \varphi_0(\gamma_0) - \varphi_t(\gamma_t) = t \frac{\text{len}(\gamma)^2}{2},
\]

and to \( \gamma^c \in G_{\varphi^c} \) with \( s = 1 \) and \( r = 1 - t \):

\[
-\varphi(\gamma_0) - (\varphi^c)_1(\gamma_1) = (\varphi^c)_1(\gamma^c_1) - (\varphi^c)_1(\gamma^c_{1-t}) = t \frac{\text{len}(\gamma^c)^2}{2} = -t \frac{\text{len}(\gamma)^2}{2},
\]

where we used that \( (\varphi^c)_1 = -(\varphi^c)^c = -\varphi \). Summing these two identities, we obtain:

\[
\varphi_t(\gamma_t) = -(\varphi^c)_{1-t}(\gamma_t),
\]

as asserted.

For the other direction, assume that \( \varphi_t(x) = -(\varphi^c)_{1-t}(x) \) for some \( x \in X \) and \( t \in (0, 1) \). By Lemma 3.5 applied to the lower semi-continuous functions \( -\varphi \) and \( -\varphi^c \), there exist \( y_t, z_t \in X \) so that

\[
-\varphi_t(x) = Q_t(-\varphi)(x) = \frac{d(x, y_t)^2}{2t} - \varphi(y_t),
\]

\[
\varphi_t(x) = -(\varphi^c)_{1-t}(x) = Q_{1-t}(-\varphi^c)(x) = \frac{d(x, z_t)^2}{2(1 - t)} - \varphi^c(z_t).
\]

Summing the two equations, the assertion follows immediately from Lemma 3.7.

\( \Box \)

We also record the following immediate corollary of Lemma 3.2:

\( \Box \) Springer
Corollary 3.8  (1) \((x, t) \mapsto \tilde{\varphi}_t(x)\) is upper semi-continuous on \(X \times [0, 1]\) and continuous on \(X \times (0, 1)\).

(2) For every \(x \in X\), \([0, 1] \ni t \mapsto \tilde{\varphi}_t(x)\) is monotone non-decreasing and continuous on \([0, 1]\).

Finally, in view of (3.3), we deduce for free:

Corollary 3.9 \(D(\hat{G}_\varphi)\) is a closed subset of \(X \times (0, 1)\).

Proof Immediate from (3.3) by the continuity of \(\varphi_t(x)\) and \(\bar{\varphi}_t(x)\) on \(X \times (0, 1)\).

\(\square\)

3.4 Length functions \(\ell_t^\pm\) and \(\tilde{\ell}_t^\pm\)

Definition (Length functions \(\ell_t^\pm, \tilde{\ell}_t^\pm\)) Given a Kantorovich potential \(\varphi : X \to \mathbb{R}\), denote

\[
\ell_t^\pm(x) := \frac{D_{\varphi}^\pm(x, t)}{t}, \quad \tilde{\ell}_t^\pm(x) := \frac{D_{\varphi_c}^\pm(x, 1 - t)}{1 - t}, \quad (x, t) \in X \times (0, 1).
\]

To provide motivation for these definitions, let us mention that we will shortly see that if \(x = \gamma_t\) with \(\gamma \in G_\varphi\) and \(t \in (0, 1)\), then

\[
\ell_t^+(x) = \ell_t^-(x) = \tilde{\ell}_t^+(x) = \tilde{\ell}_t^-(x) = \ell(\gamma).
\]

In particular, all \(\varphi\)-Kantorovich geodesics having \(x\) as their \(t\)-mid-point have the same length. These facts seem to not have been previously noted in the literature, and they will be crucially exploited in this work.

Definition For \(\ell = \ell, \tilde{\ell}\), introduce the following set:

\[
D_\ell := \left\{(x, t) \in X \times (0, 1) ; \tilde{\ell}_t^+(x) = \tilde{\ell}_t^-(x)\right\},
\]

and on it define \(\ell_t(x)\) as the common value \(\tilde{\ell}_t^+(x) = \tilde{\ell}_t^-(x)\).

Recalling that \(\varphi_t = -Q_t(-\varphi)\) and \(\bar{\varphi}_t = Q_{1-t}(-\varphi_c)\), we begin by translating Theorem 3.4 into the following corollary. We freely use standard properties of semi-convex (semi-concave) functions, like twice a.e. differentiability, non-negativity (non-positivity) of the singular part of the distributional second derivative (see e.g. Lemma A.11), etc.

Corollary 3.10 Let \(\varphi : X \to \mathbb{R}\) denote a Kantorovich potential. Then:

(1) For \(\ell = \ell, \tilde{\ell}\) and \(\varphi = \varphi, \tilde{\varphi}\), \(\ell_t^\pm(x)\) are locally finite on \(X \times (0, 1)\), and \((x, t) \mapsto \tilde{\varphi}_t(x)\) is locally Lipschitz there.
(2) For \( \tilde{\ell} = \ell, \tilde{\ell}, (x,t) \mapsto \tilde{\ell}_t^\pm(x) \) is upper \( (\tilde{\ell}_t^+(x)) \) / lower \( (\tilde{\ell}_t^-(x)) \) semi-continuous on \( X \times (0,1) \). In particular, the subset \( D_{\ell} \subset X \times (0,1) \) is Borel and \( (x,t) \mapsto \tilde{\ell}_t(x) \) is continuous on \( D_{\ell} \).

(3) For every \( x \in X \) we have

\[
\partial_t^\pm \varphi_t(x) = \frac{\ell_t^\pm(x)^2}{2}, \quad \partial_t^\pm \bar{\varphi}_t(x) = \frac{\bar{\ell}_t^\pm(x)^2}{2} \quad \forall t \in (0,1).
\]

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

\[
\bar{\varphi}_{t_2}(x) - \bar{\varphi}_{t_1}(x) = \int_{t_1}^{t_2} \frac{\tilde{\ell}_t^2(x)}{2} d\tau \quad \forall t_1, t_2 \in (0,1).
\]

(4) For every \( x \in X \):

(a) Both maps \( (0,1) \ni t \mapsto t \tilde{\ell}_t^\pm(x) \) are monotone non-decreasing. In particular, \( D_{\ell}(x) \ni t \mapsto \ell_t^2(x) \) is differentiable a.e., the singular part of its distributional derivative is non-negative, \( (0,1) \ni t \mapsto \varphi_t(x) \) is locally semi-convex, and

\[
\partial_t \frac{\ell_t^2(x)}{2} \geq -\frac{1}{t} \ell_t^2(x) \quad \forall t \in D_{\ell}(x). \tag{3.6}
\]

(b) Both maps \( (0,1) \ni t \mapsto (1-t)\tilde{\ell}_t^\pm(x) \) are monotone non-increasing. In particular, \( D_{\ell}(x) \ni t \mapsto \ell_t^2(x) \) is differentiable a.e., the singular part of its distributional derivative is non-positive, \( (0,1) \ni t \mapsto \bar{\varphi}_t(x) \) is locally semi-concave, and

\[
\partial_t \frac{\ell_t^2(x)}{2} \leq \frac{1}{1-t} \ell_t^2(x) \quad \forall t \in D_{\ell}(x). \tag{3.7}
\]

Proof The only point requiring verification is that monotonicity of \( t \mapsto t \ell_t(x) \) in (4a) and \( t \mapsto (1-t)\tilde{\ell}_t \) in (4b) implies (3.6) and (3.7), respectively. For instance, using the continuity of \( t \mapsto \ell_t(x) \) on \( D_{\ell}(x) \), (3.6) is clearly equivalent to

\[
\partial_t \ell_t(x) \geq -\frac{1}{t} \ell_t(x) \quad \forall t \in D_{\ell}(x). \tag{3.8}
\]

Now, if \( \ell_t(x) = 0 \) the monotonicity directly implies \( \partial_t \ell_t(x) \geq 0 \) and establishes (3.8), whereas otherwise, (3.8) is equivalent by the chain-rule (and again
the continuity of $t \mapsto \ell_t(x)$ on $D_\ell(x)$ to
\[
\partial_t \log(t \ell_t(x)) = \frac{1}{t} + \partial_t \log(\ell_t(x)) \geq 0 \quad \forall t \in D_\ell(x),
\]
which in turn is a consequence of the aforementioned monotonicity. The proof of \eqref{3.7} follows identically. \hfill \Box

We now arrive to the main new result of this section, which will be constantly and crucially used in this work:

**Theorem 3.11** Let $\varphi : X \to \mathbb{R}$ denote a Kantorovich potential.

1. For all $x \in e_t(G_\varphi)$ with $t \in (0, 1)$, we have
\[
\ell_t^+(x) = \ell_t^-(x) = \tilde{\ell}_t^+(x) = \tilde{\ell}_t^-(x) = \ell(\gamma),
\]
for any $\gamma \in G_\varphi$ so that $\gamma_t = x$. In other words
\[
D(\hat{G}_\varphi) = \{(x, t) \in X \times (0, 1) : x = \gamma_t \ , \ \gamma \in G_\varphi\} \subset D_\ell \cap D_{\bar{\ell}},
\]
and moreover $\ell_t(x) = \tilde{\ell}_t(x)$ there.

2. For all $x \in X$, $\hat{G}_\varphi(x) \ni t \mapsto \ell_t(x) = \tilde{\ell}_t(x)$ is locally Lipschitz:
\[
\left| \sqrt{t(1-t)}\ell_t(x) - \sqrt{s(1-s)}\ell_s(x) \right| \leq \sqrt{\ell_t(x)\ell_s(x)} \left| \sqrt{t(1-s)} - \sqrt{s(1-t)} \right| \quad \forall t, s \in \hat{G}_\varphi(x). \quad (3.9)
\]

3. For all $(x, t) \in D(\hat{G}_\varphi) \subset D_\ell \cap D_{\bar{\ell}}$ we have for both $\ast = \mathcal{P}_2\bar{\varphi}_t(x), \mathcal{P}_2\varphi_t(x)$:
\[
-\frac{1}{t}\ell_t^2(x) \leq \partial_t \ell_t^2(x) \leq \mathcal{P}_2\varphi_t(x) \leq \mathcal{P}_2\bar{\varphi}_t(x) \leq \partial_t \ell_t^2(x) \leq \frac{1}{1-t}\ell_t^2(x),
\]
where the Peano (partial) derivatives are with respect to the $t$ variable.

4. For all $(x, t) \in D(\hat{G}_\varphi) \subset D_\ell \cap D_{\bar{\ell}}$ we have:
\[
\frac{\bar{\partial}_t \ell_t^2(x)}{2} \leq \partial_t \ell_t^2(x) \leq \left( \frac{1}{1-t} + \frac{1}{t} \right) \ell_t^2(x) \leq \left( \frac{2}{1-t} + \frac{1}{t} \right) \ell_t^2(x)
\]
\[
\frac{\partial_t \ell_t^2(x)}{2} \geq \bar{\partial}_t \ell_t^2(x) \geq - \left( \frac{1}{t} + \frac{1}{1-t} \right) \ell_t^2(x) \geq - \left( \frac{2}{t} + \frac{1}{1-t} \right) \ell_t^2(x).
\]
In particular, for every \( x \in X \), we have:

\[
\partial_t \varphi_t(x) = \partial_t \tilde{\varphi}_t(x) = \frac{\ell_t^2(x)}{2} = \frac{\tilde{\ell}_t^2(x)}{2} \quad \forall t \in \hat{G}_\varphi(x),
\]

with \( t \mapsto \frac{\ell_t^2(x)}{2} \) and \( t \mapsto \frac{\tilde{\ell}_t^2(x)}{2} \) continuous on \( D_\ell(x) \cap D_{\tilde{\ell}}(x) \), differentiable a.e. there, and having locally bounded lower and upper derivatives on \( \hat{G}_\varphi(x) \subset D_\ell(x) \cap D_{\tilde{\ell}}(x) \) as in (3) and (4).

**Proof** To see (1), let \( (x, t) \in D(\hat{G}_\varphi) \). Equivalently, by Proposition 3.6 (3), we know that \( \varphi_t(x) = \tilde{\varphi}_t(x) \). In addition, Lemma 3.5 assures the existence of \( y^\pm \) and \( z^\pm \) in \( X \) so that

\[
-\varphi_t(x) = \frac{d(x, y^\pm)^2}{2t} - \varphi(y^\pm), \quad d(x, y^\pm) = t \ell_t^\pm(x)
\]

\[
-\tilde{\varphi}_t(x) = -\frac{d(x, z^\pm)^2}{2(1-t)} + \varphi^c(z^\pm), \quad d(x, z^\pm) = (1-t)\tilde{\ell}_t^\pm(x).
\]

Equating both expressions and applying Lemma 3.7, we deduce that \( x \) is the \( t \)-midpoint of a geodesic connecting \( y^\pm \) and \( z^\pm \) (for all 4 possibilities), and that

\[
\ell_t^\pm(x) = \frac{d(x, y^\pm)}{t} = \frac{d(x, z^\pm)}{1-t} = \tilde{\ell}_t^\pm(x) \tag{3.10}
\]

so that all 4 possibilities above coincide. We remark in passing that this already implies in a non-branching setting that necessarily \( y^+ = y^- \) and \( z^+ = z^- \), i.e. the uniqueness of a \( \varphi \)-Kantorovich geodesic with \( t \)-mid point \( x \).

Furthermore, if \( x = \gamma_t \) for some \( \gamma \in G_\varphi \), then by Lemma 3.3:

\[
-\varphi_t(x) = \frac{d(x, \gamma_0)^2}{2t} - \varphi(\gamma_0).
\]

It follows by definition of \( D_{-\varphi}^\pm(x, t) \) that:

\[
t\ell_t^-(x) = D_{-\varphi}^-(x, t) \leq d(x, \gamma_0) = t\ell(\gamma) \leq D_{-\varphi}^+(x, t) = t\ell_t^+(x),
\]

which together with (3.10) establishes that \( \ell(\gamma) = \ell_t(x) = \tilde{\ell}_t(x) \).

To see (2), let \( \gamma^t, \gamma^s \in G_\varphi \) be so that \( \gamma^t_t = \gamma^s_s = x \), for some \( t, s \in (0, 1) \). Then

\[
\varphi^c(\gamma^p_t) = \frac{\ell(\gamma^p)^2}{2} - \varphi(\gamma_0^p) \leq \frac{d(\gamma^p_t \cdot \gamma_0^p)^2}{2} - \varphi(\gamma_0^q)
\]
for \((p, q) = (t, s)\) and \((p, q) = (s, t)\). Summing these two inequalities, we obtain the well-known \(c\)-cyclic monotonicity of the set \(\{(\gamma_0^t, \gamma_1^t), (\gamma_0^s, \gamma_1^s)\}\):

\[
\ell(\gamma^t)^2 + \ell(\gamma^s)^2 \leq d(\gamma_0^t, \gamma_1^t)^2 + d(\gamma_0^s, \gamma_1^s)^2.
\]

To evaluate the right-hand-side, we simply pass through \(x\) and employ the triangle inequality:

\[
d(\gamma_0^p, \gamma_1^q) \leq d(\gamma_0^p, x) + d(x, \gamma_1^q) = p \ell(\gamma^p) + (1 - q) \ell(\gamma^q).
\]

Plugging this above and rearranging terms, we obtain

\[
t(1 - t)\ell(\gamma^t)^2 + s(1 - s)\ell(\gamma^s)^2 \leq (t(1 - s) + s(1 - t)) \ell(\gamma^t) \ell(\gamma^s).
\]

Completing the square by subtracting \(2\sqrt{(1 - (1 - s)t)\ell(\gamma^t)^2\ell(\gamma^s)}\) from both sides, and recalling that \(\ell(\gamma^p) = \ell_\rho(x)\) for \(p = t, s\), we readily obtain (3.9). In particular, using \(t = s\), the above argument recovers the last assertion of (1) that \(\ell(\gamma)\) is the same for all \(\gamma \in G(x)\) so that \(\gamma_t = x\).

To see (3), recall that given \(x \in X\), we know by Proposition 3.6 that \(\varphi_t(x) \leq \tilde{\varphi}_t(x)\) for all \(t \in (0, 1)\) with equality iff \(t \in \hat{G}(x)\). Since \(\hat{G}(x) \subset D(x) \cap D_\ell(x)\) by (1), we know that both maps \(t \mapsto \tilde{\varphi}_t(x)\) are differentiable at \(t_0 \in \hat{G}(x)\), and we see again that \(\frac{\ell_0^2(x)}{2} = \partial_t \varphi_{t_0}(x) = \partial_t \tilde{\varphi}_{t_0}(x) = \frac{\ell_0^2(x)}{2}\), since the derivatives of a function and its majorant must coincide at a mutual point of differentiability where they touch. Moreover, defining \(\tilde{h} = h, \tilde{h}\) as

\[
\tilde{h}(\varepsilon) := 2\left(\tilde{\varphi}_{t_0 + \varepsilon}(x) - \tilde{\varphi}_{t_0}(x) - \varepsilon \partial_t \tilde{\varphi}_{t_0}(x)\right)
\]

it follows that \(h \leq \tilde{h}\) (on \((-t_0, 1 - t_0))\). Dividing by \(\varepsilon^2\) and taking appropriate subsequential limits, we obviously obtain

\[
\underline{P}_2 \varphi_t(x) \leq \underline{P}_2 \tilde{\varphi}_t(x), \quad \overline{P}_2 \varphi_t(x) \leq \overline{P}_2 \tilde{\varphi}_t(x).
\]

Combining these inequalities with those of Lemma 2.4, (3.6) and (3.7), the chain of inequalities in (3) readily follows.

To see (4), let \(t_0 \in \hat{G}(x)\). Consider the function \(f(t) := \tilde{\varphi}_t(x) - \varphi_t(x)\) on \((0, 1)\), which is locally semi-concave by Corollary 3.10. By Proposition 3.6, we know that \(f \geq 0\) with \(f(t_0) = 0\). The function \(f\) is differentiable on \(D(x) \cap D_\ell(x)\) and satisfies \(f'(t) = \frac{\ell_0^2(x)}{2} - \frac{\ell_0^2(x)}{2}\) there. In particular, this holds at \(t_0 \in \hat{G}(x) \subset D(x) \cap D_\ell(x)\) by (1) and \(f'(t_0) = 0\). Note that by Corollary 3.10:

\[
\partial_t f'(t) \leq \frac{\ell_0^2(x)}{2} - \frac{\ell_0^2(x)}{2} \leq \frac{1}{1 - t} \ell_0^2(x) + \frac{1}{t} \ell_0^2(x).
\]
In particular, since both $D_\ell(x) \ni t \mapsto \tilde{\ell}_t(x)$ are continuous at $t = t_0 \in D_\ell(x) \cap D_{\tilde{\ell}}(x)$, for $\ell = \ell, \tilde{\ell}$, it follows that

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall t \in (t_0 - \delta, t_0 + \delta) \cap D_\ell(x) \cap D_{\tilde{\ell}}(x)$$

$$\overline{\partial}_t f'(t) \leq \frac{1}{t - t_0} \tilde{\ell}_{t_0}^2(x) + \frac{1}{t_0} \ell_{t_0}^2(x) + \varepsilon.$$

It follows that on the open interval $I_\delta \coloneqq (t_0 - \delta, t_0 + \delta) \cap (0, 1)$, $f - C_\varepsilon \frac{t^2}{2}$ is concave with $C_\varepsilon$ defined as the constant on the right-hand-side above. Applying Lemma 3.12 below to the translated function $f(\cdot + t_0)$ on the interval $I_\delta - t_0$, it follows that:

$$\frac{1}{t - t_0} \left( \frac{\tilde{\ell}_t^2(x)}{2} - \frac{\ell_{t_0}^2(x)}{2} \right) = \frac{f'(t) - f'(0)}{t - t_0} \geq -C_\varepsilon$$

$$\forall t \in (t_0 - \frac{\delta}{2}, t_0 + \frac{\delta}{2}) \cap D_\ell(x) \cap D_{\tilde{\ell}}(x).$$

As $\tilde{\ell}_{t_0}(x) = \ell_{t_0}(x)$ by (1), we obtain

$$\frac{\ell_t^2(x) - \ell_{t_0}^2(x)}{t - t_0} \leq \frac{\ell_t^2(x) - \ell_{t_0}^2(x)}{t - t_0} + C_\varepsilon$$

$$\forall t \in (t_0 - \frac{\delta}{2}, t_0 + \frac{\delta}{2}) \cap D_\ell(x) \cap D_{\tilde{\ell}}(x).$$

The assertion of (4) now follows by taking appropriate subsequential limits as $t \to t_0$ and using the fact that $\varepsilon > 0$ was arbitrary. \hfill \Box

**Lemma 3.12** Given $I \subset \mathbb{R}$ an open interval containing 0, let $f : I \to \mathbb{R}$ denote a $C$-semi-concave function, so that $I \ni t \mapsto f - C \frac{t^2}{2}$ is concave, $C \geq 0$. Assume that $f \geq 0$ on $I$, that $f$ is differentiable at 0 and that $f(0) = f'(0) = 0$. Then $\overline{\partial}_t |_{t=0} f'(t) \geq -C$, and moreover, $\frac{f'(t)}{t} \geq -C$ for all $t \in D \cap I/2$, where $D \subset I$ denotes the subset (of full measure) of differentiability points of $f$.

Note that the $C$-semi-concavity is equivalent to $\overline{\partial}_t |_{t=0} f'(t) \leq C$, while the conclusion is from the opposite direction. It is not hard to verify that the asserted lower bound is in fact best possible.

**Proof of Lemma 3.12** Set $g = f'$ on $D$. The $C$-semi-concavity is equivalent to the statement that $g(t) - Ct$ is non-increasing on $D$, so that $g(t_2) \leq g(t_1) + \varepsilon > 0$.
\[ C(t_2 - t_1) \text{ for all } t_1, t_2 \in D \text{ with } t_1 < t_2. \] It follows that necessarily \( g(t) \geq -Ct \) for all \( t \in D \cap I/2 \) with \( t \geq 0 \), since:

\[
0 \leq f(2t) - f(0) = \int_0^{2t} g(s)ds \leq \int_0^t (g(0) + Cs)ds \\
+ \int_t^{2t} (g(t) + C(s-t))ds = C\frac{t^2}{2} + tg(t) + C\frac{t^2}{2}.
\]

Repeating the same argument for \( t \mapsto f(-t) \), we see that \( -g(t) \geq Ct \) for all \( t \in D \cap I/2 \) with \( t \leq 0 \). This concludes the proof. \( \square \)

In a sense, Theorem 3.11 (2) is the temporal analogue of the spatial \( 1/2 \)-Hölder regularity proved by Villani in [77, Theorem 8.22]. Formally taking \( s \to t \) in (3.9), it is easy to check that one obtains (for both \( \ell = \ell, \bar{\ell} \)) stronger bounds than in Theorem 3.11 (3) and (4):

\[
-\frac{1}{t} \ell^2_t(x) \leq \partial_t \frac{\bar{\ell}^2_t(x)|_{\hat{G}_\varphi(x)}}{2} \leq \partial_t \frac{\bar{\ell}^2_t(x)|_{\hat{G}_\varphi(x)}}{2} \leq \frac{1}{1-t} \ell^2_t(x) \forall t \in \hat{G}_\varphi(x).
\]

(3.11)

However, we do not know how to rigorously pass from (3.9) to (3.11) or vice versa (by differentiation or integration, respectively), since we cannot exclude the possibility that the (relatively closed in \((0, 1)\)) set \( \hat{G}_\varphi(x) \) has isolated points, nor that it is disconnected. Instead, we can obtain the following stronger version of (3.11) which only holds for a.e. \( t \in \hat{G}_\varphi(x) \), but will prove to be very useful later on.

Corollary 3.13 For all \( x \in X \), for a.e. \( t \in \hat{G}_\varphi(x) \), \( \partial_t \ell^2_t(x) \) and \( \partial_t \bar{\ell}^2_t(x) \) exist, coincide, and satisfy:

\[
-\frac{1}{t} \ell^2_t(x) \leq \partial_t \frac{\ell^2_t(x)|_{\hat{G}_\varphi(x)}}{2} = \partial_t \frac{\bar{\ell}^2_t(x)|_{\hat{G}_\varphi(x)}}{2} = \partial_t \frac{\bar{\ell}^2_t(x)|_{\hat{G}_\varphi(x)}}{2} \leq \frac{1}{1-t} \ell^2_t(x).
\]

(3.12)

Proof By Corollary 3.10, for all \( x \in X \) and \( \ell = \ell, \bar{\ell}, t \mapsto \bar{\ell}^2_t(x) \) is differentiable a.e. on \( D_\ell(x) \). Consequently, the first and third equalities in (3.12) follow for a.e. \( t \in \hat{G}_\varphi(x) \subset D_\ell(x) \cap D_{\bar{\ell}}(x) \) by Remark 2.1. The second equality follows since \( \ell_t(x) = \bar{\ell}_t(x) \) for \( t \in \hat{G}_\varphi(x) \) by Theorem 3.11. The lower and upper bounds in (3.12) then follow from Theorem 3.11 (3) (or as in (3.11), by taking the limit as \( s \to t \) in Theorem 3.11 (2)). \( \square \)
3.5 Null-geodesics

**Definition 3.14 (Null-Geodesics and Null-Geodesic Points)** Given a Kantorovich potential \( \varphi : X \to \mathbb{R} \), denote the subset of null \( \varphi \)-Kantorovich geodesics by

\[
G^0_\varphi := \{ \gamma \in G_\varphi \; ; \, \ell(\gamma) = 0 \}.
\]

Its complement in \( G_\varphi \) will be denoted by \( G^+_\varphi \). The subset of \( X \) of null \( \varphi \)-Kantorovich geodesic points is denoted by

\[
X^0 := \{ x \in X \; ; \, \exists \gamma \in G^0_\varphi \; \gamma \equiv x \} = \{ x \in X \; ; \, \varphi(x) + \varphi^c(x) = 0 \}.
\]

Its complement in \( X \) will be denoted by \( X^+ \).

The following provides a convenient equivalent characterization of \( X^0 \) and \( X^+ \):

**Lemma 3.15** Given \( x \in X \), the following statements are equivalent:

1. \( x \in X^0 \), i.e. \( \varphi(x) + \varphi^c(x) = 0 \).
2. \( \forall t \in (0, 1), \varphi_t(x) = \tilde{\varphi}_t(x) = \varphi(x) = -\varphi^c(x) \).
3. \( \forall t \in (0, 1), \varphi_t(x) = c \) and \( \tilde{\varphi}_t(x) = \tilde{c} \) for some \( c, \tilde{c} \in \mathbb{R} \).
4. \( D_\varphi(x) = D_{\tilde{\varphi}}(x) = (0, 1) \) and \( \forall t \in (0, 1) \), \( \ell_t(x) = \tilde{\ell}_t(x) = 0 \).
5. \( \exists t_0 \in \hat{G}_\varphi(x) \) so that \( \varphi_{t_0}(x) = \varphi(x) \) or \( \tilde{\varphi}_{t_0}(x) = \varphi(x) \) or \( \varphi_{t_0}(x) = -\varphi^c(x) \) or \( \tilde{\varphi}_{t_0}(x) = -\varphi^c(x) \).
6. \( \exists t_0 \in \hat{G}_\varphi(x) \) so that \( \ell_{t_0}^-(x) = 0 \) or \( \ell_{t_0}^+(x) = 0 \) or \( \tilde{\ell}_{t_0}^-(x) = 0 \) or \( \tilde{\ell}_{t_0}^+(x) = 0 \).

In other words, we have the following dichotomy: all \( \varphi \)-Kantorovich geodesics having \( x \in X \) as some interior mid-point have either strictly positive length (iff \( x \in X^+ \)) or zero length (iff \( x \in X^0 \)).

**Remark 3.16** In fact, we always have \( \varphi_t(x) = \tilde{\varphi}_t(x) \) and \( \ell_t(x) = \tilde{\ell}_t(x) \) for \( t \in \hat{G}_\varphi(x) \subset D_\varphi(x) \cap D_{\tilde{\varphi}}(x) \) by Theorem 3.11, so we may simply write “\( \varphi_{t_0}(x) = \varphi(x) \) or \( \varphi_{t_0}(x) = -\varphi^c(x) \)” and “\( \ell_{t_0}(x) = \tilde{\ell}_{t_0}(x) = 0 \)” in statements (5) and (6), respectively. However, we chose to formulate these statements with the (a-priori) minimal requirements.

**Proof of Lemma 3.15** (1) \( \Rightarrow \) (2) is straightforward: for instance, (1) is by definition identical to \( \varphi_1(x) = \varphi_0(x) \) and (2) follows by the monotonicity of \( [0, 1] \ni t \mapsto \tilde{\varphi}_t(x) \) for both \( \varphi = \varphi, \tilde{\varphi} \); alternatively, apply Lemma 3.3 to the null geodesic \( \gamma^0 \equiv x \) with respect to both Kantorovich potentials \( \varphi \) and \( \varphi^c \).

(2) \( \Rightarrow \) (3) is trivial.

(3) \( \Leftrightarrow \) (4) follows by using that \( D_\varphi^+(x) \) is characterized as the subset of \( t \)-differentiability points of \( \varphi_t(x) \) on \( (0, 1) \) with \( \partial_t \varphi_t(x) = \tilde{\ell}_t^2(x)/2 \) there.
(3) $\Rightarrow$ (1): by the continuity of $t \mapsto \varphi_t(x)$ from the left at $t = 1$ it follows that $c = \varphi_1(x)$, and similarly the continuity of $t \mapsto \bar{\varphi}_t(x)$ from the right at $t = 0$ yields that $\bar{c} = \bar{\varphi}_0(x) = \varphi(x)$. Since always $\varphi \leq \bar{\varphi}$, we deduce $\varphi_1(x) = c \leq \bar{c} = \varphi(x)$. On the other hand, we always have $\varphi(x) \leq \varphi_1(x)$ by monotonicity, so we conclude that $\varphi(x) = \varphi_1(x)$, establishing statement (1).

This concludes the proof of the equivalence (1) $\iff$ (2) $\iff$ (3) $\iff$ (4).

(2) $\Rightarrow$ (5) and (4) $\Rightarrow$ (6) are trivial.

(5) $\Rightarrow$ (6) is straightforward: for instance, if $\bar{\varphi}_{t_0}(x) = \bar{\varphi}_0(x) = \varphi(x)$ for some $t_0 \in (0, 1)$ and $\bar{\varphi} \in \{\varphi, \bar{\varphi}\}$, then by monotonicity, $\bar{\varphi}_t(x) = \varphi(x)$ for all $t \in [0, t_0]$, and hence the left derivative at $t = t_0$ satisfies $\ell_{t_0}^t(x) = \partial_{t}^{-}|_{t=t_0} \varphi_t(x) = 0$ if $\bar{\varphi} = \varphi$ and $\ell_{t_0}^t(x) = \partial_{t}^{-}|_{t=t_0} \bar{\varphi}_t(x) = 0$ if $\bar{\varphi} = \bar{\varphi}_t$. If $\bar{\varphi}_{t_0}(x) = \bar{\varphi}_1(x) = -\varphi^{c}(x)$, repeat the argument using the right derivative.

The only direction requiring second-order information on $\varphi_t$ is (6) $\Rightarrow$ (3). By Corollary 3.10, $t \mapsto t \ell_t^+ \varphi_t(x)$ and $t \mapsto (1 - t)\ell_t^- \varphi_t(x)$ are monotone non-decreasing and non-increasing on $(0, 1)$, respectively. Since $t_0 \in \hat{G}_\varphi$, in view of Remark 3.16, (5) is equivalent to $\ell_{t_0}^t(x) = \ell_{t_0}^t(x) = 0$. The monotonicity implies that $\ell_{t_0}^t(x) = 0$ for all $t \in (0, t_0]$ and that $\ell_{t_0}^t(x) = 0$ for all $t \in [t_0, 1)$. It follows that $\varphi_t(x)$ is constant on $(0, t_0]$ and $\bar{\varphi}_t(x)$ is constant on $[t_0, 1)$. As $\varphi_{t_0}(x) = \bar{\varphi}_{t_0}(x)$, the monotonicity of $t \mapsto \varphi_t(x)$ and the majoration $\varphi_t \leq \bar{\varphi}_t$ forces both $t \mapsto \varphi_t(x)$ and $t \mapsto \bar{\varphi}_t(x)$ to be constant on $(0, 1)$, establishing (3) (in fact with $c = \bar{c}$).

\begin{corollary}
If $x \in X^+$ then $\ell_t(x) > 0$ for all $t \in [\inf \hat{G}_\varphi(x), 1) \cap D_\ell(x)$ and $\tilde{\ell}_t(x) > 0$ for all $t \in (0, \sup \hat{G}_\varphi(x)] \cap D_\tilde{\ell}(x)$.
\end{corollary}

\begin{proof}
Immediate by (6) and the monotonicity of $D_\ell(x) \ni t \mapsto t \ell_t(x)$ and $D_\tilde{\ell}(x) \ni t \mapsto (1 - t)\tilde{\ell}_t(x)$, together with the fact that $\hat{G}_\varphi(x)$ is relatively closed in $(0, 1)$ by Corollary 3.9.
\end{proof}

\begin{corollary}
Given $x \in X$, assume that $\exists t_1, t_2 \in \hat{G}_\varphi(x)$ with $t_1 \neq t_2$. Then $x \in X^0$ iff $\varphi_{t_1}(x) = \varphi_{t_2}(x)$ (or equivalently, $\bar{\varphi}_{t_1}(x) = \bar{\varphi}_{t_2}(x)$).
\end{corollary}

\begin{proof}
The “only if” direction follows immediately by Lemma 3.15, whereas the “if” direction follows by Corollary 3.17, after recalling that $\varphi_{t_2}(x) - \varphi_{t_1}(x) = \int_{t_1}^{t_2} \frac{\ell_t^2(x)}{2} - d\tau$ by Corollary 3.10. As usual, the equivalent condition follows by Theorem 3.11.
\end{proof}

4 Temporal theory of intermediate-time Kantorovich potentials: time-propagation

The goal of this section is to introduce and study the following function(s):
Definition (Time-Propagated Intermediate Kantorovich Potentials) Given a Kantorovich potential \( \varphi : X \to \mathbb{R} \) and \( s, t \in (0, 1) \), define the \( t \)-propagated \( s \)-Kantorovich potential \( \Phi^t_s \) on \( D_\ell(t) \), and its time-reversed version \( \bar{\Phi}^t_s \) on \( D_{\bar{\ell}}(t) \), by:

\[
\Phi^t_s := \varphi_t + (t - s) \frac{\ell^2_t}{2} \text{ on } e_t(G_\varphi), \quad \bar{\Phi}^t_s := \bar{\varphi}_t + (t - s) \frac{\bar{\ell}^2_t}{2} \text{ on } D_{\bar{\ell}}(t).
\]

Observe that for all \( s, t \in (0, 1) \):

\[
\Phi^t_s = \bar{\Phi}^t_s = \varphi_s \circ e_s \circ e_t^{-1} \text{ on } e_t(G_\varphi);
\]

indeed, while \( e_t^{-1} : e_t(G_\varphi) \to G_\varphi \) may be multi-valued, Theorem 3.11 implies that \( \ell(\gamma) = \ell_t(x) = \bar{\ell}_t(x) \) for any \( \gamma \in G_\varphi \) with \( \gamma_t = x \), and consequently Lemma 3.3 yields that \( \varphi_s \circ e_s \) is single-valued for all such \( \gamma \) and (also recalling Proposition 3.6):

\[
\Phi^t_s(\gamma_t) = \bar{\Phi}^t_s(\gamma_t) = \varphi_s(\gamma_s) \quad \forall \gamma \in G_\varphi.
\]

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

We will use the following short-hand notation. Given \( s \in [0, 1] \) and \( a_s \in \mathbb{R} \), we denote:

\[
G_{a_s} := \{ \gamma \in G_\varphi ; \varphi_s(\gamma(s)) = a_s \},
\]

suppressing the implicit dependence of \( G_{a_s} \) on \( s \). The above argument about why \( \varphi_s \circ e_s \circ e_t^{-1} \) is well-defined can be rewritten as:

**Corollary 4.1** (Inter Level-Set Propagation) For all \( s, t \in (0, 1), a_s, b_s \in \mathbb{R}, a_s \neq b_s, \) we have:

\[
e_t(G_\varphi) \cap \{ \Phi^t_s = a_s \} \cap \{ \Phi^t_s = b_s \} = e_t(G_{a_s}) \cap e_t(G_{b_s}) = \emptyset.
\]

Note that while typically disjoint sets remain disjoint under optimal-transport only under some additional non-branching assumptions, Corollary 4.1 holds true in general.
4.1 Monotonicity

**Lemma 4.2** Let \( x = \gamma^1_{t_1} = \gamma^2_{t_2} \) with \( \gamma^1, \gamma^2 \in G_\psi \) and \( 0 < t_1 < t_2 < 1 \). Then for any \( s \in (0, 1) \):

\[
\varphi_s(\gamma^2) - \varphi_s(\gamma^1) \geq 2 \min \left( \frac{s}{t_2}, \frac{1-s}{1-t_1} \right) (\varphi_{t_2}(x) - \varphi_{t_1}(x)) \geq 0. \tag{4.1}
\]

Moreover, the left-hand-side is in fact strictly positive iff \( x \in X^+ \).

**Proof** We know by Lemma 3.3 and Theorem 3.11 that:

\[
\varphi_s(\gamma^i) = \varphi_{t_i}(\gamma^i) + (t_i - s) \frac{\ell^2(\gamma^i)}{2} = \varphi_{t_i}(x) + (t_i - s) \frac{\ell^2_{t_i}(x)}{2}, \quad i = 1, 2.
\]

Recall that \( \varphi_{t_i}(x) = \tilde{\varphi}_{t_i}(x) \) and \( \ell_{t_i}(x) = \tilde{\ell}_{t_i}(x) \) by Proposition 3.6 and Theorem 3.11, as \( x = \gamma^i_{t_i} \). Now set \( \tilde{s} := (s \lor t_1) \land t_2 \). Since \( \tilde{s} \in \{t_1, t_2, s\} \), it follows that:

\[
\varphi_s(\gamma^2) - \varphi_s(\gamma^1) - (\varphi_{t_2}(x) - \varphi_{t_1}(x)) = (t_2 - s) \frac{\ell^2_{t_2}(x)}{2} - (\tilde{s} - s) \frac{\ell^2_{\tilde{s}}(x)}{2} + (s - \tilde{s}) \frac{\ell^2_{t_1}(x)}{2} - (t_1 - s) \frac{\ell^2_{t_1}(x)}{2}.
\]

By Corollary 3.10, we know for \( \tilde{\ell} = \ell, \tilde{\ell} \) that \( D_{\tilde{\ell}}(x) \ni t \mapsto \tilde{\ell}^2_{t}(x) \) is differentiable a.e., and that the singular part of its distributional derivative is non-negative for \( \tilde{\ell} = \ell \) and non-positive for \( \tilde{\ell} = \tilde{\ell} \). Consequently, we may proceed as follows:

\[
\geq \int_{\tilde{s}}^{t_2} \ell_{\tau} \left( (\tau - s) \frac{\ell^2_{\tau}(x)}{2} \right) d\tau + \int_{t_1}^{\tilde{s}} \ell_{\tau} \left( (\tau - s) \frac{\ell^2_{\tau}(x)}{2} \right) d\tau,
\]

where we used that \( \tau - s \geq 0 \) when \( \tilde{s} \leq \tau < t_2 \) and that \( \tau - s \leq 0 \) when \( \tilde{s} \geq \tau > t_1 \). Using (3.6) and (3.7) to bound the above lower and upper derivatives on the sets (having full measure) \( D_\ell(x) \) and \( D_{\tilde{\ell}}(x) \), respectively, we obtain:

\[
\geq \int_{\tilde{s}}^{t_2} \left( 1 - 2 \frac{\tau - s}{\tau} \right) \frac{\ell^2_{\tau}(x)}{2} d\tau + \int_{t_1}^{\tilde{s}} \left( 1 + 2 \frac{\tau - s}{1 - \tau} \right) \frac{\ell^2_{\tau}(x)}{2} d\tau
\]

\[
= \int_{\tilde{s}}^{t_2} \left( 2 \frac{s}{\tau} - 1 \right) \frac{\ell^2_{\tau}(x)}{2} d\tau + \int_{t_1}^{\tilde{s}} \left( 2 \frac{1 - s}{1 - \tau} - 1 \right) \frac{\ell^2_{\tau}(x)}{2} d\tau
\]

\[
\geq \left( 2 \frac{s}{t_2} - 1 \right) \int_{\tilde{s}}^{t_2} \frac{\ell^2_{\tau}(x)}{2} d\tau + \left( 2 \frac{1 - s}{1 - t_1} - 1 \right) \int_{t_1}^{\tilde{s}} \frac{\ell^2_{\tau}(x)}{2} d\tau
\]
The globalization theorem

Proof the asserted (4.1) is established in either case. Moreover, (4.1) implies that if \( x \in X \), then all geodesics having \( x \) as an interior point are null by Lemma 3.15, and hence by Corollary 3.18 that \( x \in X^0 \); and vice-versa, if \( x \in X^0 \) then all geodesics having \( x \) as an interior point are null by Lemma 3.15, and hence \( \gamma_s^1 = \gamma_s^2 = x \) and \( \varphi_s(\gamma_s^2) - \varphi_s(\gamma_s^1) = 0 \).

Summarizing, we have obtained:

\[
\varphi_s(\gamma_s^2) - \varphi_s(\gamma_s^1) \geq \left( 2 \frac{s}{t_2} - 1 \right) (\varphi_{t_2}(x) - \varphi_{\bar{t}}(x)) + \varphi_{t_2}(x) + \left( 2 \frac{1-s}{1-t_1} - 1 \right) (\varphi_{\bar{t}}(x) - \varphi_{t_1}(x)) - \varphi_{t_1}(x) .
\]

We now use the inequality \( \varphi_{\bar{t}}(x) \leq \varphi_{\bar{t}}(x) \) in the first line above when \( 2 \frac{s}{t_2} - 1 \geq 0 \), and in the second line when \( 2 \frac{1-s}{1-t_1} - 1 \geq 0 \), yielding

\[
\geq \begin{cases} 
2 \frac{s}{t_2} (\varphi_{t_2}(x) - \varphi_{\bar{t}}(x)) + 2 \frac{1-s}{1-t_1} (\varphi_{\bar{t}}(x) - \varphi_{t_1}(x)) & s \geq \frac{t_2}{2} \\
2 \frac{s}{t_2} (\varphi_{t_2}(x) - \varphi_{\bar{t}}(x)) + 2 \frac{1-s}{1-t_1} (\varphi_{\bar{t}}(x) - \varphi_{t_1}(x)) & 1-s \geq \frac{1-t_1}{2}.
\end{cases}
\]

In particular, the first estimate applies whenever \( s \geq \frac{1}{2} \) and the second one whenever \( s \leq \frac{1}{2} \). Using that \([0, 1] \ni t \mapsto \varphi_t(x)\) is monotone non-decreasing, the asserted (4.1) is established in either case. Moreover, (4.1) implies that if \( \varphi_s(\gamma_s^2) - \varphi_s(\gamma_s^1) = 0 \) then \( \varphi_{t_1}(x) = \varphi_{t_2}(x) \), and hence by Corollary 3.18 that \( x \in X^0 \); and vice-versa, if \( x \in X^0 \) then all geodesics having \( x \) as an interior point are null by Lemma 3.15, and hence \( \gamma_s^1 = \gamma_s^2 = x \) and \( \varphi_s(\gamma_s^2) - \varphi_s(\gamma_s^1) = 0 \).

We can already deduce the following important consequence, complementing Corollary 4.1, which holds for any proper geodesic space \((X, d)\), independently of any additional assumptions like various forms of non-branching:

**Corollary 4.3** (Intra Level-Set Propagation) For any \( s \in (0, 1) \), \( a_s \in \mathbb{R} \), and \( t_1, t_2 \in (0, 1) \) with \( t_1 \neq t_2 \):

\[
e_{t_1}(G_{a_s} \setminus G_0^0) \cap e_{t_2}(G_{a_s} \setminus G_0^0) = e_{t_1}(G_{a_s}) \cap e_{t_2}(G_{a_s}) \cap X^+ = \emptyset.
\]

In other words, for each \( x \in e_{(0,1)}(G_{a_s}) \cap X^+ \), there exists a unique \( t \in (0, 1) \) so that \( x \in e_t(G_{a_s}) \).

**Proof** If \( x = \gamma_{t_1}^1 = \gamma_{t_2}^2 \in X^+ \), \( 0 < t_1 < t_2 < 1 \), then Lemma 4.2 yields \( \varphi_s(\gamma^2(s)) > \varphi_s(\gamma^1(s)) \), establishing the assertion. \( \square \)
4.2 Properties of $\Phi^t_s$

The following information will be crucially used when deriving the Change-Of-Variables formula in Sect. 11:

**Proposition 4.4** For any $s \in (0, 1)$, the following properties of $\Phi^t_s$ and $\Phi^t_s$ hold:

1. The maps $(x, t) \mapsto \Phi^t_s(x)$ and $(x, t) \mapsto \Phi^t_s(x)$ are continuous on $D_{\ell}$ and $D_{\ell}$, respectively.
2. For each $x \in X$, $\Phi = \Phi$, $\Phi$ and $\ell = \ell$, $\bar{\ell}$, respectively, $D_{\ell}(x) \ni t \mapsto \Phi^t_s(x)$ is differentiable at $t$ iff $D_{\ell}(x) \ni t \mapsto \bar{\ell}^2_t(x)$ is differentiable at $t$
3. Points $t$ of differentiability:

$\partial_t \Phi^t_s(x) = \bar{\ell}^2_t(x) + (t - s) \frac{\partial_t \ell^2_t(x)}{2}$. \hspace{1cm} (4.2)

In particular, if $s \in D_{\ell}(x)$ then $\exists \partial_t|_{t=s} \Phi^t_s(x) = \bar{\ell}^2_s(x)$.

4. For each $x \in X$, the map $\hat{G}_\varphi(x) \ni t \mapsto \Phi^t_s(x) = \hat{\Phi}^t_s(x)$ is locally Lipschitz and non-decreasing (if $\# \hat{G}_\varphi(x) \geq 2$, it is strictly increasing iff $x \in X^+$).

5. For all $(x, t) \in D(\hat{G}_\varphi)$:

$$\min \left( \frac{s}{t}, \frac{1 - s}{1 - t}, \frac{t - s}{t(1 - t)} \right) \ell^2_t(x) \leq \partial_t \Phi^t_s(x) \leq \max \left( \frac{s}{t}, \frac{1 - s}{1 - t}, \frac{t - s}{t(1 - t)} \right) \ell^2_t(x),$$

$$\min \left( \frac{1 - s}{1 - t}, \frac{s}{t} \frac{t - s}{t(1 - t)} \right) \ell^2_t(x) \leq \partial_j \Phi^t_s(x) \leq \max \left( \frac{1 - s}{1 - t}, \frac{s}{t} \frac{t - s}{t(1 - t)} \right) \ell^2_t(x).$$
Proof Recall that

$$\tilde{\Phi}_s^t := \tilde{\phi}_t(x) + (t - s) \frac{\tilde{\ell}_t^2(x)}{2} \quad \text{on } D_{\tilde{\ell}}.$$ 

The first and second statements follow by Lemma 3.2 and Corollary 3.10. As $t \mapsto \tilde{\phi}_t(x)$ is differentiable on $D_{\tilde{\ell}}(x)$ with derivative $\frac{\tilde{\ell}_t^2(x)}{2}$, the points of differentiability of $t \mapsto \tilde{\Phi}_s^t(x)$ must coincide with those of $t \mapsto \tilde{\ell}_t^2(x)$ and (4.2) follows immediately, with the only possible exception being the point $t = s$ if $s \in D_{\ell}(x)$, where direct inspection and continuity of $t \mapsto \tilde{\ell}_t^2(x)$ on $D_{\tilde{\ell}}(x)$ verifies (4.2). The local Lipschitzness follows by Theorem 3.11 (2). The monotonicity follows by Lemma 4.2, since if $\gamma' \in G_\phi$ is such that $\gamma'_t = x$, then $\Phi_s^t(\gamma'_t) = \tilde{\Phi}_s^t(\gamma'_t) = \varphi_s(\gamma'_t)$. The last two assertions follow as in the proof of Lemma 4.2, after noting that:

$$\begin{cases}
\partial_t \Phi_s^t(x) = \tilde{\ell}_t^2(x) + (t - s) \frac{\partial_t \tilde{\ell}_t^2(x)}{2} & t \geq s \\
\partial_t \Phi_s^t(x) = \tilde{\ell}_t^2(x) + (t - s) \frac{\partial_t \tilde{\ell}_t^2(x)}{2} & t \leq s
\end{cases} \forall x \in D_{\tilde{\ell}}(t),$$

and similarly for $\bar{\partial}_t$. Indeed, the estimates (3.6) and (3.7) of Corollary 3.10 yield (4), which already yields half of the inequalities in (5) for all $(x, t) \in D_{\ell} \cap D_{\tilde{\ell}}$. To get the other half, we must restrict to $D(G_\phi)$ and use the estimates of Theorem 3.11 (4), thereby concluding the proof. \qed

As an immediate corollary of Proposition 4.4, Corollary 3.13 and Lemma 3.15, we obtain:

**Corollary 4.5** For all $x \in X$, for a.e. $t \in G_\phi(x)$, $\partial_t \Phi_s^t(x)$ and $\partial_t \tilde{\Phi}_s^t(x)$ exist, coincide, and satisfy:

$$\min \left( \frac{s}{t}, \frac{1 - s}{1 - t} \right) \ell_t^2(x) \leq \partial_t \Phi_s^t(x) = \partial_t \tilde{\Phi}_s^t(x) |_{G_\phi(x)} = \partial_t \Phi_s^t(x) |_{G_\phi(x)} = \partial_t \tilde{\Phi}_s^t(x) \leq \max \left( \frac{s}{t}, \frac{1 - s}{1 - t} \right) \ell_t^2(x).$$

In particular, if $x \in X^+$ then $\partial_t \Phi_s^t(x) > 0$ for a.e. $t \in G_\phi(x)$.

We will also require the following consequence of Proposition 4.4 and Theorem 3.11:

**Lemma 4.6** For any $x \in X$, $s \in (0, 1)$, and $\Phi = \Phi$, $\Phi$ and $\tilde{\ell} = \ell$, $\tilde{\ell}$, respectively:

$$\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{(s - \varepsilon, s + \varepsilon) \cap G_\phi(x)} \left( \partial_t \Phi_s^t(x) - \tilde{\ell}_s^2(x) \right) dt = 0.$$
Proof} By (4.2), the claim boils down to proving
\[
\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{(s-\varepsilon,s+\varepsilon) \cap \tilde{G}_\phi(x)} (t - s) \partial_t \tilde{\ell}_t^2(x) \, dt = 0.
\]

Using Corollary 3.13, it follows that:
\[
\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \left| \int_{(s-\varepsilon,s+\varepsilon) \cap \tilde{G}_\phi(x)} (t - s) \partial_t \tilde{\ell}_t^2(x) \, dt \right| \leq \lim_{\varepsilon \to 0} \int_{(s-\varepsilon,s+\varepsilon) \cap \tilde{G}_\phi(x)} \left| \partial_t \frac{\tilde{\ell}_t^2(x)}{2} \right| \, dt.
\]

But the latter limit is clearly 0 (e.g. by Corollary 3.10 (1)). \qed

5 Temporal theory of intermediate-time Kantorovich potentials: third order

Fix a non-null Kantorovich geodesic $\gamma \in G^+_\phi$, and denote for short $\ell := \ell(\gamma) > 0$. Recall by the results of Sect. 3 that for all $t \in (0, 1)$, $\ell_t(\gamma_t) = \tilde{\ell}_t(\gamma_t) = \ell$ and that $\partial_t \varphi_t(x) = \partial_t \bar{\varphi}_t(x) = \tilde{\ell}_t^2(x)/2$ for all $x \in e_t(G_\phi)$. Also, recall that given $x \in X$ and $\ell = \tilde{\ell}, \bar{\ell}$, the function $D_{\tilde{\ell}}(x) \ni t \mapsto \tilde{\ell}_t(x)$ is only a.e. differentiable, and even on $\tilde{G}_\phi(x) \subset D_{\tilde{\ell}}(x) \cap D_{\bar{\ell}}(x)$, we only have at the moment upper and lower bounds on $\partial_t \tilde{\ell}_t^2(x)/2$ and $\partial_t \tilde{\ell}_t^2(x)/2$, i.e. second order information on $\tilde{\varphi}_t(x)$.

The goal of this section is to rigorously make sense and prove the following formal statement, which provides second order information on $\ell_t$, or equivalently, third order information on $\varphi_t$, along $\gamma_t$:

\[
z(t) := \partial_{\tau} |_{\tau = t} \frac{\ell_t^2}{2}(\gamma_t) \quad \Rightarrow \quad z'(t) \geq \frac{z(t)^2}{\ell_t^2}.
\] (5.1)

Equivalently, this amounts to the statement that the function:

\[
L(r) = \exp \left( -\frac{1}{\ell_t^2} \int_{r_0}^r \partial_{\tau} |_{\tau = t} \frac{\ell_t^2}{2}(\gamma_t) \, dt \right)
\]
is concave in $r \in (0, 1)$, since formally:

\[
\frac{L''}{L} = (\log L)' + ((\log L)')^2 = -\frac{z'}{\ell_t^2} + \frac{z^2}{\ell_t^4} \leq 0.
\]
5.1 Formal argument

We start by providing a formal proof of (5.1) in an infinitesimally Hilbertian setting, which is rigorously justified on a Riemannian manifold if all involved functions are smooth (in time and space).

Recall that the Hopf-Lax semi-group solves the Hamilton-Jacobi equation (e.g. [6]):

$$\partial_t\varphi_t = \frac{1}{2} \ell_t^2 = \frac{1}{2} |\nabla\varphi_t|^2.$$  \hspace{1cm} (5.2)

We evaluate all subsequent functions at $x = \gamma_t$. Since:

$$z(t) = \partial_t^2\varphi_t(\gamma(t)) = \langle \nabla\partial_t\varphi_t, \nabla\varphi_t \rangle,$$

and since $\gamma'(t) = -\nabla\varphi_t$ (see e.g. [6] or Lemma 10.3),

$$z'(t) = \partial_t^3\varphi_t - \langle \nabla\partial_t^2\varphi_t, \nabla\varphi_t \rangle.$$

But taking two time derivatives in (5.2), we know that:

$$\partial_t^3\varphi_t = \langle \nabla\partial_t^2\varphi_t, \nabla\varphi_t \rangle + \langle \nabla\partial_t\varphi_t, \nabla\partial_t\varphi_t \rangle,$$

and so we conclude that:

$$z'(t) = |\nabla\partial_t\varphi_t|^2.$$

It remains to apply Cauchy–Schwarz and deduce:

$$z'(t) \geq \left( \frac{\nabla\partial_t\varphi_t}{|\nabla\varphi_t|} \right)^2 = \frac{z(t)^2}{\ell^2},$$

as asserted. Note that in a general setting, we can try and interpret $z(t)$ as minus the directional derivative of $\ell_t^2/2 = \partial_t\varphi_t$ in the direction of $\gamma'(t)$ (by taking derivative of the identity $\ell_t^2(\gamma(t)) \equiv \ell_t^2$), and thus hope to justify the Cauchy–Schwarz inequality as the statement that the local Lipschitz constant of $\partial_t\varphi_t$ is greater than any unit-directional derivative. However, a crucial point in the above argument of identifying $z'(t)$ with $|\nabla\partial_t\varphi_t|^2$ was to use the linearity of $\langle \cdot, \cdot \rangle$ in both of its arguments, and so ultimately this formal proof is genuinely restricted to an infinitesimally Hilbertian setting.

The above discussion seems to suggest that there is no hope of proving (5.1) beyond the Hilbertian setting. Furthermore, it seems that the spatial regularity of $\varphi_t$ and $\partial_t\varphi_t = \frac{1}{2} \ell_t^2$ should play an essential role in any rigorous justification.
Remarkably, we will see that this is not the case on both counts, and that an appropriate interpretation of (5.1) holds true on a general proper geodesic space \((X, \mathbf{d})\).

5.2 Notation

Recall that by the results of Sect. 3, \(\tau \mapsto \varphi_\tau(x)\) and \(\tau \mapsto \bar{\varphi}_\tau(x)\) are locally semi-convex and semi-concave on \((0, 1)\), respectively, and that \(\bar{\partial}_t^{\pm} \varphi_\tau(x) = \tilde{\ell}_t^{\pm}(x)^2/2\), \(\bar{\partial}_t^{\pm} \bar{\varphi}_\tau(x) = \tilde{\ell}_t^{\pm}(x)^2/2\) and \(\ell_t^{\pm}(\gamma_t) = \ell_t^{\pm}(\gamma_t) = \ell\) for all \(t \in (0, 1)\).

We respectively introduce \(\tilde{p} = p, \bar{p}\) by defining at \(t \in (0, 1)\):

\[
\begin{align*}
\tilde{p}^+_{\tau}(t) &= \bar{p}^+_{\tau}(t) := \partial_\tau \left|_{\tau = t} \tilde{\ell}_t^2(\gamma_t) \right| / 2 = \ell \cdot \partial_\tau \left|_{\tau = t} \tilde{\ell}_t(\gamma_t) \right| = \ell \cdot \partial_\tau \left|_{\tau = t} \tilde{\ell}_t^\pm(\gamma_t) \right|, \\
\tilde{p}^-_{\tau}(t) &= \bar{p}^-_{\tau}(t) := \partial_\tau \left|_{\tau = t} \tilde{\ell}_t^2(\gamma_t) \right| / 2 = \ell \cdot \partial_\tau \left|_{\tau = t} \tilde{\ell}_t(\gamma_t) \right| = \ell \cdot \partial_\tau \left|_{\tau = t} \tilde{\ell}_t^\pm(\gamma_t) \right),
\end{align*}
\]

where the penultimate equalities in each of the lines above follow from the continuity of \(D^\pm(\gamma_t) \ni \tau \mapsto \tilde{\ell}_t(\gamma_t)\) at \(\tau = t \in G_\varphi(\gamma_t) \subset D^\pm(\gamma_t)\), and the last ones by the monotonicity of \(\tau \mapsto \tau \ell_t^\pm(\gamma_t)\) and \(\tau \mapsto (1 - \tau) \tilde{\ell}_t^\pm(\gamma_t)\) and the density of \(D^\pm\) in \((0, 1)\). Clearly \(\tilde{p}^-_{\tau}(t) \leq \tilde{p}^+_{\tau}(t)\), and \(\tilde{p}^-_{\tau}(t) = \tilde{p}^+_{\tau}(t) = \tilde{p} \in \mathbb{R}\) iff \(D^\pm(\gamma_t) \ni \tau \mapsto \tilde{\ell}_t^2(\gamma_t)/2\) is differentiable at \(\tau = t\) with derivative \(\tilde{p}\). In addition, for \(\tilde{q} = q, \tilde{q}'\), set:

\[
\tilde{q}_+(t) := \tilde{P}_2 \tilde{\varphi}_\tau(x)|_{x = \gamma_t} \geq \tilde{P}_2 \tilde{\varphi}_\tau(x)|_{x = \gamma_t} := \tilde{q}_-(t),
\]

where the Peano (partial) derivatives are with respect to the \(t\) variable. It will be useful to recall that if we define \(\tilde{h} = h, \tilde{h}\) by:

\[
\begin{align*}
\tilde{h}(t, \varepsilon) := 2(\tilde{\varphi}_{t+\varepsilon}(\gamma_t) - \tilde{\varphi}_t(\gamma_t) - \varepsilon \partial_t \tilde{\varphi}_t(\gamma_t)) \\
= 2(\tilde{\varphi}_{t+\varepsilon}(\gamma_t) - \varphi_t(\gamma_t) - \varepsilon \ell^2/2),
\end{align*}
\]

then:

\[
\tilde{q}_+(t) = \limsup_{\varepsilon \to 0} \frac{\tilde{h}(t, \varepsilon)}{\varepsilon^2} \geq \liminf_{\varepsilon \to 0} \frac{\tilde{h}(t, \varepsilon)}{\varepsilon^2} = \tilde{q}_-(t).
\]

By definition, \(\tilde{q}_-(t) = \tilde{q}_+(t) = \tilde{q} \in \mathbb{R}\) if and only if \(\tau \mapsto \tilde{\varphi}_\tau(\gamma_t)\) has second order Peano derivative at \(\tau = t\) equal to \(\tilde{q}\), and hence by Lemma 2.3, iff \(\tilde{p}^-_{\tau}(t) = \tilde{p}^+_{\tau}(t) = \tilde{q}\), or equivalently, iff any of the other equivalent conditions for the second order differentiability of \((0, 1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_t)\) at \(\tau = t\) are satisfied. Moreover, Lemma 2.4 implies:

\[
\tilde{p}^-_{\tau}(t) \leq \tilde{q}_-(t) \leq \tilde{q}_+(t) \leq \tilde{p}^+_{\tau}(t) \quad \forall t \in (0, 1),
\]
but we will not require this here. We summarize the above discussion in:

**Corollary 5.1** The following statements are equivalent for a given \( t \in (0, 1) \):

1. \( \tilde{p}(t) = \tilde{p}(t) = \tilde{p} \in \mathbb{R} \), i.e. \( D_{\tilde{p}}(\gamma_t) \ni \tau \mapsto \tilde{p}(\gamma_t) \) is differentiable at \( \tau = t \) with derivative \( \tilde{p} \).
2. \( \tilde{q}(t) = \tilde{q}(t) = \tilde{q} \in \mathbb{R} \), i.e. \( (0, 1) \ni \tau \mapsto \tilde{q}(\gamma_t) \) has a second Peano derivative at \( \tau = t \) equal to \( \tilde{q} \).

In any of these cases \((0, 1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_t) \) is twice differentiable at \( \tau = t \), and we have:

\[
\partial^2 \varphi_{\tau=t}(\gamma_t) := \partial \tau |_{\tau=t} \tilde{p} = \tilde{q}.
\]

### 5.3 Main inequality

The following inequality and its consequences are the main results of this section.

**Theorem 5.2** For all \( s < t \) and \( \varepsilon \) so that \( s, t, s + \varepsilon, t + \varepsilon \in (0, 1) \), we have (for both possibilities for \( \pm \)):

\[
\frac{h(t, \varepsilon) - h(s, \varepsilon)}{t - s} \geq \frac{s + \varepsilon}{t + \varepsilon} (\ell_{s+\varepsilon}(\gamma_s) - \ell_s(\gamma_s))^2,
\]

and

\[
\frac{\tilde{h}(t, \varepsilon) - \tilde{h}(s, \varepsilon)}{t - s} \geq \frac{1 - t - \varepsilon}{1 - s - \varepsilon} (\ell_{t+\varepsilon}(\gamma_t) - \ell_t(\gamma_t))^2.
\]

**Proof** By Lemma 3.5, there exists \( y^\pm_s \in X \) so that

\[
-\varphi_{s+\varepsilon}(\gamma_s) = \frac{d^2(y^\pm_s, \gamma_s)}{2(s + \varepsilon)} - \varphi(y^\pm_s),
\]

with \( d(y^\pm_s, \gamma_s) = D_{\varphi}^{-}(\gamma_s, s + \varepsilon) = (s + \varepsilon)\ell_{s+\varepsilon}(\gamma_s) =: D_{s+\varepsilon}^{-}. \) By definition, note that:

\[
-\varphi_{t+\varepsilon}(\gamma_t) \leq \frac{d^2(y^\pm_s, \gamma_t)}{2(t + \varepsilon)} - \varphi(y^\pm_t).
\]

We abbreviate \( D_{r} := r \ell = d(\gamma_r, \gamma_0), r = s, t \). The proof consists of subtracting the above two expressions and applying the triangle inequality:

\[
d(y^\pm_s, \gamma_t) \leq d(y^\pm_s, \gamma_s) + d(\gamma_s, \gamma_t) = D_{s+\varepsilon}^{-} + (D_t - D_s) = D_t + (D_{s+\varepsilon}^{-} - D_s).
\]
Indeed, we obtain after subtraction, recalling the definition of $h$, and an application of Lemma 3.3:

$$0 \leq \varphi_{t+\varepsilon}(\gamma_1) - \varphi_{s+\varepsilon}(\gamma_3) + \frac{d^2(y^\pm_\varepsilon, \gamma_1)}{2(t + \varepsilon)} - \frac{d^2(y^\pm_\varepsilon, \gamma_3)}{2(s + \varepsilon)}$$

$$= \frac{1}{2} (h(t, \varepsilon) - h(s, \varepsilon)) + \varphi_t(\gamma_1) - \varphi_s(\gamma_3) + \frac{d^2(y^\pm_\varepsilon, \gamma_1)}{2(t + \varepsilon)} - \frac{(D^\pm_{s+\varepsilon})^2}{2(s + \varepsilon)}$$

$$\leq \frac{1}{2} (h(t, \varepsilon) - h(s, \varepsilon)) - \frac{\ell^2}{2} (t - s) - \frac{(D^\pm_{s+\varepsilon})^2}{2(s + \varepsilon)}$$

$$+ \frac{(D^\pm_{s+\varepsilon} - D_s)^2 + D^2_t + 2(D^\pm_{s+\varepsilon} - D_s)D_t}{2(t + \varepsilon)}.$$ 

Carefully rearranging terms, we obtain:

$$\frac{1}{2} (h(t, \varepsilon) - h(s, \varepsilon)) \geq \frac{(D^\pm_{s+\varepsilon})^2}{2(s + \varepsilon)} - \frac{D^2_s}{2s} + \frac{D^2_x}{2} \left( \frac{1}{t} - \frac{1}{t + \varepsilon} \right) - \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2(t + \varepsilon)} + D^2_t \left( \frac{1}{t} - \frac{1}{t + \varepsilon} \right)$$

$$= \frac{1}{2(s + \varepsilon)} ((D^\pm_{s+\varepsilon})^2 - D^2_s) + \frac{D^2_s}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{s} \right) + \frac{D^2_t}{2} \left( \frac{1}{t} - \frac{1}{t + \varepsilon} \right)$$

$$- \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2(t + \varepsilon)}$$

$$= (D^\pm_{s+\varepsilon} - D_s) \left( \frac{D^\pm_{s+\varepsilon} + D_s}{2(s + \varepsilon)} - \frac{D_t}{t + \varepsilon} \right) + \frac{\ell^2}{2} \left( \frac{\varepsilon t}{t + \varepsilon} - \frac{\varepsilon s}{s + \varepsilon} \right) - \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2(t + \varepsilon)}$$

$$= (D^\pm_{s+\varepsilon} - D_s) \left( \frac{D^\pm_{s+\varepsilon} + D_s}{2(s + \varepsilon)} - \frac{2D_s}{t + \varepsilon} + \frac{2s + \varepsilon}{2(s + \varepsilon)}D_t \left( \frac{1}{t} - \frac{1}{t + \varepsilon} \right) \right)$$

$$+ \varepsilon^2 \frac{\ell^2}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right) - \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2(t + \varepsilon)}$$

$$= \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right) - \varepsilon \frac{D^\pm_{s+\varepsilon} - D_s}{s + \varepsilon}$$

$$+ (D^\pm_{s+\varepsilon} - D_s)D_t \left( \frac{1}{t} - \frac{1}{t + \varepsilon} \right) + \varepsilon^2 \frac{\ell^2}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right)$$

$$= \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right) \left( \frac{(D^\pm_{s+\varepsilon} - D_s)^2}{2} - \varepsilon \ell (D^\pm_{s+\varepsilon} - D_s) + \varepsilon^2 \frac{\ell^2}{2} \right)$$

$$= \frac{1}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right) (D^\pm_{s+\varepsilon} - D_s - \varepsilon \ell)^2$$

$$= \frac{1}{2} \left( \frac{1}{s + \varepsilon} - \frac{1}{t + \varepsilon} \right) (s + \varepsilon)^2 (\ell^\pm_{s+\varepsilon}(\gamma_3) - \ell)^2$$

$$= \frac{t - s}{2} \frac{s + \varepsilon}{t + \varepsilon} (\ell^\pm_{s+\varepsilon}(\gamma_3) - \ell)^2,$$

and the first claim follows.
The second claim follows by the duality between $\varphi$ and $\varphi^c$. Indeed, exchange $\varphi, \gamma, \varepsilon, s, t$ with $\varphi^c, \gamma^c, -\varepsilon, 1-t, 1-s$, and recall that $\bar{\varphi}_t = -\varphi^c_{1-t}$. A straightforward inspection of the definitions verifies:

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

and

$$\frac{(\ell_{1-t-\varepsilon}^c(\gamma_{1-t}^c))^2}{2} = -\partial^+ \varphi^c_{1-t-\varepsilon}(\gamma_{1-t}^c) = \partial^+ \varphi^c_{1-t+\varepsilon}(\gamma_t) = \frac{(\ell_{1-t+\varepsilon}^c(\gamma_t))^2}{2},$$

and so the second claim follows from the first one. Alternatively, one may repeat the above argument by subtracting the following two expressions:

$$\bar{\varphi}_{t+\varepsilon}(\gamma_t) = \frac{d^2(z_{\varepsilon}^\pm, \gamma_t)}{2(1-t-\varepsilon)} - \varphi^c(z_{\varepsilon}^\pm),$$

$$\bar{\varphi}_{s+\varepsilon}(\gamma_s) \leq \frac{d^2(z_{\varepsilon}^\pm, \gamma_s)}{2(1-s-\varepsilon)} - \varphi^c(z_{\varepsilon}^\pm),$$

with $d(z_{\varepsilon}^\pm, \gamma_t) = D^\pm_{\varphi^c}(\gamma_t, 1-t-\varepsilon) = (1-t-\varepsilon)\ell_{t+\varepsilon}^c(\gamma_t)$ and applying the triangle inequality $d(z_{\varepsilon}, \gamma_s) \leq d(z_{\varepsilon}, \gamma_t) + d(\gamma_t, \gamma_s)$.

\section*{5.4 Consequences}

As immediate corollaries of Theorem 5.2, we obtain after diving both sides by $\varepsilon^2$ and taking appropriate subsequential limits as $\varepsilon \to 0$:

**Corollary 5.3** For both $\tilde{q} = q, \bar{q}$, the functions $t \mapsto \tilde{q}_-(t)$ and $t \mapsto \tilde{q}_+(t)$ are monotone non-decreasing on $(0, 1)$.

**Corollary 5.4** For all $0 < s < t < 1$ (and both possibilities for $\pm$):

$$\frac{q_+(t) - q_-(s)}{t-s} \geq \frac{s}{t} \left( \frac{p_+(s)}{\ell} \right)^2, \quad (5.3)$$

and

$$\frac{\bar{q}_+(t) - \bar{q}_-(s)}{t-s} \geq \frac{1-t}{1-s} \left( \frac{\bar{p}_+(t)}{\ell} \right)^2. \quad (5.4)$$

It will be convenient to use the above information in the following form:
Theorem 5.5 Assume that for a.e. \( t \in (0, 1) \):

\[(0, 1) \ni \tau \mapsto \tilde{\varphi}_\tau(\gamma_t)
\text{is twice differentiable at } \tau = t \text{ for both } \varphi, \bar{\varphi} \quad (5.5)\]

in any of the equivalent senses given by Corollary 5.1, and that moreover:

\[
\partial^2_{\tau} |_{\tau = t} \varphi(\gamma_t) = \partial^2_{\tau} |_{\tau = t} \bar{\varphi}(\gamma_t) \text{ for a.e. } t \in (0, 1).
\]

(5.6)

Furthermore, assume that the latter joint value coincides a.e. on \((0, 1)\) with some continuous function \(z_c\):

\[
\partial^2_{\tau} |_{\tau = t} \varphi(\gamma_t) = \partial^2_{\tau} |_{\tau = t} \bar{\varphi}(\gamma_t) = z_c(t) \text{ for a.e. } t \in (0, 1).
\]

(5.7)

Then (5.5) holds for all \( t \in (0, 1) \), and we have:

\[
\partial^2_{\tau} |_{\tau = t} \varphi(\gamma_t) = \partial^2_{\tau} |_{\tau = t} \bar{\varphi}(\gamma_t) = \partial_{\tau} |_{\tau = t} \frac{\ell^2}{2} (\gamma_t)
= \partial_{\tau} |_{\tau = t} \frac{\ell^2}{2} (\gamma_t) = z_c(t) \quad \forall t \in (0, 1).
\]

(5.8)

Moreover, we have the following third order information on \(\varphi_t(x)\) at \(x = \gamma_t\):

\[
\frac{z_c(t) - z_c(s)}{t - s} \geq \sqrt{\frac{s}{t} \frac{1 - t}{1 - s}} \frac{|z_c(s)||z_c(t)|}{\ell^2} \quad \forall 0 < s < t < 1.
\]

(5.9)

In particular, for any point \( t \in (0, 1) \) where \(z_c(t)\) is differentiable:

\[
z_c'(t) \geq \frac{z_c(t)^2}{\ell^2}.
\]

Proof The assumptions imply by Corollary 5.1 that \( \tilde{q}_-(t) = \tilde{q}_+(t) = z_c(t) \) for a.e. \( t \in (0, 1) \). It follows that the same is true for every \( t \in (0, 1) \) by monotonicity of \( \tilde{q}_\pm \) and the assumption that \( z_c \) is continuous, yielding (5.8).

Furthermore, Corollary 5.1 implies that \( \tilde{p}_-(t) = \tilde{p}_+(t) = z_c(t) \) for both \( \tilde{p} = p, \bar{p} \) and for all \( t \in (0, 1) \), and we obtain (5.9) by taking geometric mean of (5.3) and (5.4). The final assertion obviously follows by taking the limit in (5.9) as \( s \to t \).

\[\square\]

We do not know whether all three assumptions (5.5), (5.6) and (5.7) hold for a.e. \( t \in (0, 1) \) for a fixed Kantorovich geodesic \( \gamma \). However, we can guarantee the first two assumptions, at least for almost all Kantorovich geodesics, in the following sense:
Lemma 5.6 Let \( \nu \) denote any \( \sigma \)-finite Borel measure concentrated on \( G_\varphi \), so that for a.e. \( t \in (0, 1) \), \( \mu_t := (e_t)_*(\nu) \ll m \) for some \( \sigma \)-finite Borel measure \( m \) on \( X \). Then for \( \nu \)-a.e. geodesic \( \gamma \), (5.5) and (5.6) hold for a.e. \( t \in (0, 1) \).

Proof Recall that \( D(\hat{G}_\varphi) \) is closed in \( X \times (0, 1) \) by Corollary 3.9. Denote the following Borel subsets:

\[
P := \left\{ (x, t) \in D(\hat{G}_\varphi) : \exists \partial_t \ell_t^2(x), \exists \partial_t \tilde{\ell}_t^2(x), \partial_t \ell_t^2(x)/2 = \partial_t \tilde{\ell}_t^2(x)/2 \right\},
\]

\[
B := D(\hat{G}_\varphi) \setminus P.
\]

By Corollary 3.13, we know that \( L^1(B(x)) = 0 \) for all \( x \in X \). By Fubini

\[
0 = \int L^1(B(x))m(dx) = \int_0^1 m(B(t))L^1(dt),
\]

and so for a.e. \( t \in (0, 1) \), \( m(B(t)) = 0 \). Since \( \mu_t \ll m \) for a.e. \( t \in (0, 1) \), it follows that for a.e. \( t \in (0, 1) \), \( \nu(e_t^{-1}B(t)) = \mu_t(B(t)) = 0 \). In other words, for a.e. \( t \in (0, 1) \), the Borel set \( \{ \gamma \in G_\varphi : \gamma_t \in B(t) \} \) has zero \( \nu \)-measure. Applying Fubini again as before

\[
0 = \int \nu(\{ \gamma \in G_\varphi : \gamma_t \in B(t) \})L^1(dt) = \int L^1(\{ t \in (0, 1) ; \gamma_t \in B(t) \})\nu(\gamma),
\]

we conclude that for \( \nu \)-a.e. \( \gamma \in G_\varphi \), the set \( \{ t \in (0, 1) ; \gamma_t \in B(t) \} \) has zero Lebesgue measure, or equivalently, the set

\[
\{ t \in (0, 1) ; \exists \partial_t \|_{\tau=t} \ell_t^2(\gamma_t), \exists \partial_t \|_{\tau=t} \tilde{\ell}_t^2(\gamma_t), \partial_t \|_{\tau=t} \ell_t^2(\gamma_t)/2 = \partial_t \|_{\tau=t} \tilde{\ell}_t^2(\gamma_t)/2 \}
\]

has full Lebesgue measure. The asserted (5.5) and (5.6) now directly follow from an application of Corollary 5.1. \qed

Finally, we obtain the following concise interpretation of the 3rd order information on \( \tau \mapsto \varphi_\tau \) along \( \gamma_t \), which will play a crucial role in this work:

Lemma 5.7 Assume that for some locally absolutely continuous function \( z_{ac} \) on \( (0, 1) \) we have

\[
\exists \partial_t \|_{\tau=t} \ell_t^2/2(\gamma_t) = z_{ac}(t) \text{ for a.e. } t \in (0, 1).
\]

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

\[
L(r) = \exp\left(-\frac{1}{\ell^2} \int_{r_0}^r \partial_t \|_{\tau=t} \ell_t^2/2(\gamma_t)dt\right) = \exp\left(-\frac{1}{\ell^2} \int_{r_0}^r z_{ac}(t)dt\right).
\]
is concave on \((0, 1)\).

**Proof** Since \(L \in C^1(0, 1)\), concavity of \(L\) is equivalent to showing that the function:

\[
W(r) := -\ell^2 L'(r) = L(r) z_{ac}(r)
\]

is monotone non-decreasing. But as this function is locally absolutely continuous, this is equivalent to showing that \(W'(r) \geq 0\) for a.e. \(r \in (0, 1)\). Note that the points of differentiability of \(W\) and \(z_{ac}\) coincide. At these points (of full Lebesgue measure), we indeed have

\[
W'(r) = L'(r) z_{ac}(r) + L(r) z'_{ac}(r) = L(r) (z'_{ac}(r) - z_{ac}(r)^2 / \ell^2) \geq 0,
\]

where the last inequality follows from Theorem 5.5. This concludes the proof. \(\square\)

We will subsequently show that under synthetic curvature conditions, the above assumption is indeed satisfied for \(\nu\)-a.e. geodesic \(\gamma\).

**Part II Disintegration theory of Optimal-Transport**

**6 Preliminaries**

So far we have worked without considering any reference measure over our metric space \((X, d)\). A triple \((X, d, m)\) is called a metric measure space, m.m.s. for short, if \((X, d)\) is a complete and separable metric space and \(m\) is a non-negative Borel measure over \(X\). In this work we will only be concerned with the case that \(m\) is a probability measure, that is \(m(X) = 1\), and hence \(m\) is automatically a Radon measure (i.e. inner-regular). We refer to [3, 5, 43, 76, 77] for background on metric measure spaces in general, and the theory of Optimal-Transport on such spaces in particular.

**6.1 Geometry of Optimal-Transport on metric measure spaces**

The space of all Borel probability measures over \(X\) will be denoted by \(\mathcal{P}(X)\). It is naturally equipped with its weak topology, in duality with bounded continuous functions \(C_b(X)\) over \(X\). The subspace of those measures having finite second moment will be denoted by \(\mathcal{P}_2(X)\), and the subspace of \(\mathcal{P}_2(X)\) of those measures absolutely continuous with respect to \(m\) is denoted by \(\mathcal{P}_2(X, d, m)\). The weak topology on \(\mathcal{P}_2(X)\) is metrized by the \(L^2\)-Wasserstein distance \(W_2\).
defined as follows for any $\mu_0, \mu_1 \in \mathcal{P}(X)$:

$$W^2_2(\mu_0, \mu_1) := \inf_{\pi} \int_{X \times X} \text{d}^2(x, y) \pi(dx, dy),$$  \hspace{1cm} (6.1)

where the infimum is taken over all $\pi \in \mathcal{P}(X \times X)$ having $\mu_0$ and $\mu_1$ as the first and the second marginals, respectively; such candidates $\pi$ are called transference plans. It is known that the infimum in (6.1) is always attained for any $\mu_0, \mu_1 \in \mathcal{P}(X)$, and the transference plans realizing this minimum are called optimal transference plans between $\mu_0$ and $\mu_1$. When $W^2_2(\mu_0, \mu_1) < \infty$, it is known that given an optimal transference plan $\pi$, there exists a Kantorovich potential $\varphi : X \to \mathbb{R}$ (see Sect. 3), which is associated to $\pi$, meaning that

$$\varphi(x) + \varphi^c(y) = \frac{\text{d}(x, y)^2}{2} \quad \text{for } \pi\text{-a.e. } (x, y) \in X \times X. \hspace{1cm} (6.2)$$

In particular, when $\mu_0, \mu_1 \in \mathcal{P}_2(X)$, then necessarily $W^2_2(\mu_0, \mu_1) < \infty$ and the above discussion applies. Moreover, in this case, it is known that for any Kantorovich potential $\varphi$ associated to an optimal transference plan between $\mu_0$ and $\mu_1$, (6.2) in fact holds for all optimal transference plans $\pi$ between $\mu_0$ and $\mu_1$. In addition, in this case a transference plan $\pi$ is optimal iff it is supported on a $\text{d}^2$-cyclically monotone set. A set $\Lambda \subset X \times X$ is said to be $c$-cyclically monotone if for any finite set of points $\{(x_i, y_i)\}_{i=1}^N \subset \Lambda$ it holds

$$\sum_{i=1}^N c(x_i, y_i) \leq \sum_{i=1}^N c(x_i, y_{i+1})$$

with the convention that $y_{N+1} = y_1$.

As $(X, \text{d})$ is a complete and separable metric space then so is $(\mathcal{P}_2(X), W_2)$. Under these assumptions, it is known that $(X, \text{d})$ is geodesic if and only if $(\mathcal{P}_2(X), W_2)$ is geodesic. Recall that $e_t$ denotes the (continuous) evaluation map at $t \in [0, 1]$:

$$e_t : \text{Geo}(X) \ni \gamma \mapsto \gamma_t \in X.$$  

A measure $\nu \in \mathcal{P}(\text{Geo}(X))$ is called an optimal dynamical plan if $(e_0, e_1)_\sharp \nu$ is an optimal transference plan; it easily follows in that case that $[0, 1] \ni t \mapsto (e_t)_\sharp \nu$ is a geodesic in $\mathcal{P}_2(X, W_2)$. It is known that any geodesic $(\mu_t)_{t \in [0,1]}$ in $\mathcal{P}_2(X, W_2)$ can be lifted to an optimal dynamical plan $\nu$ so that $(e_t)_\sharp \nu = \mu_t$ for all $t \in [0, 1]$ (see for instance [3, Theorem 2.10]). We denote by $\text{OptGeo}(\mu_0, \mu_1)$ the space of all optimal dynamical plans $\nu$.
so that \((e_t)_# \nu = \mu_i\), \(i = 0, 1\). Consequently, whenever \((X, d)\) is geodesic, the set \(\ OptGeo(\mu_0, \mu_1)\) is non-empty for all \(\mu_0, \mu_1 \in \mathcal{P}_2(X)\), and for any Kantorovich potential \(\varphi\) associated to an optimal transference plan between \(\mu_0\) and \(\mu_1\), we have \(\varphi(G_{\varphi}) = 1\) for all \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\).

In order to consider restrictions of optimal dynamical plans, for any \(s, t \in [0, 1]\) with \(s \leq t\) we consider the restriction map

\[
\text{restr}^t_s : C([0, 1]; X) \ni \gamma \mapsto \gamma \circ f^t_s \in C([0, 1]; X),
\]

where \(f^t_s : [0, 1] \rightarrow [s, t]\) is defined by \(f^t_s(\tau) = s + (t - s)\tau\). During this work we will use the following facts: if \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\) then the restriction \((\text{restr}^t_s)_# \nu\) is still an optimal dynamical plan, now between \(\mu_s\) and \(\mu_t\) where \(\mu_r := (e_r)_# \nu\). Moreover, any probability measure \(\nu' \in \mathcal{P}(\text{Geo}(X))\) with \(\text{supp}(\nu') \subset \text{supp}(\nu) \subset G_{\varphi}\) is also an optimal dynamical plan, between \((e_0)_# \nu'\) and \((e_1)_# \nu'\).

On several occasions we will use the following standard lemma (whose proof is a straightforward adaptation of e.g. [29, Lemma 4.4], relying on the Arzelà–Ascoli and Prokhorov theorems):

**Lemma 6.1** Assume that \((X, d)\) is a Polish and proper space. Let \(\{\mu^0_i\}, \{\mu^1_i\} \subset \mathcal{P}_2(X)\) denote two sequences of probability measures weakly converging to \(\mu^\infty_0, \mu^\infty_1 \in \mathcal{P}_2(X)\), respectively. Assume that \(\nu^i \in \text{OptGeo}(\mu^0_i, \mu^1_i)\). Then there exists a subsequence \(\{\nu^{ij}\}\) weakly converging to \(\nu^\infty \in \text{OptGeo}(\mu^\infty_0, \mu^\infty_1)\).

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

\[
\exists t \in (0, 1) \quad \gamma^1_s = \gamma^2_s \quad \forall s \in [0, t] \implies \gamma^1_s = \gamma^2_s \quad \forall s \in [0, 1].
\]

\((X, d)\) is called non-branching if \(\text{Geo}(X)\) is non-branching. \((X, d, m)\) is called essentially non-branching [68] if for all \(\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)\), any \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\) is concentrated on a Borel non-branching set \(G \subset \text{Geo}(X)\).

Recall that a measure \(\nu\) on a measurable space \((\Omega, \mathcal{F})\) is said to be concentrated on \(A \subset \Omega\) if \(\exists B \subset A\) with \(B \in \mathcal{F}\) so that \(\nu(\Omega \setminus B) = 0\).

### 6.2 Curvature-Dimension conditions

We now turn to describe various synthetic conditions encapsulating generalized Ricci curvature lower bounds coupled with generalized dimension upper bounds.
**Definition 6.2** \((\sigma_{K,N}\text{-coefficients})\) Given \(K \in \mathbb{R}\) and \(N \in (0, \infty]\), define

\[
D_{K,N} := \begin{cases} 
\frac{\pi}{\sqrt{K/N}} & K > 0, \ N < \infty \\
\infty & \text{otherwise}
\end{cases}
\]

In addition, given \(t \in [0, 1]\) and \(0 < \theta < D_{K,N}\), define

\[
\sigma_{K,N}^{(t)}(\theta) := \begin{cases} 
\frac{\sin(t\theta \sqrt{K/N})}{\sin(\theta \sqrt{K/N})} & K > 0, \ N < \infty \\
\frac{\sinh(t\theta \sqrt{-K/N})}{\sinh(\theta \sqrt{-K/N})} & K < 0, \ N < \infty
\end{cases}
\]

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

**Definition 6.3** \((\tau_{K,N}\text{-coefficients})\) Given \(K \in \mathbb{R}\) and \(N \in (1, \infty]\), 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 \(K \leq 0\) and \(\tau_{K,1}^{(t)}(\theta) = +\infty\) if \(K > 0\).

The synthetic Curvature-Dimension condition \(\text{CD}(K, N)\) has been defined on a general m.m.s. independently in several seminal works by Sturm and Lott–Villani: the case \(N = \infty\) and \(K \in \mathbb{R}\) was defined in [73] and [51], the case \(N \in [1, \infty)\) in [74] for \(K \in \mathbb{R}\) and in [51] for \(K = 0\) (and subsequently for \(K \in \mathbb{R}\) in [50]). Our treatment in this work excludes the case \(N = \infty\) (for which the globalization result we are after is in any case known [73]). To exclude possible pathological behavior when \(N = 1\), we will always assume, unless otherwise stated, that \(K \in \mathbb{R}\) and \(N \in (1, \infty)\).

We will use the following definition introduced in [74]. Recall that given \(N \in (1, \infty)\), the \(N\)-Rényi relative-entropy functional \(E_N : \mathcal{P}(X) \to [0, 1]\) (since \(m(X) = 1\)) is defined as

\[
E_N(\mu) := \int \rho^{1-\frac{1}{N}} dm,
\]

where \(\mu = \rho m + \mu^{\text{sing}}\) is the Lebesgue decomposition of \(\mu\) with \(\mu^{\text{sing}} \perp m\). It is known [74] that \(E_N\) is upper semi-continuous with respect to the weak topology on \(\mathcal{P}(X)\).
**Definition 6.4 (CD(\(K, N\)))** A m.m.s. \((X, d, m)\) is said to satisfy \(\text{CD}(K, N)\) if for all \(\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)\), there exists \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\) so that for all \(t \in [0, 1]\), \(\mu_t := (e_t)_# \nu \ll m\), and for all \(N' \geq N\):

\[
\mathcal{E}_{N'}(\mu_t) \geq \int_{X \times X} \left( \tau_{K, N'}^{(1-t)}(d(x_0, x_1))\rho_0^{-1/N'(x_0)} + \tau_{K, N'}^{t}(d(x_0, x_1))\rho_1^{-1/N'(x_1)} \right) \pi(dx_0, dx_1), \quad (6.3)
\]

for some optimal transference plan \(\pi\) between \(\mu_0 = \rho_0 m\) and \(\mu_1 = \rho_1 m\).

**Remark 6.5** When \(m(X) < \infty\) as in our setting, it is known [74, Proposition 1.6 (ii)] that \(\text{CD}(K, N)\) implies \(\text{CD}(K, \infty)\), and hence the requirement \(\mu_t \ll m\) for all intermediate times \(t \in (0, 1)\) is in fact superfluous, as it must hold automatically by finiteness of the Shannon entropy (see [73, 74]).

The following is a local version of \(\text{CD}(K, N)\):

**Definition 6.6 (CD\(_{loc}(K, N)\))** A m.m.s. \((X, d, m)\) is said to satisfy \(\text{CD}_{loc}(K, N)\) if for any \(o \in \text{supp}(m)\), there exists a neighborhood \(X_0 \subset X\) of \(o\), so that for all \(\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)\) supported in \(X_0\), there exists \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\) so that for all \(t \in [0, 1]\), \(\mu_t := (e_t)_# \nu \ll m\), and for all \(N' \geq N\), (6.3) holds.

Note that \((e_t)_# \nu\) is not required to be supported in \(X_0\) for intermediate times \(t \in (0, 1)\) in the latter definition.

The following pointwise density inequality is a known equivalent definition of \(\text{CD}(K, N)\) on essentially non-branching spaces (the equivalence follows by combining the results of [29] and [41], see the proof of Proposition 9.1):

**Definition 6.7 (CD\((K, N)\) for essentially non-branching spaces)** An essentially non-branching m.m.s. \((X, d, m)\) satisfies \(\text{CD}(K, N)\) if and only if for all \(\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)\), there exists a unique \(\nu \in \text{OptGeo}(\mu_0, \mu_1)\), \(\nu\) is induced by a map (i.e. \(\nu = S_{\pi}(\mu_0)\) for some map \(S : X \to \text{Geo}(X)\), \(\mu_t := (e_t)_# \nu \ll m\) for all \(t \in [0, 1]\), and writing \(\mu_t = \rho_t m\), we have for all \(t \in [0, 1]\):

\[
\rho_t^{-1/N}(\gamma_t) \geq \tau_{K, N}^{(1-t)}(d(\gamma_0, \gamma_1))\rho_0^{-1/N}(\gamma_0) + \tau_{K, N}^{t}(d(\gamma_0, \gamma_1))\rho_1^{-1/N}(\gamma_1) \quad \text{for } \nu\text{-a.e. } \gamma \in \text{Geo}(X).
\]

The Measure Contraction Property \(\text{MCP}(K, N)\) was introduced independently by Ohta in [57] and Sturm in [74]. The idea is to only require the \(\text{CD}(K, N)\) condition to hold when \(\mu_1\) degenerates to \(\delta_o\), a delta-measure at \(o \in \text{supp}(m)\). However, there are several possible implementations of this idea. We start with the following one, which is a variation of the one used in [29]:
Lemma 6.12
Let
\( \text{if for any } o \in \text{supp}(m) \text{ and } \mu_0 \in \mathcal{P}_2(X, d, m) \text{ with bounded support, there exists } \nu \in \text{OptGeo}(\mu_0, \delta_o), \text{ such that for all } t \in [0, 1], \text{ if } \mu_t := (e_t)_#\nu \text{ then } \text{supp}(\mu_t) \subset \text{supp}(m), \text{ and:} \)
\[
\mathcal{E}_N(\mu_t) \geq \int_X \tau_{K,N}^{(1-t)}(d(x_0, o)) \rho_0^{-\frac{1}{2}}(x_0) m(dx_0), \tag{6.4}
\]
where \( \mu_0 = \rho_0 m. \)

The variant proposed in [57] is as follows:

Definition 6.9 (MCP(\(K, N\))) A m.m.s. \( (X, d, m) \) is said to satisfy MCP(\(K, N\)) if for any \( o \in \text{supp}(m) \) and \( \mu_0 \in \mathcal{P}_2(X, d, m) \) of the form \( \mu_0 = \frac{1}{m(A)} m_{|A} \) for some Borel set \( A \subset X \) with \( 0 < m(A) < \infty \), there exists \( \nu \in \text{OptGeo}(\mu_0, \delta_o) \) such that:
\[
\frac{1}{m(A)} m \geq (e_t)_# \left( \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1)) \right)^N \nu(d\gamma) \quad \forall t \in [0, 1]. \tag{6.5}
\]

Remark 6.10 Note that in [57] it was assumed in addition that \( \text{supp}(m) = X \) and that \((X, d)\) is a length-space, but (6.5) was only required to hold for \( A \subset B(o, D_{K,N-1}) \) if \( K > 0 \); both our version and the one from [57] imply that the diameter of \( \text{supp}(m) \) is bounded above by \( D_{K,N-1} \) (this follows in our version since \( \tau_{K,N}(\theta) = +\infty \) if \( \theta \geq D_{K,N-1} \), and by [57, Theorem 4.3] in the version from [57]), and also that \( \text{supp}(m) \) is a geodesic-space (see Lemma 6.12 below), and therefore both versions are ultimately equivalent.

When either the MCP(\(K, N\)) or MCP(\(\epsilon(K, N)\)) conditions hold for a given \( o \in \text{supp}(m) \), we will say that the space satisfies the corresponding condition with respect to \( o \).

Remark 6.11 The CD(\(K, N\), CD(\(\text{loc}(K, N\)), MCP(\(\epsilon(K, N)\)) and MCP(\(K, N\)) conditions all ensure that for all \( t \in [0, 1] \), \( \text{supp}((e_t)_#\nu) \subset \text{supp}(m) \) for the appropriate \( \nu \in \text{OptGeo}(\mu_0, \mu_1) \) appearing in the corresponding definition. Consequently, for a fixed dense countable set of times \( t \in (0, 1) \), \( \gamma_t \in \text{supp}(m) \) for \( \nu\text{-a.e. } \gamma \in \text{Geo}(X) \); since \( \text{supp}(m) \) is closed, this in fact holds for all \( t \in [0, 1] \), and hence \( \gamma \in \text{Geo} \!\!\!(\text{supp}(m)) \) for \( \nu\text{-a.e. } \gamma \in \text{Geo}(X) \), i.e. \( \text{supp}(\nu) \subset \text{Geo} \!\!\!(\text{supp}(m)) \). It follows that \((X, d, m)\) satisfies CD(\(K, N\), CD(\(\text{loc}(K, N\), MCP(\(\epsilon(K, N)\)) or MCP(\(K, N\)) iff \( \text{supp}(m), d, m \) does.

The following simple lemma will be useful for quickly establishing that \( \text{supp}(m), d \) is proper and geodesic:

Lemma 6.12 Let \((X, d, m)\) be a m.m.s. verifying CD(\(K, N\), MCP(\(\epsilon(K, N)\) or MCP(\(K, N\)). Then \( \text{supp}(m), d \) is a Polish, proper and geodesic space.
The same holds for $\text{CD}_{\text{loc}}(K, N)$ if $(\text{supp}(m), d)$ is assumed to be a length space.

Proof As $\text{supp}(m) \subset X$ is closed, $(\text{supp}(m), d)$ is Polish. It was shown in [57, Lemma 2.5, Theorem 5.1] for $\text{MCP}(K, N)$ (and hence $\text{MCP}_\varepsilon(K, N)$) and in [74, Corollary 2.4] for $\text{CD}(K, N)$ that these conditions imply a doubling condition, so that every closed bounded ball in $(\text{supp}(m), d)$ is totally bounded. Together with completeness, this already implies that the latter space is proper. By Remark 6.11, $(\text{supp}(m), d, m)$ verifies the same corresponding condition as $(X, d, m)$. In particular, if $(X, d, m)$ and hence $(\text{supp}(m), d, m)$ verifies $\text{CD}(K, N), \text{MCP}_\varepsilon(K, N)$ or $\text{MCP}(K, N)$, then for any $x, y \in \text{supp}(m)$, there is at least one geodesic in $\text{supp}(m)$ from $B(y, \varepsilon) \cap \text{supp}(m)$ to $x$; together with properness and completeness, this already implies that $(\text{supp}(m), d)$ is geodesic. On the other hand, if $(X, d, m)$ and hence $(\text{supp}(m), d, m)$ verifies $\text{CD}_{\text{loc}}(K, N)$, the above argument shows that $(\text{supp}(m), d)$ is complete and locally compact. Together with the assumption that the latter space is a length-space, the Hopf-Rinow theorem implies that it is proper and geodesic. \qed

Lemma 6.13 The following chain of implications is known:

$$\text{CD}(K, N) \Rightarrow \text{MCP}_\varepsilon(K, N) \Rightarrow \text{MCP}(K, N).$$

Proof By Remark 6.11, we may reduce to the case $\text{supp}(m) = X$. Fixing $\mu_0 \ll m$ with bounded support and $o \in X$, let $\nu^\varepsilon$ be an element of $\text{OptGeo}(\mu_0, \mu^\varepsilon_1)$ satisfying the $\text{CD}(K, N)$ condition for $\mu^\varepsilon_1 = m(B(o, \varepsilon))^{-1}m_{B(o, \varepsilon)}$. By Lemma 6.1 (which applies since the space is proper by Lemma 6.12), $\{\nu^\varepsilon\}$ has a converging subsequence to $\nu^0 \in \text{OptGeo}(\mu_0, \delta_0)$ as $\varepsilon \to 0$. The upper semi-continuity of $E_N$ and the continuity of the evaluation map $e_\varepsilon$ ensure that $\nu^0$ satisfies the $\text{MCP}_\varepsilon(K, N)$ condition (6.4). The second implication follows by the arguments of [66, Section 5] (without any types of essential non-branching assumptions). \qed

Remark 6.14 We will show in Proposition 9.1 that for essentially non-branching spaces, $\text{MCP}(K, N)$ implies back $\text{MCP}_\varepsilon(K, N)$. We remark that for non-branching spaces, the implication $\text{CD}(K, N) \Rightarrow \text{MCP}(K, N)$ was first proved in [74].

Many additional useful results on the structure of $W_2$-geodesics can be obtained just from the $\text{MCP}$ condition. The following has been shown in [29, Theorem 1.1 and Appendix] (when $\text{supp}(m) = X$; the formulation below is immediately obtained from Remark 6.11):

Theorem 6.15 ([29]) Let $(X, d, m)$ be an essentially non-branching m.m.s. satisfying $\text{MCP}(K, N)$. Given any pair $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ with $\mu_0 \ll m$ and $\text{supp}(\mu_1) \subset \text{supp}(m)$, the following holds:
– there exists a unique \( \nu \in \text{OptGeo}(\mu_0, \mu_1) \) and hence a unique optimal transference plan between \( \mu_0 \) and \( \mu_1 \);
– there exists a map \( S : X \supset \text{Dom}(S) \to \text{Geo}(X) \) such that \( \nu = S \sharp \mu_0 \);
– for any \( t \in [0, 1) \) the measure \( (e_t)_\sharp \nu \) is absolutely continuous with respect to \( m \).

The following is a standard corollary of the fact that the optimal dynamical plan is induced by a map (see e.g. the comments after [41, Theorem 1.1]); as we could not find a reference, we sketch the proof for completeness.

**Corollary 6.16** With the same assumptions as in Theorem 6.15, the unique optimal transference plan \( \nu \) is concentrated on a (Borel) set \( G \subset \text{Geo}(X) \), so that for all \( t \in [0, 1) \), the evaluation map \( e_t|_G : G \to X \) is injective. In particular, for any Borel subset \( H \subset G \):

\[
(e_t)_\sharp (\nu|_H) = (e_t)_\sharp (\nu) \circ (e_t)(H) \quad \forall t \in [0, 1).
\]

**Sketch of proof** First, we claim the existence of \( X_1 \subset X \) with \( \mu_0(X_1) = 1 \), so that for all \( x \in X_1 \), there exists a unique \( \gamma \in G_\varphi \) with \( \gamma_0 = x \). Otherwise, if \( A \subset X \) is a set of positive \( \mu_0 \)-measure where this is violated, there are at least two distinct geodesics in \( G_\varphi \) emanating from every \( x \in A \). As these geodesics must be different at some rational time in \((0, 1)\), it follows that there exists a rational \( \tilde{t} \in (0, 1) \) and \( B \subset A \) still of positive \( \mu_0 \)-measure so that both pairs of geodesics emanating from \( x \) are different at time \( \tilde{t} \) for all \( x \in B \). Consider \( \tilde{\mu}_0 = \mu_0|_B / \mu_0(B) \ll m \), and transport to time \( \tilde{t} \) half of its mass along one geodesic and the second half along the other one (see e.g. the proof of [29, Theorem 5.1]). The latter transference plan is optimal but is not induced by a map, yielding a contradiction.

Now denote \( G := S(X_1) \) (and hence \( \nu(G) = 1 \)), so that the injectivity of \( e_0|_G \) is already guaranteed. To see the injectivity of \( e_t|_G \) for all \( t \in (0, 1) \), suppose in the contrapositive the existence of \( \gamma^1, \gamma^2 \in G \) with \( \gamma^1_t = \gamma^2_t \). Denoting by \( \eta \) the gluing of \( \gamma^1 \) restricted to \([0, t]\) with \( \gamma^2 \) restricted to \([t, 1]\), it follows by \( d^2 \)-cyclic monotonicity (see e.g. the proof of [14, Lemma 2.6] or that of Lemma 3.7) that \( \eta \in G_\varphi \) with \( \eta_0 = \gamma^1_0 \) and \( \eta \neq \gamma^1 \). But this is in contradiction to the definition of \( X_1 \), thereby concluding the proof. \( \square \)

### 6.3 Disintegration theorem

We include here a version of the Disintegration Theorem that we will use. We will follow [18, Appendix A] where a self-contained approach (and a proof) of the Disintegration Theorem in countably generated measure spaces can be found. An even more general version of the Disintegration Theorem can be found in [39, Section 452].
Recall that given a measure space \((X, \mathcal{X}, m)\), a set \(A \subset X\) is called \(m\)-measurable if \(A\) belongs to the completion of the \(\sigma\)-algebra \(\mathcal{X}'\), generated by adding to it all subsets of null \(m\)-sets; similarly, a function \(f : (X, \mathcal{X}, m) \to \mathbb{R}\) is called \(m\)-measurable if all of its sub-level sets are \(m\)-measurable.

**Definition 6.17 (Disintegration on sets)** Let \((X, \mathcal{X}, m)\) denote a measure space. Given any family \(\{X_\alpha\}_{\alpha \in Q}\) of subsets of \(X\), a **disintegration of \(m\) on \(\{X_\alpha\}_{\alpha \in Q}\)** is a measure-space structure \((Q, \mathcal{Q}, q)\) and a map

\[ Q \ni \alpha \mapsto m_\alpha \in \mathcal{P}(X, \mathcal{X}) \]

so that

1. for \(q\)-a.e. \(\alpha \in Q\), \(m_\alpha\) is concentrated on \(X_\alpha\);
2. for all \(B \in \mathcal{X}'\), the map \(\alpha \mapsto m_\alpha(B)\) is \(q\)-measurable;
3. for all \(B \in \mathcal{X}'\), \(m(B) = \int m_\alpha(B) \, q(d\alpha)\).

The measures \(m_\alpha\) are referred to as **conditional probabilities**.

Given a measurable space \((X, \mathcal{X}')\) and a function \(\Omega : X \to Q\), with \(Q\) a general set, we endow \(Q\) with the **push forward \(\sigma\)-algebra** \(\mathcal{Q}\) of \(\mathcal{X}'\):

\[ C \in \mathcal{Q} \iff \Omega^{-1}(C) \in \mathcal{X}' \]

i.e. the biggest \(\sigma\)-algebra on \(Q\) such that \(\Omega\) is measurable. Moreover, given a measure \(m\) on \((X, \mathcal{X}')\), define a measure \(q\) on \((Q, \mathcal{Q})\) by pushing forward \(m\) via \(\Omega\), i.e. \(q := \Omega_# m\).

**Definition 6.18 (Consistent and Strongly Consistent Disintegration)** A disintegration of \(m\) **consistent with** \(\Omega : X \to Q\) is a map:

\[ Q \ni \alpha \mapsto m_\alpha \in \mathcal{P}(X, \mathcal{X}) \]

such that the following requirements hold:

1. for all \(B \in \mathcal{X}'\), the map \(\alpha \mapsto m_\alpha(B)\) is \(q\)-measurable;
2. for all \(B \in \mathcal{X}'\) and \(C \in \mathcal{Q}\), the following consistency condition holds:

\[ m(B \cap \Omega^{-1}(C)) = \int_C m_\alpha(B) \, q(d\alpha). \]

A disintegration of \(m\) is called **strongly consistent with respect to** \(\Omega\) if in addition:

3. for \(q\)-a.e. \(\alpha \in Q\), \(m_\alpha\) is concentrated on \(\Omega^{-1}(\alpha)\);
The above general scheme fits with the following situation: given a measure space \((X, \mathcal{X}, m)\), suppose a partition of \(X\) is given into disjoint sets \(\{X_\alpha\}_{\alpha \in Q}\) so that \(X = \bigcup_{\alpha \in Q} X_\alpha\). Here \(Q\) is the set of indices and \(\Omega : X \to Q\) is the quotient map, i.e.

\[
\alpha = \Omega(x) \iff x \in X_\alpha.
\]

We endow \(Q\) with the quotient \(\sigma\)-algebra \(\mathcal{Q}\) and the quotient measure \(q\) as described above, obtaining the quotient measure space \((Q, \mathcal{Q}, q)\). When a disintegration \(\alpha \mapsto m_\alpha\) of \(m\) is (strongly) consistent with the quotient map \(\Omega\), we will simply say that it is (strongly) consistent with the partition. Note that any disintegration \(\alpha \mapsto m_\alpha\) of \(m\) on a partition \(\{X_\alpha\}_{\alpha \in Q}\) (as in Definition 6.17) is automatically strongly consistent with the partition (as in Definition 6.18), and vice versa.

We now formulate the Disintegration Theorem (it is formulated for probability measures but clearly holds for any finite non-zero measure):

**Theorem 6.19** (Theorem A.7, Proposition A.9 of [18]) Assume that \((X, \mathcal{X}, m)\) is a countably generated probability space and that \(\{X_\alpha\}_{\alpha \in Q}\) is a partition of \(X\).

Then the quotient probability space \((Q, \mathcal{Q}, q)\) is essentially countably generated and there exists an essentially unique disintegration \(\alpha \mapsto m_\alpha\) consistent with the partition.

If in addition \(\mathcal{X}\) contains all singletons, then the disintegration is strongly consistent if and only if there exists a \(m\)-section \(S_m \in \mathcal{X}\) of the partition such that the \(\sigma\)-algebra on \(S_m\) induced by the quotient-map contains the trace \(\sigma\)-algebra \(\mathcal{X} \cap S_m := \{A \cap S_m; A \in \mathcal{X}\}\).

Let us expand on the statement of Theorem 6.19. Recall that a \(\sigma\)-algebra \(\mathcal{A}\) is *countably generated* if there exists a countable family of sets so that \(\mathcal{A}\) coincides with the smallest \(\sigma\)-algebra containing them. On the measure space \((Q, \mathcal{Q}, q)\), the \(\sigma\)-algebra \(\mathcal{Q}\) is called *essentially countably generated* if there exists a countable family of sets \(Q_n \subset Q\) such that for any \(C \in \mathcal{Q}\) there exists \(\hat{C} \in \mathcal{Q}\), where \(\mathcal{Q}\) is the \(\sigma\)-algebra generated by \(\{Q_n\}_{n \in \mathbb{N}}\), such that \(q(\hat{C} \Delta \hat{C}) = 0\).

Essential uniqueness is understood above in the following sense: if \(\alpha \mapsto m_\alpha^1\) and \(\alpha \mapsto m_\alpha^2\) are two consistent disintegrations with the partition then \(m_\alpha^1 = m_\alpha^2\) for \(q\)-a.e. \(\alpha \in Q\).

Finally, a set \(S \subset X\) is a section for the partition \(X = \bigcup_{\alpha \in Q} X_\alpha\) if for any \(\alpha \in Q\), \(S \cap X_\alpha\) is a singleton \(\{x_\alpha\}\). By the axiom of choice, a section \(S\) always exists, and we may identify \(Q\) with \(S\) via the map \(Q \ni \alpha \mapsto x_\alpha \in S\). A set \(S_m\) is an \(m\)-section if there exists \(Y \in \mathcal{X}\) with \(m(X \setminus Y) = 0\) such that the partition \(Y = \bigcup_{\alpha \in Q_m} (X_\alpha \cap Y)\) has section \(S_m\), where \(Q_m = \{\alpha \in Q; X_\alpha \cap Y \neq \emptyset\}\).
As \( q = \Omega_m \), clearly \( q(Q \setminus Q_m) = 0 \). As usual, we identify between \( Q_m \) and \( S_m \), so that now \( Q_m \) carries two measurable structures: \( \bar{\mathbf{X}} \cap Q_m \) (the push-forward of \( \mathbf{X} \cap Y \) via \( \Omega \)), and also \( \mathbf{X} \cap S_m \) via our identification. The last condition of Theorem 6.19 is that \( \bar{\mathbf{X}} \cap Q_m \supset \mathbf{X} \cap S_m \), i.e. that the restricted quotient-map \( \Omega|_Y : (Y, \mathbf{X} \cap Y) \to (S_m, \mathbf{X} \cap S_m) \) is measurable, so that the full quotient-map \( \Omega : (X, \mathbf{X}) \to (S, \mathbf{X} \cap S) \) is \( m \)-measurable.

We will typically apply the Disintegration Theorem to \( (E, \mathcal{B}(E), m_{\upharpoonright E}) \), where \( E \subset X \) is an \( m \)-measurable subset (with \( m(E) > 0 \)) of the m.m.s. \( (X, d, m) \). As our metric space is separable, \( \mathcal{B}(E) \) is countably generated, and so Theorem 6.19 applies. In particular, when \( Q \subset \mathbb{R} \), \( E \) is a closed subset of \( X \), the partition elements \( X_\alpha \) are closed and the quotient-map \( \Omega : E \to Q \) is known to be Borel (for instance, this is the case when \( \Omega \) is continuous), [72, Theorem 5.4.3] guarantees the existence of a Borel section \( S \) for the partition so that \( \Omega : E \to S \) is Borel measurable, thereby guaranteeing by Theorem 6.19 the existence of an essentially unique disintegration strongly consistent with \( \Omega \).

7 Theory of \( L^1 \)-Optimal-Transport

In this section we recall various results from the theory of \( L^1 \)-Optimal-Transport which are relevant to this work, and add some new information we will subsequently require. We refer to [2,13,19,23,37,38,47,76] for more details.

7.1 Preliminaries

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

\[
\Gamma_u := \{(x, y) \in X \times X : u(x) - u(y) = d(x, y)\}.
\] (7.1)

Its transpose is given by \( \Gamma_u^{-1} := \{(x, y) \in X \times X : (y, x) \in \Gamma_u\} \). We define the transport relation \( R_u \) and the transport set \( T_u \), as

\[
R_u := \Gamma_u \cup \Gamma_u^{-1}, \quad T_u := P_1(R_u \setminus \{x = y\}),
\] (7.2)

where \( \{x = y\} \) denotes the diagonal \( \{(x, y) \in X^2 : x = y\} \) and \( P_i \) the projection onto the \( i \)-th component. Recall that \( \Gamma_u(x) = \{y \in X \setminus \{x, y \in \Gamma_u\} \}
\] denotes the section of \( \Gamma_u \) through \( x \) in the first coordinate, and similarly for \( R_u(x) \) (through either coordinates by symmetry). Since \( u \) is 1-Lipschitz, \( \Gamma_u, \Gamma_u^{-1} \) and \( R_u \) are closed sets, and so are \( \Gamma_u(x) \) and \( R_u(x) \). Consequently \( T_u \)
is a projection of a Borel set and hence analytic; it follows that it is universally measurable, and in particular, \( m \)-measurable \([72]\).

The following is immediate to verify (see \([2, \text{Proposition 4.2}]\)):

**Lemma 7.1** Let \((\gamma_0, \gamma_1) \in \Gamma_u\) for some \(\gamma \in \text{Geo}(X)\). Then \((\gamma_s, \gamma_t) \in \Gamma_u\) for all \(0 \leq s \leq t \leq 1\).

Also recall the following definitions, introduced in \([23]\):

\[
A_+ := \{ x \in T_u : \exists z, w \in \Gamma_u(x), (z, w) \notin R_u \}, \\
A_- := \{ x \in T_u : \exists z, w \in \Gamma_u^{-1}(x), (z, w) \notin R_u \}.
\]

\(A_{\pm}\) are called the sets of forward and backward branching points, respectively. Note that both \(A_{\pm}\) are analytic sets; for instance:

\[
A_+ = P_1(\{(x, z, w) \in T_u \times X \times X : (x, z), (x, w) \in \Gamma_u, (z, w) \notin R_u \}),
\]

showing that \(A_+\) is a projection of an analytic set and therefore analytic. If \(x \in A_+\) and \((y, x) \in \Gamma_u\) necessarily also \(y \in A_+\) (as \(\Gamma_u(y) \supset \Gamma_u(x)\) by the triangle inequality); similarly, if \(x \in A_-\) and \((x, y) \in \Gamma_u\) then necessarily \(y \in A_-\).

Consider the **non-branched transport set**

\[
T^b_u := T_u \setminus (A_+ \cup A_-),
\]

which belongs to the sigma-algebra \(\sigma(A)\) generated by analytic sets and is therefore \(m\)-measurable. Define the **non-branched transport relation**:

\[
R^b_u := R_u \cap (T^b_u \times T^b_u).
\]

It was shown in \([23]\) (cf. \([19]\)) that \(R^b_u\) is an equivalence relation over \(T^b_u\) and that for any \(x \in T^b_u\), \(R_u(x) \subset (X, d)\) is isometric to a closed interval in \((\mathbb{R}, |\cdot|)\).

**Remark 7.2** Note that even if \(x \in T^b_u\), the transport ray \(R_u(x)\) need not be entirely contained in \(T^b_u\). However, we will soon prove that almost every transport ray (with respect to an appropriate measure) has interior part contained in \(T^b_u\).

It will be very useful to note that whenever the space \((X, d)\) is proper (for instance when \((X, d, m)\) verifies \(\text{MCP}(K, N)\) and \(\text{supp}(m) = X\)), \(T_u\) and \(A_{\pm}\) are \(\sigma\)-compact sets: indeed writing \(R_u \setminus \{x = y\} = \bigcup_{\varepsilon > 0} R_u \setminus \{d(x, y) > \varepsilon\}\) it follows that \(R_u \setminus \{x = y\}\) is \(\sigma\)-compact. Hence \(T_u\) is \(\sigma\)-compact. Moreover:

\[
A_+ = P_1(\{(x, z, w) \in T_u \times (R_u)^c : (x, z), (x, w) \in \Gamma_u\});
\]
since \((R_u)^c\) is open and open sets are \(F_\sigma\) in metric spaces, it follows that 
\[\{ (x, z, w) \in T_u \times (R_u)^c : (x, z), (x, w) \in \Gamma_u \}\] is \(\sigma\)-compact and therefore \(A_+\) is \(\sigma\)-compact; the same applies to \(A_-\). Consequently, \(T_u^b\) and \(R_u^b\) are Borel.

Now, from the first part of the Disintegration Theorem 6.19 applied to \((T_u^b, B(T_u^b), m_{\sim T_u^b})\), we obtain an essentially unique disintegration of \(m_{\sim T_u^b}\) consistent with the partition of \(T_u^b\) given by the equivalence classes \(\{ R_u^b(\alpha) \}_{\alpha \in Q}\) of \(R_u^b\):

\[m_{\sim T_u^b} = \int_Q m_\alpha \, q(d\alpha)\]

with corresponding quotient space \((Q, \emptyset, q)\) (\(Q \subset T_u^b\) may be chosen to be any section of the above partition). The next step is to show that the disintegration is strongly consistent. By the Disintegration Theorem, this is equivalent to the existence of a \(m_{\sim T_u^b}\)-section \(\widetilde{Q} \in B(T_u^b)\) (which by a mild abuse of notation we will call \(m\)-section), such that the quotient map associated to the partition is \(m\)-measurable, where we endow \(\widetilde{Q}\) with the trace \(\sigma\)-algebra. This has already been shown in [19, Proposition 4.4] in the framework of non-branching metric spaces; since its proof does not use any non-branching assumption, we can conclude that

\[m_{\sim T_u^b} = \int_Q m_\alpha \, q(d\alpha), \quad \text{and for } q\text{-a.e. } \alpha \in Q, \quad m_\alpha(R_u^b(\alpha)) = 1,\]

where now \(Q \supset \widetilde{Q} \in B(T_u^b)\) with \(\widetilde{Q}\) an \(m\)-section for the above partition (and hence \(q\) is concentrated on \(\widetilde{Q}\)). For a more constructive approach under the additional assumption of properness of the space, see also [25, Proposition 4.8].

A-priori the non-branched transport set \(T_u^b\) can be much smaller than \(T_u\). However, under fairly general assumptions one can prove that the sets \(A_\pm\) of forward and backward branching are both \(m\)-negligible. In [23] this was shown for a m.m.s. \((X, d, m)\) verifying \(\text{RCD}(K, N)\) and \(\text{supp}(m) = X\). The proof only relies on the following two properties which hold for the latter spaces (see also [25]):

- \(\text{supp}(m) = X\).
- Given \(\mu_0, \mu_1 \in \mathcal{P}_2(X)\) with \(\mu_0 \ll m\), there exists a unique optimal transport plan for the \(W_2\)-distance and it is induced by an optimal-transport map.

By Theorem 6.15 these properties are also verified for an essentially non-branching m.m.s. \((X, d, m)\) satisfying \(\text{MCP}(K, N)\) and \(\text{supp}(m) = X\). We summarize the above discussion in:
Corollary 7.3 Let \((X, d, m)\) be an essentially non-branching \(m.m.s.\) satisfying \(\text{MCP}(K, N)\) and \(\text{supp}(X) = m\). Then for any 1-Lipschitz function \(u : X \rightarrow \mathbb{R}\), we have \(m(T_u \setminus T_u^b) = 0\). In particular, we obtain the following essentially unique disintegration \((Q, \mathcal{Q}, q)\) of \(m|_{T_u^b} = m|_{T_u^b}\) strongly consistent with the partition of \(T_u^b\) given by the equivalence classes \(\{R_u^b(\alpha)\}_{\alpha \in Q}\) of \(R_u^b\):

\[
m|_{T_u^b} = \int_Q m_\alpha q(d\alpha), \quad \text{and for } q\text{-a.e. } \alpha \in Q, \quad m_\alpha(R_u^b(\alpha)) = 1. \tag{7.3}
\]

Here \(Q\) may be chosen to be a section of the above partition so that \(Q \supset \bar{Q} \in \mathcal{B}(T_u^b)\) with \(\bar{Q}\) an \(m\)-section with \(m\)-measurable quotient map. In particular, \(\mathcal{Q} \supset \mathcal{B}(\bar{Q})\) and \(q\) is concentrated on \(\bar{Q}\).

Remark 7.4 By modifying the definitions of \(A_+, A_-\) to only reflect branching inside \(\text{supp}(m)\), it is possible to remove the assumption that \(\text{supp}(X) = m\), but we refrain from this extraneous generality here.

Remark 7.5 If we consider \(u = d(\cdot, o)\), it is easy to check that the set \(A_+\) coincides with the cut locus \(C_o\), i.e. the set of those \(z \in X\) such that there exists at least two distinct geodesics starting at \(z\) and ending in \(o\). Hence the previous corollary implies that for any \(o \in X\), the cut locus has \(m\)-measure zero: \(m(C_o) = 0\). This in particular implies that an essentially non-branching \(m.m.s.\) verifying \(\text{MCP}(K, N)\) and \(\text{supp}(m) = X\) also supports a local \((1, 1)\)-weak Poincaré inequality, see [69].

7.2 Maximality of transport rays on non-branching transport-set

It is elementary to check that \(\Gamma_u\) induces a partial order relation on \(X\):

\[y \leq_u x \iff (x, y) \in \Gamma_u.\]

Note that by definition:

\[x \in A_+, \ y \geq_u x \quad \Rightarrow \quad y \in A_+;\]
\[x \in A_-, \ y \leq_u x \quad \Rightarrow \quad y \in A_-;\]

Recall that for any \(x \in T_u^b\), \((R_u(x), d)\) is isometric to a closed interval in \((\mathbb{R}, |\cdot|)\). This isometry induces a total ordering on \(R_u(x)\) which must coincide with either \(\leq_u\) or \(\geq_u\), implying that \((R_u(x), \leq_u, \geq_u)\) is totally ordered.

Lemma 7.6 For any \(x \in T_u^b\), \((R_u^b(x) = R_u(x) \cap T_u^b, d)\) is isometric to an interval in \((\mathbb{R}, |\cdot|)\).
Consider \( z, w \in R_u(x) \cap T^b_u \); as \((R_u(x), \leq_u)\) is totally ordered, assume without loss of generality that \( z \leq u w \). Given \( y \in R_u(x) \) with \( z \leq u y \leq u w \), we must prove that \( y \in T^b_u \). Indeed, since \( w \geq u y \) and \( w / \notin A_+ \), necessarily \( y / \notin A_+ \), and since \( z \leq u y \) and \( z / \notin A_- \), necessarily \( y / \notin A_- \). Hence \( y \in T^b_u \) and the claim follows. \( \square \)

Recall that given a partially ordered set, a chain is a totally ordered subset. A chain is called maximal if it is maximal with respect to inclusion. We introduce the following:

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

In other words, a transport ray \( R \) is the image of a closed non-null geodesic \( \gamma \) parametrized by arclength on \( I \) so that the function \( u \circ \gamma \) is affine with slope 1 on \( I \), and so that \( R \) is maximal with respect to inclusion.

**Lemma 7.8** Given \( x \in T^b_u \), \( R \) is a transport ray passing through \( x \) if and only if \( R = R_u(x) \).

**Proof** Recall that for any \( x \in T^b_u \), \((R_u(x), d, \leq_u)\) is order isometric to a closed interval in \((\mathbb{R}, |\cdot|)\). As \( R_u(x) \) is by definition maximal in \( X \) with respect to inclusion, it follows that it must be a transport ray. Conversely, note that for any transport ray \( R \) we always have \( R \subset \cap_{w \in R} R_u(w) \). Indeed, for any \( w, z \in R \), we have \( z \leq u w \) or \( z \geq u w \), and hence by definition \((w, z) \in R_u \) so that \( z \in R_u(w) \). If \( x \in R \cap T^b_u \), we already showed above that \( R_u(x) \) is a transport ray. Since \( R \subset R_u(x) \) and \( R \) is assumed to be maximal with respect to inclusion, it follows that necessarily \( R = R_u(x) \). \( \square \)

**Corollary 7.9** If \( R_1 \) and \( R_2 \) are two transport rays which intersect in \( T^b_u \) then they must coincide.

In this subsection, we reconcile between the crucial maximality property of \( R_u(\alpha) \) which we will require for the definition of \( CD^1 \) in the next section, and the fact that the disintegration in (7.3) is with respect to (the possibly non-maximal) \( R^b_u(\alpha) = R_u(\alpha) \cap T^b_u \). We will show that under \textsc{MCP}, for \( q \)-a.e. \( \alpha \), the only parts of \( R_u(\alpha) \) which are possibly not contained in \( T^b_u \) are its end points—this fact is the main new result of this section.

To rigorously state this new observation, we recall the classical definition of initial and final points, \( a \) and \( b \), respectively:

\[
a := \{ x \in T_u : \not\exists y \in T_u, (y, x) \in \Gamma_u, y \neq x \},
\]

\[
b := \{ x \in T_u : \not\exists y \in T_u, (x, y) \in \Gamma_u, y \neq x \}.
\]
Note that
\[ a = T_u \setminus P_1(\{\Gamma_u \setminus \{x = y\}\}), \]
so \(a\) is the difference of analytic sets and consequently belongs to \(\sigma(A)\); similarly for \(b\). As in the previous subsection, whenever \((X, d)\) is proper, \(a, b\) are in fact Borel sets.

**Theorem 7.10** (Maximality of transport rays on non-branched transport-set) Let \((X, d, m)\) be an essentially non-branching m.m.s. verifying MCP\((K, N)\) and \(\text{supp}(m) = X\). Let \(u : (X, d) \to \mathbb{R}\) be any 1-Lipschitz function, with (7.3) the associated disintegration of \(m_{\underline{u}}\).

Then there exists \(\hat{Q} \subset Q\) such that \(q(Q \setminus \hat{Q}) = 0\) and for any \(\alpha \in \hat{Q}\) it holds:

\[ R_u(\alpha) \setminus T^b_u \subset a \cup b. \]

**Proof** Step 1. Consider the \(m\)-section \(\bar{Q}\) from Corollary 7.3 so that \(Q \supset \bar{Q} \in B(T^b_u), \mathcal{Q} \supset B(\bar{Q})\) and \(q(Q \setminus \bar{Q}) = 0\). Consider the set

\[ Q_1 := \{\alpha \in \bar{Q} : R_u(\alpha) \setminus T^b_u \not\subset a \cup b\}. \]

The claim will be proved once we show that \(q(Q_1) = 0\). First, observe that

\[ Q_1 = \bar{Q} \cap P_1\left(R_u \cap (T^b_u \times ((A_+ \setminus a) \cup (A_- \setminus b)))\right), \]

and therefore \(Q_1 \subset \bar{Q}\) is analytic; since \(\mathcal{Q} \supset B(\bar{Q})\), it follows that \(Q_1\) is \(q\)-measurable. Now suppose by contradiction that \(q(Q_1) > 0\).

We can divide \(Q_1\) into two sets:

\[ Q_1^+ := \{\alpha \in Q_1 : \Gamma_u(\alpha) \setminus T^b_u \not\subset b\}, \quad Q_1^- := \{\alpha \in Q_1 : \Gamma^{-1}_u(\alpha) \setminus T^b_u \not\subset a\}. \]

Since \(Q_1 = Q_1^+ \cup Q_1^-\), without any loss in generality let us assume \(q(Q_1^+) > 0\), and for ease of notation assume further that \(Q_1^+ = Q_1\).

Hence, for any \(\alpha \in Q_1\), there exists \(z \in \Gamma_u(\alpha)\) such that \(z \notin T^b_u\) and \(z \notin b\); note that necessarily \(z \in A_-\). Recall that for all \(\alpha \in Q\), \(R_u(\alpha)\) and hence \(\Gamma_u(\alpha)\) are isometric via the map \(u\) to closed intervals, and hence \(\Gamma_u(\alpha) \setminus ((\alpha) \cup b)\) is isometric to an open interval. Since \(\Gamma_u(\alpha) \cap T^b_u\) is isometric to an interval and
contains $\alpha$, it follows that for $\alpha \in Q_1$, there exist distinct $a_\alpha, b_\alpha \in \Gamma_u(\alpha) \setminus T_u^b$ so that

$$(u(b_\alpha), u(a_\alpha)) \subset u(\Gamma_u(\alpha) \setminus T_u^b)$$

is a non-empty open interval. Moreover, we may select $a_\alpha$ and $b_\alpha$ to be $q$-measurable functions of $Q_1$. To see this, consider the set $\Sigma := \{(\alpha, x, y) \in Q_1 \times \Gamma_u : x \in A_-, (\alpha, x) \in \Gamma_u, d(x, y) > 0\}$, and observe that it is analytic (being the intersection of analytic sets), and that $P_1(\Sigma) = Q_1$. By von Neumann’s selection Theorem (see [72, Theorem 5.5.2]), there exists a $\sigma(A)$-measurable selection of $\Sigma$:

$Q_1 \ni \alpha \rightarrow (a_\alpha, b_\alpha),$

and so in particular these functions are $q$-measurable. It follows that

$Q_1 \ni \alpha \rightarrow u(a_\alpha), \quad Q_1 \ni \alpha \rightarrow u(b_\alpha),$

are also $\sigma(A)$-measurable and hence $q$-measurable. Possibly restricting $Q_1$, by Lusin’s Theorem we can also assume that the above functions are continuous.

**Step 2.** By Fubini’s Theorem

$$0 < \int_{Q_1} (u(a_\alpha) - u(b_\alpha)) q(d\alpha) = \int_{\mathbb{R}} q\left(\{\alpha \in Q_1 : u(b_\alpha) < t < u(a_\alpha)\}\right) dt.$$ 

Hence there exists $c \in \mathbb{R}$ and $Q_{1,c} \subset Q_1$ with $q(Q_{1,c}) > 0$, such that for any $\alpha \in Q_{1,c}$ it holds $c \in (u(b_\alpha), u(a_\alpha))$; in particular for any $\alpha \in Q_{1,c}$ there exists a unique $z_\alpha \in \Gamma_u(\alpha)$ such that $u(z_\alpha) = c$. Furthermore, we can assume that $Q_{1,c}$ is compact, and hence by continuity of $u(a_\alpha)$ it follows that

$$\exists \varepsilon > 0 \quad \forall \alpha \in Q_{1,c} \quad u(a_\alpha) - c > \varepsilon.$$ 

Then define the following set:

$$\Lambda := \{(\alpha, x, z) \in Q_{1,c} \times \Gamma_u : (\alpha, x) \in R_u^b, u(z) = c\}.$$ 

Recall that $R_u^b$ is Borel since $(X, d)$ is proper, and therefore $\Lambda$ is Borel. Note by the aforementioned discussion that $P_1(\Lambda) = Q_{1,c}$. Also note that for $(\alpha, x, z) \in \Lambda$, since $R_u^b(\alpha)$ is isometric to a closed interval, necessarily $z = z_\alpha$. Finally, we claim that $P_{2,3}(\Lambda)$ is $d^2$-cyclically monotone: for $(x_1, z_1), (x_2, z_2) \in P_{2,3}(\Lambda)$ observe that

$$d(x_1, z_1) = u(x_1) - u(z_1) = u(x_1) - c = u(x_1) - u(z_2) \leq d(x_1, z_2).$$
Hence for \( \{(x_i, z_i)\}_{i \leq n} \subset P_{2,3}(\Lambda) \), setting \( z_{n+1} = z_1 \),

\[
\sum_{i \leq n} d^2(x_i, z_i) \leq \sum_{i \leq n} d^2(x_i, z_{i+1}),
\]

and the monotonicity follows. We can then define a function \( T \) by imposing \( \text{graph}(T) = P_{2,3}(\Lambda); \) note that \( P_{2,3}(\Lambda) \) is analytic and therefore \( T \) is Borel measurable (see [72, Theorem 4.5.2]).

**Step 3.** Consider now the measure

\[
\eta_0 := \int_{Q_{1,c}} m_\alpha q(d\alpha),
\]

and since \( q(Q_{1,c}) > 0 \) it follows that \( \eta_0(X) > 0; \) note that \( \eta_0 \) is concentrated on \( \text{Dom}(T) = \bigcup_{\alpha \in Q_{1,c}} R^b_\alpha(\alpha). \) Hence there exists \( x \in X \) and \( r > 0 \) such that \( \eta_0(B_r(x)) > 0 \), and we redefine \( \eta_0 \) to be the probability measure obtained by conditioning \( \eta_0 \) to \( B_r(x) \). Clearly \( \eta_0 \ll m \). Finally we define \( \eta_1 := T^\# \eta_0 \).

By **Step 2** and Theorem 6.15, the map \( T \) is the unique Optimal-Transport map between \( \eta_0 \) and \( \eta_1 \) for the \( W_2 \)-distance (as it is supported on a \( d^2 \)-cyclically monotone set). Consider moreover \( \nu \) the unique element of \( \text{OptGeo}(\eta_0, \eta_1) \)— then \( \nu \)-a.e. \( \gamma \) it holds that

\[
\gamma_0 \in \text{Dom}(T) \cap B_r(x) \subset T^b_u, \quad u(\gamma_1) = c, \quad (\gamma_0, \gamma_1) \in \Gamma_u.
\]

It follows in particular by Lemma 7.1 that \( \gamma_s \in \Gamma_u(\gamma_0) \) for all \( s \in [0, 1] \).

Recalling that \( u(a_\alpha) - c > \varepsilon \) for all \( \alpha \in Q_{1,c} \), that \( a_\alpha \leq M \) by continuity on \( Q_{1,c} \), and that the support of \( \eta_0 \) is bounded, it follows that there exists \( \tilde{r} \in (0, 1) \) such that \( \nu \)-a.e. \( \gamma_{\tilde{r}} \in T_u \setminus T^b_u \subset A_+ \cup A_- \). Since \( m(A_+ \cup A_-) = 0 \), necessarily \( (\varepsilon_{\tilde{r}})_n \perp m \), but this is in contradiction with the assertion of Theorem 6.15 that \( (\varepsilon_{\tilde{r}})_n \ll m \) since \( \eta_0 \ll m \) and \( \tilde{r} < 1 \). The claim follows.

\[\square\]

8 **The CD1 condition**

In this section we introduce the \( \text{CD}^1(K, N) \) condition, which plays a cardinal role in this work. As a first step towards understanding this new condition, we show that it always implies \( \text{MCP}_\varepsilon(K, N) \) (and \( \text{MCP}(K, N) \)), without requiring any types of non-branching assumptions. By analogy, we also introduce the \( \text{MCP}^1(K, N) \) condition, which may be of independent interest.
8.1 Definitions of $\text{CD}^1$ and $\text{MCP}^1$

We first assume that $\text{supp}(m) = X$. Note that we do not assume that the transport rays $\{X_\alpha\}_{\alpha \in Q}$ below are disjoint or have disjoint relative interiors, in an attempt to obtain a useful definition also for m.m.s.’s which may have significant branching. However, throughout most of this work, we will typically assume in addition that the space is essentially non-branching, in which case an equivalent definition will be presented in Proposition 8.13 below.

**Definition 8.1 (CD$^1_u(K, N)$ when $\text{supp}(m) = X$)** Let $(X, d, m)$ denote a m.m.s. with $\text{supp}(m) = X$, let $K \in \mathbb{R}$ and $N \in [1, \infty]$, and let $u : (X, d) \to \mathbb{R}$ denote a 1-Lipschitz function. $(X, d, m)$ is said to verify the $\text{CD}^1_u(K, N)$ condition if there exists a family $\{X_\alpha\}_{\alpha \in Q} \subset X$, such that:

1. There exists a disintegration of $m_{\cdot | T_u}$ on $\{X_\alpha\}_{\alpha \in Q}$:

   $$m_{\cdot | T_u} = \int_Q m_\alpha q(d\alpha), \quad \text{with} \quad m_\alpha(X_\alpha) = 1, \text{ for } q\text{-a.e. } \alpha \in Q. \quad (8.1)$$

2. For $q$-a.e. $\alpha \in Q$, $X_\alpha$ is a transport ray for $\Gamma_u$ (recall Definition 7.7).
3. For $q$-a.e. $\alpha \in Q$, $m_\alpha$ is supported on $X_\alpha$.
4. For $q$-a.e. $\alpha \in Q$, the m.m.s. $(X_\alpha, d, m_\alpha)$ verifies $\text{CD}(K, N)$.

We take this opportunity to define an analogous variant of $\text{MCP}$:

**Definition 8.2 (MCP$^1_u(K, N)$ when $\text{supp}(m) = X$)** Let $(X, d, m)$ denote a m.m.s. with $\text{supp}(m) = X$, let $K \in \mathbb{R}$ and $N \in [1, \infty]$, let $o \in X$ and denote the 1-Lipschitz function $u := d(\cdot, o)$. $(X, d, m)$ is said to verify the $\text{MCP}^1_u(K, N)$ condition if there exists a family $\{X_\alpha\}_{\alpha \in Q} \subset X$, such that conditions (1)–(3) above hold, together with:

4'. For $q$-a.e. $\alpha \in Q$, the m.m.s. $(X_\alpha, d, m_\alpha)$ verifies $\text{MCP}(K, N)$ with respect to $o \in X_\alpha$.

**Remark 8.3** Note that when $u = d(\cdot, o)$ then necessarily $T_u = X$ (if $X$ is not a singleton). In addition $(x, o) \in \Gamma_u$ for any $x \in X$, and hence by maximality of a transport ray, we must have $o \in X_\alpha$ for $q$-a.e. $\alpha \in Q$, and by condition (3) we deduce that $o \in \text{supp}(m_\alpha)$ for $q$-a.e. $\alpha \in Q$. As $\text{CD}(K, N)$ implies $\text{MCP}(K, N)$ (in the one-dimensional case this is a triviality), we obviously see that $\text{CD}^1_u(K, N)$ implies $\text{MCP}^1_u(K, N)$ for all $u = d(\cdot, o)$.

We will focus on a particular class of 1-Lipschitz functions.

**Definition** (Signed Distance Function) Given a continuous function $f : (X, d) \to \mathbb{R}$ so that $\{f = 0\} \neq \emptyset$, the function

$$d_f : X \to \mathbb{R}, \quad d_f(x) := \text{dist}(x, \{f = 0\})\text{sgn}(f), \quad (8.2)$$
is called the signed distance function (from the zero-level set of \( f \)).

**Lemma 8.4** \( d_f \) 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 \( X \).

**Proof** Given \( x, y \in X \) with \( f(x)f(y) \geq 0 \), the assertion follows by the usual triangle inequality, valid for any metric space:

\[
|d_f(x) - d_f(y)| = |\text{dist}(x, \{ f = 0 \}) - \text{dist}(y, \{ f = 0 \})| \leq d(x, y).
\]

When \( f(x)f(y) < 0 \), and given \( \varepsilon > 0 \), let \( \gamma : [0, 1] \rightarrow X \) denote a continuous path with \( \gamma_0 = x, \gamma_1 = y \) and \( \ell(\gamma) \leq d(x, y) + \varepsilon \). By continuity, it follows that there exists \( t \in (0, 1) \) so that \( f(\gamma_t) = 0 \). It follows that

\[
|d_f(x) - d_f(y)| = \text{dist}(x, \{ f = 0 \}) + \text{dist}(y, \{ f = 0 \}) \\
\leq d(x, \gamma_t) + d(y, \gamma_t) \leq \ell(\gamma) \leq d(x, y) + \varepsilon.
\]

As \( \varepsilon > 0 \) was arbitrary, the assertion is proved. \( \Box \)

**Remark 8.5** To extend Remark 8.3 to more general signed distance functions, we will need to require that \((X, d)\) is proper, and in that case \( \mathcal{T}_{d_f} \supset X \setminus \{ f = 0 \} \).

Indeed, given \( x \in X \setminus \{ f = 0 \} \), consider the distance minimizing \( z \in \{ f = 0 \} \) (by compactness of bounded sets). Then \((x, z) \in R_{d_f} \) and as \( x \neq z \) it follows that \( x \in \mathcal{T}_{d_f} \).

We now remove the restriction that \( \text{supp}(m) = X \) and introduce the main new definitions of this work:

**Definition 8.6** \( CD^1_{\text{Lip}}(K, N), CD^1(K, N) \) and \( MCP^1(K, N) \) Let \((X, d, m)\) denote a m.m.s. and let \( K \in \mathbb{R} \) and \( N \in [1, \infty] \).

- \((X, d, m)\) is said to verify the \( CD^1_{\text{Lip}}(K, N) \) condition if \( \text{supp}(m), d, m \) verifies \( CD^1_u(K, N) \) for all 1-Lipschitz functions \( u : \text{supp}(m), d \rightarrow \mathbb{R} \).
- \((X, d, m)\) is said to verify the \( CD^1(K, N) \) condition if \( \text{supp}(m), d, m \) verifies \( CD^1_{d_f}(K, N) \) for all continuous functions \( f : \text{supp}(m), d \rightarrow \mathbb{R} \) so that \( \{ f = 0 \} \neq \emptyset \) and \( d_f : \text{supp}(m), d \rightarrow \mathbb{R} \) is 1-Lipschitz.
- \((X, d, m)\) is said to verify \( MCP^1(K, N) \) if \( \text{supp}(m), d, m \) verifies \( MCP^1_u(K, N) \) for all functions \( u(x) = d(x, o) \) with \( o \in \text{supp}(m) \).

**Remark 8.7** Clearly \( CD^1_{\text{Lip}}(K, N) \Rightarrow CD^1(K, N) \Rightarrow MCP^1(K, N) \) in view of Remark 8.3. Note that we do not a-priori know that \( d_f \) is 1-Lipschitz, since we do not know that \( \text{supp}(m), d \) is a length-space (see Lemma 8.4); nevertheless, we will shortly see that the \( CD^1(K, N) \) condition implies that \( \text{supp}(m), d \) must be a geodesic space, and hence the sentence “so that \( d_f \) is 1-Lipschitz” is in fact redundant.
Remark 8.8 By definition, the \( \text{CD}_{\text{Lip}}^1 \), \( \text{CD}^1 \) and \( \text{MCP}^1 \) conditions hold for \((X, d, m)\) iff they hold for \((\text{supp}(m), d, m)\). It is also possible to introduce a definition of \( \text{CD}_u^1 \) and \( \text{MCP}_u^1 \) which applies to \((X, d, m)\) directly, without passing through \((\text{supp}(m), d, m)\)—this would involve requiring that the transport rays \( \{X_\alpha\} \) are maximal inside \(\text{supp}(m)\), and in the case of \( \text{CD}^1_u \) would only apply to functions \(u\) which are 1-Lipschitz on \(\text{supp}(m)\) (these may be extended to the entire \(X\) by McShane’s theorem). Our choice to use a tautological approach is motivated by the analogous situation for the more classical \(W_2\) definitions of curvature-dimension (see Remark 6.11) and is purely for convenience, so as not to overload the definitions.

8.2 \( \text{MCP}^1 \) implies \( \text{MCP}_\varepsilon \)

Proposition 8.9 Let \((X, d, m)\) be a m.m.s. verifying \( \text{MCP}^1(K, N) \) with \(K \in \mathbb{R}\) and \(N \in (1, \infty)\) (in particular, this holds if it verifies \( \text{CD}_{\text{Lip}}^1(K, N) \) or \( \text{CD}^1(K, N) \)). Then it verifies \( \text{MCP}_\varepsilon(K, N) \).

Proof We will show that \((\text{supp}(m), d, m)\) satisfies \( \text{MCP}_{\varepsilon}(K, N) \), and consequently so will \((X, d, m)\). By Remark 8.8, we may therefore assume that \(\text{supp}(m) = X\). Fix any \(o \in X\) and consider the 1-Lipschitz function \(u(x) := d(x, o)\). From \( \text{MCP}^1(K, N) \) and Remark 8.5 we deduce the existence of a disintegration of \(m\) on \(T_u = X\) along a family of Borel sets \(\{X_\alpha\}_{\alpha \in Q}\):

\[
m = \int_Q m_\alpha q(d\alpha), \quad m_\alpha(X_\alpha) = 1, \text{ for } q\text{-a.e. } \alpha \in Q,
\]

so that \(X_\alpha\) is a transport ray for \(\Gamma_u\), \(m_\alpha\) is supported on \(X_\alpha\) and \((X_\alpha, d, m_\alpha)\) verifies \( \text{MCP}(K, N) \) with respect to \(o \in X_\alpha\), for \(q\)-a.e. \(\alpha \in Q\).

Now consider any \(\mu_0 \in \mathcal{P}(X)\) with \(\mu_0 \ll m\), so that \(\rho_0 := \frac{d\mu_0}{dm}\) has bounded support. By measurability of the disintegration, the function \(Q \ni \alpha \mapsto z_\alpha := \int \rho_0(x)m_\alpha(dx)\) is \(q\)-measurable, and hence \(\tilde{Q} := \{\alpha \in Q \mid z_\alpha \in (0, \infty)\}\) is \(q\)-measurable. Clearly \(\int_{\tilde{Q}} z_\alpha q(d\alpha) = \int_{\tilde{Q}} z_\alpha q(d\alpha) = 1\) since \(z_\alpha < \infty\) for \(q\)-a.e. \(\alpha \in \tilde{Q}\).

Define \(\mu_0^\alpha := \frac{1}{z_\alpha} \rho_0 m_\alpha \in \mathcal{P}(X_\alpha)\) for all \(\alpha \in \tilde{Q}\). Since for \(q\)-a.e. \(\alpha \in \tilde{Q}\), the one-dimensional (non-branching) \((X_\alpha, d)\) contains \(o\), there exists a unique element \(\nu^\alpha\) of \(\text{OptGeo}(\mu_0^\alpha, \delta_o) \cap \mathcal{P}(\text{Geo}(X_\alpha))\) where \(\text{Geo}(X_\alpha)\) denotes the space of geodesics in \(X_\alpha\). Define then

\[
\nu := \int_{\tilde{Q}} \nu^\alpha z_\alpha q(d\alpha), \quad (8.3)
\]

and observe that \((e_0)_z^* \nu = \rho_0 m = \mu_0\) and \((e_1)_z^* \nu = \delta_o\). To conclude that \(\nu \in \text{OptGeo}(\mu_0, \delta_o)\) we must show that \(t \mapsto (e_t)_z^* \nu =: \mu_t\) is a \(W_2\)-geodesic.
Indeed, for any $0 \leq s < t \leq 1$, consider the transference plan $(e_s, e_t)_*\nu$ between $\mu_s$ and $\mu_t$, yielding:

$$
W_2^2(\mu_s, \mu_t) \leq \int_{\tilde{Q}} \int_{X_\alpha \times X_\alpha} d^2(x, y)(e_s, e_t)_*\nu^\alpha(dxdy)z_\alpha q(d\alpha)
= \int_{\tilde{Q}} (t-s)^2 \int_{X_\alpha \times X_\alpha} d^2(x, y)(e_0, e_1)_*\nu^\alpha(dxdy)z_\alpha q(d\alpha)
= (t-s)^2 \int_{\tilde{Q}} \int_{X_\alpha} d^2(x, o)\mu_0^\alpha(dx)z_\alpha q(d\alpha)
= (t-s)^2 \int_{\tilde{Q}} \int_{X_\alpha} d^2(x, o)\rho_0(x)m_\alpha(dx)q(d\alpha)
= (t-s)^2 \int_{X} d^2(x, o)\rho_0(x)m(dx)
= (t-s)^2 W_2^2(\mu_0, \delta_o).
$$

By the triangle inequality, it follows that $t \mapsto \mu_t$ must indeed be a geodesic in $(P_2(X), W_2)$. Note that this property is particular to transportation to a delta measure.

It remains to establish the MCP$\varepsilon$ inequality of Definition 6.8. Fix $t \in (0, 1)$, and recall that for $q$-a.e. $\alpha \in \tilde{Q}$, the (one-dimensional, non-branching) $(X_\alpha, d, m_\alpha)$ verifies MCP$(K, N)$ (and hence MCP$\varepsilon(K, N)$), and as $\mu_0^\alpha \ll m_\alpha$ and $o \in \text{supp}(m_\alpha)$, in particular $\mu_t^\alpha := (e_t)_*\nu^\alpha \ll m_\alpha$. Applying $e_t$ to both sides of (8.3), it follows that $\mu_t = (e_t)_*\nu \ll m$. Writing $\mu_t = \rho_t m$ and $\mu_t^\alpha = \rho_t^\alpha m_\alpha$ for $q$-a.e. $\alpha \in \tilde{Q}$, the MCP$\varepsilon$ condition implies that:

$$
\int_X (\rho_t^\alpha(x))^1-N m_\alpha(dx) \geq \int_X \tau_{K,N}^{1-N}(d(x, o)) \left(\frac{\rho_0(x)}{z_\alpha}\right)^{1-N} m_\alpha(dx)
\forall q - \text{a.e. } \alpha \in \tilde{Q}.
$$

(8.4)

In addition, the application of $e_t$ to both sides of (8.3) yields the following disintegration:

$$
\rho_t m = \int_{\tilde{Q}} \rho_t^\alpha z_\alpha m_\alpha q(d\alpha).
$$

(8.5)

Now consider the set $Y = \{\rho_t > 0\}$, and note that by (8.5):

$$
\int_{X \setminus Y} \rho_t^\alpha(x)m_\alpha(dx) = 0 \quad \forall q\text{-a.e. } \alpha \in \tilde{Q}.
$$

(8.6)
Integrating (8.5) against $\rho_t^{-\frac{1}{N}}$ on $Y = \{ \rho_t > 0 \}$, applying Hölder’s inequality on the interior integral for $q$-a.e. $\alpha \in \tilde{Q}$, using (8.6), employing the one-dimensional MCP inequality (8.4) and canceling $z_\alpha$, and finally applying Hölder’s inequality again on the exterior integral, we obtain

\[
\int_X \rho_t(x)^{1-\frac{1}{N}} m(dx) = \int_Y \rho_t(x)^{1-\frac{1}{N}} m(dx)
\]

\[
= \int_{\tilde{Q}} \int_Y \rho_t^q(x) \rho_t(x)^{1-\frac{1}{N}} m_{\alpha}(dx) \zeta_q(d\alpha)
\]

\[
\geq \int_{\tilde{Q}} \left( \int_Y \rho_t^q(x) \right)^{1-\frac{1}{N}} m_{\alpha}(dx) \left( \int_Y \rho_t(x)^{\frac{N-1}{N}} m_{\alpha}(dx) \right)^{-\frac{1}{N-1}} \zeta_q(d\alpha)
\]

\[
= \int_{\tilde{Q}} \left( \int_X \tau^{(1-i)}_{K,N} (d(x,o)) \rho_0(x)^{1-\frac{1}{N}} m_{\alpha}(dx) \right)^{\frac{N}{N-1}}
\]

\[
\times \left( \int_X \rho_t(x)^{\frac{N-1}{N}} m_{\alpha}(dx) \right)^{-\frac{1}{N-1}} q(d\alpha)
\]

\[
\geq \left( \int_{\tilde{Q}} \int_X \tau^{(1-i)}_{K,N} (d(x,o)) \rho_0(x)^{1-\frac{1}{N}} m_{\alpha}(dx) q(d\alpha) \right)^{\frac{N}{N-1}}
\]

\[
\left( \int_{\tilde{Q}} \int_X \rho_t(x)^{\frac{N-1}{N}} m_{\alpha}(dx) q(d\alpha) \right)^{-\frac{1}{N-1}}
\]

\[
\geq \left( \int_{\tilde{Q}} \int_X \tau^{(1-i)}_{K,N} (d(x,o)) \rho_0(x)^{1-\frac{1}{N}} m_{\alpha}(dx) q(d\alpha) \right)^{\frac{N}{N-1}}
\]

\[
\left( \int_{\tilde{Q}} \int_X \rho_t(x)^{\frac{N-1}{N}} m_{\alpha}(dx) q(d\alpha) \right)^{-\frac{1}{N-1}}
\]

\[
= \left( \int_X \tau^{(1-i)}_{K,N} (d(x,o)) \rho_0(x)^{1-\frac{1}{N}} m(dx) \right)^{\frac{N}{N-1}} \left( \int_X \rho_t(x)^{1-\frac{1}{N}} m(dx) \right)^{-\frac{1}{N-1}}
\]

where the last inequality above follows since $\rho_0 m_{\alpha} = 0$ for $\alpha \in Q \setminus \tilde{Q}$ and since the exponent on the second term is negative. Note that we applied Hölder’s inequality above in reverse form

\[
\int |f|^{\alpha} |g|^{\beta} \, d\omega \geq \left( \int |f| \, d\omega \right)^{\alpha} \left( \int |g| \, d\omega \right)^{\beta},
\]

which is valid as soon as $\alpha + \beta = 1$, $\beta < 0$, regardless of whether or not $|g| > 0$ $\omega$-a.e..
Rearranging terms above and raising to the power of $\frac{N-1}{N}$, the desired inequality follows:

$$\int_X \rho_t(x)^{1-\frac{1}{N}} m(dx) \geq \int_X \tau_{K,N}^{(1-t)}(d(x,o)) \rho_0(x)^{1-\frac{1}{N}} m(dx).$$

\[\square\]

**Remark 8.10** Note that the above proof shows that, not only does it hold that $\text{supp}(\mu_t) \subset \text{supp}(m)$ for all $t \in [0, 1)$, as required in the definition of $\text{MCP}_\varepsilon(K, N)$, but in fact $\mu_t \ll m$.

**Remark 8.11** Recalling that $\text{MCP}_\varepsilon(K, N)$ always implies $\text{MCP}(K, N)$, we deduce that $\text{MCP}^1(K, N)$ implies $\text{MCP}(K, N)$. In fact, a direct proof of the latter implication is elementary. Indeed, let $A \subset X$ be any Borel set with $0 < m(A) < \infty$, and denote $\mu_t = \frac{1}{m(A)} m_A$. Recall that for $\tau$-a.e. $\alpha \in \tilde{Q}$, $o \in X_\alpha$, $\text{supp}(m_\alpha) = X_\alpha$ and $(X_\alpha, d, m_\alpha)$ verifies $\text{MCP}(K, N)$. Defining $\nu$ as in (8.3) and continuing with the notation used there, it follows by uniqueness of $\nu^\alpha$ and the $\text{MCP}$ condition with respect to the point $o \in X_\alpha$, that for any Borel set $B \subset X$:

$$m_\alpha(B) \geq \int_{\mathcal{e}_t^{-1}(B)} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1))^N m_\alpha(A) \nu^\alpha(d\gamma),$$

for $\tau$-a.e. $\alpha \in \tilde{Q}$. Integrating over $\tilde{Q}$ we obtain

$$m(B) \geq \int_{\tilde{Q}} m_\alpha(B) q(d\alpha)$$

\[\geq \int_{\mathcal{e}_t^{-1}(B)} \int_{\tilde{Q}} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1))^N m_\alpha(A) \nu^\alpha(d\gamma) q(d\alpha)\]

$$= \int_{\mathcal{e}_t^{-1}(B)} \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1))^N m(A) \nu(d\gamma),$$

and the claim follows.

As a consequence, we immediately obtain from Lemmas 6.12 and 8.4:

**Corollary 8.12** Let $(X, d, m)$ be a m.m.s. verifying $\text{CD}^1(K, N)$ with $K \in \mathbb{R}$ and $N \in (1, \infty)$. Then $(\text{supp}(m), d)$ is a Polish, proper and geodesic space. In particular, for any continuous function $f : (\text{supp}(m), d) \to \mathbb{R}$ with $\{f = 0\} \neq \emptyset$, the function $d_f : (\text{supp}(m), d) \to \mathbb{R}$ is $1$-Lipschitz.
8.3 On essentially non-branching spaces

Having at our disposal \( \text{MCP}(K, N) \), we can now invoke the results of Sect. 7 concerning the theory of \( L^1 \)-Optimal-Transport, and obtain the following important equivalent definitions of \( \text{CD}^1_{\text{Lip}}(K, N) \), \( \text{CD}^1(K, N) \) and \( \text{MCP}^1(K, N) \) assuming that \((X, d, m)\) is essentially non-branching.

**Proposition 8.13** Let \((X, d, m)\) be an essentially non-branching m.m.s. with \( \text{supp}(m) = X \). Given \( K \in \mathbb{R} \) and \( N \in (1, \infty) \), the following statements are equivalent:

1. \((X, d, m)\) verifies \( \text{CD}^1_{\text{Lip}}(K, N) \).
2. For any 1-Lipschitz function \( u : (X, d) \to \mathbb{R} \), let \( \{R^b_u(\alpha)\}_{\alpha \in Q} \) denote the partition of \( T^b_u \) given by the equivalence classes of \( R^b_u \). Denote by \( X_\alpha \) the closure \( \overline{R^b_u(\alpha)} \). Then all the conditions (1)–(4) of Definition 8.1 hold for the family \( \{X_\alpha\}_{\alpha \in Q} \). In particular, \( X_\alpha = R_u(\alpha) \) is a transport-ray for \( q \)-a.e. \( \alpha \in Q \).

Moreover, the sets \( \{X_\alpha\}_{\alpha \in Q} \) have disjoint interiors \( \{\overline{R^b_u(\alpha)}\}_{\alpha \in Q} \) contained in \( T^b_u \), and the disintegration \((Q, \mathcal{Q}, q)\) of \( m|_{T^b_u} \) on \( \{X_\alpha\}_{\alpha \in Q} \) given by (8.1) is essentially unique.

Furthermore, \( Q \) may be chosen to be a section of the above partition so that \( Q \supset \overline{Q} \in \mathcal{B}(T^b_u) \) with \( \overline{Q} \) an \( m\)-section with \( m\)-measurable quotient map, so that in particular \( \mathcal{Q} \supset \mathcal{B}(\overline{Q}) \) and \( q \) is concentrated on \( \overline{Q} \).

An identical statement holds for \( \text{CD}^1(K, N) \) when only considering signed distance functions \( u = d_f \).

An identical statement also holds for \( \text{MCP}^1(K, N) \) when only considering the functions \( u = d(\cdot, o) \), after replacing above condition (4) of Definition 8.1 with condition (4’) of Definition 8.2.

**Proof** The only direction requiring proof is (1) \( \Rightarrow \) (2). Given a 1-Lipschitz function \( u \) as above, we may assume that \( m(T_u) > 0 \), otherwise there is nothing to prove. The \( \text{CD}^1(K, N) \) condition ensures there exists a family \( \{Y_\beta\}_{\beta \in P} \) of sets and a disintegration:

\[
m|_{T_u} = \int_P m^P_\beta p(d\beta), \quad \text{with} \quad m^P_\beta(Y_\beta) = 1, \quad \text{for} \ p\text{-a.e.} \ \beta \in P,
\]

so that for \( p\text{-a.e.} \ \beta \in P \), \( Y_\beta \) is a transport ray for \( \Gamma_u \), \((Y_\beta, d, m^P_\beta)\) satisfies \( \text{CD}(K, N) \) and \( \text{supp}(m^P_\beta) = Y_\beta \). By removing a \( p\)-null-set from \( P \), let us assume without loss of generality that the above properties hold for all \( \beta \in P \).

As \( \text{CD}^1_{\text{Lip}}(K, N) \Rightarrow \text{CD}^1(K, N) \Rightarrow \text{MCP}^1(K, N) \Rightarrow \text{MCP}(K, N) \), and as our space is essentially non-branching with full-support, Corollary 7.3
implies that $m(A_+ \cup A_-) = 0$ and that there exists an essentially unique disintegration $(Q, \mathcal{Q}, q)$ of $m_{U_u} = m_{\mathcal{T}_u}$ strongly consistent with the partition of $\mathcal{T}_u$ given by $\{R_u^b(\alpha)\}_{\alpha \in Q}$:

$$m_{U_u} = \int_Q m_{\alpha} q(d\alpha) \quad \text{with} \quad m_{\alpha}(R_u^b(\alpha)) = 1, \quad \text{for q-a.e. } \alpha \in Q. \quad (8.7)$$

By Corollary 7.3, $Q$ may be chosen to be a section of the above partition satisfying the statement appearing in the formulation of Proposition 8.13. Again, let us assume without loss of generality that $m_{\alpha}(R_u^b(\alpha)) = 1$ for all $\alpha \in Q$.

By Theorem 7.10, there exists $Q_1 \subset Q$ of full $q$-measure so that $R_u(\alpha) = \overline{R_u^b(\alpha)} \supset R_u^b(\alpha) \supset \hat{R}_u(\alpha)$ for all $\alpha \in Q_1$. In addition, since $m(U_u \setminus \mathcal{T}_u) = 0$, there exists $P_1 \subset P$ of full $p$-measure so that $m_{P}^b(T_u^b) = 1$ for all $\beta \in P_1$. By Lemmas 7.6 and 7.8, $(Y_{\beta} \cap T_u^b, d)$ is isometric to an interval in $(\mathbb{R}, |\cdot|)$, and therefore $(Y_{\beta} \cap T_u^b, d, (m_{P}^b)_{\mathcal{T}_u})$ still satisfies $CD(K, N)$, is of total measure 1 and satisfies $\text{supp}(m_{P}^b) = Y_{\beta} \cap T_u^b$, for all $\beta \in P_1$.

Now by Lemma 7.8, since $Y_{\beta} \cap T_u^b \neq \emptyset$ for all $\beta \in P_1$, $Y_{\beta} = R_u(x)$ for all $x \in Y_{\beta} \cap T_u^b$. In particular, for all $\beta \in P_1$, there exists a unique (since $R_u^b$ is an equivalence relation on $T_u^b$ and by uniqueness of the section map) $\alpha = \alpha(\beta) \in Q$ so that $Y_{\beta} = R_u(\alpha)$. Denoting by $\tilde{Q} \subset Q$ the set of indices $\alpha$ obtained in this way, it is clear that $\tilde{Q}$ if of full $q$-measure, since:

$$0 = p(P \setminus P_1) = m(T_u^b \setminus \bigcup_{\beta \in P_1} Y_{\beta})$$

$$= m(T_u^b \setminus \bigcup_{\alpha(\beta) : \beta \in P_1} R_u(\alpha(\beta))) = q(Q \setminus \tilde{Q}).$$

Consequently, $Q_2 := \tilde{Q} \cap Q_1$ is of full $q$-measure as well. Denoting $P_2 := \alpha^{-1}(Q_2)$ and repeating the above argument, it follows that $P_2 \subset P_1$ is of full $p$-measure and satisfies that for all $\beta \in P_2$, $Y_{\beta} = R_u(\alpha)$ for $\alpha = \alpha(\beta) \in Q_2$.

We conclude that there is a one-to-one correspondence:

$$\eta : P_2 \ni \beta \leftrightarrow \alpha \in Q_2 \quad \text{whenever} \quad Y_{\beta} \cap T_u^b = R_u^b(\alpha)(= R_u(\alpha) \cap T_u^b),$$

so both of these representations yield an identical partition (up to relabeling) of the set

$$C := \bigcup_{\beta \in P_2} (Y_{\beta} \cap T_u^b) = \bigcup_{\alpha \in Q_2} R_u^b(\alpha).$$
Clearly \( m(T_u \setminus C) = 0 \) and so \( C \) is \( m \)-measurable. Therefore, by the above two disintegration formulae:

\[
m(T_u \setminus C) = m(T_u \setminus C) = \int_{P_2} (m^P_\beta)_{\cdot \cap T_u^b} p(d\beta) = \int_{Q_2} m_\alpha q(d\alpha).
\]

After identifying between \( P_2 \) and \( Q_2 \) via \( \eta \), it follows necessarily that \( q_{\cdot \cap Q_2} = p_{\cdot \cap P_2} \) as they are both the push-forward of \( m_{\cdot \cap C} \) under the partition map (since \( (m^P_\beta)_{\cdot \cap T_u^b} \) and \( m_\alpha \) are both probability measures on \( T_u \)). Applying the Disintegration Theorem 6.19 to \((C, B(C), m_{\cdot \cap C})\), we conclude that there is an essentially unique disintegration of \( m_{\cdot \cap C} \) on the above partition of \( C \). Consequently, there exist \( P_3 \subset P_2 \) of full \( p \)-measure and \( Q_3 = \eta(P_3) \subset Q_2 \) of full \( q \)-measure so that

\[
(m^P_\beta)_{\cdot \cap T_u^b} = m_\alpha
\]

for all pairs \( (\beta, \alpha) \in P_3 \times Q_3 \) related by the correspondence \( \eta \).

Recall that \( X_\alpha := R^b_{u}(\alpha) \). It follows that for all \( \alpha \in Q_3 \) (with corresponding \( \beta \in P_3 \)):

1. \( X_\alpha = R^b_{u}(\alpha) = R_u(\alpha) \) is a transport ray.
2. \( (Y_\beta \cap T_u^b, d, (m^P_\beta)_{\cdot \cap T_u^b}) = (R^b_{u}(\alpha) = X_\alpha, d, m_\alpha) \) satisfies \( \text{CD}(K, N) \) with total measure 1.
3. Consequently

\[
m_{T_u \setminus T_u^b} = \int_{Q} m_\alpha q(d\alpha), \tag{8.8}
\]

is a disintegration on \( \{X_\alpha\}_{\alpha \in Q} \).

4. \( m_\alpha = (m^P_\beta)_{\cdot \cap T_u^b} \) is supported on \( Y_\beta \cap T_u^b = R^b_{u}(\alpha) = X_\alpha \).

This confirms the 4 conditions of Definition 8.1, and the essential uniqueness of the disintegration (8.8) readily follows from that of the disintegration (8.7) and the arguments above.

Finally, by Lemma 7.6, since \( (R^b_{u}(\alpha) = R_u(\alpha) \cap T_u^b, d) \) is isometric to an interval in \( (\mathbb{R}, |\cdot|) \), then \( \hat{X}_\alpha = \hat{R}^b_{u}(\alpha) \) for all \( \alpha \in Q \). As \( \{\hat{R}^b_{u}(\alpha)\}_{\alpha \in Q} \) are equivalence classes, it follows that \( \{\hat{X}_\alpha\}_{\alpha \in Q} \) is a family of disjoint subsets of \( T_u^b \). This concludes the proof for the case of \( \text{CD}^1_{Lip} \) and \( \text{CD}^1 \).

For \( \text{MCP}^1 \), one just needs to note that if \( u = d(\cdot, o) \) then \( o \in Y_\beta \) for all \( \beta \in P \) (by Remark 8.3, since \( Y_\beta \) is a transport ray). Recalling the definition of \( P_1 \subset P \), since \( (Y_\beta \cap T_u^b, d) \) is isometric to an interval and \( m^P_\beta(Y_\beta \cap T_u^b) = 1 \) for all \( \beta \in P_1 \), it follows necessarily that for those \( \beta, o \in Y_\beta \cap T_u^b \) and
Recall moreover that we already derived several properties of $W_2$-geodesics in essentially non-branching m.m.s.’s verifying $\text{MCP}(K, N)$. Hence from Proposition 8.9 we also obtain all the claims of Theorem 6.15 and Corollary 6.16, as well as all of the results of the next section, provided the m.m.s. is essentially non-branching and verifies $\text{CD}^1(K, N)$ for $N \in (1, \infty)$.

9 Temporal-regularity under $\text{MCP}$

In this section we deduce from the Measure Contraction and essentially non-branching properties various temporal-regularity results for the map $t \mapsto \rho_t(\gamma_t)$ and related objects, which we will require for this work. By Proposition 8.9, these results also apply under the $\text{CD}^1$ condition. While these properties are essentially standard consequences of recently available results and tools, they appear to be new and may be of independent interest.

As usual, we assume that $K \in \mathbb{R}$ and $N \in (1, \infty)$. We begin with:

\begin{proposition}
Let $(X, d, m)$ denote an essentially non-branching m.m.s. Then the following are equivalent:

1. $(X, d, m)$ verifies $\text{MCP}(K, N)$.
2. $(X, d, m)$ verifies $\text{MCP}_\varepsilon(K, N)$.
3. For all $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ with $\mu_0 \ll m$ and $\text{supp}(\mu_1) \subset \text{supp}(m)$, there exists a unique $\nu \in \text{OptGeo}(\mu_0, \mu_1)$, $\nu$ is induced by a map (i.e. $\nu = S_n(\mu_0)$ for some map $S : X \to \text{Geo}(X)$), $\mu_t := (e_t)^\# \nu \ll m$ for all $t \in [0, 1]$, and writing $\mu_t = \rho_t m$, we have for all $t \in [0, 1]$:

$$\rho_t^{-\frac{1}{N}}(\gamma_t) \geq \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1))\rho_t^{-\frac{1}{N}}(\gamma_0) \text{ for } \nu\text{-a.e. } \gamma \in \text{Geo}(X),$$

(9.1)

and (integrating with respect to $\nu$):

$$\mathcal{E}_N(\mu_t) \geq \int \tau_{K,N}^{(1-t)}(d(\gamma_0, \gamma_1))\rho_0^{-\frac{1}{N}}(\gamma_0)\nu(d\gamma).$$

(9.2)

4. For all $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ of the form $\mu_1 = \delta_o$ for some $o \in \text{supp}(m)$ and $\mu_0 = \frac{1}{m(A)}m_{\ll A}$ for some Borel set $A \subset X$ with $0 < m(A) < \infty$, there exists a $\nu \in \text{OptGeo}(\mu_0, \mu_1)$ so that for all $t \in [0, 1]$, $\mu_t := (e_t)^\# \nu \ll m$ and (9.1), (9.2) hold.

Moreover, the equivalence (1) $\iff$ (4) does not require the essentially non-branching assumption.
Lemma 6.12. Given Proof of Proposition 9.1

Indeed, the proof of (assuming or a-posteriori concluding some type of non-branching assumption. of, where one can obtain information on the density along geodesics without work, it is worthwhile remarking that this is the only instance we are aware in addition the assertions of Theorem 6.15 and Corollary 6.16 are in force.

MCP

Indeed, we have already seen in the previous section that MCP(K, N) implies MCP(K, N)), note that the MCP(K, N) condition is precisely equivalent to the validity of (9.2) for all measures μ₀, μ₁ ∈ P₂(X) of the form μ₁ = δ₀, with o ∈ supp(m) and μ₀ ≪ m with bounded support.

Remark 9.2 In fact, for essentially non-branching spaces, it is also possible to add the MCP₁(K, N) condition to the above list of equivalent statements. Indeed, we have already seen in the previous section that MCP₁(K, N) ⇒ MCPᵦ(K, N) without any non-branching assumptions. The converse implication for non-branching spaces follows from [19, Proposition 9.5] (without identifying the MCP₁(K, N) condition by this name), and it is possible to extend this to essentially non-branching spaces by following the arguments of [23, Proposition A.1].

Remark 9.3 Note that in (3), one is allowed to test any μ₁ with supp(μ₁) ⊂ supp(m), not only μ₁ = δ₀ as in the other statements. By Theorem 6.15 (recall that MCPᵦ(K, N) implies MCP(K, N)), note that the MCPᵦ(K, N) condition is precisely equivalent to the validity of (9.2) for all measures μ₀, μ₁ ∈ P₂(X) of the form μ₁ = δ₀ with o ∈ supp(m) and μ₀ ≪ m with bounded support.

Remark 9.4 While the equivalence (1) ⇔ (4) will not be directly used in this work, it is worthwhile remarking that this is the only instance we are aware of, where one can obtain information on the density along geodesics without assuming or a-posteriori concluding some type of non-branching assumption. Indeed, the proof of (1) ⇒ (4) relies on the (newly available) Theorem 3.11.

Proof of Proposition 9.1 (1) ⇒ (4). (supp(m), d) is proper and geodesic by Lemma 6.12. Given μ₀ and μ₁ = δ₀ as in (4), any ν ∈ OptGeo(μ₀, μ₁) is concentrated on Gᵦ where φ is the associated Kantorovich potential, and so Theorem 3.11 implies that d(γ₀, γ₁) = ℓ₁(γ₁) for ν-a.e. γ. It follows that with the notation of Sect. 3:

\[
\frac{1}{m(A)} m ≥ (e_t)₁(τ₁(K, N)(d(γ₀, γ₁))^N ν(dγ)) = ρ₁(x)τ₁(K, N)(ℓ₁(x))^N m(dx).
\]

The pointwise inequality between densities follows for m-a.e. x, and since ℓ₁ < ∞ (and hence τ₁(K, N)(ℓ₁(x)) > 0) for t ∈ (0, 1), this in fact implies that (e_t)₁(ν) ≪ m (without relying on Theorem 6.15, which is unavailable without the essentially non-branching assumption). Since (e_t)₁(ν) ≪ m, the inequality between densities is verified at x = γ₁ for ν-a.e. γ. Noting that \( \frac{1}{m(A)} = ρ₀(γ₀) \) for ν-a.e. γ, (9.1) and hence (9.2) are established for μ₀, μ₁ as above.

(4) ⇒ (1). This follows by applying (9.1) to μ₀ = \( \frac{1}{m(A)} m_{≤ A} \) and μ₁ = δ₀, raising the resulting inequality to the power of N, and integrating it against ν_{≤ γ₁} for all Borel sets B ⊂ supp(μ₁), thereby verifying the MCP(K, N) inequality (6.5).

(4) ⇒ (2). Let o ∈ supp(m) and let μ₀ = ρ₀m ∈ P(X) with bounded support. As (4) ⇒ (1), Lemma 6.12 implies that (supp(m), d) is proper, and in addition the assertions of Theorem 6.15 and Corollary 6.16 are in force.
Now, there exists an non-decreasing sequence \( \{ f^i \}_{i \in \mathbb{N}} \) of simple functions, that is
\[
f^i = \sum_{k \leq n(i)} \alpha_k^i \chi_{A_k^i}, \quad \alpha_k^i > 0, \quad m(A_k^i) > 0, \quad A_k^i \cap A_j^i = \emptyset, \text{ if } k \neq j,
\]
such that \( \mu_0^i := \rho_0^i m := \frac{1}{z^i} f^i m \in \mathcal{P}(X) \) is of bounded support, \( z^i := \int f^i dm \uparrow 1, \ f^i \uparrow \rho_0 \) pointwise, and \( \mu_0^i \rightharpoonup \mu_0 \) weakly, as \( i \to \infty \). By Theorem 6.15 there exists a unique \( \nu^i \in \text{OptGeo}(\mu_0^i, \delta_o) \), it is induced by a map, and can be written as
\[
\nu^i = \sum_{k \leq n(i)} \frac{1}{z^i} \alpha_k^i m(A_k^i) \nu_k^i,
\]
with each \( \nu_k^i \) the unique optimal dynamical plan between \( \mu_0^i, k := \rho_0^i m \subseteq A_k^i \) and \( \delta_o \). Moreover, \( (e_t)_\# \nu_k^i \perp (e_t)_\# \nu_j^i \) whenever \( k \neq j \), for all \( t \in [0, 1) \) by Corollary 6.16. Lastly, \( \text{supp}(\nu^i) \subset \text{Geo}(\text{supp}(m)) \) by Remark 6.11.

It follows by (9.2) applied to \( \nu_k^i \) that
\[
\mathcal{E}_N((e_t)_\# \nu_k^i) \geq \int \tau_{K,N}^{(1-t)}(d(x, o)) \left( \rho_{0,k}^i(x) \right)^{1-\frac{1}{N}} m(dx).
\]

Passing to a subsequence if necessary, Lemma 6.1 implies that \( \nu^i \rightharpoonup \nu^\infty \in \text{OptGeo}(\mu_0, \delta_o) \), and hence \( (e_t)_\# \nu^i \rightharpoonup (e_t)_\# \nu^\infty \). It follows by upper semi-continuity of \( \mathcal{E}_N \) on the left-hand side of (9.3), and monotone convergence (and \( z_i \to 1 \)) on the right hand side, that taking \( i \to \infty \) yields the \( \text{MCP}_\varepsilon(K, N) \) inequality (6.4). (2) \( \Rightarrow \) (3). By Remark 6.11, we may reduce to the case \( \text{supp}(m) = X \). In view of Remark 9.3, we first extend the validity of (9.2) by removing the (immaterial) restriction that \( \mu_0 \) has bounded support. When \( K > 0 \), \( \text{supp}(\mu_0) \) is automatically bounded since \( \text{MCP}_\varepsilon(K, N) \) implies \( \text{MCP}(K, N) \) which by Remark 6.10 implies a Bonnet-Myers diameter estimate. When \( K \leq 0 \), we may weakly approximate a general \( \mu_0 \in \mathcal{P}_2(X, d, m) \) by measures \( \mu_0^i \ll m \) having bounded support and repeat the argument presented above in the proof of (4) \( \Rightarrow \) (2).
The case of a general \( \mu_1 \in \mathcal{P}_2(X) \) with \( \text{supp}(\mu_1) \subset \text{supp}(m) \) follows by approximating \( \mu_1 \) by a convex combination of delta-measures:

\[
\mu_i^1 = \sum_{k \leq n(i)} a_k^i \delta_{o_k^i}, \quad o_k^i \in \text{supp}(m) \text{ for } k \leq n(i), \quad \text{and } \sum_{k \leq n(i)} a_k^i = 1;
\]

with \( W_2(\mu_i^1, \mu_1) \to 0 \) as \( i \to \infty \). By Theorem 6.15 (recall again that \( \text{MCP}_\varepsilon(K, N) \) implies \( \text{MCP}(K, N) \)), for each \( i \) there exists a unique \( \nu_i \in \text{OptGeo}(\mu_0, \mu_i^1) \), and we may write \( \nu_i = \sum_{k \leq n(i)} a_k^i \nu_k^i \) so that

\[
\nu_k^i \in \text{OptGeo}((e_0)_# \nu_k^i, \delta_{o_k^i}).
\]

Moreover, as explained above, \((e_t)_# \nu_k^i \perp (e_t)_# \nu_j^i\) whenever \( k \neq j \), for all \( t \in [0, 1) \). Furthermore, as \((e_0)_# \nu_k^i \ll m \) (since \((e_0)_# \nu^1 = \mu_0 = \rho_0 m \ll m)\), Theorem 6.15 implies that \((e_t)_# \nu_k^i \ll m \) for all \( t \in [0, 1) \). Writing \((e_t)_# \nu_k^i = \rho_{k, t}^i m\), the \( \text{MCP}_\varepsilon(K, N) \) condition implies for all \( t \in [0, 1) \):

\[
\int (\rho_{k, t}^i)^{1 - \frac{1}{N}}(x) m(dx) \geq \int \tau^{(1-t)}_{K, N}(d(x, o_k^i))(\rho_{0, k}^i)^{1 - \frac{1}{N}}(x) m(dx).
\]

Multiplying by \((a_k^i)^{1-1/N}\), summing over \( k \) and using the mutual singularity of the corresponding measures, we obtain

\[
\mathcal{E}_N((e_t)_# \nu^i) \geq \int_X \tau^{(1-t)}_{K, N}(d(x, y)) \rho_{0}^{-\frac{1}{N}}(x) (e_0, e_1)_# \nu^i(dx dy).
\]

Passing as usual to a subsequence if necessary, Lemma 6.1 implies that \( \nu^i \rightharpoonup \nu^\infty \in \text{OptGeo}(\mu_0, \mu_1) \), and hence \((e_t)_# \nu^i \rightharpoonup (e_t)_# \nu^\infty \). Invoking the upper semi-continuity of \( \mathcal{E}_N \) on the left-hand-side, and lower semi-continuity of the right-hand-side (see [74, Lemma 3.3], noting that the first marginal of \( \nu^i \) is fixed to be \( \mu_0 = \rho_0 m \)), (9.2) finally follows in full generality.

The density estimate (9.1) then follows using a straightforward variation of [41, Proposition 3.1], where it was shown how the existence of (a necessarily unique) transport map \( S \) may be used to obtain a pointwise density inequality such as (9.1) from an integral inequality such as (9.2) (the statement of [41, Proposition 3.1] involves an assumption on infinitesimal Hilbertianity of the space, but the only property used in the proof is the existence of a transport map \( S \) inducing a unique optimal dynamical plan).

Finally, (3) \( \Rightarrow \) (4) is trivial. This concludes the proof. \( \square \)
Corollary 9.5  Let \((X, d, m)\) be an essentially non-branching m.m.s. verifying MCP\((K, N)\). Then with the same assumptions and notation as in Proposition 9.1 (3), there exist versions of the densities \(\rho_t = \frac{d\mu_t}{dm}\), \(t \in [0, 1)\), so that for \(v\)-a.e. \(\gamma \in \text{Geo}(X)\), for all \(0 \leq s \leq t < 1\):

\[
\rho_s(\gamma_s) > 0, \quad \left(\tau_{K,N}^{(\frac{s}{t})}(d(\gamma_0, \gamma_t))\right)^N \leq \frac{\rho_t(\gamma_t)}{\rho_s(\gamma_s)} \leq \left(\tau_{K,N}^{(\frac{t-1}{t})}(d(\gamma_s, \gamma_1))\right)^{-N}
\]

\[(9.4)\]

(with \(\frac{s}{t} = \frac{0}{0}\) interpreted as 1 above). In particular, for \(v\)-a.e. \(\gamma\), the map \(t \mapsto \rho_t(\gamma_t)\) is locally Lipschitz on \((0, 1)\) and upper semi-continuous at \(t = 0\).

Proof Step 1. Given \(0 \leq s \leq t < 1\), observe that \((\text{restr}_{s}^{t})_{\mu}v\) is the unique element of OptGeo\((\mu_s, \mu_t)\); indeed \(\mu_s\) is absolutely continuous with respect to \(m\) and so Theorem 6.15 applies. In particular, we deduce that for each \(0 \leq s \leq t < 1\) and \(v\)-a.e. \(\gamma\):

\[
\rho_t(\gamma_t)^{-1/N} \geq \rho_s(\gamma_s)^{-1/N} \tau_{K,N}^{(\frac{s}{t})}(d(\gamma_0, \gamma_t)),
\]

with the exceptional set depending on \(s\) and \(t\). Reversing time and the roles of \(\mu_s, \mu_t\), we similarly obtain for each \(0 \leq s \leq t < 1\) and \(v\)-a.e. \(\gamma\) that:

\[
\rho_s(\gamma_s)^{-1/N} \geq \rho_t(\gamma_t)^{-1/N} \tau_{K,N}^{(\frac{s}{t})}(d(\gamma_0, \gamma_t)),
\]

with the exceptional set depending on \(s\) and \(t\) (the case \(s = 0\) is also included as the conclusion is then trivial). Note that given \(s \in [0, 1)\), as \(\rho_s(x) > 0\) for \(\mu_s\)-a.e. \(x\), we have that \(\rho_s(\gamma_s) > 0\) for \(v\)-a.e. \(\gamma\). Altogether, we see that for each \(0 \leq s \leq t < 1\), for \(v\)-a.e. \(\gamma\):

\[
\rho_s(\gamma_s) > 0,
\]

\[
\rho_s(\gamma_s) \left(\tau_{K,N}^{(\frac{s}{t})}(d(\gamma_0, \gamma_t))\right)^N \leq \rho_t(\gamma_t) \leq \rho_s(\gamma_s) \left(\tau_{K,N}^{(\frac{t-1}{t})}(d(\gamma_s, \gamma_1))\right)^{-N},
\]

\[(9.5)\]

with the exceptional set depending on \(s\) and \(t\).

Together with an application of Corollary 6.16, we deduce the existence of a Borel set \(H \subset \text{Geo}(X)\) with \(v(H) = 1\) such that \(c_t|_H : H \to X\) is injective for all \(t \in [0, 1)\), and such that for every \(\gamma \in H\), the double sided estimate \((9.5)\) holds for all \(s, t \in [0, 1) \cap \mathbb{Q}\). We then define for \(t \in [0, 1)\) and \(\gamma \in H\):

\[
\hat{\rho}_t(\gamma_t) := \begin{cases} 
\lim_{(0,1) \cap \mathbb{Q} \ni s \to t} \rho_s(\gamma_s) & t \in (0, 1) \\
\rho_0(\gamma_0) & t = 0
\end{cases}
\]
and $\hat{\rho}_t = 0$ outside of $e_t(H)$. By (9.5) we see that for any $\gamma \in H$ and $t \in (0, 1)$ the above limit always exists, and so by injectivity of $e_t|_H$, $\hat{\rho}_t$ is well-defined. Furthermore, (9.5) implies that for all $\gamma \in H$, $\hat{\rho}_t(\gamma)$ satisfies (9.5) itself for all $0 \leq s \leq t < 1$. Finally, for each $t \in [0, 1)$ consider any sequence $\{s_n\} \subset \mathbb{Q}$ converging to $t$; then (9.5) is valid for $\nu$-a.e. $\gamma$ at $t$ and $s_n$, with the exceptional set not depending on $n$. Taking the limit as $n \to \infty$ implies $\rho_t(\gamma_t) = \hat{\rho}_t(\gamma_t)$. Hence we have obtained that for each $t \in [0, 1)$, for $\nu$-a.e. $\gamma$:

$$\rho_t(\gamma_t) = \hat{\rho}_t(\gamma_t),$$

with the exceptional set depending only on $t$.

It follows that for all $t \in [0, 1)$, $\rho_t(x) = \hat{\rho}_t(x)$ for $\mu_t$-a.e. $x$. As $\mu_t$ and $\mathfrak{m}$ are mutually absolutely continuous on $\{\rho_t > 0\}$, it follows that $\rho_t \mathfrak{m} = \hat{\rho}_t 1_{\{\rho_t > 0\}} \mathfrak{m}$ for all $t \in [0, 1)$.

**Step 2.** We now claim that for all $t \in [0, 1)$, $\mathfrak{m}(\{\rho_t = 0\} \cap e_t(H)) = 0$. This will establish that $\mu_t = \rho_t = \hat{\rho}_t \mathfrak{m}$, so that $\hat{\rho}_t$ is indeed a density of $\mu_t$, thereby concluding the proof.

Suppose in the contrapositive that the above is false, so that there exists $t \in [0, 1)$ with $\mathfrak{m}(\{\rho_t = 0\} \cap e_t(H)) > 0$. As $e_t|_H$ is injective, there exist $K \subset H$ such that $K_t := e_t(K) = \{\rho_t = 0\} \cap e_t(H)$.

Set $K_s := e_s(K)$ for all $s \in [0, 1)$. We claim that $\mathfrak{m}(K_s) > 0$ for all $s \in (0, 1)$. Indeed, define $\eta_t := \mathfrak{m}_{\leq K_t}/\mathfrak{m}(K_t)$ and set $\bar{\nu} := (e_t|_H)^{-1} \eta_t$ and $\eta_s := (e_s)\# \bar{\nu}$. As $\bar{\nu}$ is concentrated on $K \subset H \subset \text{supp}(\nu)$, it follows that $(\text{restr}_{\cdot}^t) \bar{\nu}$ must be an optimal dynamical plan between $\eta_t$ and $\eta_1$. As $\eta_t \ll \mathfrak{m}$, Theorem 6.15 implies that the latter plan is in fact the unique element of $\text{OptGeo}(\eta_t, \eta_1)$, and that $\eta_s \ll \mathfrak{m}$ for all $s \in [t, 1)$. As $\eta_t(K_s) = 1$, it follows that $\mathfrak{m}(K_s) > 0$. If $t > 0$, a similar argument applies to the range $s \in (0, t]$.

However, by definition, for all $s \in [0, 1) \cap \mathbb{Q}$ we have $0 < \hat{\rho}_s = \rho_s$ on $e_s(H)$, and in particular on $e_s(K) = K_s$. Choosing any $s \in (0, 1) \cap \mathbb{Q}$, we obtain the desired contradiction:

$$0 < \int_{K_s} \rho_s \mathfrak{m} = \mu_s(K_s) = \mu_t(K_t) = \int_{K_t} \rho_t \mathfrak{m} = 0.$$

This concludes the proof.

**Proposition 9.6** Let $(X, d, \mathfrak{m})$ be an essentially non-branching m.m.s. verifying $\text{MCP}(K, N)$. Consider any $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ with $\mu_0 \ll \mathfrak{m}$ and $\text{supp}(\mu_1) \subset \text{supp}(\mathfrak{m})$, and let $\nu$ denote the unique element of $\text{OptGeo}(\mu_0, \mu_1)$. Then for any compact set $G \subset \text{Geo}(X)$ with $\nu(G) > 0$, such that (9.4) holds for all $\gamma \in G$ and $0 \leq s \leq t < 1$, we have for all $s \in [0, 1)$, $\mathfrak{m}(e_s(G)) > 0,$
and for all $0 \leq s \leq t < 1$:

$$
\left( \frac{1 - t}{1 - s} \right)^N e^{-d(G)(t-s)\sqrt{(N-1)K^-}} \\
\leq \frac{m(e_t(G))}{m(e_s(G))} \leq \left( \frac{t}{s} \right)^N e^{d(G)(t-s)\sqrt{(N-1)K^-}},
$$

(9.6)

where $d(G) = \sup\{\ell(\gamma) : \gamma \in G\} < \infty$ and $K^- = \max\{0, -K\}$ (and with $t \geq \frac{0}{\ast}$ interpreted as 1 above). In particular, the map $t \mapsto m(e_t(G))$ is locally Lipschitz on $(0, 1)$ and lower semi-continuous at $t = 0$.

Proof We proceed with the usual notation repeatedly used above. Fix $s \in [0, 1)$. Since $\mu_s(e_s(G)) \geq \nu(G) > 0$ and $\mu_s \ll m$, it follows that $m(e_s(G)) > 0$. Define $\bar{\mu}_0 := m_{\ll e_s(G)/m(e_s(G))}$.

By Corollary 6.16, there exists a Borel set $H \subset G$ such that $e_s^{-1}: e_s(H) \to G$ is a single valued map and:

$$
\nu(G \setminus H) = 0, \quad m(e_s(G) \setminus e_s(H)) = 0,
$$

(9.7)

where the second assertion above follows since $m$ and $\mu_s$ are mutually absolutely continuous on $\{\rho_s > 0\}$, and since our assumption (9.4) guarantees that $e_s(G) \subset \{\rho_s > 0\}$. Now consider:

$$
\tilde{v} := (\text{rest}_s^1 \circ e_s^{-1})_{\ast}^\sharp(\bar{\mu}_0 \ll e_s(H)) = \int_{e_s(H)} \delta_{\text{rest}_s^1(e_s^{-1}(x))}\bar{\mu}_0(dx) \in \mathcal{P}(\text{Geo}(X)).
$$

By construction and (9.7), $(e_0)_{\ast}^\sharp \tilde{v} = \bar{\mu}_0$; define $\bar{\mu}_1 := (e_1)_{\ast}^\sharp \tilde{v}$ and note that necessarily $\tilde{v} \in \text{OptGeo}(\bar{\mu}_0, \bar{\mu}_1)$ (since $\tilde{v}$ is still supported on a $d^2/2$-cyclically monotone set) and that it is induced by the map $T := e_1 \circ e_t^{-1}$. Theorem 6.15 then implies that $\bar{\mu}_r = \bar{\rho}_r m \ll m$ for all $r \in [0, 1)$. Note that $\bar{\mu}_r$ is concentrated on the compact set $e_t(G)$ with $t := s + r(1 - s)$, and therefore $m(\text{supp}(\bar{\mu}_r)) \leq m(e_t(G))$. It follows by Jensen’s inequality together with the MCP($K , N$) assumption that:

$$
m(e_t(G))^{1/N} \geq m(\text{supp}(\bar{\mu}_r))^{1/N} \geq \int_{e_t(G)} \bar{\rho}_r^{1/N}(x) m(dx) \\
\geq m(e_s(G))^{1/N} \int_{e_s(G)} \tau_{k,n}^{(1-r)}(d(x, T(x))) m(dx) \\
\geq m(e_s(G))^{1/N} (1 - r)e^{-(1-s)d(G)r\sqrt{(N-1)K^-}/N},
$$
where the last inequality follows from the lower bound (see e.g. [29, Remark 2.3]):
\[
\tau_{K,N}^{(1-r)}(\theta) = (1 - r) \left( \frac{\sigma_{K,N-1}^{(1-r)}(\theta)}{1 - r} \right)^{N-1} \geq (1 - r)e^{-\theta r \sqrt{(N-1)K^{-1}/N}}.
\]

Substituting \( r = \frac{t-s}{1-s} \), the left-hand side of (9.6) is established. Reversing the time, the right-hand side of (9.6) immediately follows, thereby concluding the proof. \( \square \)

The following two consequences of Proposition 9.6 will be required for the proof of the change-of-variables formula in Sect. 11. Recall that for any \( G \subset \text{Geo}(X) \),
\[
D(G) := \{(x, t) \in X \times [0, 1]: x = \gamma_t, \gamma \in G\},
\]
and that \( D(G)(x) = \{t \in [0, 1]: x = \gamma_t, \gamma \in G\} \) and \( D(G)(t) = \{x \in X: x = \gamma_t, \gamma \in G\} = e_t(G) \). To simplify the notation, we directly write \( G(x) \) instead of \( D(G)(x) \).

**Proposition 9.7** With the same assumptions as in Proposition 9.6, we have for any \( t \in (0, 1) \):
\[
\lim_{\epsilon \to 0+} \frac{\mathcal{L}^1(G(x) \cap (t - \epsilon, t + \epsilon))}{2\epsilon} = 1 \quad \text{in } L^1(e_t(G), m).
\]
The same result also holds for \( t = 0 \) if we dispense with the factor of 2 in the denominator.

The proof follows the same line as the proof of [26, Theorem 2.1]. We include it for the reader’s convenience.

**Proof** Fix \( t \in (0, 1) \). Suppose in the contrapositive that the claim is false:
\[
\limsup_{\epsilon \to 0} \int_{e_t(G)} \left| 1 - \frac{\mathcal{L}^1(G(x) \cap (t - \epsilon, t + \epsilon))}{2\epsilon} \right| m(dx) > 0.
\]
Consider the complement \( G(x)^c = \{t \in [0, 1]: x \notin e_t(G)\} \), and deduce the existence of a sequence \( \epsilon_n \to 0 \) such that
\[
\lim_{n \to \infty} \int_{e_t(G)} \frac{\mathcal{L}^1(G(x)^c \cap (t - \epsilon_n, t + \epsilon_n))}{2\epsilon_n} m(dx) > 0. \quad (9.8)
\]
Now let
\[ E := \{(x, s) \in e_t(G) \times (0, 1) : s \in G(x)^c\} \]
with \(E(x), E(s)\) the corresponding sections. By Fubini’s Theorem and (9.8) we obtain that:
\[
\lim_{n \to \infty} \frac{1}{2\varepsilon_n} \int_{(t-\varepsilon_n, t+\varepsilon_n)} \mathcal{L}^1(E(s)) \, ds
\]
\[
= \lim_{n \to \infty} \frac{1}{2\varepsilon_n} \mathcal{M} \times \mathcal{L}^1(E \cap (e_t(G) \times (t - \varepsilon_n, t + \varepsilon_n)))
\]
\[
= \lim_{n \to \infty} \frac{1}{2\varepsilon_n} \int_{e_t(G)} \mathcal{L}^1(G(x)^c \cap (t - \varepsilon_n, t + \varepsilon_n)) \, m(dx) > 0,
\]
so there must be a sequence of \(\{s_n\}_{n \in \mathbb{N}}\) converging to \(t\) so that \(m(E(s_n)) \geq \kappa\), for some \(\kappa > 0\). Repeating the above argument for the case \(t = 0\) with the appropriate obvious modifications, the latter conclusion also holds in that case as well. Note that
\[ E(s_n) = \{x \in e_t(G) : x \notin e_{s_n}(G)\} = e_t(G) \setminus e_{s_n}(G). \]

The compact sets \(e_{s_n}(G)\) converge to \(e_t(G)\) in Hausdorff distance: indeed, \(d(\gamma_t, \gamma_{s_n}) \leq C|t - s_n|\) where \(C := \sup_{\gamma \in G} \ell(\gamma) < \infty\) by compactness of \(G\). Hence, for each \(\varepsilon > 0\) there exists \(n(\varepsilon)\) such that for all \(n \geq n(\varepsilon)\) it holds \(e_t(G) \supset e_{s_n}(G)\) (and vice-versa), where \(A^\varepsilon := \{y \in X : d(y, A) \leq \varepsilon\}\). It follows that
\[ m(e_t(G)^\varepsilon) \geq m(e_t(G) \setminus e_{s_n}(G)) + m(e_{s_n}(G)) \geq \kappa + m(e_{s_n}(G)). \]

Taking the limit as \(n \to \infty\), the continuity property of Proposition 9.6 (lower semi-continuity if \(t = 0\)) implies that for each \(\varepsilon > 0\):
\[ m(e_t(G)^\varepsilon) \geq \kappa + m(e_t(G)) \]
with \(\kappa\) independent of \(\varepsilon\). Since \(m(e_t(G)) = \lim_{\varepsilon \to 0} m(e_t(G)^\varepsilon)\) we obtain a contradiction, and the claim is proved. \(\square\)

**Corollary 9.8** With the same assumptions as in Proposition 9.6, and assuming that \(\text{supp}(m) = X\), we have
\[ \nu(e_0^{-1}(X^0) \cap G_\varphi^+) = 0, \]
where $\varphi$ is an associated Kantorovich potential to the $c$-optimal-transport problem from $\mu_0$ to $\mu_1$ with $c = d^2/2$. In particular:

$$\mu_t \ll \chi^0 = \mu_0 \ll \chi^0 \quad \forall t \in [0, 1).$$

Recall from Sect. 3 that $G_\varphi \subset \text{Geo}(X)$ denotes the set of $\varphi$-Kantorovich geodesics, $G^+_\varphi$ denotes the subset of geodesics in $G_\varphi$ having positive length, and $X^0 = e_{[0, 1]}(G^0_\varphi)$ denotes the subset of null geodesic points in $X$. Necessarily $\nu(G_\varphi) = 1$. The assumption $\text{supp}(m) = X$ guarantees by Lemma 6.12 that $(X, d)$ is proper and geodesic, so that the results of Part I are in force; by Remark 6.11 this poses no loss in generality.

**Proof of Corollary 9.8** Suppose by contradiction that $\nu(e_{0}^{-1}(X^0) \cap G^+_\varphi) > 0$. By inner regularity, there exists a compact $G \subset e_{0}^{-1}(X^0) \cap G^+_\varphi$ with $\nu(G) > 0$ verifying the hypothesis of Proposition 9.6 and therefore also the conclusion of Proposition 9.7 for $t = 0$. In particular, for $m$-a.e. $x \in e_0(G) \subset X^0$ there exists $\gamma \in G \subset G^+_\varphi$ and $t \in (0, 1)$ (sufficiently small) such that $x = \gamma_t$. But $\mu_0(e_0(G)) = \nu(e_{0}^{-1}(e_0(G))) \geq \nu(G) > 0$, and hence $m(e_0(G)) > 0$ as $\mu_0 \ll m$. It follows that there exists at least one $x \in e_0(G)$ as above, in direct contradiction to the characterization of $X^0$ given in Lemma 3.15. Hence we can conclude that $\nu$-almost-surely, $e_t^{-1}(X^0)$ is contained in the set of null geodesics $G^0_\varphi$. For $t \in (0, 1)$, $e_t^{-1}(X^0) \subset G^0_\varphi$ by Lemma 3.15, and so we conclude that $\mu_t \ll \chi^0 = \mu_0 \ll \chi^0$ for all $t \in [0, 1)$. \hfill $\square$

**Remark 9.9** When applying the results of this section, note that when both $\mu_0, \mu_1 \ll m$, then by reversing the roles of $\mu_0$ and $\mu_1$, we in fact obtain all the above results also at the right end-point $t = 1$.

10 Two families of conditional measures

The next two sections will be devoted to the study of $W_2$-geodesics over $(X, d, m)$, when $(X, d, m)$ is assumed to be essentially non-branching and verifies $\text{CD}^1(K, N)$. By Remark 8.8, we also assume $\text{supp}(m) = X$. We will use Proposition 8.13 as an equivalent definition for $\text{CD}^1(K, N)$. By Proposition 8.9 and Remark 8.11, $X$ also verifies $\text{MCP}(K, N)$, and so Theorem 6.15 applies. In addition, it follows by Lemma 6.12 that $(X, d)$ is geodesic and proper, and so the results of Part I apply.

Fix $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$, and denote by $\nu$ the unique element of $\text{OptGeo}(\mu_0, \mu_1)$. As usual, we denote $\mu_t := (e_t)_#^\nu \ll m$ for all $t \in [0, 1]$, and set

$$\mu_t := \rho_t m \quad \forall t \in [0, 1].$$
Fix also an associated Kantorovich potential \( \varphi : X \to \mathbb{R} \) for the \( c \)-optimal-transport problem from \( \mu_0 \) to \( \mu_1 \), with \( c = d^2/2 \). Recall that \( G_\varphi \subset \text{Geo}(X) \) denotes the set of \( \varphi \)-Kantorovich geodesics and that necessarily \( \nu(G_\varphi) = 1 \). We further recall from Sect. 3 that the interpolating Kantorovich potential and its time-reversed version at time \( t \in (0, 1) \) are defined for any \( x \in X \) as

\[
\begin{align*}
-\varphi_t(x) &= \inf_{y \in X} \frac{d^2(x, y)}{2t} - \varphi(y), \\
\bar{\varphi}_t(x) &= \inf_{y \in X} \frac{d^2(x, y)}{2(1 - t)} - \varphi^c(y) \quad \forall t \in (0, 1),
\end{align*}
\]

with \( \varphi_0 = \bar{\varphi}_0 = \varphi \) and \( \varphi_1 = \bar{\varphi}_1 = -\varphi^c \). By Proposition 3.6 we have, for all \( t \in (0, 1) \), \( \varphi_t(x) \leq \bar{\varphi}_t(x) \), with equality iff \( x \in e_t(G_\varphi) \).

It will be convenient from a technical perspective to first restrict \( \nu \), by inner regularity of Radon measures, Corollary 9.5 (applied to both pairs \( \mu_0, \mu_1 \) and \( \mu_1, \mu_0 \), Proposition 9.7 and Corollary 6.16, to a suitable good compact subset \( G \subset G_\varphi^+ \) with \( \nu(G) \geq \nu(G_\varphi^+) - \varepsilon \). Recall that \( G_\varphi^+ \) was defined in Sect. 3 as the subset of geodesics in \( G_\varphi \) having positive length, and note that the length function \( \ell : \text{Geo}(X) \to [0, \infty) \) is continuous and hence is bounded away from 0 and \( \infty \) on a compact \( G \subset G_\varphi^+ \).

**Definition 10.1 (Good Subset of Geodesics)** A subset \( G \subset G_\varphi^+ \) is called good if the following properties hold:

- \( G \) is compact;
- there exists \( c > 0 \) so that for every \( \gamma \in G \):
  \[
  c \leq \ell(\gamma) \leq 1/c; \tag{10.1}
  \]
- for every \( \gamma \in G \), \( \rho_s(\gamma_s) > 0 \) for all \( s \in [0, 1] \) and \( (0, 1) \ni s \mapsto \rho_s(\gamma_s) \) is continuous;
- the claim of Proposition 9.7 holds true for \( G \);
- The map \( e_t|_G : G \to X \) is injective (and we will henceforth restrict \( e_t \) to \( G \) or its subsets).

**Assumption 10.2** We will assume in this section and in Sect. 11.1 that \( \nu \) is concentrated on a good \( G \subset G_\varphi^+ \).

We will dispose of this assumption in the Change-of-Variables Theorem 11.4.
10.1 $L^1$ partition

For $s \in [0, 1]$ and $a_s \in \mathbb{R}$, we recall the following notation (introduced in Sect. 4 for $G = G_\phi$, but now we treat a general $G \subset G_\phi$ as above):

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

As $G$ is compact and $e_s : G \rightarrow X$ is continuous, $e_s(G)$ is compact. When $s \in (0, 1)$, $\varphi_s : X \rightarrow \mathbb{R}$ is continuous by Lemma 3.2, and hence $G_{a_s}$ is compact as well.

The structure of the evolution of $G_{a_s}$, i.e. $e_{[0, 1]}(G_{a_s}) = \{ \gamma_t : t \in [0, 1], \gamma \in G_{a_s} \}$, will be the topic of this subsection, so the properties we prove below are only meaningful for $a_s \in \varphi_s(e_s(G))$ (and moreover typically when $m(e_{[0, 1]}(G_{a_s})) > 0$). It will be convenient to use a short-hand notation for the signed-distance function from a level set of $\varphi_s$:

$$d_s \varphi := d \varphi_s - a_s$$

(see (8.2)).

**Lemma 10.3** For any $s \in [0, 1]$ and $a_s \in \varphi_s(e_s(G))$ the following holds: for each $\gamma \in G_{a_s}$ and $0 \leq r \leq t \leq 1$, $(\gamma_r, \gamma_t) \in \Gamma_{d_s}$. In particular, the evolution of $G_{a_s}$ is a subset of the transport set associated to $d_{a_s}$:

$$e_{[0, 1]}(G_{a_s}) \subset T_{d_{a_s}}.$$  

**Proof** Fix $\gamma \in G_{a_s}$. If $s \in [0, 1)$ then for any $p \in \{ \varphi_s = a_s \}$:

$$\frac{d^2(\gamma_s, \gamma_1)}{2(1-s)} = \varphi_s(\gamma_s) + \varphi^c(\gamma_1) = \varphi_s(p) + \varphi^c(\gamma_1) \leq \bar{\varphi}_s(p) + \varphi^c(\gamma_1) \leq \frac{d^2(p, \gamma_1)}{2(1-s)}$$

by Lemma 3.3 and Proposition 3.6 (2), and hence $d(\gamma_s, \gamma_1) \leq d(p, \gamma_1)$; the latter also holds for $s = 1$ trivially. Similarly, if $s \in (0, 1]$ then for any $q \in \{ \varphi_s = a_s \}$:

$$\frac{d^2(\gamma_0, \gamma_s)}{2s} = \varphi(\gamma_0) - \varphi_s(\gamma_s) = \varphi(\gamma_0) - \varphi_s(q) \leq \frac{d^2(\gamma_0, q)}{2s},$$

and therefore $d(\gamma_0, \gamma_s) \leq d(\gamma_0, q)$, with the latter also holding for $s = 0$ trivially. Consequently, for any $p, q \in \{ \varphi_s = a_s \}$:

$$d(\gamma_0, \gamma_1) \leq d(\gamma_0, p) + d(q, \gamma_1).$$

Taking infimum over $p$ and $q$ it follows that:

$$d(\gamma_0, \gamma_1) \leq d_{a_s}(\gamma_0) - d_{a_s}(\gamma_1),$$

\(\square\) Springer
where the sign of $d_{a_s}$ was determined by the fact that $s \mapsto \varphi_s(\gamma_s)$ is decreasing (e.g. by Lemma 3.3). On the other hand

$$d_{a_s}(\gamma_0) - d_{a_s}(\gamma_1) \leq d(\gamma_0, \gamma_1),$$

thanks to the 1-Lipschitz regularity of $d_{a_s}$ ensured by Lemma 8.4 since $(X, d)$ is geodesic. Therefore equality holds and $(\gamma_0, \gamma_1) \in \Gamma_{d_{a_s}}$. The assertion then follows by Lemma 7.1. □

Next, recall by Proposition 8.13 applied to the function $u = d_{a_s}$, that according to the equivalent characterization of $\text{CD}_1^K N$, the following disintegration formula holds:

$$m_{\operatorname{T}_{d_{a_s}}} = \int_Q \hat{m}^{a_s}_\alpha \hat{q}^{a_s}(d\alpha), \quad (10.2)$$

where $Q$ is a section of the partition of $\operatorname{T}_{d_{a_s}}^b$ given by the equivalence classes $\{R_{d_{a_s}}^b(\alpha)\}_{\alpha \in Q}$, and for $\hat{q}^{a_s}$-a.e. $\alpha \in Q$, the probability measure $\hat{m}^{a_s}_\alpha$ is supported on the transport ray $X_\alpha = R_{d_{a_s}}^b(\alpha) = R_{d_{a_s}}(\alpha)$ and $(X_\alpha, d, \hat{m}^{a_s}_\alpha)$ verifies $\text{CD}(K, N)$. It follows by Lemma 10.3 that

$$m_{\operatorname{e}_{[0,1]}^{e_s}(G_{a_s})} = \int_Q \hat{m}^{a_s}_{\alpha \leftarrow e_s} \hat{q}^{a_s}(d\alpha). \quad (10.3)$$

It will be convenient to make the previous disintegration formula a bit more explicit. We refer to the “Appendix” for the definition of $\text{CD}(K, N)$ density and the (suggestive) relation to one-dimensional $\text{CD}(K, N)$ spaces. Recall that $\ell_s(\gamma_s) = \ell(\gamma)$ for all $\gamma \in G$.

**Proposition 10.4** For any $s \in (0, 1)$ and $a_s \in \varphi_s(e_s(G))$, the following disintegration formula holds:

$$m_{\operatorname{e}_{[0,1]}^{e_s}(G_{a_s})} = \int_{e_s(G_{a_s})} g^{a_s}(\beta, \cdot) \# \left( h^{a_s}_\beta \cdot L^1_{\operatorname{L}[0,1]} \right) q^{a_s}(d\beta), \quad (10.4)$$

with $q^{a_s}$ a Borel measure concentrated on $e_s(G_{a_s})$ of mass $m(e_s(G_{a_s}))$, $g^{a_s} : e_s(G_{a_s}) \times [0, 1] \rightarrow X$ is defined by $g^{a_s}(\beta, t) = e_t(e_s^{-1}(\beta))$ and is Borel measurable, for $q^{a_s}$-a.e. $\beta \in e_s(G_{a_s})$, $h^{a_s}_\beta$ is a $\text{CD}(\ell_s(\beta)^2 K, N)$ probability density on $[0, 1]$ vanishing at the end-points, and the map $e_s(G_{a_s}) \times [0, 1] \ni (\beta, t) \mapsto h^{a_s}_\beta(t)$ is $q^{a_s} \otimes L^1_{\operatorname{L}[0,1]}$-measurable.

**Proof** We will abbreviate $u = d_{a_s}$.
**Step 1.** We claim that

\[ \forall \gamma \in G_{\alpha_s}, \forall \alpha \in Q, \ e_{[0,1]}(\gamma) \cap R_u^b(\alpha) \neq \emptyset \Rightarrow R_u(\alpha) \supset e_{[0,1]}(\gamma). \]

Indeed, if \( x \in e_{[0,1]}(\gamma) \), then \( R_u(x) \supset e_{[0,1]}(\gamma) \) by Lemma 10.3. But on the other hand, \( R_u(x) = R_u(\alpha) \) for all \( x \in R_u^b(\alpha) \), since any two transport rays intersecting in \( T_u^b \) must coincide by Corollary 7.9. Hence, if \( \exists x \in e_{[0,1]}(\gamma) \cap R_u^b(\alpha) \), the assertion follows.

**Step 2.** We also claim that

\[ \forall \gamma^1, \gamma^2 \in G_{\alpha_s}, \forall \alpha \in Q, \ e_{[0,1]}(\gamma^i) \cap R_u^b(\alpha) \neq \emptyset, \ i = 1, 2 \Rightarrow \gamma^1 = \gamma^2. \]

Indeed, since \( \alpha \in Q \subset T_u^b \) then \( R_u(\alpha) \) is a transport ray by Lemma 7.8, and since \( u = d_{\alpha_s} \) is affine (with slope 1) on a transport ray, \( R_u(\alpha) \) must intersect \( \{d_{\alpha_s} = 0\} = \{\varphi_s = a_s\} \), and hence \( e_s(G_{\alpha_s}) \), at most once. It follows by **Step 1** that \( \gamma^1_s = \gamma^2_s \), and so by injectivity of \( e_s|G : G \rightarrow X \), that \( \gamma^1 = \gamma^2 \).

**Step 3.** Denote

\[ G_{\alpha_s}^1 := \left\{ \gamma \in G_{\alpha_s} : T_u^b \cap e_{[0,1]}(\gamma) \neq \emptyset \right\}, \]

\[ Q^1 := \left\{ \alpha \in Q : R_u^b(\alpha) \cap e_{[0,1]}(G_{\alpha_s}) \neq \emptyset \right\}. \]

We claim that there exists a bijective map:

\[ \eta : Q^1 \ni \alpha \mapsto \gamma^\alpha \in G_{\alpha_s}^1, \]

for which

\[ R_u^b(\alpha) \cap e_{[0,1]}(G_{\alpha_s}) = T_u^b \cap e_{[0,1]}(\gamma^\alpha) = R_u^b(\alpha) \cap e_{[0,1]}(\gamma^\alpha). \]

Indeed, for all \( \alpha \in Q^1 \), there exists precisely one \( \gamma \in G_{\alpha_s} \) (and hence \( \gamma \in G_{\alpha_s}^1 \)) so that \( R_u^b(\alpha) \cap e_{[0,1]}(\gamma) \neq \emptyset \) by **Step 2**. And vice versa, given any \( \gamma \in G_{\alpha_s}^1 \), there is at least one \( \alpha \in Q \) (and hence \( \alpha \in Q^1 \)) so that \( R_u^b(\alpha) \cap e_{[0,1]}(\gamma) \neq \emptyset \), and it follows by **Step 1** that \( e_{[0,1]}(\gamma) \subset R_u(\alpha) \) and hence \( T_u^b \cap e_{[0,1]}(\gamma) \subset R_u^b(\alpha) \); but this means that for all \( \alpha \neq \beta \in Q \), \( R_u^b(\beta) \cap e_{[0,1]}(\gamma) = \emptyset \), since \( \{R_u^b(\beta)\}_{\beta \in Q} \) is a partition of \( T_u^b \), implying the uniqueness of \( \alpha \in Q^1 \).

Moreover, we claim that the map \( \eta : (Q^1, \mathcal{B}(Q^1)) \rightarrow (G_{\alpha_s}^1, \mathcal{B}(G_{\alpha_s})) \) is measurable. Indeed, recall that \( G_{\alpha_s} \) is compact, and since \((X, d)\) is proper, \( T_u^b \) and \( R_u^b \) are Borel, and hence \( G_{\alpha_s}^1 \) is analytic. Then write:

\[ \Lambda := P_{1,2}((\gamma, x, t) \in T_u^b \times G_{\alpha_s} \times X \times [0, 1] : (y, x) \in R_u^b, x = \gamma_t), \]

\( \ominus \) Springer
and

$$\text{graph}(\eta) = \Lambda \cap (Q^1 \times G_{a_s}) = \Lambda \cap (Q^1 \times G_{a_s}).$$

Note that $\Lambda$ is analytic and that $\Lambda(x)$ is either an empty set or a singleton for all $x \in T_u^b$ by Step 2 (and the fact that $R_u^b$ is an equivalence relation on $T_u^b$). It follows that for any $B \in B(G_{a_s})$, both $A_1 = P_1(\Lambda \cap (T_u^b \times B))$ and $A_2 = P_1(\Lambda \cap (T_u^b \times (G_{a_s} \setminus B)))$ are analytic, disjoint and $Q^1 = (Q^1 \cap A_1) \cup (Q^1 \cap A_2)$. By the Lusin separability principle [72, Theorem 4.4.1], there exists a Borel subset $B_1 \subset T_u^b$ containing $A_1$ which is still disjoint from $A_2$. Consequently $\eta^{-1}(B \cap G_{a_s}) = \eta^{-1}(B) = Q^1 \cap A_1 = Q^1 \cap B_1 \in B(Q^1)$, concluding the proof that $\eta$ is Borel measurable on $Q^1$.

**Step 4.** Recall that for all $\alpha \in \tilde{Q}$ of full $\hat{q}^{a_s}$ measure, $\hat{m}_a^{a_s}$ is supported on the transport ray $R_u(\alpha) = R_u^b(\alpha)$ and $(R_u(\alpha), d, \hat{m}_a^{a_s})$ verifies CD($K, N$). Consequently, for such $\alpha$’s, $\hat{m}_a^{a_s}$ gives positive mass to any relatively open subset of $R_u(\alpha)$ and does not charge points. It follows that for $\alpha \in \tilde{Q}$, since $e_{[0, 1]}(\gamma^{\alpha}) \subset R_u(\alpha)$ has non-empty relative interior, it holds that:

$$R_u^b(\alpha) \cap e_{[0, 1]}(G_{a_s}) \neq \emptyset \iff \hat{m}_a^{a_s}(e_{[0, 1]}(G_{a_s})) = \hat{m}_a^{a_s}(R_u^b(\alpha) \cap e_{[0, 1]}(G_{a_s}))$$

$$= \hat{m}_a^{a_s}(R_u^b(\alpha) \cap e_{[0, 1]}(\gamma^{\alpha}))$$

$$= \hat{m}_a^{a_s}(e_{[0, 1]}(\gamma^{\alpha})) > 0.$$ 

In particular, $Q^1$ coincides up to a $\hat{q}^{a_s}$-null set with the $\hat{q}^{a_s}$-measurable set $Q^2 := \{\alpha \in \tilde{Q} : \hat{m}_a^{a_s}(e_{[0, 1]}(G_{a_s})) > 0\}$, and thus $Q^1$ is itself $\hat{q}^{a_s}$-measurable. In fact, it is easy to see that $Q^1$ coincides with an analytic set up to a $\hat{q}^{a_s}$-null-set.

**Step 5.** Recalling that $e_{[0, 1]}(G_{a_s}) \subset T_u$ by Lemma 10.3 and that $m(T_u \setminus T_u^b) = 0$ by Corollary 7.3, we obtain from (10.2) the following disintegration of $m_{\cdot \cdot}e_{[0, 1]}(G_{a_s})$:

$$m_{\cdot \cdot}e_{[0, 1]}(G_{a_s}) = m_{\cdot \cdot}T_u^b \cap e_{[0, 1]}(G_{a_s}) = \int_{\tilde{Q}} \hat{m}_a^{a_s}(\cdot \cdot)_{\cdot \cdot}T_u^b \cap e_{[0, 1]}(G_{a_s}) \hat{q}^{a_s}(d\alpha)$$

$$= \int_{\tilde{Q} \cap Q^1} \hat{m}_a^{a_s}(\cdot \cdot \cdot)_{\cdot \cdot}e_{[0, 1]}(\gamma^{\alpha}) \hat{q}^{a_s}(d\alpha) = \int_{\tilde{Q} \cap Q^1} \frac{\hat{m}_a^{a_s}(\cdot \cdot \cdot)_{\cdot \cdot}e_{[0, 1]}(\gamma^{\alpha})}{\hat{m}_a^{a_s}(e_{[0, 1]}(\gamma^{\alpha}))} \hat{m}_a^{a_s}(e_{[0, 1]}(\gamma^{\alpha})) \hat{q}^{a_s}(d\alpha),$$

where the last two transitions and the measurability of $\alpha \mapsto \hat{m}_a^{a_s}(e_{[0, 1]}(\gamma^{\alpha})) > 0$ follow from Step 4. For all $\alpha \in \tilde{Q} \cap Q^1$, define the probability measure:

$$\tilde{m}_a^{a_s} := \frac{\hat{m}_a^{a_s}(\cdot \cdot \cdot)_{\cdot \cdot}e_{[0, 1]}(\gamma^{\alpha})}{\hat{m}_a^{a_s}(e_{[0, 1]}(\gamma^{\alpha}))}.$$
Since $e_{[0,1]}(\gamma^\alpha)$ is a convex subset of $R_u(\alpha)$, it follows that the one-dimensional m.m.s. $(e_{[0,1]}(\gamma^\alpha), d, m^{\alpha_s})$ verifies $\text{CD}(K, N)$ and is of full support for all $\alpha \in \hat{Q} \cap Q^1$. Similarly, define:

$$\tilde{q}^{\alpha_s} := \tilde{m}_{\alpha_s}^{\alpha_s}(e_{[0,1]}(\gamma^\alpha)) \hat{q}^{\alpha_s} \ll Q^1(d\alpha).$$

**Step 6.** Recall that our original disintegration (10.2) was on $(Q, \mathcal{Q}, \hat{\tilde{q}}^{\alpha_s})$, so that there exists $\hat{Q} \subset Q$ of full $\hat{\tilde{q}}^{\alpha_s}$ measure so that $\hat{Q} \in B(T_u^b)$ and $\mathcal{Q} \supset B(\hat{Q})$. It follows that we may find $\mathcal{Q} \ni \hat{Q}^1 \subset Q^1$ with $\hat{\tilde{q}}^{\alpha_s}(Q^1 \setminus \hat{Q}^1) = 0$ so that $\mathcal{Q} \supset B(\hat{Q}^1)$. Let us now push-forward the measure space $(Q^1, \mathcal{Q} \cap Q^1, \hat{\tilde{q}}^{\alpha_s})$ via the Borel measurable map $e_{\alpha} \circ \eta$ (by Step 3), yielding the measure space $(e_{\alpha}(G_{\alpha_s}), \mathcal{Q}, q^{\alpha_s})$, which is thus guaranteed to satisfy $\mathcal{Q} \supset B(\tilde{S})$, where $\tilde{S} := e_{\alpha} \circ \eta(\hat{Q}^1)$ is of full $q^{\alpha_s}$ measure. Restricting the space to $\tilde{S}$ and abusing notation, we obtain $(\tilde{S}, \mathcal{Q}, q^{\alpha_s})$ with $\mathcal{Q} \supset B(\tilde{S})$, implying that $q^{\alpha_s}$ is a Borel measure concentrated on $\tilde{S} \subset e_{\alpha}(G^1_{\alpha_s}) \subset e_{\alpha}(G_{\alpha_s})$. Note that $\hat{\tilde{q}}^{\alpha_s}, \tilde{q}^{\alpha_s}$ and $q^{\alpha_s}$ all have total mass $m(e_{[0,1]}(G_{\alpha_s})).$

Denoting $m^{\alpha_s}_{\alpha_s} := m^{\alpha_s}_{\alpha_s}$, the disintegration from Step 5 translates to

$$m_{\mathcal{Q} \cap Q^1}(e_{\alpha}(G_{\alpha_s})) = \int_{e_{\alpha}(G_{\alpha_s})} m^{\alpha_s}_{\beta} q^{\alpha_s}(d\beta).$$

Furthermore, for $q^{\alpha_s}$-a.e. $\beta$, the m.m.s. $(e_{[0,1]}(e_{\alpha}^{-1}(\beta))), d, m^{\alpha_s}_{\beta})$ verifies $\text{CD}(K, N)$ and is of full support, and is therefore isometric to $(I^\alpha_{\beta}, \| \cdot \|, \hat{h}^{\alpha_s}_{\beta} L^1 \ll I^\alpha_{\beta})$, where $I^\alpha_{\beta} := [0, \ell_s(\beta)]$ and $\hat{h}^{\alpha_s}_{\beta}$ is a $\text{CD}(K, N)$ probability density on $I^\alpha_{\beta}$ (see Definition A.1). To prevent measurability issues, we will use the convention that $\hat{h}^{\alpha_s}_{\beta}$ vanishes at the end-points of $I^\alpha_{\beta}$.

**Step 7.** Next, we observe that $g^{\alpha_s}$ is Borel. Indeed, note that by injectivity of $e_{\alpha}:

$$\text{graph}(g^{\alpha_s}) = P_{1,2,3}([\{ (\beta, t, x, \gamma) \in e_{\alpha}(G_{\alpha_s}) \times [0, 1] \times X \times G_{\alpha_s} ; 
\gamma_t = \beta , \ \gamma_t = x \}]).$$

As $G_{\alpha_s}$ is compact, it follows that $\text{graph}(g^{\alpha_s})$ is analytic, and hence (see [72, Theorem 4.5.2]) $g^{\alpha_s}$ is Borel measurable.

**Step 8.** It follows that $m^{\alpha_s}_{\beta} = g^{\alpha_s}(\beta, \cdot)_{\beta}(h^{\alpha_s}_{\beta} L^1 \ll [0, 1])$, where:

$$[0, 1] \ni t \mapsto h^{\alpha_s}_{\beta}(t) := \ell_s(\beta) \hat{h}^{\alpha_s}_{\beta}(t \ell_s(\beta)).$$

Clearly $h^{\alpha_s}_{\beta}$ is now a $\text{CD}(\ell_s(\beta)^2 K, N)$ probability density on the interval $[0, 1]$. The only remaining task is to prove that the map $e_{\alpha}(G_{\alpha_s}) \times [0, 1] \ni (\beta, t) \mapsto$
$h_{\beta}^{a_s}(t)$ is $q^{a_s} \otimes L^1_{|[0,1]}$-measurable. By measurability of the disintegration (10.3) (recall Definition 6.18), the map $Q \ni \alpha \mapsto \hat{m}_{\alpha}^{a_s}(B)$ is $\hat{q}^{a_s}$-measurable for any Borel set $B \subset X$. It follows that for any compact $I \subset (0,1)$, the map:

$$e_s(G_{a_s}) \supset \tilde{S} \ni \beta \mapsto F(\beta) := \int_I h_{\beta}^{a_s}(\tau) d\tau = \frac{\hat{m}_{\alpha(\beta)}^{a_s}(e_I(G_{a_s}))}{\hat{m}_{\alpha(\beta)}^{a_s}(e_{[0,1]}(G_{a_s}))}$$

is $q^{a_s}$-measurable, where $\alpha(\beta) := (e_s \circ \eta)^{-1}(\beta)$ is $q^{a_s}$-measurable as a map $(\tilde{S}, \mathcal{S}, q^{a_s}) \to (Q, \mathcal{Q}, \hat{q}^{a_s})$ by the construction from Step 6. As $h_{\beta}^{a_s}$ is continuous on $(0,1)$ for $q^{a_s}$-a.e. $\beta$, we know that for such $\beta$ and all $t \in (0,1)$:

$$h_{\beta}^{a_s}(t) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{[t-\varepsilon, t+\varepsilon]} h_{\beta}^{a_s}(\tau) d\tau.$$

It follows by [72, Proposition 3.1.27] that for all $t \in (0,1)$, the map

$$\tilde{S} \ni \beta \mapsto h_{\beta}^{a_s}(t)$$

is $q^{a_s}$-measurable. As for $q^{a_s}$-a.e. $\beta$, the map $(0,1) \ni t \mapsto h_{\beta}^{a_s}(t)$ is continuous, [72, Theorem 3.1.30] confirms the required measurability.

This concludes the proof. \hfill \Box

It will be convenient to invert the order of integration in (10.4) using Fubini’s Theorem

$$m_{e_{[0,1]}(G_{a_s})} = \int_{[0,1]} g^{a_s}(\cdot, t)_\# (h_{\beta}^{a_s}(t) \cdot q^{a_s}) L^1(dt).$$

We thus define

$$m^a_t := g^{a_s}(\cdot, t)_\# (h_{\beta}^{a_s}(t) \cdot q^{a_s}),$$

so that the final formula is

$$m_{e_{[0,1]}(G_{a_s})} = \int_{[0,1]} m^a_t L^1(dt). \quad (10.5)$$

Remark 10.5 Since for $q^{a_s}$-a.e. $\beta$, the CD$(\ell^2_s(\beta) K, N)$ density $h_{\beta}^{a_s}$ must be strictly positive on $(0,1)$ (see “Appendix”), by multiplying and dividing $q^{a_s}$ by the positive $q^{a_s}$-measurable function $\beta \mapsto h_{\beta}^{a_s}(s)$ (recall that $s \in (0,1)$), we may always renormalize and assume that $h_{\beta}^{a_s}(s) = 1$. Note that this does not
affect the definition of \( m^a_s \) above. This normalization ensures that \( m^a_s = q^a_s \) so that

\[
m^a_t : = g^a_s (\cdot, t) \left( h^a_s (t) \cdot m^a_s \right).
\] (10.6)

**Remark 10.6** Note that since \( q^a_s \) is concentrated on \( e_s (G_{a_s}) \), by definition \( m^a_s \) is concentrated on \( e_t (G_{a_t}) \) for all \( t \in (0, 1) \). By Corollary 4.3, the latter sets are disjoint for different \( t \)'s in \( (0, 1) \) (recall that \( s \in (0, 1) \) and that \( G \subset G^{+}_\varphi \)). Formula (10.5) can thus be seen again as a disintegration formula over a partition. In particular, for any \( s \in (0, 1) \) and \( 0 < t, \tau < 1 \) with \( t \neq \tau \), the measures \( m^a_t \) and \( m^a_\tau \) are mutually singular.

**Proposition 10.7** For any \( s \in (0, 1) \) and \( a_s \in \varphi_s (e_s (G)) \), the map

\[
(0, 1) \ni t \mapsto m^a_t
\]

is continuous in the weak topology, we have

\[
m (e_{[0,1]} (G_{a_s})) > 0 \ \Rightarrow \ \forall t \in (0, 1) \ m^a_s (e_t (G_{a_t})) > 0,
\]

and

\[
\forall t \in [0, 1] \ m^a_t (e_t (G_{a_t})) = \| m^a_t \| \leq C \ m (e_{[0,1]} (G_{a_s})),
\]

for some \( C > 0 \) depending only on \( K, N \) and \( c > 0 \) from assumption (10.1).

**Proof** Recall that the definition of \( m^a_t \) does not depend on the last normalization we performed, when we imposed that \( h^a_s (s) = 1 \), so we revert to the normalization that \( h^a_s \) is a \( \text{CD}(\ell_s (\beta)^2 K, N) \) probability density on \( [0, 1] \), and hence \( \| q^a_s \| = m (e_{[0,1]} (G_{a_s})) \). The second assertion follows since whenever the latter mass is positive, by positivity of a \( \text{CD}(K, N) \) density in the interior of its support (see “Appendix”):

\[
\forall t \in (0, 1) \ m^a_s (e_t (G_{a_t})) = \| m^a_t \| = \int h^a_s (t) q^a_s (d \beta) > 0.
\]

Similarly, it follows by Lemma A.8, the lower semi-continuity of \( h^a_s \) at the end-points (see “Appendix”), and assumption (10.1), that \( \max_{t \in [0, 1]} h^a_s (t) \) is uniformly bounded in \( a_s \) and \( \beta \) for \( q^a_s \)-a.e. \( \beta \) by a constant \( C > 0 \) as above, implying that

\[
\forall t \in [0, 1] \quad \| m^a_t \| = \| h^a_s (t) \cdot q^a_s \| \leq C \| q^a_s \| = C \ m (e_{[0,1]} (G_{a_s})),
\]

yielding the third assertion.
Now note that the density $(0, 1) \ni t \mapsto h^a_t(t)$ is continuous (see “Appendix”) for $q^{a_s}$-a.e. $\beta$, and the same trivially holds for the map $[0, 1] \ni t \mapsto g^a(\beta, t)$. We conclude by Dominated Convergence that for any $f \in C_b(X)$ and any $t \in (0, 1)$:

$$\lim_{\tau \to t} \int f(x) m^a_{\tau}(dx) = \lim_{\tau \to t} \int f(g^a(\alpha, \tau)) h^a_\beta(\tau) q^a(\tau) d\beta = \int f(g^a(\beta, t)) h^a_\alpha(t) q^a(\alpha) d\beta = \int f(x) m^a_{t}(dx),$$

yielding the first assertion, and concluding the proof. 

\[ \Box \]

### 10.2 $L^2$ partition

For each $t \in (0, 1)$, we can find a natural partition of $e_t(G) \subset e_t(G_\psi)$ consisting of level sets of the time-propagated intermediate Kantorovich potentials $\Phi^I_s$ introduced in Sect. 4. Recall that the function $\Phi^I_s(s, t \in (0, 1))$ was defined as:

$$\Phi^I_s = \varphi_t + (t-s) \frac{\ell^2}{2},$$

and interpreted on $e_t(G_\psi)$ as the propagation of $\varphi_s$ from time $s$ to $t$ along $G_\psi$, i.e. $\Phi^I_s = \varphi_s \circ e_s \circ e_{t}^{-1}$. In particular, for any $\gamma \in G$, $\Phi^I_s(\gamma_t) = \varphi_s(\gamma_s)$, and $e_t(G_{a_s}) \cap e_t(G_{b_s}) = \emptyset$ as soon as $a_s \neq b_s$ (see Corollary 4.1). It follows that for any $s, t \in (0, 1)$, we can consider the partition of the compact set $e_t(G)$ given by its intersection with the family $\{ \Phi^I_s = a_s \}_{a_s \in \mathbb{R}}$; as usual, it will be sufficient to take $a_s \in \Phi^I_s(e_t(G)) = \varphi_s(e_t(G))$.

Since $\Phi^I_s$ is continuous, the Disintegration Theorem 6.19 yields the following essentially unique disintegration of $m_{e_t(G)}$ strongly consistent with respect to the quotient-map $\Phi^I_s$:

$$m_{e_t(G)} = \int_{\Phi^I_s(e_t(G))} \hat{m}_{a_s}^I q^I_s(da_s)$$

(10.7)

so that for $q^I_s$-a.e. $a_s$, $\hat{m}_{a_s}^I$ is a probability measure concentrated on the set $e_t(G) \cap \{ \Phi^I_s = a_s \} = e_t(G_{a_s})$. By definition, $q^I_s = (\Phi^I_s)^#m_{e_t(G)}$. To make this disintegration more explicit, we show:

**Proposition 10.8** (1) For any $s, t, \tau \in (0, 1)$, the quotient measures $q^I_s$ and $q^I_{\tau}$ are mutually absolutely continuous.

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

\[ \Box \]
Proof Recall that $q^t_s = (\Phi^t_s)^#m_{e_t(G)}$. 

(1) For any Borel set $I \subset \mathbb{R}$, note that:
\[
q^t_s(I) = m(\{ \gamma_t : \varphi_s(\gamma_s) \in I, \gamma \in G \}) > 0
\]
\[
\Leftrightarrow \mu_t(\{ \gamma_t : \varphi_s(\gamma_s) \in I, \gamma \in G \}) > 0,
\]
since $\mu_t \ll m$ and its density $\rho_t$ is assumed to be positive on $e_t(G)$ where $\mu_t$ is supported (see Definition 10.1). But
\[
\mu_t = (e_t \circ e_t^{-1})#\mu_t,
\]
and so:
\[
\mu_t(\{ \gamma_t : \varphi_s(\gamma_s) \in I, \gamma \in G \}) = \mu_t(\{ \gamma_t : \varphi_s(\gamma_s) \in I, \gamma \in G \}).
\]
It follows that $q^t_s(I) > 0$ iff $q^t_s(I) > 0$, thereby establishing the first assertion.

(2) Thanks to the first assertion, it is enough to only consider the case $t = s$ in the second one. Recall that $\Phi^t_s = \varphi_s$. Then the claim boils down to showing that $m(\varphi_s^{-1}(I) \cap e_s(G)) = 0$ whenever $I \subset \varphi_s(e_s(G))$ is a compact set with $L^1(I) = 0$.

By compactness, we fix a ball $B_r(o)$ containing $e_s(G)$. Since $\varphi_s$ is Lipschitz continuous on bounded sets (Corollary 3.10 (1)), possibly using a cut-off Lipschitz function over $B_r(o)$, we may assume that $\varphi_s$ has bounded total variation measure $\| D\varphi_s \|$ (we refer to [54] and [9] for all missing notions and background regarding BV-functions on metric-measure spaces). From the local Poincaré inequality (see Remark 7.5 and [54, page 992]) and the doubling property (see Lemma 6.12 and recall that $\text{supp}(m) = X$), it follows that the total variation measure of $\varphi_t$ is absolutely continuous with respect to $m$, and that
\[
\exists c > 0 \quad c|\nabla \varphi_s| m \leq \| D\varphi_s \| \leq |\nabla \varphi_s| m
\]
(see [54, page 992] or [12, Section 4]), where
\[
|\nabla \varphi_s|(x) := \liminf_{\delta \to 0} \sup_{y \in B_\delta(x)} \frac{|\varphi_s(y) - \varphi_s(x)|}{\delta}.
\]
By [31, Theorem 6.1], the previous quantity in fact coincides in our setting with the pointwise Lipschitz constant of $\varphi_s$ at $x$, which in turn coincides with $\ell_s^+(x)$ by [6, Theorem 3.6]; hence for $x = \gamma_s$ we have $|\nabla \varphi_s|(x) = \ell_s(x)$. By the co-area formula (see [54, Proposition 4.2]), for any Borel set $A \subset B_r(o)$:
\[
\int_{-\infty}^{+\infty} \| \partial(\varphi_s > \tau) \|(A) d\tau = \| D\varphi_s \|(A),
\]
where $\| \partial(\varphi_s > \tau) \|$ denotes the total variation measure associated to the set of finite perimeter $\{ \varphi_s > \tau \}$. From [1, Theorem 5.3] it follows that $\| \partial(\varphi_s > \tau) \|$
is concentrated on \( \{ \varphi_s = \tau \} \) and therefore, for any Borel set \( I \subset \varphi_s(e_s(G)) \) with \( \mathcal{L}^1(I) = 0 \), it follows by (10.9) and (10.8):

\[
\| D\varphi_s \| (\varphi_s^{-1}(I)) = 0, \quad |\nabla \varphi_s| \, m(\varphi_s^{-1}(I)) = 0.
\]

Since \( |\nabla \varphi_s| = \ell_s(x) > 0 \) on \( e_s(G) \), it follows that \( m(\varphi_s^{-1}(I) \cap e_s(G)) = 0 \), thereby concluding the proof. \( \square \)

**Remark 10.9** Inspecting the proof of Proposition 10.8, from the co-area formula ([54, Proposition 4.2]) and the Hausdorff representation of the perimeter measure ([1, Theorem 5.3]), it follows that for \( q^s_t \)-a.e. \( a_s \in \varphi_s(e_s(G)) \) the measure \( m^t_{as} \) is absolutely continuous with respect to the Hausdorff measure of codimension one (see [1] for more details).

Employing the previous proposition, we define:

\[
m^t_{as} := (dq^t_s/d\mathcal{L}^1) \cdot \hat{m}^t_{as},
\]

obtaining from (10.7) the following disintegration (for every \( s, t \in (0, 1) \)):

\[
m_{e_s(G)} = \int_{\varphi_s(e_s(G))} m^t_{as} \, d\mathcal{L}^1(as),
\]

with \( m^t_{as} \) concentrated on \( e_t(G_{as}) \), for \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \).

We now shed light on the relation of the above disintegration to \( L^2 \)-Optimal-Transport, by relating it to another disintegration formula for \( \nu \), the unique element of \( \text{OptGeo}(\mu_0, \mu_1) \). Observe that the family of sets \( \{ G_{as} \}_{a_s \in \mathbb{R}} \) is a partition of \( G \) and that \( G_{as} = \{ \varphi_s \circ e_s = a_s \} \). Since the quotient-map \( \varphi_s \circ e_s : \text{Geo}(X) \to \mathbb{R} \) is continuous and \( G \) is compact, the Disintegration Theorem 6.19 ensures the existence of an essentially unique disintegration of \( \nu \) strongly consistent with \( \varphi_s \circ e_s :\)

\[
\nu = \int_{\varphi_s(e_s(G))} \nu_{as} \, q^v_s(da_s),
\]

so that for \( q^v_s \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), the probability measure \( \nu_{as} \) is concentrated on \( G_{as} \). Clearly \( q^v_s(\varphi_s(e_s(G))) = \| \nu \| = 1 \).

**Corollary 10.10** (1) For any \( s \in (0, 1) \), the quotient measure \( q^v_s \) is mutually absolutely continuous with respect to \( q^s_s \), and 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. \( a_s \in \varphi_s(e_s(G)) \):

\[
\rho_t \cdot m^t_{as} = q^v_s(a_s) \cdot (e_t)_#v_{as},
\]

\( \checkmark \) Springer
where \( q^v_s := d q^v_s / d \mathcal{L}^1 \). In particular, \( m^t_{a_s} \) and \((e_t)_# v_{a_s}\) are mutually absolutely-continuous for \( q^v_s \)-a.e. \( a_s \in \varphi_s(e_s(G)) \).

(3) In particular, for any \( s \in (0, 1) \) and \( q_s^v \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), the map:

\[
[0, 1] \ni t \mapsto \rho_t \cdot m^t_{a_s}
\]

coincides for \( \mathcal{L}^1 \)-a.e. \( t \in [0, 1] \) with the \( W_2 \)-geodesic \( t \mapsto (e_t)_# v_{a_s} \) up to a positive multiplicative constant depending only on \( a_s \).

**Proof** Recall that \( \mu_s \ll m \) is supported on \( e_s(G) \) and \( \rho_s > 0 \) there (see Definition 10.1), so that \( \mu_s \) and \( m_{e_s(G)} \) are mutually absolutely-continuous. It immediately follows that the same holds for \( (\varphi_s)_# \mu_s \) and \( q_s^v = (\varphi_s)_# m_{e_s(G)} \).

But:

\[
(\varphi_s)_# (\mu_s) = (\varphi_s)_# ((e_s)_# v) = (\varphi_s \circ e_s)_# (v) = q_s^v,
\]

establishing (1).

Denoting the resulting probability density \( q_s^v := d q_s^v / d \mathcal{L}^1 \), (10.11) translates to:

\[
\nu = \int_{\varphi_s(e_s(G))} q_s^v(a_s) v_{a_s} \mathcal{L}^1(da_s).
\]

Pushing forward both sides via the evaluation map \( e_t \) given \( t \in (0, 1) \), we obtain:

\[
\rho_t m = \int_{\varphi_s(e_s(G))} q_s^v(a_s) \cdot (e_t)_# v_{a_s} \mathcal{L}^1(da_s),
\]

with \( q_s^v(a_s) \cdot (e_t)_# v_{a_s} \) concentrated on \( e_t(G_{a_s}) \) for \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \). On the other hand, multiplying both sides of (10.10) by \( \rho_t \) (which is supported on \( e_t(G) \)), we obtain

\[
\rho_t m = \int_{\varphi_s(e_s(G))} \rho_t \cdot m^t_{a_s} \mathcal{L}^1(da_s),
\]

with \( \rho_t \cdot m^t_{a_s} \) concentrated on \( e_t(G_{a_s}) \) for \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \). By the essential uniqueness of the disintegration (Theorem 6.19), noting that \( \varphi_s(e_s(G)) \) is compact, (10.12) immediately follows. As \( \rho_t > 0 \) on \( e_t(G) \) (see Definition 10.1) and \( q_s^v(a_s) \in (0, \infty) \) for \( q_s^v \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), the “in particular” part of (2) is also established.

Finally, by Fubini’s theorem, it follows that for each \( s \in (0, 1) \) and \( q_s^v \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), (10.12) holds with \( q_s^v(a_s) \in (0, \infty) \) for \( \mathcal{L}^1 \)-a.e. \( t \in [0, 1] \).
(0, 1). Note that for \( q^\nu \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), the curve \( t \mapsto (e_t)_* \nu_{a_s} \) is a \( W_2 \)-geodesic (since \( \nu_{a_s} \) is concentrated on \( G_{a_s} \subset G \)). This establishes (3), thereby concluding the proof. \( \square \)

11 Comparison between conditional measures

So far we have proved, under Assumption 10.2, that for each \( s \in (0, 1) \) we have the following two families of disintegrations:

\[
m_{e_t(G)} = \int_{\varphi_s(e_s(G))} m'_{a_s} \mathcal{L}^1(da_s) \quad \text{and} \quad m_{[0,1]}(G_{a_s}) = \int_{[0,1]} m^{a_s}_{t} \mathcal{L}^1(dt)
\]

(11.1) for each \( t \in (0, 1) \) and each \( a_s \in \varphi_s(e_s(G)) \), respectively, corresponding to the partitions:

\[
\{e_t(G_{a_s})\}_{a_s \in \varphi_s(e_s(G))} \quad \text{and} \quad \{e_t(G_{a_s})\}_{t \in (0,1)}.
\]

Moreover, both \( m'_{a_s} \) and \( m^{a_s}_{t} \) are concentrated on \( e_t(G_{a_s}) \), for each \( t \in (0, 1) \) for \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), and for each \( a_s \in \varphi_s(e_s(G)) \) and all \( t \in (0, 1) \), respectively, so that the above disintegrations are strongly consistent with respect to the corresponding partition. In addition, we have by (10.6) and (10.12) for all \( s, t \in (0, 1) \) and a.e. \( a_s \in \varphi_s(G_{a_s}) \):

\[
m^{a_s}_{t} = (e_t \circ e_s^{-1})_*(h^{a_s}(t)m^{a_s}_{s}) \quad \rho_t m'_{a_s} = (e_t \circ e_s^{-1})_*(\rho_s m^{a_s}_{a_s}).
\]

(11.2)

The goal of the first subsection, in which we retain Assumption 10.2, is to prove that \( m'_{a_s} \) and \( m^{a_s}_{t} \) are in fact equivalent measures. We will prove in particular that for all \( s \in (0, 1) \):

\[
m^{a_s}_{t} = \partial_t \Phi^t_s \ m'_{a_s} \quad \text{for a.e.} \ t \in (0,1), \ a_s \in \varphi_s(G_{a_s}).
\]

(11.3)

A heuristic formal argument for establishing (11.3) may be seen as follows. Writing \( \Phi^t_s(x) = \Phi_s(t, x) \), we have:

\[
e_t(G_{a_s}) = e_t(G) \cap \{x \in X ; \ \Phi_s(t, x) = a_s\} = e_t(G) \cap \{x \in X ; \ \Phi_s(\cdot, x)^{-1}(a_s) = t\}.
\]

Formally applying the coarea formula (assuming spatial regularity), we have:

\[
\frac{m^{a_s}_{t}}{m'_{a_s}} = \frac{\left|\nabla_x \Phi_s(t, x)\right|}{\left|\nabla_x \Phi_s(\cdot, x)^{-1}(a_s)\right|} = \left|\partial_t \Phi_s(t, x)\right|
\]
where the last transition follows by the implicit function theorem \( \nabla_x \Phi_s + \partial_t \Phi_s \cdot \nabla_x \Phi_s^{-1} = 0 \).

In the second subsection, we deduce the change-of-variables formula (1.6) for the density along geodesics, discarding Assumption 10.2. An insightful heuristic argument may be seen by combining (11.2) and (11.3) as follows:

\[
\frac{\partial}{\partial \tau} \bigg|_{\tau = \Phi_\gamma(\gamma)} \frac{m^a_s}{\rho_s} = h^{a_s}_\gamma(t) m^a_s \Bigg|_{\rho_s(m^a_s)} = \frac{\partial}{\partial \tau} \bigg|_{\tau = \Phi_\gamma(\gamma)} \frac{h^{a_s}_\gamma(t)}{\rho_s(\gamma)} \ell(\gamma)^2.
\]

### 11.1 Equivalence of conditional measures

Recall that Assumption 10.2 is still in force in this subsection. We start with the following auxiliary:

**Lemma 11.1** For every \( s, t \in (0, 1) \) and \( a_s \in \varphi_s(e_s(G)) \), the following limit:

\[
m^a_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_X f(z)m_{e_{[t-\varepsilon,t+\varepsilon]}(G_{a_s})} dz
\]

holds true in the weak topology. Moreover, for any \( f \in C_b(X) \), the map \( \varphi_s(e_s(G)) \ni a_s \mapsto \int_X f m^a_t \) is Borel.

**Proof** By Proposition 10.7, \( (0, 1) \ni t \mapsto m^a_t \) is continuous in the weak topology, and so together with (11.1), we see that for any \( f \in C_b(X) \):

\[
\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_X f(z)m_{e_{[t-\varepsilon,t+\varepsilon]}(G_{a_s})} dz = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} \left( \int_X f(z)m^a_t(z) dz \right) \mathcal{L}^1(d\tau) = \int_X f(z)m^a_t(z) dz,
\]

thereby concluding the proof of the first assertion. For the second assertion, given a compact set \( I \subset [0, 1] \), consider the compact set

\[K := \{(x, t, \gamma, a_s) \in X \times I \times G \times \varphi_s(e_s(G)) : x = \gamma_t, \varphi_s(\gamma_s) = a_s\}.
\]

Hence \( B := P_{1,4}(K) = \{(e_I(G_{a_s}), a_s) : a_s \in \varphi_s(e_s(G))\} \) is compact as well. It follows by Fubini’s theorem that the map \( \varphi_s(e_s(G)) \ni a_s \mapsto \int_{e_I(G_{a_s})} f m \) is Borel. Taking \( I = [t-\varepsilon, t+\varepsilon] \), employing the first assertion, and recalling that the pointwise limit of Borel functions is Borel, the second assertion follows. \( \square \)
Remark 11.2 One may similarly show (employing an additional density argument) that for every \( s, t \in (0, 1) \) and \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \), the following limit:

\[
m_{a_s}^t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} m(\Phi_s^{-1}[a_s-\varepsilon, a_s+\varepsilon] \cap e_t(G))
\]

holds true in the weak topology, but this will not be required.

We now find explicit expressions for the densities.

**Theorem 11.3** For any \( s \in (0, 1) \),

\[
m_s^{a_s} = \ell_s^2 \cdot m_{a_s}^t \quad \text{for } \mathcal{L}^1 \text{-a.e. } a_s \in \varphi_s(e_s(G)). \tag{11.4}
\]

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 for \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(e_s(G)) \) and \( m_{a_s}^t \)-a.e. \( x \), and we have:

\[
m_{a_s}^t = \partial_t \Phi_s^t \cdot m_{a_s}^t \quad \text{for } \mathcal{L}^1 \text{-a.e. } a_s \in \varphi_s(e_s(G)). \tag{11.5}
\]

For the ensuing proof, it will be convenient to introduce the following notation. For all \( t_0 \in \mathbb{R} \) and \( x_0 \in X \), denote

\[
1_{t_0}^1 : X \ni x \mapsto (t_0, x) \in (\mathbb{R}, X), \quad 1_{x_0}^2 : \mathbb{R} \ni t \mapsto (t, x_0) \in (\mathbb{R}, X).
\]

Recall that \( G(x) \) denotes the section \( \{ t \in [0, 1] ; \exists \gamma \in G , \; \gamma_t = x \} \) and \( \hat{G}(x) = G(x) \cap (0, 1) \).

**Proof of Theorem 11.3** **Step 1.** Fix \( s, t \in (0, 1) \). By Lemma 11.1 and the boundedness of \( \| m_{a_s}^t \| \) uniformly in \( a_s \) and \( t \in [0, 1] \) (see Proposition 10.7), it is easy to deduce (e.g. by Dominated Convergence Theorem) the following limit of measures on \( \varphi_s(e_s(G)) \times X \) in the weak topology (i.e. in duality with \( C_b(\varphi_s(e_s(G)) \times X) \)):

\[
\int_{\varphi_s(e_s(G))} (1_{a_s}^1_{x}) \epsilon (m_{a_s}^t) \mathcal{L}^1(da_s) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{t+\varepsilon} \int_{\varphi_s(e_s(G))} (1_{a_s}^1_{x}) \epsilon (m_{a_s}^t) \mathcal{L}^1(da_s) \mathcal{L}^1(d\tau).
\]

Using Fubini’s Theorem and (11.1), we proceed as follows:

\[
= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_{\varphi_t(e_t(G))} (1_{a_t}^1_{x}) \epsilon (m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

Using Fubini’s Theorem and (11.1), we proceed as follows:

\[
= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]

= \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} (\mathcal{L}^1 \otimes m_{a_t}^{(t-\varepsilon, t+\varepsilon)}(G_{a_t})) \mathcal{L}^1(da_s)
\]
= \lim_{\varepsilon \to 0} \int_{|t| < \varepsilon} \frac{1}{2\varepsilon} (t^2)_{\varepsilon} (L^1_{\varepsilon}(\Phi^r_s(x); \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x))) m(dx). \quad (11.6)

Moreover, we claim that it is enough to integrate on $e_{\varepsilon}(G)$ above:

$$= \lim_{\varepsilon \to 0} \int_{e_{\varepsilon}(G)} \frac{1}{2\varepsilon} (t^2)_{\varepsilon} (L^1_{\varepsilon}(\Phi^r_s(x); \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x))) m(dx). \quad (11.7)$$

To see this, recall that by Proposition 4.4 (3) (relying on Theorem 3.11 (2)), the map $(t - \varepsilon, t + \varepsilon) \cap \hat{G}(x) \ni \tau \mapsto \Phi^r_s(x)$ is Lipschitz with Lipschitz constant bounded uniformly in $\varepsilon \in (0, t/2 \wedge (1 - t)/2)$ and $x \in \cup_{|t| < \varepsilon} e_{\varepsilon}(G)$ (recall that for any $\gamma \in G$, $\ell(\gamma) \leq 1/c$); we denote the latter Lipschitz bound by $L$. Hence the family of measures

$$\frac{1}{2\varepsilon} L^1_{\varepsilon}(\Phi^r_s(x); \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x))$$

is bounded in the total-variation norm by $L$, uniformly in $\varepsilon$ and $x$ as above. But by continuity:

$$\lim_{\varepsilon \to 0} m(\cup_{|t| < \varepsilon} e_{\varepsilon}(G) \setminus e_{\varepsilon}(G)) = 0,$$

and so we can modify the domain of integration in (11.6) yielding (11.7).

**Step 2.** Fixing $x \in e_{\varepsilon}(G)$, we now focus on the weak limit:

$$\lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} L^1_{\varepsilon}(\Phi^r_s(x); \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x)).$$

Recall that $(t - \varepsilon, t + \varepsilon) \cap \hat{G}(x) \ni \tau \mapsto \Phi^r_s(x)$ has Lipschitz constant bounded by $L$, and moreover, is increasing by Proposition 4.4 (3). Now extend it to the entire $(0, 1)$ while preserving (non-strict) monotonicity and the bound on the Lipschitz constant, e.g. $\hat{\Phi}^r_s(x) := \inf_{r \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x)} \Phi^r_s(x) + L(t - r)^+$. Then for any $f \in C_b(\mathbb{R})$, by the change-of-variables formula for (monotone) Lipschitz functions:

$$\frac{1}{2\varepsilon} \int_{e_{\varepsilon}(G)} f(a) L^1_{\varepsilon}(da) = \frac{1}{2\varepsilon} \int_{(t - \varepsilon, t + \varepsilon) \cap \hat{G}(x)} f(\Phi^r_s(x)) \partial_\tau \hat{\Phi}^r_s(x) L^1_{\varepsilon}(d\tau) = \frac{1}{2\varepsilon} \int_{(t - \varepsilon, t + \varepsilon) \cap \hat{G}(x)} f(\Phi^r_s(x)) \partial_\tau \Phi^r_s(x) L^1_{\varepsilon}(d\tau);$$
the last transition follows since \( \tau \mapsto \Phi^\tau_s(x) \) is differentiable a.e. on \( D_\ell(x) \) and hence \( \partial_\tau \Phi^\tau_s(x) = \partial_\tau \Phi^\tau_s(x)|_{\{t-\varepsilon, t+\varepsilon\}} \) \( \Phi^\tau_s(x) \) for a.e. \( \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x) \) by Remark 2.1, and in addition since \( \partial_\tau \Phi^\tau_s(x)|_{\{t-\varepsilon, t+\varepsilon\}} \) \( \Phi^\tau_s(x) \) for a.e. \( \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x) \) by Remark 2.2. Recall that Proposition 4.4 ensures that for all \( x \in X, \partial_\tau \Phi^\tau_s(x) \) exists for \( L^1 \)-a.e. \( \tau \in \hat{G}(x) \), including at \( \tau = s \) if \( s \in \hat{G}(x) \) (in which case \( \partial_\tau \Phi^\tau_s|_{\tau = s} = \ell^2_s(x) \)). Moreover, Corollary 4.5 and our assumption that \( G \subset G_\varphi^+ \) ensure that \( \partial_\tau \Phi^\tau_s(x) > 0 \) for \( L^1 \)-a.e. \( \tau \in \hat{G}(x) \), including at \( \tau = s \). Applying Fubini’s theorem, we have

\[
0 = \int_X L^1(\hat{G}(x) \setminus \{t \in \hat{G}(x) : \exists \partial_\tau \Phi^\tau_s(x) > 0\}) m(dx)
\]

\[
= \int_0^1 m(e_\tau(G) \setminus \{x \in e_\tau(G) : \exists \partial_\tau \Phi^\tau_s(x) > 0\}) L^1(dt).
\]

It follows that for \( L^1 \)-a.e. \( \tau \in (0, 1) \), \( \partial_\tau \Phi^\tau_s(x) \) exists and is positive for m-a.e. \( x \in e_\tau(G) \) (including at \( \tau = s \) for all \( x \in e_s(G) \)).

**Step 3.** We now claim that for \( L^1 \)-a.e. \( \tau \in (0, 1) \) including \( \tau = s \), if \( f \in C_b(\mathbb{R}) \) and \( \Psi \in C_b(X) \) then

\[
\lim_{\varepsilon \to 0} \int_{e_\tau(G)} \left[ \frac{1}{2\varepsilon} \int_{\{t-\varepsilon, t+\varepsilon\}} f(\Phi^\tau_s(x)) \partial_\tau \Phi^\tau_s(x) L^1(d\tau) \right.
\]

\[
- f(\Phi^\tau_s(x)) \partial_\tau \Phi^\tau_s(x) \] \( \Psi(x) m(dx) = 0. \]

To this end, we will show that for such \( t \)'s, both

\[
I_\varepsilon(x) := \frac{1}{2\varepsilon} \int_{\{t-\varepsilon, t+\varepsilon\}} \left( f(\Phi^\tau_s(x)) - f(\Phi^\tau_s(x)) \right) \partial_\tau \Phi^\tau_s(x) L^1(d\tau),
\]

and

\[
II_\varepsilon(x) := f(\Phi^\tau_s(x)) \left[ \frac{1}{2\varepsilon} \int_{\{t-\varepsilon, t+\varepsilon\}} \partial_\tau \Phi^\tau_s(x) L^1(d\tau) - \partial_\tau \Phi^\tau_s(x) \right],
\]

tend to 0 in \( L^1(e_\tau(G), m) \) as \( \varepsilon \to 0 \).

**Step 4.** To see the claim about \( I_\varepsilon \), since \( |\partial_\tau \Phi^\tau_s(x)| \leq L \) (uniformly in \( \tau \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x) \) and \( x \in e_\tau(G) \)), it is clear that \( \lim_{\varepsilon \to 0} I_\varepsilon(x) = 0 \) pointwise by continuity of \( f \) and \( \hat{G}(x) \ni \tau \mapsto \Phi^\tau_s(x) \) (see Proposition 4.4). To obtain convergence in \( L^1(e_\tau(G), m) \), it is therefore enough to show by Dominated Convergence that

\[
\frac{1}{2\varepsilon} \int_{\{t-\varepsilon, t+\varepsilon\}} \left( f(\Phi^\tau_s(x)) - f(\Phi^\tau_s(x)) \right) L^1(d\tau) \leq C,
\]

(11.8)
uniformly in $x \in e_t(G)$. Since $f$ is uniformly continuous on the compact set $\varphi_s(e_s(G))$, the uniform estimate (11.8) follows since $\hat{G}(x) \ni \tau \mapsto \Phi^T_s(x)$ is Lipschitz on $[\delta, 1 - \delta]$, with Lipschitz constant depending only on $\delta > 0$ and an upper bound on $\{\ell(\gamma) \; ; \; \gamma \in G\}$ (see Proposition 4.4 (3) and Theorem 3.11 (2)).

**Step 5.** To see the claim about $\Pi_\varepsilon$, it is clearly enough to show that

$$\tilde{\Pi}_\varepsilon(x) := \frac{1}{2\varepsilon} \int_{(t-\varepsilon, t+\varepsilon) \cap \hat{G}(x)} \partial_\tau \Phi^T_s(x) L^1(d\tau) - \partial_t \Phi^T_s(x)$$

$$\rightarrow 0 \text{ in } L^1(e_t(G), m). \quad (11.9)$$

**Step 5a.** We first establish (11.9) for $L^1$-a.e. $t \in (0, 1)$ (independently of $f$ and $\Psi$). Since $\partial_\tau \Phi^T_s(x) \leq L$ uniformly in $t \in (t - \varepsilon, t + \varepsilon) \cap \hat{G}(x)$ and $x \in e_t(G)$, by Dominated Convergence, it is enough to establish pointwise convergence in (11.9) for $m$-a.e. $x \in e_t(G)$.

For every $x \in X$, denote

$$Leb(x) := \{t \in \hat{G}(x) \; ; \; t \text{ is a Lebesgue point of } \tau \mapsto \partial_\tau \Phi^T_s(x) 1_{\hat{G}(x)}(\tau)\}.$$ 

By Proposition 4.4 (based on Theorem 3.11), we know that for every $x \in X$, the map $\tau \mapsto \partial_\tau \Phi^T_s(x)$ is in $L^\infty_{loc}(\hat{G}(x))$, and so by Lebesgue’s Differentiation Theorem, $L^1(\hat{G}(x) \setminus Leb(x)) = 0$. Integrating over $m$ and applying Fubini’s Theorem, it follows that for $L^1$-a.e. $t \in (0, 1)$:

$$m(d_t(G) \setminus \{x \in d_t(G) \; ; \; t \text{ is a Lebesgue point of } \tau \mapsto \partial_\tau \Phi^T_s(x) 1_{\hat{G}(x)}(\tau)\}) = 0,$$

thereby establishing (by definition) the pointwise convergence in (11.9) for $m$-a.e. $x \in e_t(G)$.

**Step 5b.** We next establish (11.9) at $t = s$. Write

$$\tilde{\Pi}_\varepsilon(x) = \frac{1}{2\varepsilon} \int_{(s-\varepsilon, s+\varepsilon) \cap \hat{G}(x)} \left( \partial_\tau \Phi^T_s(x) - \ell^2_s(x) \right) L^1(d\tau)$$

$$+ \ell^2_s(x) \left[ \frac{1}{2\varepsilon} \int_{(s-\varepsilon, s+\varepsilon) \cap \hat{G}(x)} L^1(d\tau) - 1 \right].$$

The first expression tends to 0 pointwise for all $x \in X$ by Lemma 4.6, and hence by Dominated Convergence also in $L^1(e_t(G), m)$ (since $|\partial_\tau \Phi^T_s(x)| \leq L$ and $\ell_s(x) \leq 1/c$ uniformly). The second expression tends to 0 in $L^1(e_t(G), m)$ by Proposition 9.7 and the uniform boundedness of $\ell^2_s(x)$. 

\[ \Theta \] Springer
Step 6. In other words, we have verified in Steps 3-5 the following weak convergence, for $L^1$-a.e. $t \in (0, 1)$ including at $t = s$:

$$\lim_{\varepsilon \to 0} \int_{e_t(G)} \frac{1}{2\varepsilon} (t_x^2)_{\varepsilon} (L^1 \land \{ \Phi^t_s(x) ; \tau \in (t - \varepsilon, t + \varepsilon) \cap \dot{G}(x) \}) \, m(dx)$$

$$= \int_{e_t(G)} (t_x^2)_{\varepsilon} (\delta \Phi^t_s(x)) \partial_t \Phi^t_s(x) \, m(dx),$$

where recall $\Phi^t_s(x) = \varphi_s(x)$ and $\partial_t \Phi^t_s |_{t=s} = \ell^2_s(x)$. Combining this with Step 1, we deduce that:

$$\int \varphi_s(e_t(G)) (t_s^1)_{\varepsilon} (m_s^{a_s}) L^1(da_s) = \int_{e_t(G)} (t_x^2)_{\varepsilon} (\delta \Phi^t_s(x)) \partial_t \Phi^t_s(x) \, m(dx).$$

Integrating this identity against $1 \otimes \psi$ with $1 \in C_b(\mathbb{R})$ and $\psi \in C_b(X)$, we obtain:

$$\int \varphi_s(e_t(G)) \int_{e_t(G)} \psi(x) m_s^{a_s}(dx) L^1(da_s) = \int_{e_t(G)} \psi(x) \partial_t \Phi^t_s(x) \, m(dx)$$

$$= \int \varphi_s(e_t(G)) \int_{e_t(G)} \psi(x) \partial_t \Phi^t_s(x) m_s^{a_s}(dx) L^1(da_s),$$

where we used that $m_s^{a_s}$ is concentrated on $e_t(G)_{a_s} \subset e_t(G)$ for all $t \in (0, 1)$ and $a_s \in \varphi_s(e_t(G))$ in the first expression, and the disintegration (11.1) of $m_{e_t(G)}$ in the last transition. In other words, we obtained for $L^1$-a.e. $t \in (0, 1)$ including at $t = s$:

$$\int \varphi_s(e_t(G)) m_s^{a_s} L^1(da_s) = \int \varphi_s(e_t(G)) \partial_t \Phi^t_s m_s^{a_s} L^1(da_s).$$

Since $m_s^{a_s}$ is also concentrated on $e_t(G)_{a_s}$ for all $t \in (0, 1)$ and $L^1$-a.e. $a_s \in \varphi_s(e_t(G))$, the assertion follows by essential uniqueness of consistent disintegrations (Theorem 6.19). Note that by Step 2, $\partial_t \Phi^t_s(x)$ exists and is positive for $L^1$-a.e. $t \in (0, 1)$ including at $t = s$ for $m$-a.e. $x \in e_t(G)$, and so by (11.1), the same holds for $L^1$-a.e. $a_s \in \varphi_s(e_t(G))$ and $m_s^{a_s}$-a.e. $x$.

11.2 Change-of-variables formula

We now obtain the following main result of Sects. 10 and 11. At this time, we dispense of Assumption 10.2.

**Theorem 11.4** (Change-of-Variables) Let $(X, d, m)$ be an essentially non-branching m.m.s. verifying $CD^1(K, N)$ with supp$(m) = X$, and let $\mu_0, \mu_1 \in$
\(\mathcal{P}_2(X, d, m)\). Let \(v\) denote the unique element of \(\text{OptGeo}(\mu_0, \mu_1)\), and set \(\mu_t := (e_t)_\# v \ll m\) for all \(t \in (0, 1)\).

Then there exist versions of the densities \(\rho_t := d\mu_t/dm, t \in [0, 1]\), so that for \(v\text{-a.e. } \gamma \in \text{Geo}(X), (9.4)\) holds for all \(0 \leq s \leq t \leq 1\), and in particular, for \(v\text{-a.e. } \gamma, t \mapsto \rho_t(\gamma_t)\) is positive and locally Lipschitz on \((0, 1)\), and upper semi-continuous at \(t = 0, 1\).

Moreover, for any \(s \in (0, 1)\), for \(\mathcal{L}^1\text{-a.e. } t \in (0, 1)\) and \(v\text{-a.e. } \gamma \in G^\gamma, \partial_t|_{t=s}\Phi^\gamma_s(\gamma_t)\) exists, is positive, and the following change-of-variables formula holds:

\[
\frac{\rho_t(\gamma_t)}{\rho_s(\gamma_s)} = \frac{\partial_t|_{t=s}\Phi^\gamma_s(\gamma_t)}{\ell^2(\gamma)} \cdot \frac{1}{h_{\gamma_s}^{\Phi_s}(t)}.
\]  

(11.10)

Here \(\phi\) denotes a Kantorovich potential associated to the c-optimal-transport problem between \(\mu_0\) and \(\mu_1\) with cost \(c = d^2/2\), and \(\Phi^\gamma_t\) denotes the time-propagated intermediate Kantorovich potential introduced in Sect. 4; \(h_{\gamma_s}^{\Phi_s}(s)\) is the \(\text{CD}(\ell(\gamma)^2, K, N)\) density on \([0, 1]\) from Proposition 10.4, after applying the re-normalization from Remark 10.5, so that \(h_{\gamma_s}^{\Phi_s}(s) = 1\). In particular, for \(v\text{-a.e. } \gamma \in G^\gamma, \) the above change-of-variables formula holds for \(\mathcal{L}^1\text{-a.e. } t, s \in (0, 1)\).

Lastly, for all \(\gamma \in G^\gamma, \) we have

\[
\rho_t(\gamma_t) = \rho_s(\gamma_s) \quad \forall t, s \in [0, 1].
\]

(11.11)

Recall that \(v\) is concentrated on \(G^\gamma = G^\gamma_\phi \cup G^0_\phi\), where \(G^\gamma_\phi\) and \(G^0_\phi\) denote the subsets of positive and zero length \(\phi\text{-Kantorovich geodesics}, \) respectively. Note that \(\partial_t|_{t=s}\Phi^\gamma_s(\gamma_s) = \ell^2_s(\gamma_s) = \ell^2(\gamma)\) by Proposition 4.4, so that together with our normalization that \(h_{\gamma_s}^{\Phi_s}(s) = 1\), we see that both sides of (11.10) are indeed equal to 1 for \(t = s\).

**Proof of Theorem 11.4**

**Step 0.** As usual, by Proposition 8.9 and Remark 8.11, \((X, d, m)\) also verifies \(\text{MCP}(K, N)\), and so Theorem 6.15 and all the results of Sect. 9 apply. We will use the versions of the densities given by Corollary 9.5. On \(X^0 = e_{[0, 1]}(G^0_\phi), \) we know by Corollary 9.8 that \(\mu_0 \ll \mu_0 \ll X^0 = \mu_0 \ll X^0 = \mu_0 \ll X^0\) for all \(t \in [0, 1]\), and so if necessary, we simply redefine \(\rho_t|_{X^0} := \rho_0|_{X^0} \) for all \(t \in (0, 1)\), so that (11.11) holds. Note that by Lemma 3.15, this will not affect \((0, 1) \ni t \mapsto \rho_t(\gamma_t)\) for all \(\gamma \in G^\gamma_\phi\), and Corollary 9.8 (applied to the pair \(\mu_1, \mu_0\)) ensures that the same is true for \(v\text{-a.e. } \gamma \in G^\gamma_\phi\) at \(t = 1\).

**Step 1.** As explained in the beginning of Sect. 10, by inner regularity of Radon measures, Corollary 9.5 (applied to both pairs \(\mu_0, \mu_1\) and \(\mu_1, \mu_0\)), Proposition 9.7 and Corollary 6.16, there exists a *good* compact subset \(G^\varepsilon \subset G^\gamma_\phi\) with \(v(G^\varepsilon) \geq v(G^\gamma_\phi) - \varepsilon\) for any \(\varepsilon > 0\) (recall Definition 10.1). Of
course, we may assume that \( G^\varepsilon \) is increasing as \( \varepsilon \) decreases to 0 (say, along a fixed sequence). Fixing \( \varepsilon > 0 \) and a good \( G^\varepsilon \), denote \( \nu^\varepsilon = \frac{1}{v(G^\varepsilon)} \nu \ll G^\varepsilon \) and \( \mu^\varepsilon_t := (e_t)_\sharp \nu^\varepsilon \ll \mathcal{M} \), so that all of the results of Sects. 10 and 11.1 apply to \( \nu^\varepsilon \). Note that by Corollary 6.16, we have that

\[
\mu^\varepsilon_t = \frac{1}{v(G^\varepsilon)} (\mu^\varepsilon_t)_{\ll t} \quad \forall t \in [0,1],
\]

and therefore

\[
\mu^\varepsilon_t = \rho^\varepsilon_t \mathcal{M}, \quad \rho^\varepsilon_t := \frac{1}{v(G^\varepsilon)} \rho_t |_{e_t(G^\varepsilon)} \quad \forall t \in [0,1].
\]

Also note that as \( \nu \) is concentrated on \( G^\varepsilon \subset G_\varphi \), \( \varphi \) is still a Kantorovich potential for the associated transport-problem.

**Step 2.** Recall that by Corollary 10.10 (3), for each \( s \in (0,1) \) and \( q^\varepsilon_s \)-a.e. \( a_s \in \varphi_s(e_s(G^\varepsilon)) \), the map

\[
[0,1] \ni t \mapsto \rho^\varepsilon_t \cdot m^\varepsilon_{a_s}
\]

coincides for \( \mathcal{L}^1 \)-a.e. \( t \in [0,1] \) with the geodesic \( t \mapsto (e_t)_\sharp \nu^\varepsilon_{a_s} \) up to a (positive) constant \( C^\varepsilon_{a_s} \) depending on \( a_s \), where \( \nu^\varepsilon_{a_s} \) is the conditional measure from the disintegration in (10.11). Consequently, for such \( s \) and \( a_s \), for \( \mathcal{L}^1 \)-a.e. \( t \in [0,1] \) and any Borel \( H \subset G^\varepsilon_{a_s} \), the quantity

\[
\int_{e_t(H)} \rho^\varepsilon_t(x)m^\varepsilon_{a_s}(dx) = C^\varepsilon_{a_s} \int_{e_t(H)} (e_t)_\sharp \nu^\varepsilon_{a_s}(dx) = C^\varepsilon_{a_s} \nu^\varepsilon_{a_s}(H)
\]

is constant (where we used the fact that \( e_t|_{G^\varepsilon} : G^\varepsilon \to X \) is injective).

By Theorem 11.3, for \( \mathcal{L}^1 \)-a.e. \( t \in (0,1) \) and \( \mathcal{L}^1 \)-a.e. \( a_s \in \varphi_s(G^\varepsilon_s) \) (and hence for \( q^\varepsilon_{a_s} \)-a.e. \( a_s \in \varphi_s(G^\varepsilon_s) \) by Proposition 10.8), \( \partial_t \Phi^t_s(x) \) exists and is positive for \( m^\varepsilon_{a_s} \)-a.e. \( x \), and \( m^\varepsilon_{a_s} = \partial_t \Phi^t_s \cdot m^\varepsilon_{a_s} \). It follows that for those \( t \) and \( a_s \) for which this representation and (11.12) hold true:

\[
C^\varepsilon_{a_s} \nu^\varepsilon_{a_s}(H) = \int_{e_t(H)} \rho^\varepsilon_t(x)m^\varepsilon_{a_s}(dx) = \int_{e_t(H)} \frac{\rho^\varepsilon_t(x)}{\partial_t \Phi^t_s(x)} m^\varepsilon_{a_s}(dx) \quad (11.13)
\]

\[
= \int_{e_t(H)} \frac{\rho^\varepsilon_t(g^\varepsilon_{a_s}(\beta, t))}{\partial_t |_{\tau=t} \Phi^t_s(g^\varepsilon_{a_s}(\beta, t))} h^\varepsilon_{a_s}(t)m^\varepsilon_{a_s}(d\beta)
\]

\[
= \int_{e_t(H)} \frac{\rho^\varepsilon_t(g^\varepsilon_{a_s}(\beta, t))}{\partial_t |_{\tau=t} \Phi^t_s(g^\varepsilon_{a_s}(\beta, t))} h^\varepsilon_{a_s}(t)\ell^2_s(\beta)m^\varepsilon_{a_s}(d\beta),
\]

where the second transition follows from our normalization and Remark 10.5, ensuring that \( m^\varepsilon_{a_s} = (g^\varepsilon_{a_s}(\cdot, t))_\sharp (h^\varepsilon_{a_s}(t)m^\varepsilon_{a_s}) \), and the last transition follows from Theorem 11.3.
Note that \( g \) and \( h \) above do not depend on \( \varepsilon > 0 \). For \( g \), this follows by its very definition as \( g^{\alpha_s}(\beta, t) = e_t(e_s^{-1}(\beta)) \) (and the injectivity of \( e_s|G_\varepsilon \) for all \( \varepsilon > 0 \)). For \( h \), this immediately follows by inspecting the proof of Proposition 10.4, where \( h^{\alpha_s}_{\gamma_s}(t) \) was uniquely defined (for \( t \in (0, 1) \)) as the continuous version of the density of \( \hat{\gamma} \) and the continuity of \( m \) simultaneously for a countable sequence \( \alpha \).

**Step 3.** As the left-hand-side of (11.13) does not depend on \( \tau \), it follows that for all \( s \in (0, 1) \) and for \( \delta^{e,s}_s \)-a.e. \( \alpha_s \in \varphi_s(e_s(G_\varepsilon)) \) (both of which we fix for the time being), there exists a subset \( T \subset (0, 1) \) of full \( L^1 \) measure, so that for all \( H \subset G^{\alpha_s}_s \):

\[
T \ni t \mapsto \int_{e_s(H)} \frac{\rho^e_t(g^{\alpha_s}(\beta, t))}{\partial_t|_{\tau=t} \Phi^s_\gamma(g^{\alpha_s}(\beta, t))} h^{\alpha_s}_{\gamma_s}(t) \mu^{e,s}_s(d\beta)
\]

is constant. As any Borel subset of \( e_s(G^{\alpha_s}_s) \) may be written as \( e_s(H) \), equality of measures follows, and hence equality of densities for \( m^{e,s}_s \)-a.e. \( \beta \). We have therefore proved that for \( t, t' \in T \):

\[
\rho^e_t(\gamma_{t'})(\partial_t|_{\tau=t'} \Phi^s_{\gamma_{t'}}(\gamma_{t'}))^{-1} h^{\alpha_s}_{\gamma_{t'}}(t') = \rho^e_t(\gamma_{t})(\partial_t|_{\tau=t} \Phi^s_{\gamma_{t}}(\gamma_{t}))^{-1} h^{\alpha_s}_{\gamma_{s}}(t),
\]

(11.14)

for \( m^{e,s}_s \)-a.e. \( \beta \in e_s(G^{\varepsilon}_{\alpha_s}) \), where \( \gamma = \gamma^{\beta} = e_s^{-1}(\beta) = g^{\alpha_s}(\beta, \cdot) \in G^{\alpha_s}_s \), with the exceptional set depending on \( t, t' \). Note that given \( t' \in T \), \( \partial_t|_{\tau=t'} \Phi^s_{\gamma_{t'}}(\gamma_{t'}) \) indeed exists for \( m^{e,s}_s \)-a.e. \( \beta \in e_s(G^{\varepsilon}_{\alpha_s}) \) by Corollary 10.10 (2).

It follows that for all \( t \in T \), for \( m^{e,s}_s \)-a.e. \( \beta \in e_s(G^{\varepsilon}_{\alpha_s}) \), (11.14) holds simultaneously for a countable sequence \( t' \in T' \subset T \) which is dense in \((0, 1)\). Taking the limit in (11.14) as \( T' \ni t' \to s \), using Proposition 4.4 (5) which entails:

\[
\lim_{T' \ni t' \to s} \partial_t|_{\tau=t'} \Phi^s_{\gamma_{t'}}(\gamma_{t'}) = \ell_s(\gamma_{t'}^{\beta})^2 = \ell(\gamma^{\beta})^2,
\]

employing the continuity of \((0, 1) \ni t' \mapsto h^{\alpha_s}_{\gamma_{t'}}(t')\), our normalization \( h^{\alpha_s}_{\gamma_s}(s) = 1 \), and the continuity of \((0, 1) \ni t' \mapsto \rho^e_t(\gamma_{t'})\) (as \( G^\varepsilon \) is good), it follows that for all \( s \in (0, 1) \), for \( \delta^{e,s}_s \)-a.e. \( \alpha_s \in \varphi_s(e_s(G^\varepsilon)) \) and \( L^1 \)-a.e. \( t \in (0, 1) \):

\[
\rho^e_s(\gamma_s)\ell(\gamma)^{-2} = \rho^e_t(\gamma_{t})(\partial_t|_{\tau=t} \Phi^s_{\gamma}(\gamma_{t}))^{-1} h^{\alpha_s}_{\gamma_{s}}(t)
\]

(11.15)

for \( m^{e,s}_s \)-a.e. \( \beta \in e_s(G^{\varepsilon}_{\alpha_s}) \), with \( \gamma = e_s^{-1}(\beta) \in G^{\varepsilon}_{\alpha_s} \).
Step 4. Recall that by Corollary 10.10 (2), \( m^{ε,s}_a \) and \((e_ε)_t \nu_{a_s}^ε \) are mutually absolutely continuous for \( q^{ε,s}_s \)-a.e. \( a_s \in \phi_s(e_s(G^ε)) \). It follows that for all \( s \in (0, 1) \), for \( q^{ε,s}_s \)-a.e. \( a_s \in \phi_s(e_s(G^ε)) \) and \( L^1 \)-a.e. \( t \in (0, 1) \), \( 11.15 \) holds for \( \nu_{a_s} \)-a.e. \( \gamma \). By Corollary 10.10 (1), note that \( q^{ε,s}_s \) and \( q^{ε,v}_s \) are mutually absolutely continuous, and hence the disintegration formula \( (10.11) \) implies that for all \( s \in (0, 1) \) and \( L^1 \)-a.e. \( t \in (0, 1) \):

\[
\rho^ε_s(\gamma_s)\ell(\gamma)^{-2} = \rho^ε_t(\gamma_t)(\partial_{\tau}\mid_{\tau=t}\Phi^ε_s(\gamma_t))^{-1}h^ε_s(\gamma_s)(t),
\]

for \( \nu \)-a.e. \( \gamma \in G^ε \), and in particular that \( \partial_{\tau}\mid_{\tau=t}\Phi^ε_s(\gamma_t) \) exists and is positive for those \( s, t \) and \( \gamma \). Taking the limit as \( \varepsilon \to 0 \) along a countable sequence, it follows for all \( s \in (0, 1) \), \( L^1 \)-a.e. \( t \in (0, 1) \) and \( \nu \)-a.e. \( \gamma \in G^+_{\phi} \), that

\[
\rho_s(\gamma_s)\ell(\gamma)^{-2} = \rho_t(\gamma_t)(\partial_{\tau}\mid_{\tau=t}\Phi^ε_s(\gamma_t))^{-1}h^ε_s(\gamma_s)(t),
\]

thereby concluding the proof of \( 11.10 \). As a consequence, an application of Fubini’s Theorem verifies that for \( \nu \)-a.e. \( \gamma \in G^+_{\phi} \), \( 11.10 \) holds for \( L^1 \)-a.e. \( s, t \in (0, 1) \).

\[\square\]

Remark 11.5 Observe that all of the results of this section also equally hold for \( \Phi^ε_s \) in place of \( \Phi^ε_s \). Indeed, recall that for all \( x \in X \), \( \Phi^ε_s(x) = \Phi^ε(x) \) for \( t \in G^s(x) \), and that by Corollary 4.5, \( \partial_t\Phi^ε_s(x) = \partial_t\Phi^ε(x) \) for a.e. \( t \in G^s(x) \). As these were the only two properties used in the above derivation (in particular, in Step 2 of the proof of Theorem 11.3), the assertion follows.

Part III Putting it all together

12 Combining change-of-variables formula with Kantorovich 3rd order information

Let \( (X, d, m) \) denote an essentially non-branching m.m.s. verifying \( \text{CD}^1(K, N) \). Let \( \mu_0, \mu_1 \in P_2(X, d, m) \), and let \( \nu \) be the unique element of \( \text{OptGeo}(\mu_0, \mu_1) \) (by Proposition 8.9, Remark 8.11 and Theorem 6.15). Recall that \( \mu_t := (e_t)_\sharp \nu \ll m \) for all \( t \in [0, 1] \), and we subsequently denote by \( \rho_t \) the versions of the corresponding densities given by Theorem 11.4 (resulting from Corollary 9.5). Finally, denote by \( \varphi \) a Kantorovich potential associated to the corresponding optimal transference plan, so that \( \nu(G^\varphi) = 1 \).

12.1 Change-of-variables rigidity

Recall that by the Change-of-Variables Theorem 11.4, we know that for \( \nu \)-a.e. geodesic \( \gamma \in G^+_{\phi} \) and for a.e. \( t, s \in (0, 1) \), \( \partial_{\tau}\mid_{\tau=t}\Phi^s(\gamma_t) \) exists, is positive,
and it holds that:

$$\rho_s(\gamma_s) \rho_t(\gamma_t) = \frac{h_{\gamma_s}(\gamma_s)(t)}{\partial_{\tau}\mid_{\tau=t_{\Phi_1}}(\gamma_t)/\ell(\gamma)^2}. \quad (12.1)$$

In fact, by Remark 11.5, the same also holds with $\bar{\Phi}$ in place of $\Phi$, so that in particular:

$$\partial_{\tau}\mid_{\tau=t_{\Phi_1}}(\gamma_t) = \partial_{\tau}\mid_{\tau=t_{\bar{\Phi}_1}}(\gamma_t) \quad \text{for } \nu\text{-a.e. } \gamma \in G^+_{\rho} \quad \text{for a.e. } t, s \in (0, 1). \quad (12.2)$$

Recall that given $t, s \in (0, 1)$, $\Phi = \Phi, \bar{\Phi}$ and $\bar{\ell} = \ell, \bar{\ell}$, respectively, $\Phi_s^t$ was defined on $D_{\bar{\ell}}$ as:

$$\Phi_s^t = \bar{\Phi}_t + (t-s)\bar{\ell}^2/2,$$

and that by Proposition 4.4 (2), the differentiability points of $t \mapsto \Phi_s^t(x)$ and $t \mapsto \bar{\ell}_i^2(x)$ coincide for all $t \neq s$, and at those points:

$$\partial_t \Phi_s^t(x) = \bar{\ell}_i^2(x) + (t-s)\partial_t \bar{\ell}_i^2(x). \quad (12.3)$$

It follows from (12.2) that for $\nu$-a.e. geodesic $\gamma \in G^+_{\rho}$ and for a.e. $t \in (0, 1)$:

$$\exists \partial_{\tau}\mid_{\tau=t_{\Phi_1}}(\gamma_t) = \exists \partial_{\tau}\mid_{\tau=t_{\bar{\Phi}_1}}(\gamma_t), \quad \partial_{\tau}\mid_{\tau=t_{\Phi_1}}(\gamma_t) = \partial_{\tau}\mid_{\tau=t_{\bar{\Phi}_1}}(\gamma_t). \quad (12.4)$$

Alternatively, (12.4) follows directly by Lemma 5.6, in fact for $\nu$-a.e. $\gamma$ (not just $\gamma \in G^+_{\rho}$).

Plugging (12.3) and (12.4) into (12.1), it follows that we may express the Change-of-Variables Theorem 11.4 as the statement that for $\nu$-a.e. geodesic $\gamma \in G^+_{\rho}$, we have:

$$\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{h_{\gamma_s}(\gamma_s)(t)}{1 + (t-s)\partial_{\tau}\mid_{\tau=t_{\Phi_1}}(\gamma_t)/\ell(\gamma)^2} = \frac{h_{\gamma_s}(\gamma_s)(t)}{1 + (t-s)\partial_{\tau}\mid_{\tau=t_{\bar{\Phi}_1}}(\gamma_t)/\ell(\gamma)^2} \quad \text{for a.e. } t, s \in (0, 1). \quad (12.5)$$

Note that the denominators on the right-hand-side of (12.5) are always positive (when defined) for all $t, s \in (0, 1)$ by Theorem 3.11 (3). Fixing the geodesic $\gamma,$
we denote for brevity $\rho(t) := \rho_t(\gamma_t), h_s(t) := h_{\gamma_s}(\gamma_s)(t)$ and $K_0 := K \cdot \ell(\gamma)^2$. We then have the following additional information for $\nu$-a.e. $\gamma \in G^+_\psi$, by Corollary 9.5 and Proposition 10.4, respectively:

(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 $\text{CD}(K_0, N)$ density on $[0, 1]$, satisfying $h_s(s) = 1$. In particular, it is locally Lipschitz continuous on $(0, 1)$ and strictly positive there.

**Remark 12.1** It is in fact possible to deduce (A) just from the Change-of-Variables formula (12.5) and without referring to Corollary 9.5. This may be achieved by a careful bootstrap argument, exploiting the separation of variables on the left-hand-side of (12.5) and the a-priori estimates of Lemma A.9 in the “Appendix” on the logarithmic derivative of $\text{CD}(K_0, N)$ densities. But since we already know (A), and since (A) was actually (mildly) used in the proof of the Change-of-Variables Theorem 11.4, we only mention this possibility in passing. Note that Corollary 9.5 applies to all $\text{MCP}(K, N)$ essentially non-branching spaces, whereas the Change-of-Variables formula requires knowing the stronger $\text{CD}^1(K, N)$ condition.

Fix a geodesic $\gamma \in G^+_\psi$ satisfying (12.5), (A) and (B) above. Let $I \subset (0, 1)$ be the set of full measure where $(12.5)$ holds for all $s \in I$. It follows from (12.5) that for all $s \in I$, $t \mapsto \frac{\partial_t \big|_{\tau=t} \ell^2_\tau / 2(\gamma_t)}{\ell(\gamma)^2}$ coincide a.e. on $(0, 1)$ for both $\tilde{\ell} = \ell, \hat{\ell}$ with the same locally Lipschitz function $t \mapsto z_s(t)$ defined on $(0, 1) \setminus \{s\}$:

$$z_s(t) := \frac{1}{\rho_s(\gamma_s)} h_{\gamma_s}(\gamma_s) \rho_t(\gamma_t) - 1}{t - s}.$$  

By continuity, it follows that the functions $\{z_s\}_{s \in I}$ must all coincide on their entire domain of definition with a single function $t \mapsto z(t)$ defined on $(0, 1)$; the latter function must therefore be locally Lipschitz continuous, and satisfy

$$z(t) = \frac{\partial_t \big|_{\tau=t} \ell^2_\tau / 2(\gamma_t)}{\ell(\gamma)^2} = \frac{\partial_t \big|_{\tau=t} \tilde{\ell}^2_\tau / 2(\gamma_t)}{\ell(\gamma)^2} \text{ for a.e. } t \in (0, 1).$$  

By Theorem 5.5, which provides us with 3rd order information on intermediate-time Kantorovich potentials, we obtain the following additional information on $z$:

(C) $(0, 1) \ni t \mapsto z(t)$ is locally Lipschitz.

For any $\delta \in (0, 1/2)$, there exists $C_\delta > 0$ so that
\[
\frac{z(t) - z(s)}{t - s} \geq (1 - C\delta(t - s)) |z(s)||z(t)|
\]
\[
\forall 0 < \delta \leq s < t \leq 1 - \delta < 1.
\]

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

**Remark 12.2** By Theorem 5.5, we obtain the following interpretation for \(z(t)\)—it 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 \(\tau = t\). In particular, these second Peano derivatives are guaranteed to exist for all \(t \in (0, 1)\) and are a continuous function thereof.

We have already seen above how (12.5) enabled us to deduce (12.6), thereby gaining (by Theorem 5.5) an additional order of regularity for \(\partial_\tau|_{\tau = t} \rho^2/2(\gamma(t))\). The purpose of this section is to show that the combination of the Change-of-Variables Formula:

\[
\frac{\rho(s)}{\rho(t)} = \frac{h_s(t)}{1 + (t - s)z(t)} \quad \text{for a.e. } t, s \in (0, 1), \tag{12.7}
\]

together with properties (A), (B) and (C) above, forms a very rigid condition, and already implies the following representation for \(\frac{1}{\rho(t)}\); we formulate this independently of the preceding discussion as follows:

**Theorem 12.3** (Change-of-Variables Rigidity) Assume that (12.7) holds, where \(\rho, \{h_s\}\) and \(z\) satisfy (A), (B) and (C) above. Then

\[
\frac{1}{\rho(t)} = L(t)Y(t) \quad \forall t \in (0, 1),
\]

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

**12.2 Formal argument**

To better motivate the ensuing proof of Theorem 12.3, we begin with a formal argument.

Assume that the functions \(\rho\) and \(z\) are \(C^2\) smooth and that equality holds in (12.7) for all \(t, s \in (0, 1)\). It follows that the mapping \((s, t) \mapsto h_s(t)\) is also \(C^2\) smooth. Fix any \(r_0 \in (0, 1)\), and define the functions \(L\) and \(Y\) by

\[
\log L(r) := -\int_{r_0}^r z(s)ds, \quad \log Y(r) := \int_{r_0}^r \partial_t|_{t=s} \log h_s(t)ds.
\]
Note that by (12.7):

\[
\log \frac{\rho(r_0)}{\rho(r)} = \int_{r_0}^r \partial_t |_{t=s} \log \frac{\rho(s)}{\rho(t)} \, ds
\]

\[
= \int_{r_0}^r \partial_t |_{t=s} \log h_s(t) \, ds - \int_{r_0}^r \partial_t |_{t=s} \log(1 + (t - s)z(t)) \, ds
\]

\[
= \log Y(r) + \log L(r).
\]

As already noted in Lemma 5.7, the concavity of \( L \) follows from (C), since

\[
\frac{L''}{L} = (\log L)'' + ((\log L)')^2 = -z' + z^2 \leq 0.
\]

The more interesting function is \( Y \). We have for all \( r \in (0, 1) \):

\[
(\log Y)'(r) = \partial_t |_{t=r} \log h_r(t),
\]

\[
(\log Y)''(r) = \partial_t^2 |_{t=r} \log h_r(t) + \partial_s \partial_t |_{t=s=r} \log h_s(t).
\]

To handle the last term on right-hand-side above, note that by the separation of variables on the left-hand-side of (12.7), we have by (C) again, after taking logarithms and calculating the partial derivatives in \( t \) and \( s \):

\[
\partial_s \partial_t |_{t=s=r} \log h_s(t) = \partial_s \partial_t |_{t=s=r} \log(1 + (t - s)z(t))
\]

\[
= -z'(r) + z^2(r) \leq 0.
\]

(12.8)

We therefore conclude that for all \( r \in (0, 1) \):

\[
(\log Y)''(r) + \frac{(\log Y)'(r))^2}{N - 1} \leq \partial_t^2 |_{t=r} \log h_r(t) + \frac{(\partial_t |_{t=r} \log h_r(t))^2}{N - 1} \leq -K_0,
\]

where the last inequality follows from (B) and the differential characterization of \( \text{CD}(K_0, N) \) densities (applied to \( h_r(t) \) at \( t = r \)). Applying the characterization again, we deduce that \( Y \) is a \( (C^2\text{-smooth}) \) \( \text{CD}(K_0, N) \) density on \( (0, 1) \). This concludes the formal proof that

\[
\frac{\rho(r_0)}{\rho(r)} = L(r)Y(r) \quad \forall r \in (0, 1),
\]

with \( L \) and \( Y \) satisfying the desired properties. In a sense, the latter argument has been tailored to “reverse-engineer” the smooth Riemannian argument, where the separation to orthogonal and tangential components of the Jacobian.
is already encoded in the Jacobi equation, (B) is a consequence of the corresponding Riccati equation, and (C) is a consequence of Cauchy–Schwarz (cf. [74, Proof of Theorem 1.7]).

12.3 Rigorous argument

It is surprisingly very tedious to modify the above formal argument into a rigorous one. It seems that an approximation argument cannot be avoided, since the definition of \( Y \) above is inherently differential, and so on one hand we do not know how to check the \( \text{CD}(K_0, N) \) condition for \( Y \) synthetically, but on the other hand \( Y \) is not even differentiable, so it is not clear how to check the \( \text{CD}(K_0, N) \) condition by taking derivatives. The main difficulty in applying an approximation argument here stems from the fact that we do not know how to approximate \( \{h_s\} \) and \( z \) by smooth functions \( \{h^\varepsilon_s\} \) and \( z^\varepsilon \), so that simultaneously:

- \( \{h^\varepsilon_s\} \) are \( \text{CD}(K_0 - \varepsilon, N) \) densities;
- \( z^\varepsilon \) is a function of \( t \) only, and not of \( s \);
- and the separation of variables structure of (12.7) is preserved.

Our solution is to note that the main role of the separation of variables in the above formal argument was to ensure that (12.8) holds, and so we will replace the rigid third requirement with the following relaxed one:

- \( \partial_s \partial_t|_{t=s=r} \log h^\varepsilon_s(t) \leq B\delta \varepsilon \) for all \( r \in [\delta, 1 - \delta] \) and \( \delta > 0 \).

**Proof of Theorem 12.3**

**Step 1: Redefining \( h_s(t) \).**

First, observe that there exists \( I_y \subset (0, 1) \) of full measure so that for all \( s \in I_y \), (12.7) is satisfied for a.e. \( t \in (0, 1) \), and hence for all \( t \in (0, 1) \), since all the functions \( \rho, \{h_s\} \) and \( z \) are assumed to be continuous on \((0, 1)\). Unfortunately, we cannot extend this to all \( s \in (0, 1) \) as well, since there may be a null set of \( s \)’s for which the densities \( h_s(t) \) do not comply at all with the equation (12.7). To remedy this, we simply force (12.7) to hold for all \( s, t \in (0, 1) \) by defining

\[
\tilde{h}_s(t) := \frac{\rho(s)}{\rho(t)}(1 + (t - s)z(t)) \quad s, t \in (0, 1),
\]

and claim that for all \( s \in (0, 1) \), \( \tilde{h}_s \) is a \( \text{CD}(K_0, N) \) density on \((0, 1)\). Indeed, for \( s \in I_y \), \( \tilde{h}_s = h_s \) and there is nothing to check. If \( s_0 \in (0, 1) \setminus I_y \), simply note that \( \tilde{h}_s(t) \) is locally Lipschitz in \( s \in (0, 1) \) (since \( \rho(s) \) is), and hence

\[
\tilde{h}_{s_0}(t) = \lim_{s \to s_0} \tilde{h}_s(t) = \lim_{I_y \ni s \to s_0} \tilde{h}_s(t) = \lim_{I_y \ni s \to s_0} h_s(t) \quad \forall t \in (0, 1).
\]
But the family of $\text{CD}(K_0, N)$ densities on $(0, 1)$ is clearly closed under pointwise limits (it is characterized by a family of inequalities between 3 points), and so $\tilde{h}_{s_0}$ is a $\text{CD}(K_0, N)$ density, as asserted.

**Step 2: Properties of $z$ and $\{\tilde{h}_s\}$.**

We next collect several additional observations regarding the functions $z$ and $\{\tilde{h}_s\}$. Recall that $\rho$ (by assumption) and $\tilde{h}_s$ (as $\text{CD}(K_0, N)$ densities) are strictly positive in $(0, 1)$. Together with (12.9) (or directly from (12.7)), this implies that $1 + (t - s)z(t) > 0$ for all $t, s \in (0, 1)$, and hence

\[(D) \quad -\frac{1}{t} \leq z(t) \leq \frac{1}{1-t} \quad \forall t \in (0, 1).\]

In fact, we already knew this by Theorem 3.11 (3) but refrained from including this into our assumption (C) since this is a consequence of the other assumptions. Furthermore

\[(E) \quad I_x := \{t \in (0, 1) ; \tau \mapsto \tilde{h}_s(\tau) \text{ is differentiable at } \tau = t \text{ for all } s \in (0, 1)\} \text{ is of full measure.}\]

Indeed, this follows directly from the definition (12.9) by considering the set all points $t$ where $\rho(t)$ and $z(t)$ are differentiable. In addition, we clearly have:

\[(F) \quad \forall t \in I_x, (0, 1) \ni s \mapsto \partial_t \tilde{h}_s(t) \text{ is continuous.}\]

**Step 3: Defining $L$ and $Y$.**

Now fix $r_0 \in (0, 1)$, and define the functions $L, Y$ on $(0, 1)$ as follows:

\[\log L(r) := -\int_{r_0}^r z(s)ds, \quad \log Y(r) := \int_{r_0}^r \partial_t |_{t=s} \log \tilde{h}_s(t)ds.\]

Clearly, the function $L$ is well defined for all $r \in (0, 1)$ as $z$ is assumed locally Lipschitz. As for the function $Y$, (E) implies that $\partial_t |_{t=s} \log \tilde{h}_s(t)$ exists for a.e. $s \in (0, 1)$, and the fact that the latter integrand is locally integrable on $(0, 1)$ is a consequence of Lemma A.9 in the “Appendix”, which guarantees a-priori locally-integrable estimates on the logarithmic derivative of $\text{CD}(K_0, N)$ densities.

Consequently, as in our formal argument, we may write (since $\log \rho$ is locally absolutely continuous on $(0, 1)$):

\[
\log \frac{\rho(r_0)}{\rho(r)} = \int_{r_0}^r \partial_t |_{t=s} \log \frac{\rho(s)}{\rho(t)}ds \\
= \int_{r_0}^r \partial_t |_{t=s} \log \tilde{h}_s(t)ds - \int_{r_0}^r \partial_t |_{t=s} \log (1 + (t - s)z(t))ds \\
= \log Y(r) + \log L(r),
\]
and hence
\[ \frac{\rho(r_0)}{\rho(r)} = L(r)Y(r) \quad \forall r \in (0, 1). \]

We have already verified in Lemma 5.7 that the property \( z'(s) \geq z^2(s) \) a.e. in \( s \in (0, 1) \) implies that \( L \) is concave on \( (0, 1) \), so it remains to show that \( Y \) is a \( \mathbb{CD}(K_0, N) \) density on \( (0, 1) \).

**Step 4: Approximation argument**

We now arrive to our approximation argument. Given \( \varepsilon_1, \varepsilon_2 > 0 \), \( t \in (\varepsilon_1, 1 - \varepsilon_1) \) and \( s \in (\varepsilon_2, 1 - \varepsilon_2) \), define the double logarithmic mollification of \( \tilde{h}_s(t) \) by
\[
\log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) := \int \int \log \tilde{h}_y(x) \psi_{\varepsilon_1}(t - x) \psi_{\varepsilon_2}(s - y) dx dy,
\]

where \( \psi_{\varepsilon}(x) = \frac{1}{\varepsilon} \psi(x/\varepsilon) \) and \( \psi \) is a \( C^2 \)-smooth non-negative function on \( \mathbb{R} \) supported on \([-1, 1] \) and integrating to 1. Since for all \( \eta \in (0, 1/2) \), we clearly have by (12.9) (and, say, (D))
\[
\int_{1-\eta}^{1-\varepsilon_1} \int_{1-\eta}^{1-\varepsilon_2} \left| \log \tilde{h}_y(x) \right| dx dy < \infty,
\]

it follows by Proposition A.12 in the “Appendix” on logarithmic convolutions that \( \{\tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t)\}_{s \in (\varepsilon_2, 1 - \varepsilon_2)} \) is a \( C^2 \)-smooth \( (t, s) \) family of \( \mathbb{CD}(K_0, N) \) densities on \((\varepsilon_1, 1 - \varepsilon_1)\).

**Step 5: Concluding the proof assuming (H1) and (H2)**

We will subsequently show the following two additional properties of the family \( \{h^{\varepsilon_1, \varepsilon_2}_s(t)\} \):

(H1) \( \lim_{\varepsilon_2 \to 0} \lim_{\varepsilon_1 \to 0} \partial_t |_{t=s} \log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) = \partial_t |_{t=s} \log \tilde{h}_s(t) \) for a.e. \( s \in (0, 1) \).

(H2) \( \forall \delta \in (0, 1/2) \exists C_{\delta} > 0 \forall \varepsilon \in (0, \frac{\delta}{8}] \forall \varepsilon_1, \varepsilon_2 \in (0, \varepsilon] \):
\[
\partial_s \partial_t |_{t=s=r} \log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) \leq 2C_{\delta} \varepsilon \quad \forall r \in [\delta, 1 - \delta].
\]

Assuming these additional properties, let us show how to conclude the proof of Theorem 12.3. Set \( \varepsilon = \max(\varepsilon_1, \varepsilon_2) \), and assuming that \( \varepsilon < \min(r_0, 1 - r_0) \), define the function \( Y^{\varepsilon_1, \varepsilon_2} \) on \((\varepsilon, 1 - \varepsilon)\) given by
\[
\log Y^{\varepsilon_1, \varepsilon_2}(r) := \int_{r_0}^{r} \partial_t |_{t=s} \log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) ds.
\]
First, we claim to have the following pointwise convergence for all $r \in (0, 1)$:

$$
\lim_{\varepsilon_2 \to 0} \lim_{\varepsilon_1 \to 0} \log Y^{\varepsilon_1, \varepsilon_2}(r) = \lim_{\varepsilon_2 \to 0} \lim_{\varepsilon_1 \to 0} \int_{r_0}^{r} \partial_t |_{t=s} \log \tilde{h}_{s}^{\varepsilon_1, \varepsilon_2}(t) \, ds
$$

$$
= \int_{r_0}^{r} \partial_t |_{t=s} \log \tilde{h}_{s}(t) \, ds = \log Y(r). \quad (12.10)
$$

Indeed, the pointwise convergence of the integrands is ensured by property (H1), and as soon as $r_0, r \in (\eta, 1-\eta)$ for some $\eta > 0$, we obtain by the a-priori estimates of Lemma A.9 in the “Appendix” (since $\tilde{h}_{s}^{\varepsilon_1, \varepsilon_2}$ is a CD($K_0, N$) density on $(\eta, 1-\eta)$ for all $\varepsilon_1, \varepsilon_2 \in (0, \eta]$ and $s \in (\eta, 1-\eta)$):

$$
\forall t, s \in [r_0, r] \ \forall \varepsilon_1, \varepsilon_2 \in (0, \eta] \quad \left| \partial_t \log \tilde{h}_{s}^{\varepsilon_1, \varepsilon_2}(t) \right| \leq C(r, r_0, \eta, K_0, N).
$$

Consequently, (12.10) follows by Lebesgue’s Dominated Convergence theorem.

Now $Y^{\varepsilon_1, \varepsilon_2}$ is $C^2$-smooth, and so as in our formal argument, we have for all $r \in (\varepsilon, 1-\varepsilon)$:

$$(\log Y^{\varepsilon_1, \varepsilon_2})'(r) = \partial_t |_{t=r} \log \tilde{h}_{r}^{\varepsilon_1, \varepsilon_2}(t),$$

$$(\log Y^{\varepsilon_1, \varepsilon_2})''(r) = \partial_t^2 |_{t=r} \log \tilde{h}_{r}^{\varepsilon_1, \varepsilon_2}(t) + \partial_s \partial_t |_{t=r} \log \tilde{h}_{s}^{\varepsilon_1, \varepsilon_2}(t).$$

As $\tilde{h}_{r}^{\varepsilon_1, \varepsilon_2}$ is a CD($K_0, N$) density on $(\varepsilon, 1-\varepsilon)$, we know by the differential characterization of such densities that

$$
\partial_t^2 |_{t=r} \log \tilde{h}_{r}^{\varepsilon_1, \varepsilon_2}(t) + \frac{1}{N-1} (\partial_t |_{t=r} \log \tilde{h}_{r}^{\varepsilon_1, \varepsilon_2}(t))^2 \leq -K_0.
$$

Combining this with property (H2), we conclude that for any $\delta \in (0, 1/2)$, whenever $\varepsilon = \max(\varepsilon_1, \varepsilon_2) \in (0, \min(r_0, 1-r_0, \frac{\delta}{8}))$:

$$(\log Y^{\varepsilon_1, \varepsilon_2})''(r) + \frac{1}{N-1} ((\log Y^{\varepsilon_1, \varepsilon_2})'(r))^2 \leq -K_0 + 2C_\delta \varepsilon \ \forall r \in [\delta, 1-\delta],$$

and hence $Y^{\varepsilon_1, \varepsilon_2}$ is a $C^2$-smooth CD($K_0 - 2C_\delta \varepsilon, N$) density on $[\delta, 1-\delta]$.

Combining all of the preceding information, since (as before) the family of CD($K'_0, N$) densities is closed under pointwise limits, we conclude from (12.10) that $Y$ is a CD($K_0 - 2C_\delta \varepsilon, N$) density on $[\delta, 1-\delta]$, for any $\delta \in (0, 1/2)$ and $\varepsilon \in (0, \min(r_0, 1-r_0, \frac{\delta}{8}))$. Taking the limit as $\varepsilon \to 0$ and then as $\delta \to 0$,
we confirm that $Y$ must be a $CD(K_0, N)$ density on $(0, 1)$, concluding the proof.

It remains to establish properties (H1) and (H2).

**Step 6: proof of (H1)**

Given $y \in (0, 1)$ and $t \in (\varepsilon_1, 1 - \varepsilon_1)$, denote

$$\log \tilde{h}^{\varepsilon_1}_y(t) := \int \log \tilde{h}_y(x) \psi\varepsilon_1(t - x)dx$$

so that for every $s \in (\varepsilon_2, 1 - \varepsilon_2)$:

$$\log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) = \int \log \tilde{h}^{\varepsilon_1}_y(t) \psi\varepsilon_2(s - y)dy. \quad (12.11)$$

By Proposition A.10 in the “Appendix”, $\tilde{h}^{\varepsilon_1}_y$ is a $CD(K_0, N)$ density on $(\varepsilon_1, 1 - \varepsilon_1)$ for all $y \in (0, 1)$. Consequently, Lemma A.9 implies that $t \mapsto \log \tilde{h}^{\varepsilon_1}_y(t)$ is locally Lipschitz on $(\varepsilon_1, 1 - \varepsilon_1)$, uniformly in $y \in (0, 1)$:

$$\sup_{y \in (0, 1)} \left| \frac{\partial}{\partial t} \log \tilde{h}^{\varepsilon_1}_y(t) \right| \leq C(t, \varepsilon_1, K_0, N). \quad (12.12)$$

In particular, it follows that we may differentiate in $t$ under the integral in (12.11) at any $t_0 \in (\varepsilon_1, 1 - \varepsilon_1)$:

$$\left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) = \int \left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}^{\varepsilon_1}_y(t) \psi\varepsilon_2(s - y)dy. \quad (12.13)$$

Now, by a standard argument (see Lemma 12.5 at the end of this section), we know that the derivative of an $\varepsilon$-mollification of a Lipschitz function converges to the derivative itself, at all points where the derivative exists, namely:

$$\forall t_0 \in I_x \forall y \in (0, 1) \lim_{\varepsilon_1 \to 0} \left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}^{\varepsilon_1}_y(t) = \left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}_y(t).$$

Together with (12.12) and (12.13), it follows by Dominated Convergence theorem that:

$$\forall t_0 \in I_x \forall s \in (\varepsilon_2, 1 - \varepsilon_2) \lim_{\varepsilon_1 \to 0} \left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}^{\varepsilon_1, \varepsilon_2}_s(t) = \int \left. \frac{\partial}{\partial t} \right|_{t=t_0} \log \tilde{h}_y(t) \psi\varepsilon_2(s - y)dy.$$
But by property (F), we know that \((0, 1) \ni y \mapsto \partial_I|_{t=0} \log \tilde{h}_y(t)\) is continuous for all \(t_0 \in I_x\), and therefore taking the limit as \(\varepsilon_2 \to 0\):

\[
\forall t_0 \in I_x \quad \forall s \in (0, 1) \quad \lim_{\varepsilon_2 \to 0} \lim_{\varepsilon_1 \to 0} \partial_I|_{t=0} \log \tilde{h}_s^{\varepsilon_1, \varepsilon_2}(t) = \partial_I|_{t=0} \log \tilde{h}_s(t).
\]

By property (E), \(I_x\) has full measure, thereby concluding the proof of (an extension of) property (H1).

**Step 7: proof of (H2)**

We will require the following:

**Lemma 12.4** Let \(z\) satisfy (C) and (D). Then for all \(\delta \in (0, 1/2)\), there exists \(C_\delta > 0\), so that for all \(\varepsilon \in (0, \frac{\delta}{4}]\), \(r \in [\delta, 1 - \delta]\), \(r - \varepsilon \leq t_1 < t_2 \leq r + \varepsilon\) and \(r - \varepsilon \leq s_1 < s_2 \leq r + \varepsilon\), we have:

\[
(1 + (t_1 - s_1)z(t_1))(1 + (t_2 - s_2)z(t_2)) \\
\leq (1 + C_\delta \varepsilon (t_2 - t_1)(s_2 - s_1))(1 + (t_2 - s_1)z(t_2))(1 + (t_1 - s_2)z(t_1)).
\]

**Proof** Opening the various brackets, the assertion is equivalent to the statement:

\[
z(t_1)(s_2 - s_1) - z(t_2)(s_2 - s_1) + z(t_1)z(t_2)(t_2 - t_1)(s_2 - s_1) \\
\leq C_\delta \varepsilon (t_2 - t_1)(s_2 - s_1)(1 + (t_2 - s_1)z(t_2))(1 + (t_1 - s_2)z(t_1)),
\]

and after dividing by \((t_2 - t_1)(s_2 - s_1)\), we see that our goal is to establish:

\[
z(t_1)z(t_2) - \frac{z(t_2) - z(t_1)}{t_2 - t_1} \leq C_\delta \varepsilon (1 + (t_2 - s_1)z(t_2))(1 + (t_1 - s_2)z(t_1)),
\]

for an appropriate \(C_\delta\). Note that the right-hand-side of (12.14) is always positive by (D). As \(\min(t_i, 1 - t_i) \geq \delta - \varepsilon \geq \frac{3}{4}\delta\), by our assumption (C), (12.14) would follow from:

\[
|z(t_1)||z(t_2)| B_{\frac{3}{4}\delta} 2\varepsilon \leq C_\delta \varepsilon (1 - 2\varepsilon |z(t_2)|)(1 - 2\varepsilon |z(t_1)|),
\]

or equivalently (assuming \(|z(t_1)||z(t_2)| > 0\), otherwise there is nothing to prove):

\[
2B_{\frac{3}{4}\delta} \leq C_\delta \left(\frac{1}{|z(t_1)|} - 2\varepsilon\right) \left(\frac{1}{|z(t_2)|} - 2\varepsilon\right).
\]

\(\square\) Springer
But \( \frac{1}{|\xi(t_i)|} \geq \min(t_i, 1-t_i) \geq \frac{3}{4} \delta \) by (D), and as \( \varepsilon \in (0, \frac{\delta}{4}] \), we see that (12.15) is ensured by setting:

\[
C_\delta := \frac{32}{\delta^2} B_{\frac{3}{4} \delta}.
\]

Translating the statement of Lemma 12.4 into a statement for \( \tilde{h}_s(t) \) using (12.9), we obtain that for all \( \delta \in (0, 1/2) \), there exists \( C_\delta > 0 \), so that for all \( \varepsilon \in (0, \delta/4] \), \( r \in [\delta, 1-\delta] \), \( r-\varepsilon \leq t, s \leq r+\varepsilon \) and \( \Delta t, \Delta s \in [0, \varepsilon] \), we have:

\[
\log \tilde{h}_s(t) + \log \tilde{h}_s+\Delta s(t + \Delta t) \\
\leq \log \tilde{h}_s(t + \Delta t) + \log \tilde{h}_s+\Delta s(t) + 2C_\delta \varepsilon \Delta t \Delta s.
\]

Integrating the above in \( t \) against \( \psi_\varepsilon_1(r-t) \) and in \( s \) against \( \psi_\varepsilon_2(r-s) \) with \( \varepsilon_1, \varepsilon_2 \in (0, \varepsilon] \), we obtain that under the same assumptions as above:

\[
\log \tilde{h}_{r\varepsilon_1,\varepsilon_2}(r) + \log \tilde{h}_{r+\Delta s\varepsilon_1,\varepsilon_2}(r + \Delta t) \\
\leq \log \tilde{h}_{r\varepsilon_1,\varepsilon_2}(r + \Delta t) + \log \tilde{h}_{r+\Delta s\varepsilon_1,\varepsilon_2}(r) + C_\delta \varepsilon \Delta t \Delta s.
\]

Exchanging sides, dividing by \( \Delta t > 0 \) and taking limit as \( \Delta t \to 0 \), and then dividing by \( \Delta s > 0 \) and taking limit as \( \Delta s \to 0 \), we obtain precisely:

\[
\partial_s \partial_t |_{t=s=r} \log \tilde{h}_{r\varepsilon_1,\varepsilon_2}(t) \leq 2C_\delta \varepsilon,
\]

thereby confirming (H2).

For completeness, we provide a proof of the following lemma, used in Step 6 above.

**Lemma 12.5** Let \( f \) be a locally Lipschitz function on an open interval \( I \subset \mathbb{R} \). Let \( \psi \) denote a \( C^1 \)-smooth compactly supported function on \( \mathbb{R} \) which integrates to 1. Denote by \( \psi_\varepsilon(x) = \frac{1}{\varepsilon} \psi(x/\varepsilon), \varepsilon > 0 \), the corresponding family of mollifiers. Then:

\[
\lim_{\varepsilon \to 0} (f \ast \psi_\varepsilon)'(x) = f'(x),
\]

at all points \( x \in I \) where \( f \) is differentiable.

**Proof** Without loss of generality, assume that \( 0 \in I \), that \( f \) is differentiable at 0 and that \( f(0) = 0 \). Assume that \( \psi \) is supported in \([-M, M]\), and let \( \varepsilon > 0 \) be small enough so that \([-M \varepsilon, M \varepsilon] \subset I \). Then:

\[
(f \ast \psi_\varepsilon)'(0) = \frac{d}{dx} \left|_{x=0} \right. \int f(x + y)\psi_\varepsilon(y)dy = \int f'(y)\psi_\varepsilon(y)dy,
\]
The globalization theorem

where the differentiation under the integral is justified since \( f \) is locally Lipschitz. Integrating by parts (which is justified as \( f \psi_{\varepsilon} \) is absolutely continuous), we obtain

\[
(f * \psi_{\varepsilon})'(0) = - \int_{-M}^{M} f(y)\psi_{\varepsilon}'(y)dy = - \int_{-M}^{M} \frac{f(\varepsilon z)}{\varepsilon} z\psi'(z)dz.
\]

But for each \( z \in [-M, M] \setminus \{0\} \), \( \lim_{\varepsilon \to 0} \frac{f(\varepsilon z)}{\varepsilon} = f'(0) \), and since \( f \) is Lipschitz on \([-\varepsilon M, \varepsilon M]\), we obtain by Lebesgue’s Dominated Convergence Theorem that

\[
\lim_{\varepsilon \to 0} (f * \psi_{\varepsilon})'(0) = - \int_{-M}^{M} f'(0)z\psi'(z)dz = f'(0) \int_{-M}^{M} \psi(z)dz = f'(0),
\]

as asserted. \( \square \)

13 Final results

In this final section, we combine the results obtained in Parts I, II and the previous section, establishing at last the Main Theorem 1.1 and the globalization theorem for the \( \text{CD}(K, N) \) condition. We also treat the case of an infinitesimally Hilbertian space.

Throughout this section, recall that we assume \( K \in \mathbb{R} \) and \( N \in (1, \infty) \).

13.1 Proof of the Main Theorem 1.1

**Theorem 13.1** Let \((X, d, m)\) be an essentially non-branching m.m.s., so that \((\text{supp}(m), d)\) is a length space. Then:

\[ \text{CD}_{loc}(K, N) \Rightarrow \text{CD}^{1}_{Lip}(K, N). \]

**Proof** By Remark 6.11, \((X, d, m)\) satisfies \( \text{CD}_{loc}(K, N) \) if and only if \((\text{supp}(m), d, m)\) does. By Remark 8.8, the same is true for \( \text{CD}^{1}_{Lip}(K, N) \). Consequently, we may assume that \( \text{supp}(m) = X \). By Lemma 6.12 we deduce that \((X, d)\) is proper and geodesic (note that this would be false without the length space assumption above). Note that for geodesic essentially non-branching spaces, it is known that \( \text{CD}_{loc}(K, N) \) implies \( \text{MCP}(K, N) \)—see [30] for a proof assuming non-branching, but the same proof works under essentially non-branching, see the comments after [29, Corollary 5.4]. Consequently, the results of Sect. 7 apply.

Recall that given a 1-Lipschitz function \( u : X \to \mathbb{R} \), the equivalence relation \( R_{u}^{b} \) on the transport set \( T_{u}^{b} \) induces a partition \( \{R_{u}^{b}(\alpha)\}_{\alpha \in Q} \) of \( T_{u}^{b} \). By Corol-
lary 7.3, we know that \( m(T_u \setminus T_u^b) = 0 \) with associated strongly consistent disintegration:

\[
m_{T_u} = m_{T_u^b} = \int_Q m_\alpha q(d\alpha), \quad \text{with } m_\alpha(R_u^b(\alpha)) = 1, \text{ for } q\text{-a.e. } \alpha \in Q.
\]

It was proved in [27] that the \( CD_{loc}(K, N) \) condition ensures that for \( q\text{-a.e. } \alpha \in Q, R_u^b(\alpha), d, m_\alpha \) verifies \( CD(K, N) \) with \( supp(m_\alpha) = R_u^b(\alpha) \). Denoting by \( X_\alpha \) the closure \( R_u^b(\alpha) \), Theorem 7.10 ensures that \( X_\alpha \) coincides with the transport ray \( R_u(\alpha) \) for \( q\text{-a.e. } \alpha \in Q \). Consequently, all 4 conditions of the \( CD^1_u(K, N) \) Definition 8.1 are verified, and the assertion follows. \( \square \)

**Theorem 13.2** Let \((X, d, m)\) be an essentially non-branching m.m.s. Then:

\[
CD^1(K, N) \Rightarrow CD(K, N).
\]

**Proof** By Remark 8.8, \((X, d, m)\) satisfies \( CD^1(K, N) \) if and only if \((supp(m), d, m)\) does. By Remark 6.11, the same is true for \( CD(K, N) \). Consequently, we may assume that \( supp(m) = X \).

By Proposition 8.9 and Remark 8.11, \( X \) also verifies \( MCP(K, N) \), and so Theorem 6.15 applies. Given \( \mu_0, \mu_1 \in \mathcal{P}_2(X, d, m) \), consider the unique \( \nu \in OptGeo(\mu_0, \mu_1) \), and denote \( \mu_t := (e_\ell)_t^\nu (v) \ll m \) for all \( t \in [0, 1] \). Let \( \rho_t := d\mu_t/dm \) denote the versions of the densities guaranteed by Corollary 9.5.

Denote an associated Kantorovich potential by \( \varphi \), and recall that \( \nu \) is concentrated on \( G_\varphi = G^+_\varphi \cup G^0_\varphi \), where \( G^+_\varphi \) and \( G^0_\varphi \) denote the subsets of positive and zero length \( \varphi \)-Kantorovich geodesics, respectively. The change-of-variables Theorem 11.4 and Proposition 4.4 yield that for \( \nu\text{-a.e. geodesic } \gamma \in G^+_\varphi \):

\[
\frac{\rho_s(\gamma_s)}{\rho_t(\gamma_t)} = \frac{h_{\gamma_s}^{\varphi_s(\gamma_s)}(t)}{1 + (t - s) \frac{\partial_{\ell_t}(t - s)(\ell^2_\gamma / 2(\gamma_t))}{\ell(\gamma)^2}} = \frac{h_{\gamma_s}^{\varphi_s(\gamma_s)}(t)}{1 + (t - s) \frac{\partial_{\ell_t}(t - s)(\ell^2_\gamma / 2(\gamma_t))}{\ell(\gamma)^2}} \quad \text{for a.e. } t, s \in (0, 1). \tag{13.1}
\]

where for all \( s \in (0, 1) \), \( h_s = h_{\gamma_s}^{\varphi_s(\gamma_s)} \) is a \( CD(K_0, N) \) density, with \( K_0 = \ell^2(\gamma) K \) and \( h_s(s) = 1 \). Together with Corollary 9.5, which ensures the Lipschitz regularity (and positivity) of \( (0, 1) \ni t \mapsto \rho(\gamma_t) \), this verifies assumptions (A) and (B) of Theorem 12.3. As explained in Sect. 12, the 3rd order information on the Kantorovich potential \( \varphi \) asserted by Theorem 5.5 verifies assumption (C) of Theorem 12.3. It follows by Theorem 12.3 (and the discussion preceding it) that the rigidity of (13.1) necessarily implies that for
those \( \gamma \in G^+_{\varphi} \) satisfying (13.1), it holds:

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

where \( L \) is concave and \( Y \) is a \( \mathbb{CD}(K_0, N) \) density on \((0, 1)\). Noting that \( \sigma_{K_0, N}^{(\alpha)}(\theta) = \sigma_{K, N}^{(\alpha)}(\theta \ell(\gamma)) \), we obtain by a standard application of Hölder’s inequality that for any \( t_0, t_1 \in (0, 1) \), \( \alpha \in [0, 1] \) and \( t_\alpha = \alpha t_1 + (1 - \alpha)t_0 \):

\[
\rho_{t_\alpha}^{-\frac{1}{N}}(\gamma_{t_\alpha}) = L\frac{1}{N}(t_\alpha)Y\frac{1}{N}(t_\alpha) \\
\geq \left( \alpha L(t_1) + (1 - \alpha)L(t_0) \right)^{\frac{1}{N}}.
\]

\[
= \left( \sigma_{K_0, N}^{(\alpha)}(\ell(\gamma))\right)^{-\frac{N-1}{N}} \rho_{t_\alpha}^{-\frac{1}{N}}(\gamma_{t_\alpha}) \\
+ (1 - \alpha)\frac{1}{N} \sigma_{K, N}^{(1-\alpha)}(\ell(\gamma))^{-\frac{N-1}{N}} \rho_{t_0}^{-\frac{1}{N}}(\gamma_{t_0}).
\]

Using the upper semi-continuity of \( t \mapsto \rho_t(\gamma_t) \) at the end-points \( t = 0, 1 \) ensured by Corollary 9.5 (as both \( \mu_0, \mu_1 \ll \nu \)), we conclude that for \( \nu \)-a.e. \( \gamma \in G^+_{\varphi} \), the previous inequality in fact holds for all \( t_0, t_1 \in [0, 1] \). In particular, for \( t_0 = 0, t_1 = 1 \) and all \( \alpha \in [0, 1] \):

\[
\rho_{\alpha}^{-\frac{1}{N}}(\gamma_{\alpha}) \geq \tau_{K, N}^{(\alpha)}(d(\gamma_0, \gamma_1)) \rho_1^{-\frac{1}{N}}(\gamma_1) + \tau_{K, N}^{(1-\alpha)}(d(\gamma_0, \gamma_1)) \rho_0^{-\frac{1}{N}}(\gamma_0).
\]

As for null-geodesics \( \gamma \in G^0_{\varphi} \) (having zero length), note that \( \tau_{K, N}^{(s)}(0) = s \) and that \( [0, 1] \ni t \mapsto \rho_t(\gamma_t) \) remains constant by Theorem 11.4, and therefore (13.3) holds trivially with equality for all \( \gamma \in G^0_{\varphi} \). In conclusion, (13.3) holds for \( \nu \)-a.e. geodesic \( \gamma \), thereby confirming the validity of Definition 6.7 and verifying \( \mathbb{CD}(K, N) \).

As an immediate consequence of the previous two theorems, we obtain the Local-to-Global Theorem for the Curvature-Dimension condition.

\[ \square \] Springer
**Theorem 13.3** Let \((X, d, m)\) be an essentially non-branching m.m.s. so that \((\text{supp}(m), d)\) is a length space. Then:

\[
\text{CD}_{\text{loc}}(K, N) \iff \text{CD}(K, N).
\]

**Remark 13.4** It is clear that the above globalization theorem is false without some global assumption ultimately ensuring that \((\text{supp}(m), d)\) is geodesic. Indeed, simply consider a \(\text{CD}(K, N)\) space, and restrict it to two disjoint geodesically-convex closed subsets of \((\text{supp}(m), d)\) (each having positive measure)—the resulting space clearly satisfies \(\text{CD}_{\text{loc}}(K, N)\) but not \(\text{CD}(K, N)\); it is also easy to construct similar examples where \((\text{supp}(m), d)\) is connected. In addition, as already mentioned in the Introduction, the globalization theorem is known to be false without some type of non-branching assumption (see [67]).

As an interesting byproduct, we also obtain that \(\text{CD}^1\) and \(\text{CD}_{\text{Lip}}^1\) are equivalent conditions on essentially non-branching spaces:

**Corollary 13.5** Let \((X, d, m)\) be an essentially non-branching m.m.s. Then:

\[
\text{CD}(K, N) \iff \text{CD}^1(K, N) \iff \text{CD}_{\text{Lip}}^1(K, N).
\]

**Proof** \(\text{CD}_{\text{Lip}}^1(K, N)\) is by definition stronger than \(\text{CD}^1(K, N)\), which in turn implies \(\text{CD}(K, N)\) by Theorem 13.2. But \(\text{CD}(K, N)\) implies its local version \(\text{CD}_{\text{loc}}(K, N)\), as well as that \((\text{supp}(m), d)\) is geodesic by Lemma 6.12. The cycle is then closed by Theorem 13.1.

Finally, we deduce a complete equivalence between the reduced and the classic Curvature-Dimension conditions on essentially non-branching spaces. Recall that the reduced version \(\text{CD}^*(K, N)\), introduced in [14] (in the non-branching setting), is defined exactly in the same manner as \(\text{CD}(K, N)\), with the only (crucial) difference being that one employs the slightly smaller \(s_{K,N}(\theta)\) coefficients instead of the \(\tau_{K,N}(\theta)\) ones in Definition 6.4.

**Corollary 13.6** Let \((X, d, m)\) be an essentially non-branching m.m.s. Then:

\[
\text{CD}^*(K, N) \iff \text{CD}(K, N).
\]

**Proof** By definition \(\text{CD}(K, N)\) is stronger than \(\text{CD}^*(K, N)\) (see [14, Proposition 2.5 (i)]). For the converse implication, note that \(\text{CD}^*(K, N)\) implies that \((\text{supp}(m), d)\) is proper and geodesic, by verbatim repeating the proof of Lemma 6.12. Then we observe that \(\text{CD}^*(K, N) \Rightarrow \text{CD}_{\text{loc}}(K^-, N)\), where \(\text{CD}_{\text{loc}}(K^-, N)\) denotes that \((X, d, m)\) verifies \(\text{CD}_{\text{loc}}(K', N)\) for every
$K' < K$ (with the open neighborhoods possibly depending on $K'$). For non-branching spaces, this was proved in [14, Proposition 5.5] (see also [34, Lemma 2.1]), but the proof does not rely on any non-branching assumptions. Then, by Theorem 13.3, we obtain $\text{CD}(K', N)$ for any $K' < K$. Finally, by uniqueness of dynamical plans (see Theorem 6.15 and Lemma 6.13) and continuity of $\tau_{K', N}^{(t)}(\theta)$ in $K'$, the claim follows.

13.2 $\text{RCD}(K, N)$ spaces

We also mention the more recent Riemannian Curvature Dimension condition $\text{RCD}^*(K, N)$. In the infinite dimensional case $N = \infty$, it was introduced in [7] for finite measures $m$ and in [4] for $\sigma$-finite ones. The class $\text{RCD}^*(K, N)$ with $N < \infty$ has been proposed in [40] and extensively investigated in [8, 11, 35]. We refer to these papers and references therein for a general account on the synthetic formulation of the latter Riemannian-type Ricci curvature lower bounds. Here we only briefly recall that it is a strengthening of the reduced Curvature Dimension condition: a m.m.s. verifies $\text{RCD}^*(K, N)$ if and only if it satisfies $\text{CD}^*(K, N)$ and is infinitesimally Hilbertian [40, Definition 4.19 and Proposition 4.22], meaning that the Sobolev space $W^{1,2}(X, m)$ is a Hilbert space (with the Hilbert structure induced by the Cheeger energy). Recall also that the local-to-global property for the $\text{RCD}^*(K, N)$ condition (say for length spaces of full support) has already been established for $N = \infty$ in [7, Theorem 6.22] for non-branching spaces with finite second moment, for $N < \infty$ in [35, Theorems 3.17 and 3.25] for strong $\text{RCD}^*(K, N)$ spaces, and for all $N \in [1, \infty]$ in [10, Theorems 7.2 and 7.8] for proper spaces without any non-branching assumptions.

We are now in a position to introduce the following (expected) definition:

**Definition** We will say that a m.m.s. $(X, d, m)$ satisfies $\text{RCD}(K, N)$ if it verifies $\text{CD}(K, N)$ and is infinitesimally Hilbertian.

We can now immediately deduce:

**Corollary 13.7**

$$\text{RCD}(K, N) \iff \text{RCD}^*(K, N).$$

Note that $\text{CD}^*(K, \infty)$ and $\text{CD}(K, \infty)$ are the same condition, so the above also holds for $N = \infty$.

**Proof** Since $\text{CD}(K, N)$ is stronger than $\text{CD}^*(K, N)$, one implication is straightforward. For the other implication, recall that $\text{RCD}^*(K, N)$ forces the space to be essentially non-branching (see [68, Corollary 1.2]), and so the assertion follows by Corollary 13.6.
Corollary 13.8 Let \((X, d, m)\) be an m.m.s. so that \((\text{supp}(m), d)\) is a length space. Then:

\[
\text{RCD}_{\text{loc}}(K, N) \iff \text{RCD}(K, N).
\]

Proof One implication is trivial. For the converse, as usual, we may assume that \(\text{supp}(m) = X\) by Remark 6.11. By Lemma 6.12, we know that \((X, d)\) is proper and geodesic (as usual, this would be false without the length space assumption above). As the local-to-global property has been proved for proper geodesic \(\text{RCD}^*(K, N)\) spaces without any non-branching assumptions in [10], it follows that:

\[
\text{RCD}_{\text{loc}}(K, N) \Rightarrow \text{RCD}^*_{\text{loc}}(K, N) \Rightarrow \text{RCD}^*(K, N) \Rightarrow \text{RCD}(K, N),
\]

where the last implication follows by Corollary 13.7.

\[\square\]

13.3 Concluding remarks

We conclude this work with several brief remarks and suggestions for further investigation.

– Note that the proof of Theorem 13.2 in fact yields more than stated: not only does the synthetic inequality (13.2) hold (for all \(t_0, t_1 \in [0, 1]\)), but in fact we obtain for \(\nu\)-a.e. geodesic \(\gamma\) the a-priori stronger disentanglement (or “L-Y” decomposition):

\[
\frac{1}{\rho_t(\gamma_t)} = L_{\gamma}(t)Y_{\gamma}(t) \quad \forall t \in (0, 1),
\]

where \(L_{\gamma}\) is concave and \(Y_{\gamma}\) is a \(\text{CD}(\ell(\gamma)^2, K, N)\) density on \((0, 1)\). As explained in the Introduction, it follows from [34] that for a fixed \(\gamma\), (13.4) is indeed strictly stronger than (13.2). In view of Main Theorem 1.1, this constitutes a new characterization of essentially non-branching \(\text{CD}(K, N)\) spaces.

– According to [35, p. 1026], it is possible to localize the argument of [68] and deduce from a strong \(\text{CD}_{\text{loc}}(K, \infty)\) condition (when \(K\)-convexity of the entropy is assumed along any \(W_2\)-geodesic with end-points inside the local neighborhood), that the space is globally essentially non-branching. In combination with our results, it follows that the strong \(\text{CD}(K, N)\) condition enjoys the local-to-global property, without \(a\)-\(priori\) requiring any additional non-branching assumptions.

– It would still be interesting to clarify the relation between the \(\text{CD}(K, N)\) condition and the property \(\text{BM}(K, N)\) of satisfying a Brunn-Minkowski

\[\text{ Springer}\]
inequality (with sharp dependence on $K, N$ as in [74]). Note that by Main Theorem 1.1, it is enough to understand this locally on essentially non-branching spaces.

- It would also be interesting to study the $\text{CD}^1(K, N)$ condition on its own, when no non-branching assumptions are assumed, and to verify the usual list of properties desired by a notion of Curvature-Dimension (see [28, 51, 74]).

- A natural counterpart of $\text{RCD}(K, N)$ would be $\text{RCD}^1(K, N)$: we will say that a m.m.s. verifies $\text{RCD}^1(K, N)$ if it verifies $\text{CD}^1(K, N)$ and it is infinitesimally Hilbertian. Recall that an $\text{RCD}(K, N)$ space is always essentially non-branching [68], and hence Main Theorem 1.1 immediately yields:

$$\text{RCD}(K, N) \Rightarrow \text{RCD}^1(K, N).$$

The converse implication would be implied by the following claim which we leave for a future investigation: an $\text{RCD}^1(K, N)$-space is always essentially non-branching.

- In regards to the novel third order temporal information on the intermediate-time Kantorovich potentials $\varphi_t$ we obtain in this work—it would be interesting to explore whether it has any additional consequences pertaining to the spatial regularity of solutions to the Hamilton-Jacobi equation in general, and of the transport map $T_{s,t} = e_t \circ e_s|_G^{-1}$ from an intermediate time $s \in (0, 1)$ in particular (where $G \subset G_\varphi$ is the subset of injectivity guaranteed by Corollary 6.16). In the smooth Riemannian setting, the map $T_{s,t}$ is known to be locally Lipschitz by Mather’s regularity theory (see [77, Chapter 8] and cf. [77, Theorem 8.22]). A starting point for this investigation could be the following bound on the (formal) Jacobian of $T_{s,t}$, which follows immediately from (12.5), Theorem 3.11 (3) and Lemma A.9: for $\mu_s$-a.e. $x$, the Jacobian is bounded above by a function of $s, t, K, N, l_s(x)$ only.

Acknowledgements We would like to thank Theo Sturm and Cédric Villani for numerous discussions and for encouraging us to pursue the globalization problem. We also thank the referees for their careful reading of the manuscript and helpful comments.

Funding Open access funding provided by Scuola Internazionale Superiore di Studi Avanzati - SISSA within the CRUI-CARE Agreement. E. Milman: The research leading to these results is part of a project that has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 637851).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide
A Appendix: One dimensional $\text{CD}(K, N)$ densities

Definition A.1 A non-negative function $h$ defined on an interval $I \subset \mathbb{R}$ is called a $\text{CD}(K, N)$ density on $I$, for $K \in \mathbb{R}$ and $N \in (1, \infty)$, if for all $x_0, x_1 \in I$ and $t \in [0, 1]$:

$$h(t x_1 + (1 - t) x_0) \frac{1}{N - 1} \geq \sigma^{(t)}_{K, N - 1}(|x_1 - x_0|) h(x_1) \frac{1}{N - 1} + \sigma^{(1-t)}_{K, N - 1}(|x_1 - x_0|) h(x_0) \frac{1}{N - 1},$$

(recalling the coefficients $\sigma$ from Definition 6.2). While we avoid in this work the case $N = \infty$, it will be useful in this section to also treat the case $N = \infty$, whence the latter condition is interpreted by subtracting 1 from both sides, multiplying by $N - 1$, and taking the limit as $N \to \infty$, namely:

$$\log h(t x_1 + (1 - t) x_0) \geq t \log h(x_1) + (1 - t) \log h(x_0) + \frac{K}{2} t (1 - t)(x_1 - x_0)^2.$$

For completeness, we will say that $h$ is a $\text{CD}(K, 1)$ density on $I$ iff $K \leq 0$ and $h$ is constant on the interior of $I$.

Unless otherwise stated, we assume in this “Appendix” that $K \in \mathbb{R}$ and $N \in (1, \infty]$. The following is a specialization to dimension one of a well-known result in the theory of $\text{CD}(K, N)$ mm-spaces, which explains the terminology above. Here we do not assume that a m.m.s. is necessarily equipped with a probability measure.

Theorem A.2 If $h$ is a $\text{CD}(K, N)$ density on an interval $I \subset \mathbb{R}$ then the m.m.s. $(I, |\cdot|, h(t) dt)$ verifies $\text{CD}(K, N)$. Conversely, if the m.m.s. $(\mathbb{R}, |\cdot|, \mu)$ verifies $\text{CD}(K, N)$ and $I = \text{supp}(\mu)$ is not a point, then $\mu \ll \mathcal{L}^1$ and there exists a version of the density $h = d\mu/d\mathcal{L}^1$ which is a $\text{CD}(K, N)$ density on $I$.

Proof The first assertion follows from e.g. [74, Theorem 1.7 (ii)], and the second follows by considering the $\text{CD}(K, N)$ condition for uniform measures $\mu_0, \mu_1$ on intervals of length $\varepsilon$ and $\alpha \varepsilon$, respectively, letting $\varepsilon \to 0$, employing Lebesgue’s differentiation theorem, and optimizing on $\alpha > 0$ (e.g. as in the proof of [30, Theorem 4.3]). \qed
Let $h$ be a $\text{CD}(K, N)$ density on an interval $I \subset \mathbb{R}$. A few standard and easy consequences of Definition A.1 are:

- $h$ is also a $\text{CD}(K_2, N_2)$ density for all $K_2 \leq K$ and $N_2 \in [N, \infty]$ (this follows from the corresponding monotonicity of the coefficients $\sigma^{(i)}_{K, N-1}(\theta)$ in $K$ and $N$, see e.g. [51, 74]).
- $h$ is lower semi-continuous on $I$ and locally Lipschitz continuous in its interior (this is easily reduced to a standard identical statement for concave functions on $I$).
- $h$ is strictly positive in the interior whenever it does not identically vanish (follows immediately from the definition).
- $h$ is locally semi-concave in the interior, i.e. for all $x_0$ in the interior of $I$, there exists $C_{x_0} \in \mathbb{R}$ so that $h(x) - C_{x_0} x^2$ is concave in a neighborhood of $x_0$ (easily checked for $\text{CD}(K, \infty)$ densities). In particular, it is twice differentiable (in the sense of Lemma 2.3) a.e. in $I$.

### A.1 Differential characterization

The following is a well-known differential characterization of $C^2$-smooth $\text{CD}(K, N)$ densities:

**Lemma A.3** Let $h \in C^2_{\text{loc}}(I)$ on some open interval $I \subset \mathbb{R}$. The following are equivalent:

1. $h$ is a $\text{CD}(K, N)$ density on $I$.
2. For all $x \in I$:

$$
(\log h)^\prime\prime(x) + \frac{1}{N-1} (\log h)^\prime(x)^2 = (N - 1) \frac{(h^{\frac{1}{N-1}})^\prime(x)}{h^{\frac{1}{N-1}}(x)} \leq -K.
$$

(A.1)

where the left hand side is interpreted as $(\log h)^\prime\prime(x)$ when $N = \infty$.

**Remark A.4** The equality in (A.1) holds for any $N \in (1, \infty)$ by the Leibniz and chain rules at any point $x$ where $h(x)$ is positive and twice differentiable (and in particular, $h^{\frac{1}{N-1}}$ and $\log h$ are also twice differentiable at such a point $x$). The condition (A.1) is the one-dimensional specialization of the Bakry–Émery $\text{CD}(K, N)$ condition for smooth weighted Riemannian manifolds [15, 16].

In fact, we will require a couple of extensions of the above standard claim, which in particular, together imply Lemma A.3; to avoid unnecessary generality, we only treat the case $N \in (1, \infty)$.
Lemma A.5 Let $h$ denote a CD$(K, N)$ density on an interval $I \subset \mathbb{R}$, $N \in (1, \infty)$. Then $h$ satisfies (A.1) at any point $x$ in the interior where it is twice differentiable (in particular, (A.1) holds for a.e. $x \in I$).

Proof Let $x$ be a point as above. Observe that:

$$\sigma_{K,N-1}^{(1/2)}(\theta) = \frac{1}{2} + \frac{\theta^2}{16} \frac{K}{N-1} + o(\theta^2) \quad \text{as } \theta \to 0,$$

and so denoting $g = h^{\frac{1}{N-1}}$, the CD$(K, N)$ condition with $x_0 = x - \varepsilon$, $x_1 = x + \varepsilon$ and $t = 1/2$ implies:

$$2g(x) \geq \left(1 + \frac{\varepsilon^2}{2} \frac{K}{N-1} + o(\varepsilon^2)\right)(g(x + \varepsilon) + g(x - \varepsilon)) \quad \text{as } \varepsilon \to 0.$$

It follows by Taylor’s theorem and continuity of $g$ in the interior of $I$ that:

$$g''(x) = \lim_{\varepsilon \to 0} \frac{g(x + \varepsilon) + g(x - \varepsilon) - 2g(x)}{\varepsilon^2} \leq \lim_{\varepsilon \to 0} - \frac{K}{N-1} \frac{g(x + \varepsilon) + g(x - \varepsilon)}{2} = -\frac{K}{N-1} g(x),$$

confirming (A.1) and concluding the proof.

Lemma A.6 Let $h$ be a positive differentiable function on an open interval $I \subset \mathbb{R}$ whose derivative is locally absolutely continuous there (and hence $h$ is twice differentiable a.e. in $I$). If $h$ satisfies (A.1) for a.e. $x \in I$ and $N \in (1, \infty)$, then $\ell(I) \leq D_{K,N-1}$ and $h$ is a CD$(K, N)$ density on $I$.

Remark A.7 The differentiability assumption at every point cannot be relaxed, as witnessed by the convex function $h(x) = |x|$, which satisfies $h''(x) = 0$ for a.e. $x$ but nevertheless is not concave.

Proof Given $x_0, x_1 \in I$ with $|x_1 - x_0| < D_{K,N-1}$, consider the function $\Delta$ on $[0, 1]$ given by:

$$\Delta(t) := h(tx_1 + (1-t)x_0)^{\frac{1}{N-1}} - \sigma_{K,N-1}^{(t)}(|x_1 - x_0|)h(x_1)^{\frac{1}{N-1}} - \sigma_{K,N-1}^{(1-t)}(|x_1 - x_0|)h(x_0)^{\frac{1}{N-1}}.$$

As $\Delta$ is positive and bounded away from zero on $[0, 1]$, and since $y^{\frac{1}{N-1}}$ is Lipschitz on compact sub-intervals of $(0, \infty)$, it follows that $\Delta$ is differentiable with absolutely continuous derivative on $[0, 1]$. In addition, clearly $\Delta(0) =$
\( \Delta(1) = 0 \). Abbreviating \( \sigma(t) = \sigma^{(t)}_{K,N-1}(|x_1 - x_0|) \), it is immediate to verify that:

\[
\frac{d^2}{dt^2} \sigma(t) = -\frac{K}{N-1} (x_1 - x_0)^2 \sigma(t),
\]  

(A.2)

and therefore our assumption (A.1) for a.e. \( x \in I \) implies:

\[
\frac{d^2}{dt^2} \Delta(t) \leq -\frac{K}{N-1} (x_1 - x_0)^2 \Delta(t) \quad \text{for a.e.} \ t \in [0, 1].
\]  

(A.3)

Now set \( \Delta_0(t) = \Delta(t) \) and \( \Delta_1(t) = \Delta(1 - t) \), and for each \( i \in \{0, 1\} \), denote by \( \beta_i \) the absolutely continuous function on \([0, 1] \) given by:

\[
\beta_i(t) := \Delta_i'(t)\sigma(t) - \Delta_i(t)\sigma'(t).
\]  

(A.4)

It follows by the Leibniz rule that for any \( i \in \{0, 1\} \):

\[
\beta_i'(t) = \Delta_i''(t)\sigma(t) - \Delta_i(t)\sigma''(t) \leq 0 \quad \text{for a.e.} \ t \in [0, 1],
\]

and since \( \sigma(0) = 0 \) we also have \( \beta_i(0) = 0 \). The absolute continuity implies that \( \beta_i \) is monotone non-increasing, and hence \( \beta_i(t) \leq 0 \) for all \( t \in [0, 1] \).

We are ready to conclude that \( \Delta \geq 0 \) on \([0, 1] \), by showing that \( \Delta(t_0) \geq 0 \) for any local extremum point \( t_0 \in (0, 1) \) of \( \Delta \). Indeed, when \( K \leq 0 \), this is immediate, since \( \sigma' > 0 \) and:

\[
0 \geq \beta_0(t_0) = \Delta_0'(t_0)\sigma(t_0) - \Delta_0(t_0)\sigma'(t_0) = -\Delta(t_0)\sigma'(t_0).
\]

When \( K > 0 \), set \( t_1 = 1 - t_0 \) which is a local extremal point of \( \Delta_1 \) in \((0, 1) \), and note that \( t_i^* \in (0, 1/2) \) for some \( i^* \in \{0, 1\} \). Since \( |x_1 - x_0| < D_{K,N-1} \), it follows that \( \sigma' > 0 \) on \([0, 1/2] \), and so the same argument as for the case \( K \leq 0 \) but applied to \( \Delta_i^* \) yields that \( \Delta(t_0) = \Delta_i^*(t_i^*) \geq 0 \), as asserted.

Finally, when \( K > 0 \), assume in the contrapositive that there exist \( x_0, x_1 \in I \) with \( x_1 - x_0 = D_{K,N-1} \). Denote \( \Delta_0(t) = \Delta(t) := h(tx_1 + (1-t)x_0) \) and set \( \sigma(t) := \sin(\pi t) \) for \( t \in [0, 1] \). Note that as before, (A.2) and (A.3) are satisfied, and so defining the function \( \beta_0 \) by (A.4), \( \beta_0 \) is again monotone non-increasing on \([0, 1] \). But:

\[
\beta_0(0) = -\Delta_0(0)\sigma'(0) = -\pi h(x_0) < \pi h(x_1) = -\Delta_0(1)\sigma'(1) = \beta_0(1),
\]

yielding a contradiction to the monotonicity, and concluding the proof. \( \square \)
A.2 A-priori estimates

We will also require the following a-priori estimates on the supremum and logarithmic derivative of $\mathbf{CD}(K, N)$ densities. Here it is crucial that $N \in (1, \infty)$.

**Lemma A.8** Let $h$ denote a $\mathbf{CD}(K, N)$ density on a finite interval $(a, b)$, $N \in (1, \infty)$, which integrates to 1. Then:

$$\sup_{x_0 \in (a, b)} h(x_0) \leq \frac{1}{b - a} \left\{ N \left( \int_0^1 (\sigma_{K,N-1}^{(t)}(b-a))^{N-1} dt \right)^{-1} \right\} K \geq 0$$

$$\left\{ \int_0^1 (\sigma_{K,N-1}^{(t)}(b-a))^{N-1} dt \right\} K < 0.$$

In particular, for fixed $K$ and $N$, $h$ is uniformly bounded from above as long as $b - a$ is uniformly bounded away from 0 (and from above if $K < 0$).

**Proof** Given $x_0 \in (a, b)$, we have by the $\mathbf{CD}(K, N)$ condition:

$$1 = (x_0 - a) \int_0^1 h(tx_0 + (1 - t)a) dt$$

$$+ (b - x_0) \int_0^1 h((1 - t)b + tx_0) dt$$

$$\geq h(x_0) \left( (x_0 - a) \int_0^1 (\sigma_{K,N-1}^{(t)}(x_0 - a))^{N-1} dt 
+ (b - x_0) \int_0^1 (\sigma_{K,N-1}^{(t)}(b - x_0))^{N-1} dt \right).$$

When $K \geq 0$, the monotonicity of $K \mapsto \sigma_{K,N-1}^{(t)}(\theta)$ implies that $\sigma_{K,N-1}^{(t)}(\theta) \geq \sigma_{0,N-1}^{(t)}(\theta) = t$, and we obtain:

$$1 \geq h(x_0) \frac{b - a}{N}.$$

When $K < 0$, one may show that the function $\theta \mapsto \sigma_{K,N-1}^{(t)}(\theta)$ is decreasing on $\mathbb{R}_+$, as this is equivalent to showing that the function $x \mapsto \log \sinh \exp(x)$ is convex on $\mathbb{R}_+$, and the latter may be verified by direct differentiation (and using that $\sinh(x) \cosh(x) \geq x$). Consequently, we obtain:

$$1 \geq h(x_0)(b - a) \int_0^1 (\sigma_{K,N-1}^{(t)}(b-a))^{N-1} dt.$$
as asserted. We remark that when \( K > 0 \), one may similarly show that the function \( \theta \mapsto \sigma_{K,N-1}(\theta) \) is increasing on \([0, D_{K,N-1}]\), and since \( \sigma_{K,N-1}(0) = t \), we obtain the previous estimate we employed.

**Lemma A.9** Let \( h \) denote a \( \text{CD}(K, N) \) density on a finite interval \((a, b)\), \( N \in (1, \infty) \). Then:

\[
-\sqrt{K(N-1)} \cot((b-x)\sqrt{K/(N-1)}) \\
\leq (\log h)'(x) \leq \sqrt{K(N-1)} \cot((x-a)\sqrt{K/(N-1)}),
\]

for any point \( x \in (a, b) \) where \( h \) is differentiable. In particular, \( \log h(x) \) is locally Lipschitz on \( x \in (a, b) \) with estimates depending continuously only on \( x, a, b, K, N \).

**Proof** Denote \( \Psi_t = h \frac{1}{N-1} \). The inequality on the right-hand-side follows since:

\[
\Psi(tx + (1-t)a) \geq \sigma_{K,N-1}(x-a) \Psi(x) \quad \forall t \in [0, 1]
\]

with equality at \( t = 1 \), and hence we may compare derivatives at \( t = 1 \):

\[
(x-a)\Psi'(x) \leq \partial_t |_{t=1} \sigma_{K,N-1}(x-a) \Psi(x),
\]

whenever \( \Psi \) is differentiable at \( x \). The inequality on the left-hand-side follows similarly.

**A.3 Logarithmic convolutions**

We will require the following:

**Proposition A.10** Let \( h \) denote a \( \text{CD}(K, N) \) density on an interval \((a, b)\). Let \( \psi_\varepsilon \) denote a non-negative \( C^2 \) function supported on \([-\varepsilon, \varepsilon]\) with \( \int \psi_\varepsilon = 1 \). For any \( \varepsilon \in (0, \frac{b-a}{2}) \), define the function \( h^\varepsilon \) on \((a + \varepsilon, b - \varepsilon)\) by:

\[
\log h^\varepsilon := \log h * \psi_\varepsilon.
\]

Then \( h^\varepsilon \) is a \( C^2 \)-smooth \( \text{CD}(K, N) \) density on \((a + \varepsilon, b - \varepsilon)\).

For the proof, we will require the following general:

**Lemma A.11** Let \( g \) denote a semi-concave function on an open interval \( I \) (i.e. \( g(x) - M \frac{x^2}{2} \) is concave for some \( M \geq 0 \)). Let \( \psi \) denote a \( C^2 \)-smooth non-negative test function with compact support in \( I \). Then:

\[
\int_I g(x)\psi''(x)dx \leq \int_I g''(x)\psi(x)dx.
\]
In other words, the singular part of \( g \)’s distributional second derivative is non-positive.

The argument is identical to the one used by D. Cordero–Erausquin in the proof of [32, Lemma 1]. For completeness, we present the proof.

**Proof** Extend \( g \) and \( \psi \) to the entire \( \mathbb{R} \) by defining them as equal to zero outside of \( I \). Given \( \varepsilon > 0 \) and \( x \in I \), denote:

\[
D_\varepsilon^2 g(x) := \frac{g(x + \varepsilon) + g(x - \varepsilon) - 2g(x)}{\varepsilon^2},
\]

and similarly for \( D_\varepsilon^2 \psi(x) \). By Taylor’s theorem, for any point \( x \in I \) where \( g \) is twice differentiable we have \( \lim_{\varepsilon \to 0} D_\varepsilon^2 g(x) = g''(x) \). In fact, this holds at any point where \( g \) has a second Peano derivative, see Sect. 2.2; in the context of convex functions on \( \mathbb{R}^n \), such points are called points possessing a Hessian in the sense of Aleksandrov. Now since for small enough \( \varepsilon > 0 \), \( D_\varepsilon^2 g \leq M \) on the support of \( \psi \) by semi-concavity (and since \( \psi \geq 0 \)), we obtain by Fatou’s lemma:

\[
\int_I g''(x)\psi(x)dx \geq \limsup_{\varepsilon \to 0} \int_I D_\varepsilon^2 g(x)\psi(x)dx = \limsup_{\varepsilon \to 0} \int_I g(x)D_\varepsilon^2 \psi(x)dx = \int_I g(x)\psi''(x)dx,
\]

where the last equality follows by Lebesgue’s Dominated Convergence theorem using the fact that \( |D_\varepsilon^2 \psi(x)| \leq \max |\psi''| \) for all \( x \in I, \varepsilon > 0 \), and the fact that \( g \) is locally integrable. \( \square \)

**Proof of Proposition A.10** Note that \( \log h \) is locally integrable on \( (a, b) \), so that the integral:

\[
\log h^\varepsilon(x) = \int \log h(y)\psi_\varepsilon(x - y)dy,
\]

is well-defined for all \( x \in (a + \varepsilon, b - \varepsilon) \), and we may take two derivatives in \( x \) under the integral (as \( \psi_\varepsilon \) is \( C^2 \)-smooth with bounded corresponding derivatives), implying the asserted smoothness. In addition:

\[
(\log h^\varepsilon)'(x) = \int \log h(y) \frac{d}{dx}\psi_\varepsilon(x - y)dy = -\int \log h(y) \frac{d}{dy}\psi_\varepsilon(x - y)dy = \int (\log h)'(y)\psi_\varepsilon(x - y)dy,
\]

\( \square \) Springer
where the last equality follows from the usual integration by parts formula and Leibniz rule since \((\log h(y))\psi_\varepsilon(x - y)\) is absolutely continuous. Furthermore:

\[
(\log h^\varepsilon)'(x) = \int \log h(y) \frac{d^2}{(dy)^2} \psi_\varepsilon(x - y) dy
\]

\[
\leq \int (\log h''(y))\psi_\varepsilon(x - y) dy,
\]

where the last inequality follows by Lemma A.11 applied to \(g = \log h\), since \(h\) is a \(\text{CD}(K, \infty)\) density (by monotonicity in \(N\)), and hence \(\log h(x) + K \frac{x^2}{2}\) is concave on \((a, b)\).

Putting everything together and applying Jensen’s inequality, we obtain:

\[
(\log h^\varepsilon)''(x) + \frac{1}{N - 1} (\frac{d}{dx})^2 (\log h^\varepsilon)'(x) \leq \int (\log h''(y))\psi_\varepsilon(x - y) dy + \frac{1}{N - 1} \left( \int (\log h'(y))\psi_\varepsilon(x - y) dy \right)^2
\]

\[
\leq \int (\log h''(y))\psi_\varepsilon(x - y) dy + \frac{1}{N - 1} (\frac{d}{dy})^2 (\log h)'(y)^2 \psi_\varepsilon(x - y) dy \leq 0,
\]

where the last inequality follows since the integrand is non-positive (where it is defined) by Lemma A.5. A final application of Lemma A.3 concludes the proof.

We will use Proposition A.10 in the following form:

**Proposition A.12** Let \(\{h_s(t)\}_{s \in (c, d)}\) denote a Borel measurable family of \(\text{CD}(K, N)\) densities on \((a, b)\) (so that for every \(t \in (a, b), (c, d) \ni s \mapsto h_s(t)\) is Borel measurable). Assume in addition that:

\[
\int_c^d \int_a^b |\log h_y(x)| \, dx \, dy < \infty. \tag{A.5}
\]

Given \(\varepsilon_1, \varepsilon_2 > 0\) and \(s \in (c + \varepsilon_2, d - \varepsilon_2)\), denote the following function:

\[
\log h_s^{\varepsilon_1, \varepsilon_2}(t) := \int \int \log h_y(x) \psi_{\varepsilon_1}(t - x) \psi_{\varepsilon_2}(s - y) \, dx \, dy, \tag{A.6}
\]

\(t \in (a + \varepsilon_1, b - \varepsilon_1)\).
where as usual, $\psi_{\varepsilon_i}$ denotes a non-negative $C^2$ function supported on $[-\varepsilon_i, \varepsilon_i]$ with $\int \psi_{\varepsilon_i} = 1$. Then $\{h_{s, \varepsilon_2}(t)\}_{s \in (c+\varepsilon_2, d-\varepsilon_2)}$ is a $C^2$-smooth (in $(s, t)$) family of $\text{CD}(K, N)$ densities on $(a + \varepsilon_1, b - \varepsilon_1)$.

**Proof** The proof is a repetition of the proof of the previous proposition, so we will be brief. Our assumption (A.5) implies that (A.6) is well-defined, and justifies taking two derivatives in $t$ and $s$ under the integral, implying the assertion on smoothness. The first derivative in $t$ under the integral may be integrated by parts, whereas for the second derivative we apply Lemma A.11. A final application of Jensen’s inequality as in Proposition A.10 establishes the asserted differential characterization of $\text{CD}(K, N)$, concluding the proof.

\[\Box\]

**References**

1. Ambrosio, L.: Fine properties of sets of finite perimeter in doubling metric measure spaces. Set Valued Anal. **10**, 111–128 (2002)
2. Ambrosio, L.: Lecture notes on optimal transport problems. In: Mathematical Aspects of Evolving Interfaces (Funchal, 2000). Lecture Notes in Mathematics, vol. 1812, pp. 1–52. Springer, Berlin (2003)
3. Ambrosio, L., Gigli, N.: A user’s guide to optimal transport. In: Piccoli, B., Rascele, M. (eds.) Modelling and Optimisation of Flows on Networks. Lecture Notes in Mathematics, vol. 2062, pp. 1–155. Springer, Heidelberg (2013)
4. Ambrosio, L., Gigli, N., Mondino, A., Rajala, T.: Riemannian Ricci curvature lower bounds in metric measure spaces with $\sigma$-finite measure. Trans. Am. Math. Soc. **367**(7), 4661–4701 (2015)
5. Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH-Zürich. Birkhäuser, Basel (2005)
6. Ambrosio, L., Gigli, N., Savaré, G.: Calculus and heat flow in metric measure spaces and application to spaces with Ricci curvature bounded from below. Invent. Math. **195**, 289–391 (2014)
7. Ambrosio, L., Gigli, N., Savaré, G.: Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163**, 1405–1490 (2014)
8. Ambrosio, L., Gigli, N., Savaré, G.: Bakry–Émery curvature-dimension condition and Riemannian Ricci curvature bounds. Ann. Probab. **43**, 339–404 (2015)
9. Ambrosio, L., Di Marino, S.: Equivalent definitions of BV space and of total variation on metric measure spaces. J. Funct. Anal. **266**, 4150–4188 (2014)
10. Ambrosio, L., Mondino, A., Savaré, G.: On the Bakry–Émery condition, the gradient estimates and the local-to-global property of $\text{RCD}^*(K, N)$ metric measure spaces. J. Geom. Anal. **26**, 24–56 (2016)
11. Ambrosio, L., Mondino, A., Savaré, G.: Nonlinear diffusion equations and curvature conditions in metric measure spaces. Mem. Am. Math. Soc. **262**(1270), v+121 (2019)
12. Ambrosio, L., Miranda Jr., M., Pallara, D.: Special functions of bounded variation in doubling metric measure spaces. In: Calculus of variations: topics from the mathematical heritage of E. De Giorgi, vol. 14, pp. 1–45. Dept. Math., Seconda Univ. Napoli, Caserta (2004)
13. Ambrosio, L., Pratelli, A.: Existence and stability results in the $L^1$ theory of optimal transportation. In: Optimal Transportation and Applications (Martina Franca, 2001), Lecture Notes in Mathematics, vol. 183, pp. 123–160. Springer, Berlin (2003)
14. Bacher, K., Sturm, K.T.: Localization and tensorization properties of the curvature-dimension condition for metric measure spaces. J. Funct. Anal. 259(1), 28–56 (2010)
15. Bakry, D.: L’hypercontractivité et son utilisation en théorie des semigroupes. In: Lectures on Probability Theory (Saint-Flour, 1992), Lecture Notes in Mathematics, vol. 1581, pp. 1–114. Springer, Berlin (1994)
16. Bakry, D., Émery, M.: Diffusions hypercontractives. In: Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Mathematics, vol. 1123, pp. 177–206. Springer (1985)
17. Bakry, D., Gentil, I., Ledoux, M.: Analysis and geometry of Markov diffusion operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348. Springer, Cham (2014)
18. Bianchini, S., Caravenna, L.: On the extremality, uniqueness and optimality of transference plans. Bull. Inst. Math. Acad. Sin. (N.S.) 4(4), 353–454 (2009)
19. Bianchini, S., Cavalletti, F.: The Monge problem for distance cost in geodesic spaces. Commun. Math. Phys. 318, 615–673 (2013)
20. Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. 44(4), 375–417 (1991)
21. Burago, D., Burago, Y., Ivanov, S.: A course in metric geometry. Graduate Studies in Mathematics, vol. 33, pp. xiv+415. American Mathematical Society, Providence, RI (2001)
22. Cavalletti, F.: Monge problem in metric measure spaces with Riemannian curvature-dimension condition. Nonlinear Anal. 99, 136–151 (2014)
23. Cavalletti, F.: Decomposition of geodesics in the Wasserstein space and the globalization property. Geom. Funct. Anal. 24, 493–551 (2014)
24. Cavalletti, F.: An overview of $L^1$ optimal transportation on metric measure spaces. In: Measure Theory in Non-Smooth Spaces. Partial Differ. Equ. Meas. Theory, pp. 98–144. De Gruyter Open, Warsaw (2017)
25. Cavalletti, F., Huesmann, M.: Self-intersection of optimal geodesics. Bull. Lond. Math. Soc. 46, 653–656 (2014)
26. Cavalletti, F., Mondino, A.: Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. Invention. Math. 208(3), 803–849 (2017)
27. Cavalletti, F., Mondino, A.: Sharp geometric and functional inequalities in metric measure spaces with lower Ricci curvature bounds. Geom. Topol. 21(1), 603–645 (2017)
28. Cheeger, J.: Differentiability of Lipschitz functions on metric measure spaces. Geom. Funct. Anal. 3(9), 428–517 (1999)
29. Cordero-Erausquin, D.: Some applications of mass transport to Gaussian-type inequalities. Arch. Ration. Mech. Anal. 161(3), 257–269 (2002)
30. Cordero-Erausquin, D., McCann, R.J., Schmuckenschläger, M.: A Riemannian interpolation inequality à la Borell. Brascamp Lieb. Invention. Math. 146, 219–257 (2001)
31. Deng, Q., Sturm, K.-T.: Localization and tensorization properties of the curvature-dimension condition for metric measure spaces. II. J. Funct. Anal. 260, 3718–3725 (2011)
32. Erbar, M., Kuwada, K., Sturm, K.-T.: On the equivalence of the entropic curvature-dimension condition and Bochner’s inequality on metric measure spaces. Invent. Math. 201(3), 993–1071 (2015)
33. Evans, L.C.: Partial differential equations and Monge–Kantorovich mass transfer. Current developments in mathematics. 1997 (Cambridge, MA), pp. 65–126. International Press, Boston, MA (1999)
34. Evans, L.C.: Partial differential equations and Monge–Kantorovich mass transfer. Current developments in mathematics. 1997 (Cambridge, MA), pp. 65–126. International Press, Boston, MA (1999)
37. Evans, L.C., Gangbo, W.: Differential equations methods for the Monge–Kantorovich mass transfer problem. Mem. Am. Math. Soc. 137(653), viii+66 (1999)
38. Feldman, M., McCann, R.J.: Monge’s transport problem on a Riemannian manifold. Trans. Am. Math. Soc. 354(4), 1667–1697 (2002)
39. Fremlin, D.H.: Measure Theory, vol. 4. Torres Fremlin, Colchester (2006)
40. Gigli, N.: On the differential structure of metric measure spaces and applications. Mem. Am. Math. Soc. 236(1113), vi+91
41. Gigli, N., Rajala, T., Sturm, K.-T.: Optimal Maps and exponentiation on finite-dimensional spaces with Ricci curvature bounded from below. J. Geom. Anal. 26, 2914–2929 (2016)
42. Gromov, M.: Paul Lévy’s isoperimetric inequality. Preprint, I.H.E.S. (1980)
43. Gromov, M.: Metric Structures for Riemannian and Non-Riemannian Spaces. Birkhäuser, Basel (2001)
44. Gromov, M., Milman, V.D.: Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces. Compositio Math. 62(3), 263–282 (1987)
45. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex analysis and minimization algorithms. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 305. Springer, Berlin (1993)
46. Kannan, R., Lovász, L., Simonovits, M.: Isoperimetric problems for convex bodies and a localization lemma. Discrete Comput. Geom. 13(3–4), 541–559 (1995)
47. Klartag, B.: Needle decompositions in Riemannian geometry. Mem. Am. Math. Soc. 249(1180), v+77 (2017)
48. Ledoux, M.: The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs, vol. 89. American Mathematical Society, Providence, RI (2001)
49. Lott, J., Villani, C.: Hamilton–Jacobi semigroup on length spaces and applications. J. Math. Pures Appl. 88, 219–229 (2007)
50. Lott, J., Villani, C.: Weak curvature conditions and functional inequalities. J. Funct. Anal. 245(1), 311–333 (2007)
51. Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. 169(3), 903–991 (2009)
52. McCann, R.J.: A convexity principle for interacting gases. Adv. Math. 128(1), 153–179 (1997)
53. McCann, R.J., Guillen, N.: Five lectures on optimal transportation: geometry, regularity and applications. In: Analysis and Geometry of Metric Measure Spaces, CRM Proc. Lecture Notes, vol. 56, pp. 145–180. American Mathematical Society, Providence, RI (2013)
54. Miranda Jr., M.: Functions of bounded variation on “good” metric spaces. J. Math. Pures Appl. 82, 975–1004 (2003)
55. Milman, E.: Sharp isoperimetric inequalities and model spaces for the curvature-dimension-diameter condition. J. Eur. Math. Soc. (JEMS) 17(5), 1041–1078 (2015)
56. Milman, E.: Beyond traditional curvature-dimension I: new model spaces for isoperimetric and concentration inequalities in negative dimension. Trans. Am. Math. Soc. 369(5), 3605–3637 (2017)
57. Ohta, S.-I.: On the measure contraction property of metric measure spaces. Comment. Math. Helv. 82, 805–828 (2007)
58. Ohta, S.-I.: Finsler interpolation inequalities. Calc. Var. Partial Differ. Equ. 36(2), 211–249 (2009)
59. Ohta, S.-I.: Needle decompositions and isoperimetric inequalities in Finsler geometry. J. Math. Soc. Japan 70(2), 651–693 (2018)
60. Ohta, S.-I.: $(K, N)$-convexity and the curvature-dimension condition for negative $N$. J. Geom. Anal. 26(3), 2067–2096 (2016)
61. Oliver, H.W.: The exact Peano derivative. Trans. Am. Math. Soc. 76, 444–456 (1954)
62. Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. 173(2), 361–400 (2000)
63. Payne, L.E., Weinberger, H.F.: An optimal Poincaré inequality for convex domains. Arch. Ration. Mech. Anal. 5, 286–292 (1960)
64. Petrunin, A.: Alexandrov meets Lott–Villani–Sturm. Münster J. Math. 4, 53–64 (2011)
65. Rachev, S.T., Rüschendorf, L.: Mass Transportation Problem, vol. I. Probability and its Applications (New York). Springer, New York (1998)
66. Rajala, T.: Interpolated measures with bounded densities in metric spaces satisfying the curvature-dimension conditions of Sturm. J. Funct. Anal. 263, 896–924 (2012)
67. Rajala, T.: Failure of the local-to-global property for $\mathbb{CD}(K,N)$ spaces. Ann. Sc. Norm. Super. Pisa Cl. Sci. 16, 45–68 (2016)
68. Rajala, T., Sturm, K.-T.: Non-branching geodesics and optimal maps in strong $\mathbb{CD}(K,\infty)$-spaces. Calc. Var. Partial Differ. Equ. 50, 831–846 (2014)
69. von Renesse, M.-K.: On local Poincaré via transportation. Math. Z. 259, 21–31 (2008)
70. von Renesse, M.-K., Sturm, K.-T.: Transport inequalities, gradient estimates, entropy and Ricci curvature. Commun. Pure Appl. Math. 58, 923–940 (2005)
71. Schneider, R.: Convex Bodies: The Brunn–Minkowski Theory. Encyclopedia of Mathematics and its Applications, vol. 44. Cambridge University Press, Cambridge (1993)
72. Srivastava, S.M.: A Course on Borel Sets. Graduate Texts in Mathematics. Springer, Berlin (1998)
73. Sturm, K.-T.: On the geometry of metric measure spaces. I. Acta Math. 196(1), 65–131 (2006)
74. Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. 196(1), 133–177 (2006)
75. Urbas, J.: Mass Transfer Problems. Lecture notes, University of Bonn (1998)
76. Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence, RI (2003)
77. Villani, C.: Optimal Transport—Old and New. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338. Springer, Berlin (2009)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.