Optimal transport and timelike lower Ricci curvature bounds on Finsler spacetimes

Mathias Braun*, Shin-ichi Ohta†

January 2024

Abstract We prove that a Finsler spacetime endowed with a smooth reference measure whose induced weighted Ricci curvature Ric_N is bounded from below by a real number K in every timelike direction satisfies the timelike curvature-dimension condition $\mathrm{TCD}_q(K,N)$ for all $q \in (0,1)$. The converse and a nonpositive-dimensional version ($N \leq 0$) of this result are also shown. Our discussion is based on the solvability of the Monge problem with respect to the q-Lorentz–Wasserstein distance as well as the characterization of q-geodesics of probability measures. One consequence of our work is the sharp timelike Brunn–Minkowski inequality in the Lorentz–Finsler case.

1 Introduction

The investigation of lower Ricci curvature bounds on generalizations of Lorentzian manifolds is making breathtaking progress in recent time. Most importantly, McCann [45] pioneered the equivalence between the lower weighted timelike Ricci curvature bound and the convexity of an entropy functional in terms of optimal transport theory (see [52] as well), and then in [21] Cavalletti–Mondino developed the synthetic theory of non-smooth measured Lorentzian spaces with timelike Ricci curvature bounded below on top of the framework of Lorentzian length spaces introduced by Kunzinger–Sämann [37] (see also [13, 51] for recent alternatives to either approaches).

This development is reminiscent of (and motivated by) the highly successful theory of *curvature-dimension condition* $\mathrm{CD}(K,N)$ for metric measure spaces developed by Lott, Sturm, Villani, and many others (we refer to [65] for a good overview). This condition is defined by a convexity property of an entropy functional along geodesics in the Wasserstein space of probability measures (in other words, along optimal transports with respect to a specific cost function). For a weighted Riemannian manifold $(M, g, \mathrm{e}^{-\psi}\operatorname{vol}_g)$, where ψ is a weight

*Department of Mathematics, University of Toronto Bahen Centre, Room 6290, Toronto, Ontario M5S 2E4, Canada. braun@math.toronto.edu

†Department of Mathematics, Osaka University, Osaka 560-0043, Japan. s.ohta@math.sci.osaka-u.ac.jp

function, $\mathrm{CD}(K,N)$ is equivalent to the lower bound $\mathrm{Ric}_N \geq K$ for the weighted Bakry–Émery Ricci tensor Ric_N [7]. Thus, a metric measure space satisfying $\mathrm{CD}(K,N)$ can be regarded as a non-smooth space with $\mathrm{Ric}_N \geq K$.

In the Lorentzian setting, the characterizations [45, 52] of McCann and Mondino–Suhr yield in particular a non-smooth analogue of the Hawking–Penrose strong energy condition [32, 34]. This property is regarded as the reason for which gravity is attractive, and not repulsive [19]. A second physically relevant consequence is the prediction of singularities in the presence of e.g. trapped surfaces. In fact, through their approach Cavalletti–Mondino obtained a synthetic Hawking singularity theorem which generalizes classical [32, 33] and recent [20, 27, 30, 38] results for high and low regularity spacetimes, cf. the recent work [15] of Braun–Calisti. This shows the strength of these tools and makes them interesting both in smooth and non-smooth settings.

In the Lorentzian framework of *weighted Finsler spacetimes*, some singularity theorems, comparison theorems, and a timelike splitting theorem have been established by Lu–Minguzzi–Ohta in [40, 41, 42]. As seen before, such results are implied by timelike curvature bounds. It is thus natural to expect that the known characterization of the lower weighted timelike Ricci curvature bound $\mathrm{Ric}_N \geq K$ by the *timelike curvature-dimension condition* $\mathrm{TCD}(K,N)$ (here we suppress the dependence on a parameter $q \in (0,1)$ for simplicity) can be generalized to weighted Finsler spacetimes (equipped with an appropriate reference measure).

The aim of this paper is to establish this equivalence (Theorems 5.9, 6.1), followed by the timelike Brunn–Minkowski inequality as a geometric application (Theorem 5.11). Our characterization not only covers $N \in [n,+\infty]$, but also the case of non-positive "dimension" $N \in (-\infty,0]$ (going back to [56, 57] in the Riemannian or Finsler framework), which is new even in the Lorentzian case. We refer to [23, 66, 67] for some works on this range of N.

Our characterization of $\mathrm{Ric}_N \geq K$ by $\mathrm{TCD}(K,N)$ has two major contributions. On the one hand, it enlarges the availability of $\mathrm{TCD}(K,N)$ to Lorentz–Finsler manifolds; the "sharp" TCD condition from Definition 5.7, in contrast to its entropic or reduced counterparts, is new even in the genuine Lorentzian setting. Lorentz–Finsler manifolds form a much wider class than Lorentzian manifolds and allow for more room in physical interpretations. In fact, they play a key role in several approaches to gravity, e.g. fields in media, or models for dark matter or dark energy, and have resolved some important difficulties; see [10, 35, 43, 47, 64] and the references therein. Then the validity of $\mathrm{TCD}(K,N)$ is obviously meaningful progress. On the other hand, this validity strongly motivates the study of some appropriate "Lorentzian" condition in addition to $\mathrm{TCD}(K,N)$ in order to discuss genuinely Lorentzian properties, namely an analogue of the *Riemannian curvature-dimension condition* $\mathrm{RCD}(K,N)$ for metric measure spaces in the Lorentzian context. One motivation behind the introduction of $\mathrm{RCD}(K,N)$ was that Finsler manifolds can satisfy $\mathrm{CD}(K,N)$. Precisely, the equivalence between $\mathrm{CD}(K,N)$ and $\mathrm{Ric}_N \geq K$ holds true also in the Finsler context, as shown by Ohta [55, 58]. To discuss genuinely Riemannian properties such as isometric splitting theorems, as obtained by Gigli in [28], the linearity of the heat flow was added to $\mathrm{CD}(K,N)$ by Gigli [29], see also [4, 5, 25]. On

Finsler manifolds, this linearity is equivalent to the manifold being Riemannian.

Besides the above characterization of $\mathrm{TCD}(K,N)$, we show that synthetic timelike sectional curvature bounds (from below or above) in terms of triangle comparison [2, 31, 37, 51] rule out non-Lorentzian Lorentz–Finsler manifolds (Proposition 2.8). Moreover, we provide a simpler proof of the absolute continuity of intermediate measures along optimal transports following the idea of Figalli–Juillet [26] relying on the measure contraction property.

This article is organized as follows. In Section 2 we review the basics of Lorentz–Finsler geometry (including variation formulae of arclength and the invalidity of triangle comparison). Section 3 is concerned with the Lorentz–Finsler distance function; we especially show the semi-convexity and the failure of semi-concavity on timelike cut loci. In Section 4 we introduce the q-Lorentz–Wasserstein distance for $q \in (0,1]$ and study the associated optimal transports (q-geodesics). The useful notion of q-separation will be introduced following McCann [45]. Then Section 5 is devoted to the first half of our main result, namely we establish that $\mathrm{Ric}_N \geq K$ in the timelike directions implies $\mathrm{TCD}(K,N)$. We also show the timelike Brunn–Minkowski inequality as an application. Finally, in Section 6 we show the converse implication from $\mathrm{TCD}(K,N)$ to $\mathrm{Ric}_N \geq K$.

Acknowledgments MB was supported by the Fields Institute for Research in Mathematical Sciences. He thanks Osaka University for its kind hospitality during the workshop "Geometry and Probability 2022". SO was supported in part by JSPS Grant-in-Aid for Scientific Research (KAKENHI) 19H01786, 22H04942. He thanks the Fields Institute for its kind hospitality during the workshop "Aspects of Ricci Curvature Bounds". The authors are sincerely grateful to Robert McCann and Ettore Minguzzi for several helpful discussions about this project, and to the Erwin Schrödinger International Institute for Mathematics and Physics for its kind hospitality during the workshop "Non-regular Spacetime Geometry".

2 Finsler spacetimes

We review the theory of Finsler spacetimes required to develop our program. We refer the readers to [11, 33, 59] for the basics of Lorentzian geometry, and to the works [46, 48, 49, 50] of Minguzzi for some generalizations including Lorentz–Finsler manifolds.

Throughout the paper, let M be a connected C^∞-manifold without boundary of dimension $\dim M = n \geq 2$. Given a local coordinate system $(x^\alpha)_{\alpha=1}^n$ on an open set $U \subset M$, we will represent a tangent vector $v \in T_xM$ with $x \in U$ as $v = \sum_{\alpha=1}^n v^\alpha \, (\partial/\partial x^\alpha)|_x$.

2.1 Preliminaries for Lorentz–Finsler manifolds

The following notations follow the conventions used in Lu–Minguzzi–Ohta [40, 41, 42] except that $\dim M$ has been set to $n+1$ therein.

Definition 2.1 (Lorentz–Finsler structure)**.** A *Lorentz–Finsler structure* of M is a function $L : TM \longrightarrow \mathbb{R}$ satisfying the following conditions.

(1) $L \in C^\infty(TM \setminus 0)$, where 0 denotes the zero section.

(2) $L(c\,v) = c^2\,L(v)$ for all $v \in TM$ and all $c > 0$.

(3) For any $v \in TM \setminus 0$, the symmetric matrix

$$\big(g_{\alpha\beta}(v)\big)_{\alpha,\beta=1}^{n} := \left(\frac{\partial^2 L}{\partial v^\alpha \partial v^\beta}(v)\right)_{\alpha,\beta=1}^{n} \tag{2.1}$$

is non-degenerate with signature $(-,+,\dots,+)$.

Then we call (M, L) a *Lorentz–Finsler manifold*.

We say that (M, L) is *reversible* if $L(-v) = L(v)$ for all $v \in TM$. Given any $v \in T_xM \setminus \{0\}$, we define a Lorentzian metric g_v on T_xM by using (2.1) as

$$g_v\left(\sum_{\alpha=1}^{n} a^\alpha \left.\frac{\partial}{\partial x^\alpha}\right|_x, \sum_{\beta=1}^{n} b^\beta \left.\frac{\partial}{\partial x^\beta}\right|_x\right) := \sum_{\alpha,\beta=1}^{n} g_{\alpha\beta}(v)\,a^\alpha\,b^\beta. \tag{2.2}$$

Then we have $g_v(v, v) = 2\,L(v)$ by Euler's homogeneous function theorem.

A tangent vector $v \in TM$ is said to be *timelike* (or *null*, respectively) if $L(v) < 0$ (or $L(v) = 0$, respectively). We say that v is *lightlike* if it is null and nonzero, and *causal* (or *non-spacelike*) if it is timelike or lightlike (i.e. $L(v) \le 0$ and $v \neq 0$). *Spacelike* vectors are those for which $L(v) > 0$ or $v = 0$. We denote by $\Omega'_x \subset T_xM$ the set of timelike vectors, and we set $\Omega' := \bigcup_{x\in M} \Omega'_x$. For later use, we define, for causal vectors v,

$$F(v) := \sqrt{-2\,L(v)} = \sqrt{-g_v(v, v)}. \tag{2.3}$$

Next we introduce the covariant derivative and the Ricci curvature. Given any $v \in TM \setminus 0$ and any $\alpha, \beta, \delta = 1, 2, \dots, n$, define

$$\overline{\Gamma}^\alpha_{\beta\delta}(v) := \frac{1}{2}\sum_{\lambda=1}^{n} g^{\alpha\lambda}(v)\left(\frac{\partial g_{\lambda\delta}}{\partial x^\beta}(v) + \frac{\partial g_{\beta\lambda}}{\partial x^\delta}(v) - \frac{\partial g_{\beta\delta}}{\partial x^\lambda}(v)\right),$$

where $(g^{\alpha\beta}(v))^n_{\alpha,\beta=1}$ is the inverse matrix of $(g_{\alpha\beta}(v))^n_{\alpha,\beta=1}$,

$$G^\alpha(v) := \frac{1}{2}\sum_{\beta,\delta=1}^{n} \overline{\Gamma}^\alpha_{\beta\delta}(v)\,v^\beta\,v^\delta, \qquad N^\alpha_\beta(v) := \frac{\partial G^\alpha}{\partial v^\beta}(v),$$

and

$$\begin{aligned}\Gamma^\alpha_{\beta\delta}(v) := \overline{\Gamma}^\alpha_{\beta\delta}(v) - \frac{1}{2}\sum_{\lambda,\mu=1}^{n} g^{\alpha\lambda}(v)\Bigg(&\frac{\partial g_{\lambda\delta}}{\partial v^\mu}(v)\,N^\mu_\beta(v) \\ &+ \frac{\partial g_{\beta\lambda}}{\partial v^\mu}(v)\,N^\mu_\delta(v) - \frac{\partial g_{\beta\delta}}{\partial v^\mu}(v)\,N^\mu_\lambda(v)\Bigg).\end{aligned} \tag{2.4}$$

Then the *covariant derivative* of a vector field $X = \sum_{\alpha=1}^{n} X^\alpha \, (\partial/\partial x^\alpha)$ in the direction $v \in T_xM$ with reference vector $w \in T_xM \setminus \{0\}$ is defined as

$$D_v^w X := \sum_{\alpha,\beta=1}^{n} \left(v^\beta \, \frac{\partial X^\alpha}{\partial x^\beta}(x) + \sum_{\delta=1}^{n} \Gamma^\alpha_{\beta\delta}(w) \, v^\beta \, X^\delta(x) \right) \frac{\partial}{\partial x^\alpha}\bigg|_x .$$

We remark that the functions $\Gamma^\alpha_{\beta\delta}$ in (2.4) are the coefficients of the *Chern connection* (or *Chern–Rund connection*). In the Lorentzian case, $g_{\alpha\beta}$ is constant in each tangent space (thus $\Gamma^\alpha_{\beta\delta} = \overline{\Gamma}^\alpha_{\beta\delta}$) and the covariant derivative does not depend on the choice of a reference vector.

The *geodesic equation* for L is written as $D^{\dot\gamma}_{\dot\gamma}\dot\gamma = 0$. A C^∞-vector field J along a geodesic γ is called a *Jacobi field* if it satisfies $D^{\dot\gamma}_{\dot\gamma} D^{\dot\gamma}_{\dot\gamma} J + R_{\dot\gamma}(J) = 0$. Here the *curvature tensor* is defined by setting, for $v, w \in T_xM$,

$$R_v(w) := \sum_{\alpha,\beta=1}^{n} R^\alpha_\beta(v) \, w^\beta \, \frac{\partial}{\partial x^\alpha},$$

where

$$R^\alpha_\beta(v) := 2\,\frac{\partial G^\alpha}{\partial x^\beta}(v) - \sum_{\delta=1}^{n} \left(\frac{\partial N^\alpha_\beta}{\partial x^\delta}(v)\, v^\delta - 2\,\frac{\partial N^\alpha_\beta}{\partial v^\delta}(v)\, G^\delta(v) \right) - \sum_{\delta=1}^{n} N^\alpha_\delta(v)\, N^\delta_\beta(v).$$

A Jacobi field is also characterized as the variational vector field of a geodesic variation. Note that $R_v(w)$ is linear in w, thereby $R_v \colon T_xM \longrightarrow T_xM$ is an endomorphism for each $v \in T_xM$.

Definition 2.2 (Ricci curvature)**.** For $v \in T_xM$, we define the *Ricci curvature* (or *Ricci scalar*) of v as the trace of R_v, i.e. $\mathrm{Ric}(v) := \mathrm{trace}(R_v)$.

We remark that $\mathrm{Ric}(c\,v) = c^2\,\mathrm{Ric}(v)$ for every $c > 0$. For later use, we also recall that

$$R_v(v) = 0, \qquad g_v\big(w_1, R_v(w_2)\big) = g_v\big(R_v(w_1), w_2\big) \tag{2.5}$$

for any $v \in \Omega'_x$ and $w_1, w_2 \in T_xM$ (see [48, Proposition 2.4], [40, Proposition 3.4]).

2.2 Preliminaries for Finsler spacetimes

Definition 2.3 (Finsler spacetime)**.** If a Lorentz–Finsler manifold (M, L) admits a smooth timelike vector field X, then (M, L) is said to be *time oriented* (by X). A time oriented Lorentz–Finsler manifold is called a *Finsler spacetime*.

In such a Finsler spacetime, a causal vector $v \in T_xM$ is said to be *future-directed* if it lies in the same connected component of $\overline{\Omega}'_x \setminus \{0\}$ as $X(x)$. We denote by $\Omega_x \subset \Omega'_x$ the set of future-directed timelike vectors, and define

$$\Omega := \bigcup_{x\in M} \Omega_x, \qquad \overline{\Omega} := \bigcup_{x\in M} \overline{\Omega}_x, \qquad \overline{\Omega} \setminus 0 := \bigcup_{x\in M} (\overline{\Omega}_x \setminus \{0\}).$$

A C^1-curve in (M, L) is said to be *timelike* (or *causal*, respectively) if its tangent vector is always timelike (or causal, respectively). All causal curves will be future-directed unless otherwise indicated.

We recall some fundamental concepts in causality theory. Given $x, y \in M$, we write $x \ll y$ (or $x < y$, respectively) if there is a future-directed timelike (or causal, respectively) curve from x to y, and $x \leq y$ means that $x = y$ or $x < y$. Then the *chronological past* and the *chronological future* of x are defined by

$$I^-(x) := \{y \in M \mid y \ll x\}, \qquad I^+(x) := \{y \in M \mid x \ll y\},$$

and its *causal past* and its *causal future* are defined by

$$J^-(x) := \{y \in M \mid y \leq x\}, \qquad J^+(x) := \{y \in M \mid x \leq y\}.$$

For a general set $S \subset M$, we define $I^-(S)$, $I^+(S)$, $J^-(S)$, and $J^+(S)$ analogously. For a Borel probability measure μ on M, we abbreviate $I^-(\mu) := I^-(\operatorname{spt}\mu)$ in terms of its support $\operatorname{spt}\mu \subset M$, and analogously for $I^+(\mu)$, $J^-(\mu)$, and $J^+(\mu)$.

Definition 2.4 (Causality conditions)**.** Let (M, L) be a Finsler spacetime.

(1) (M, L) is said to be *chronological* if $x \notin I^+(x)$ for all $x \in M$.

(2) We say that (M, L) is *causal* if there is no closed causal curve.

(3) (M, L) is said to be *strongly causal* if, for all $x \in M$, every neighborhood U of x contains another neighborhood V of x such that no causal curve intersects V more than once.

(4) We say that (M, L) is *globally hyperbolic* if it is strongly causal and, for any $x, y \in M$, $J^+(x) \cap J^-(y)$ is compact.

We also define the *Lorentz–Finsler distance* $l(x, y)$ for $x, y \in M$ (also called the *time separation function*) by

$$l(x, y) := \sup_{\gamma} \mathfrak{L}(\gamma), \qquad \mathfrak{L}(\gamma) := \int_0^1 F\big(\dot\gamma(t)\big)\,\mathrm{d}t,$$

where $\gamma\colon [0, 1] \longrightarrow M$ runs over all causal curves from x to y (recall (2.3) for the definition of F). A causal curve attaining the above supremum is said to be *maximizing*. Similarly to [45], we set $l(x, y) := -\infty$ if there is no causal curve from x to y (namely $x \not< y$). (Note that we set $l(x, y) = 0$ for $x \not< y$ in [40, 41, 42].) In general, the distance function l is only lower semi-continuous on $\{l \geq 0\}$ [48, Proposition 6.7] and can be infinite. In globally hyperbolic Finsler spacetimes, l is finite and continuous on $\{l \geq 0\}$ and upper semi-continuous on $M \times M$ [48, Proposition 6.8], and any pair of points $x, y \in M$ with $x < y$ admits a maximizing geodesic from x to y [48, Proposition 6.9]. More details about the regularity of l under global hyperbolicity are deferred to Section 3.

To consider the dual structure to L and the Legendre transform (see [47], [49, §3.1], or [41, §4.4] for further discussions), define the *polar cone* to Ω_x by

$$\Omega_x^* := \big\{\zeta \in T_x^*M \mid \zeta(v) < 0 \text{ for all } v \in \overline{\Omega}_x \setminus \{0\}\big\}.$$

This is an open convex cone in T_x^*M. For $\zeta \in \Omega_x^*$, we define

$$L^*(\zeta) := -\frac{1}{2}\left(\sup \zeta(v)\right)^2 = -\frac{1}{2}\inf\bigl(\zeta(v)\bigr)^2,$$

where the supremum and the infimum run over $v \in F^{-1}(1) \cap \Omega_x$. Then we have the *reverse Cauchy–Schwarz inequality*

$$4\,L^*(\zeta)\,L(v) \le \bigl(\zeta(v)\bigr)^2 \tag{2.6}$$

for every $v \in \Omega_x$ and every $\zeta \in \Omega_x^*$, as well as the following variational definition of the Legendre transform.

Definition 2.5 (Legendre transform)**.** The *Legendre transform* $\mathscr{L}^*\colon \Omega_x^* \longrightarrow \Omega_x$ is defined as the map sending $\zeta \in \Omega_x^*$ to the unique element $v \in \Omega_x$ satisfying $L(v) = L^*(\zeta) = \zeta(v)/2$. We also define $\mathscr{L}^*(0) := 0$.

A coordinate expression of the Legendre transform is given by

$$\mathscr{L}^*(\zeta) = \sum_{\alpha=1}^{n} \frac{\partial L^*}{\partial \zeta_\alpha}(\zeta)\,\frac{\partial}{\partial x^\alpha}, \qquad \text{where } \zeta = \sum_{\alpha=1}^{n} \zeta_\alpha\,\mathrm{d}x^\alpha.$$

The inverse map $\mathscr{L} := (\mathscr{L}^*)^{-1}$ is written in the same way as

$$\mathscr{L}(v) = \sum_{\alpha=1}^{n} \frac{\partial L}{\partial v^\alpha}(v)\,\mathrm{d}x^\alpha.$$

We refer to Lu–Minguzzi–Ohta [41] and Minguzzi [47, 49] for some basic properties of the Legendre transforms.

2.3 Variations of arclength

Here we present variation formulas for arclength in the Lorentz–Finsler setting. Although they are obtained by the standard arguments, we give outlines for completeness. We refer to O'Neill's book [59] for the semi-Riemannian case (including Lorentzian manifolds) and to Bao–Chern–Shen [9, Chapter 5] or Ohta [58, Chapter 7] for the Finsler case.

Let $\sigma\colon [a,b] \times (-\varepsilon,\varepsilon) \longrightarrow M$ be a C^∞-variation such that $\sigma_s(t) := \sigma(t,s)$ is a timelike curve for each $s \in (-\varepsilon,\varepsilon)$. We set

$$\mathfrak{L}(s) := \mathfrak{L}(\sigma_s) = \int_a^b F(\dot\sigma_s)\,\mathrm{d}t.$$

For brevity, we define the tangent and variation vector fields of σ as $T := \partial_t\sigma$ and $V := \partial_s\sigma$, respectively. We assume that T vanishes nowhere.

Lemma 2.6 (First variation)**.** *For σ and $\mathfrak{L}$ as above, we have*

$$\begin{aligned}\mathfrak{L}'(s) &= -\left[g_T\left(V, \frac{T}{F(T)}\right)(\cdot, s)\right]_a^b + \int_a^b g_T\left(V, D_T^T\left[\frac{T}{F(T)}\right]\right)(t,s)\,\mathrm{d}t \\ &= -\int_a^b g_T\left(D_T^T V, \frac{T}{F(T)}\right)(t,s)\,\mathrm{d}t.\end{aligned}$$

In particular, if σ_s is a geodesic, then

$$\mathfrak{L}'(s) = -\frac{1}{F(\dot\sigma_s)}\left[g_T(V,T)(\cdot,s)\right]_a^b.$$

Proof. This is shown in the same way as the Finsler case, cf. e.g. [58, Proposition 7.1], by recalling $F^2(v) = -g_v(v,v)$ for every $v \in \Omega$ in (2.3). □

To write down the second variation, we define the *index form* for C^1-vector fields V, W along a timelike geodesic $\gamma\colon [a,b] \longrightarrow M$ as

$$I(V,W) := -\frac{1}{F(\dot\gamma)}\int_a^b \left(g_{\dot\gamma}(D_{\dot\gamma}^{\dot\gamma}V, D_{\dot\gamma}^{\dot\gamma}W) - g_{\dot\gamma}\big(R_{\dot\gamma}(V),W\big)\right)\mathrm{d}t,$$

see Beem–Ehrlich–Easley's book [11, Chapter 10] for the Lorentzian case. Note that $I(V,W) = I(W,V)$ and, if V is a Jacobi field, then we have

$$I(V,W) = -\frac{1}{F(\dot\gamma)}\left[g_{\dot\gamma}(D_{\dot\gamma}^{\dot\gamma}V, W)\right]_a^b. \tag{2.7}$$

Proposition 2.7 (Second variation)**.** *For σ and $\mathfrak{L}$ as above, suppose that $\gamma := \sigma_0$ is a geodesic. Then we have*

$$\begin{aligned}\mathfrak{L}''(0) &= I(V_0,V_0) - \left[g_{\dot\gamma}\left(D_V^T V(\cdot,0), \frac{\dot\gamma}{F(\dot\gamma)}\right)\right]_a^b - \int_a^b \frac{(\partial_s[F(T)](t,0))^2}{F(\dot\gamma(t))}\,\mathrm{d}t \\ &= I(V_0^\perp, V_0^\perp) - \left[g_{\dot\gamma}\left(D_V^T V(\cdot,0), \frac{\dot\gamma}{F(\dot\gamma)}\right)\right]_a^b,\end{aligned}$$

where we set $V_0 := V(\cdot,0)$ and

$$V_0^\perp := V_0 + g_{\dot\gamma}\left(V_0, \frac{\dot\gamma}{F(\dot\gamma)}\right)\frac{\dot\gamma}{F(\dot\gamma)}$$

is the projection of V_0 to the $g_{\dot\gamma}$-orthogonal subspace to $\dot\gamma$.

Proof. We can again argue similarly to the Finsler case by taking care of the signature, let us explain along the lines of [58, Theorem 7.6] and [9, §5.2]. By differentiating $\mathfrak{L}'(s)$ in Lemma 2.6 and by $F^2(T) = -g_T(T,T)$, we find

$$\begin{aligned}\mathfrak{L}''(0) = -\frac{1}{F(\dot\gamma)}\int_a^b &\left(g_{\dot\gamma}\big(D_V^T[D_T^T V], \dot\gamma\big) + g_{\dot\gamma}\big(D_{\dot\gamma}^{\dot\gamma}V_0, D_{\dot\gamma}^{\dot\gamma}V_0\big)\right)\mathrm{d}t \\ &- \int_a^b \frac{(\partial_s[F(T)](t,0))^2}{F(\dot\gamma(t))}\,\mathrm{d}t.\end{aligned} \tag{2.8}$$

Then, choosing local coordinates $(x^\alpha)_{\alpha=1}^n$ such that $\dot\gamma = \partial/\partial x^1$, $g_{11}(\dot\gamma) = -1$, $g_{ii}(\dot\gamma) = 1$ for $i = 2, \dots, n$, and $g_{\alpha\beta}(\dot\gamma) = 0$ for distinct $\alpha, \beta = 1, \dots, n$, we have

$$g_{\dot\gamma}\big(D_V^T[D_T^T V] - D_T^T[D_V^T V], \dot\gamma\big) = -\sum_{\alpha,\beta=1}^n R^1_{\alpha\beta 1}(\dot\gamma)\, V_0^\alpha\, V_0^\beta.$$

Now we apply the skew-symmetry [58, (5.11)] to see

$$\sum_{\alpha,\beta=1}^n R^1_{\alpha\beta 1}(\dot\gamma)\, V_0^\alpha\, V_0^\beta = \sum_{\alpha,\beta,\delta=1}^n g_{\alpha\delta}(\dot\gamma)\, R^\delta_\beta(\dot\gamma)\, V_0^\alpha\, V_0^\beta = g_{\dot\gamma}\big(R_{\dot\gamma}(V_0), V_0\big)$$

(thus [58, (7.2)] holds as it is). Plugging these into (2.8) yields the first equation.

The second equation can be derived following [9, Exercise 5.2.6]. Precisely, we have

$$g_{\dot\gamma}(D_{\dot\gamma}^{\dot\gamma} V_0^\perp, D_{\dot\gamma}^{\dot\gamma} V_0^\perp) = g_{\dot\gamma}(D_{\dot\gamma}^{\dot\gamma} V_0, D_{\dot\gamma}^{\dot\gamma} V_0) + \big(\partial_s[F(T)](\cdot, 0)\big)^2$$

by calculation and

$$g_{\dot\gamma}\big(R_{\dot\gamma}(V_0^\perp), V_0^\perp\big) = g_{\dot\gamma}\big(R_{\dot\gamma}(V_0), V_0\big)$$

by [40, Proposition 3.4]. This completes the proof. □

2.4 Invalidity of triangle comparisons

Besides Ricci curvature bounds, synthetic lower and upper sectional curvature bounds can be formulated in terms of triangle comparisons, akin to celebrated theories of Alexandrov spaces and CAT(k)-spaces arising from the Alexandrov–Toponogov triangle comparison theorems in Riemannian geometry, cf. e.g. Burago–Burago–Ivanov [18]. However, similarly to Alexandrov spaces and CAT(k)-spaces, those curvature bounds by triangle comparisons rule out genuinely non-Lorentzian Lorentz–Finsler structures. This means that, in particular, the splitting theorem in Beran–Ohanyan–Rott–Solis [12] is not applicable to this framework.

A Lorentzian analogue to the triangle comparison property can be given as follows along [37, Definition 4.7] by Kunzinger–Sämann. We denote by $\mathbb{M}_k^2$ the two-dimensional model space of constant sectional curvature $k \in \mathbb{R}$, and by $\bar l$ the distance function on it [37, Page 426]. When $k = 0$, this is simply the two-dimensional Minkowski spacetime $\mathbb{R}^2$. Then, for any $a, b, c > 0$ with $a + b \le c$, additionally satisfying $c < \pi/\sqrt{k}$ (or $c < \pi/\sqrt{-k}$, respectively) if $a + b = c$ and $k > 0$ (or $a + b < c$ and $k < 0$, respectively), there exist $\bar x, \bar y, \bar z \in \mathbb{M}_k^2$ such that $\bar x \ll \bar y \ll \bar z$ and

$$\bar l(\bar x, \bar y) = a, \qquad \bar l(\bar y, \bar z) = b, \qquad \bar l(\bar x, \bar z) = c.$$

(See the realizability lemmas in [2, Lemma 2.1] and [37, Lemma 4.6]). Such a triangle $\bar\triangle(\bar x, \bar y, \bar z)$ is unique up to isometry and called a *comparison triangle* of $\triangle(x, y, z)$ in (M, L) when $l(x, y) = a$, $l(y, z) = b$, and $l(x, z) = c$.

We say that (M, L) has *timelike curvature bounded below by $k \in \mathbb{R}$ in the sense of triangle comparison* if every point in M admits a neighborhood U satisfying the following: For any $x, y, z \in U$ with $x \ll y \ll z$, a comparison triangle $\triangle(\bar{x}, \bar{y}, \bar{z}) \subset \mathbb{M}^2_k$ and for any p, q on (the sides of) $\triangle(x, y, z)$ and the corresponding points $\bar{p}, \bar{q}$ on (the sides of) $\triangle(\bar{x}, \bar{y}, \bar{z})$, we have $l(p, q) \le \bar{l}(\bar{p}, \bar{q})$. By reversing the latter inequality, we define (M, L) to have *timelike curvature bounded above by k in the sense of triangle comparison*.

The investigation of triangle comparison theorems in the Lorentzian setting goes back to Harriss [31]. We refer to Alexander–Bishop [2] for the equivalence between triangle comparison properties and the corresponding sectional curvature bounds for general semi-Riemannian manifolds (see also [37, Example 4.9]).

Proposition 2.8 (Triangle comparison implies Lorentzianity)**.** *Let (M, L) be globally hyperbolic. If (M, L) has timelike curvature bounded below or above by some $k \in \mathbb{R}$ in the sense of triangle comparison, then $g_v = 2L$ holds on Ω_x for every $v \in \Omega_x$ and every $x \in M$. In particular, L is necessarily quadratic in Ω_x for every $x \in M$.*

Proof. We first consider the lower curvature bound. Fix arbitrary $x \in M$ and $v, w \in \Omega_x$. By taking a scaling limit of the triangle comparison inequality on M, we find that the curvature bound also holds for triangles in the tangent space T_xM with $k = 0$ (i.e. (T_xM, L) is nonnegatively curved). Now, as a comparison triangle for $\triangle(0, v, v + w)$, take $\triangle(0, \bar{v}, \bar{v} + \bar{w})$ in the Minkowski space $(\mathbb{M}^2_0, \langle \cdot, \cdot \rangle)$. Then for every $t \in (0, 1)$ we obtain

$$\langle \bar{v} + t\,\bar{w}, \bar{v} + t\,\bar{w} \rangle = (1 - t)\,\langle \bar{v}, \bar{v} \rangle + t\,\langle \bar{v} + \bar{w}, \bar{v} + \bar{w} \rangle - (1 - t)\,t\,\langle \bar{w}, \bar{w} \rangle.$$

Therefore, it follows from the hypothesis that

$$\begin{aligned} 2\,L(v + t\,w) &\ge (1 - t)\,\langle \bar{v}, \bar{v} \rangle + t\,\langle \bar{v} + \bar{w}, \bar{v} + \bar{w} \rangle - (1 - t)\,t\,\langle \bar{w}, \bar{w} \rangle \\ &= 2\,(1 - t)\,L(v) + 2\,t\,L(v + w) - 2\,(1 - t)\,t\,L(w). \end{aligned}$$

On the other hand, considering the triangle $\triangle(0, t\,w, v + t\,w)$ and noticing

$$t\,w + t\,\big((v + t\,w) - t\,w\big) = t\,(v + w),$$

we observe

$$L\big(t\,(v + w)\big) \ge (1 - t)\,L(t\,w) + t\,L(v + t\,w) - (1 - t)\,t\,L(v),$$

which is rewritten as

$$L(v + t\,w) \le (1 - t)\,L(v) + t\,L(v + w) - (1 - t)\,t\,L(w).$$

This yields

$$L(v + t\,w) = (1 - t)\,L(v) + t\,L(v + w) - (1 - t)\,t\,L(w),$$

and hence we have

$$g_v(w, w) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\bigg|_0 L(v + t\,w) = 2L(w).$$

This completes the proof for the lower curvature bound.

The case of upper curvature bound can be established in the same way by reversing the inequalities. □

3 Behavior of the Lorentz–Finsler distance function

From now on, we assume (M, L) to be a globally hyperbolic Finsler spacetime. Under this standing assumption, in this section we study the behavior of the Lorentz–Finsler distance function l, and introduce the Lagrangian and the Hamiltonian induced from the q-th power of $F = \sqrt{-2\,L}$ for $q \in (0, 1]$. We follow the exposition of McCann [45, Sections 2, 3].

3.1 Lagrangians and Hamiltonians

We first introduce the *q-dependent Lagrangian* $\mathcal{L}_q$ as in McCann [45] (see Suhr [63] for the case of $q = 1$). Let $q \in (0, 1]$, and define $\mathcal{L}_q \colon TM \longrightarrow (-\infty, 0] \cup \{+\infty\}$ as

$$\mathcal{L}_q(v) := \begin{cases} -q^{-1} \left(-2\,L(v)\right)^{q/2} & \text{if } v \in \overline{\Omega}, \\ +\infty & \text{otherwise.} \end{cases}$$

Observe that $\mathcal{L}_q(v) = -q^{-1}\,F(v)^q$ for $v \in \overline{\Omega}$. Its convex dual is the *q-dependent Hamiltonian* $\mathcal{H}_q \colon T^*M \longrightarrow [0, +\infty]$ given by

$$\mathcal{H}_q(\zeta) := \sup_{v \in TM} \big(\zeta(v) - \mathcal{L}_q(v)\big), \tag{3.1}$$

where $\zeta(v)$ is the usual duality pairing. When $q \in (0, 1)$ and $p < 0$ denotes its dual exponent, i.e. $q^{-1} + p^{-1} = 1$, we have

$$\mathcal{H}_q(\zeta) = \begin{cases} -p^{-1} \left(-2\,L^*(\zeta)\right)^{p/2} & \text{if } \zeta \in \Omega^*, \\ +\infty & \text{otherwise,} \end{cases} \tag{3.2}$$

while for $q = 1$ it is the $\{0, +\infty\}$-valued indicator function of a solid hyperboloid $\{L^* \le -1/2\} \subset \Omega^*$; see Corollary 3.2.

In view of (3.1), we can show an analogue of $\mathscr{L} = (\mathscr{L}^*)^{-1}$. Furthermore, choosing $q \in (0, 1)$ makes the Lagrangian $\mathcal{L}_q$ strictly convex on Ω_x. Precisely, we have the following result.

Lemma 3.1 (Duality and convexity of $\mathcal{L}_q$)**.** *Fix $q \in (0, 1)$ and $x \in M$.*

(i) *The maps*

$$v \longmapsto \sum_{\alpha=1}^{n} \frac{\partial \mathcal{L}_q}{\partial v^\alpha}(v)\,\mathrm{d}x^\alpha, \qquad \zeta \longmapsto \sum_{\alpha=1}^{n} \frac{\partial \mathcal{H}_q}{\partial \zeta_\alpha}(\zeta)\,\frac{\partial}{\partial x^\alpha}$$

are the inverses of each other between Ω_x and Ω_x^*.

(ii) *The Lagrangian $\mathcal{L}_q$ is convex on T_xM and strictly convex on Ω_x.*

Proof. (i) Observe that

$$\sum_{\alpha=1}^{n} \frac{\partial \mathcal{L}_q}{\partial v^\alpha}(v)\,\mathrm{d}x^\alpha = F(v)^{q-2} \sum_{\alpha=1}^{n} \frac{\partial L}{\partial v^\alpha}(v)\,\mathrm{d}x^\alpha = F(v)^{q-2}\,\mathscr{L}(v),$$

and similarly

$$\sum_{\alpha=1}^{n} \frac{\partial \mathcal{H}_q}{\partial \zeta_\alpha}(\zeta)\,\frac{\partial}{\partial x^\alpha} = F^*(\zeta)^{p-2}\,\mathscr{L}^*(\zeta)$$

for $v \in \Omega_x$ and $\zeta \in \Omega_x^*$, where $F^*(\zeta) := \sqrt{-2\,L^*(\zeta)}$. Since $\mathscr{L}^* = \mathscr{L}^{-1}$ and both involved maps are positively 1-homogeneous, we obtain the claim by noting that

$$\big(F(v)^{q-1}\big)^{p-2}\,F(v)^{q-2} = 1.$$

(ii) Let $v \in \Omega_x$. Recalling the definition of $g_{\alpha\beta}(v)$ in (2.1), we observe

$$\frac{\partial \mathcal{L}_q}{\partial v^\alpha\,\partial v^\beta}(v) = F(v)^{q-4}\left((2-q)\,\frac{\partial L}{\partial v^\alpha}(v)\,\frac{\partial L}{\partial v^\beta}(v) + F(v)^2\,g_{\alpha\beta}(v)\right).$$

Now note that, by writing $w \in T_xM$ as $w = \sum_{\alpha=1}^n w^\alpha(\partial/\partial x^\alpha)|_x$,

$$\sum_{\alpha=1}^{n} \frac{\partial L}{\partial v^\alpha}(v)\,w^\alpha = g_v(v,w), \qquad \sum_{\alpha,\beta=1}^{n} g_{\alpha\beta}(v)\,w^\alpha\,w^\beta = g_v(w,w).$$

Thus, if $g_v(w,w) \geq 0$ we immediately find

$$\sum_{\alpha,\beta=1}^{n} \frac{\partial \mathcal{L}_q}{\partial v^\alpha \partial v^\beta}(v)\,w^\alpha\,w^\beta \geq 0.$$

If instead $g_v(w,w) < 0$, i.e. w is g_v-timelike, then the reverse Cauchy–Schwarz inequality for g_v yields

$$\big(g_v(v,w)\big)^2 \geq g_v(v,v)\,g_v(w,w) = -F(v)^2\,g_v(w,w).$$

Since the identity $g_v(v,w) = 0$ holds only when $w = 0$ or $g_v(w,w) > 0$, we get the positive-definiteness. Hence $\mathcal{L}_q$ is strictly convex on Ω_x and convex on T_xM (by the convexity of the set Ω_x). □

In the case of $q = 1$, note that $\mathcal{L}_1 = -F$ is positively 1-homogeneous and fails to be strictly convex in radial directions. In order to formulate the corresponding next result, recall that a function $u\colon T_xM \longrightarrow [-\infty,+\infty]$ is said to be *subdifferentiable* at $v \in T_xM$ with a *subgradient* $\zeta \in T_x^*M$ if $u(v) \in \mathbb{R}$ and

$$u(v+w) \geq u(v) + \zeta(w) + o(|w|_h)$$

holds for $w \in T_xM$ close to 0, where h is an auxiliary Riemannian metric. We denote by $\partial_\bullet u(v)$ the set of all subgradients. (See (3.7) below for the case of functions on M and the concept of superdifferentiability.) Since the subdifferentiability is a local notion, this property and subgradients do not depend on the choice of h.

Corollary 3.2 (The case of $q = 1$)**.** *Fix $x \in M$.*

(i) *The Lagrangian $\mathcal{L}_1$ is convex on T_xM.*

(ii) *We have*

$$\mathcal{H}_1(\zeta) = \begin{cases} 0 & \text{if } \zeta \in \Omega_x^* \text{ and } L^*(\zeta) \le -1/2, \\ +\infty & \text{otherwise.} \end{cases}$$

(iii) *$\mathcal{L}_1|_{T_xM}$ is not subdifferentiable on $(\mathcal{L}_1|_{T_xM})^{-1}(0) \setminus \{0\}$. At $0 \in T_xM$, $\mathcal{L}_1|_{T_xM}$ is subdifferentiable with*

$$\partial_\bullet(\mathcal{L}_1|_{T_xM})(0) = \mathcal{H}_1^{-1}(0) \cap T_x^*M.$$

Proof. (i) This follows from the convexity of $\mathcal{L}_q$ by letting $q \to 1$.

(ii) This is a standard fact following from (3.1) together with the reverse Cauchy–Schwarz inequality, or from (3.2) by letting $p \to -\infty$.

(iii) The failure of the subdifferentiability of $\mathcal{L}_1$ (or the superdifferentiability of F) on $(\mathcal{L}_1|_{T_xM})^{-1}(0) \setminus \{0\}$ is a consequence of the fact that, for $s > 0$,

$$\frac{\partial}{\partial t}\sqrt{t^2 - s^2} = \frac{t}{\sqrt{t^2 - s^2}} \to +\infty \qquad \text{as } t \downarrow s.$$

At $0 \in T_xM$, on the one hand, $\zeta \in \Omega_x^*$ with $L^*(\zeta) \le -1/2$ satisfies, by the reverse Cauchy–Schwarz inequality (2.6),

$$\zeta(v) \le -\sqrt{4\,L(v)\,L^*(\zeta)} = \mathcal{L}_1(v)\,\sqrt{-2\,L^*(\zeta)} \le \mathcal{L}_1(v)$$

for arbitrary $v \in T_xM$, which implies $\zeta \in \partial_\bullet(\mathcal{L}_1|_{T_xM})(0)$. On the other hand, if $\zeta \in \partial_\bullet(\mathcal{L}_1|_{T_xM})(0)$, then

$$\zeta(v) = \frac{\zeta(c\,v)}{c} \le \frac{\mathcal{L}_1(c\,v) + o(c\,|v|_h)}{c} \to -1 \qquad \text{as } c \downarrow 0$$

for any $v \in F^{-1}(1) \cap \Omega_x$. Thus $L^*(\zeta) \le -1/2$ by the definition of $L^*(\zeta)$. □

3.2 The q-Lorentz–Finsler distance functions

For $q \in (0, 1]$, the *q-action* of a Lipschitz curve $\gamma\colon [0, 1] \longrightarrow M$ induced by $\mathcal{L}_q$ is

$$\mathcal{A}_q(\gamma) := \int_0^1 \mathcal{L}_q(\dot\gamma)\,\mathrm{d}t,$$

and we define $l_q\colon M \times M \longrightarrow [0, +\infty] \cup \{-\infty\}$ by

$$l_q(x, y) := -\inf_\gamma \mathcal{A}_q(\gamma), \tag{3.3}$$

where $\gamma\colon [0, 1] \longrightarrow M$ runs over all Lipschitz curves with $\gamma(0) = x$ and $\gamma(1) = y$. We call l_q the *q-Lorentz–Finsler distance function*; for $q = 1$, l_1 coincides with the Lorentz–Finsler distance function l introduced in Subsection 2.2.

For notational simplicity, we will adopt the conventions:

$$(-\infty) + \infty := \infty + (-\infty) := -\infty, \qquad (-\infty)^q = (-\infty)^{1/q} := -\infty. \tag{3.4}$$

Then l satisfies the *reverse triangle inequality*

$$l(x, y) \geq l(x, z) + l(z, y), \tag{3.5}$$

and, for every $x, y \in M$ with $l(x, y) > -\infty$ and every $q \in (0, 1)$, we find

$$l(x, y) = \big(q\, l_q(x, y)\big)^{1/q} \tag{3.6}$$

(by re-parametrizing an appropriate curve γ so that $\mathcal{L}_1(\dot\gamma)$ is constant almost everywhere). In particular, the sets $\{l_q > 0\}$ and $\{l_q \geq 0\}$ do not depend on the exponent $q \in (0, 1]$.

The hypothesized global hyperbolicity of (M, L) entails finer properties of the Lorentz–Finsler distance function l. First of all, by the relation (3.6), we immediately observe the following.

Lemma 3.3 (Continuity of l_q)**.** *For each $q \in (0, 1]$, l_q is finite and continuous on $\{l_q \geq 0\}$, and upper semi-continuous on $M \times M$.*

For analyzing the behavior of l, we introduce the singular set.

Definition 3.4 (Singular set of l)**.** We define $\mathrm{sing}(l)$ as the set of all pairs $(x, y) \in M \times M$ such that $l(x, y) \leq 0$ or no maximizing geodesic contains x and y in its relative interior.

Then, thanks to Lemma 3.3, it is not difficult to show that $\mathrm{sing}(l)$ is a closed set; see McCann [45, Theorem 3.6] for instance.

The following two lemmas are shown in exactly the same way as [45, Lemmas 2.4, 2.5]. We remark that the limit curve theorems in [11] generalize to our Lorentz–Finsler framework by Minguzzi [48, Proposition 6.1].

Lemma 3.5 (Continuity of intermediate points)**.** *For every $t \in [0, 1]$ and every $(x, y) \in (M \times M) \setminus \mathrm{sing}(l)$, there exists a unique point $z \in M$, henceforth denoted by $z_t(x, y)$, such that*

$$l(x, z) = t\, l(x, y), \qquad l(z, y) = (1 - t)\, l(x, y).$$

Moreover, $z_t(x, y)$ depends smoothly on $(t, x, y) \in [0, 1] \times ((M \times M) \setminus \mathrm{sing}(l))$.

Lemma 3.6 (Compactness of intermediate point sets)**.** *For $x, y \in M$, $t \in [0, 1]$, and $E \subset M \times M$, we set*

$$Z_t(x, y) := \big\{z \in M \mid l(x, z) = t\, l(x, y),\, l(z, y) = (1 - t)\, l(x, y)\big\}$$

if $l(x, y) \geq 0$ and $Z_t(x, y) := \emptyset$ otherwise, as well as

$$Z_t(E) := \bigcup_{(x,y)\in E} Z_t(x, y), \qquad Z(E) := \bigcup_{t\in[0,1]} Z_t(E).$$

Then $Z(E)$ is precompact if E is. If E is compact, so are $Z(E)$ and $Z_t(E)$.

Remark 3.7. Take $x, y \in M$ with $l(x,y) > 0$. For $t \in (0,1)$, by concatenating timelike geodesics, note that $z \in Z_t(x,y)$ if and only if z is the t-intermediate point of a timelike maximizing geodesic from x to y. Similarly, we have $Z_0(x,y) = \{x\}$ and $Z_1(x,y) = \{y\}$ since maximizing curves do not change the causal type.

3.3 Convexity, concavity, and smoothness

Definition 3.8 (Semi-convexity and semi-concavity)**.** Let $U \subset M$ be open. A function $u\colon U \longrightarrow \mathbb{R}$ is said to be *semi-convex* if for some smooth Riemannian metric h on M there exists $c \in \mathbb{R}$ such that

$$\liminf_{w \to 0} \frac{u(\exp_x^h(w)) + u(\exp_x^h(-w)) - 2\,u(x)}{2\,|w|_h^2} \geq c$$

for every $x \in U$, where $w \in T_xM$. The largest possible such constant c is called the *semi-convexity constant* of u on U. Lastly, we call u *semi-concave* if $-u$ is semi-convex.

We remark that, if U is relatively compact, the semi-convexity is independent of the choice of h, whereas the semi-convexity constant depends on h.

Proposition 3.9 (Semi-convexity of l)**.** *Let h be a smooth Riemannian metric on M, and let $(x,y) \in \{l > 0\}$. Then*

$$c(x,y) := \liminf_{w \to 0} \frac{l(\exp_x^h(w), y) + l(\exp_x^h(-w), y) - 2\,l(x,y)}{2\,|w|_h^2}$$

is real-valued, where $w \in T_xM$. Moreover, the quantity $c(x,y)$ depends continuously on $(x,y) \in \{l > 0\}$.

Proof. Let $\gamma\colon [0,1] \longrightarrow M$ be a maximizing geodesic from x to y. Take $w \in T_xM$ with $|w|_h = 1$ and extend it to a vector field P along γ which is parallel with respect to $g_{\dot\gamma}$, namely $D_{\dot\gamma}^{\dot\gamma} P = 0$ and $P(0) = w$. Consider the vector field

$$W(t) := (1-t)P(t)$$

along γ and define a smooth variation $\sigma\colon [0,1] \times (-\varepsilon, \varepsilon) \longrightarrow M$ of γ, where $\varepsilon > 0$ is a fixed small parameter, by

$$\sigma_s(t) = \sigma(t,s) := \exp_{\gamma(t)}^h \big(sW(t)\big).$$

Note that $\sigma_0 = \gamma$ and $\sigma_s(1) = \gamma(1) = y$ for all $s \in (-\varepsilon, \varepsilon)$. Set

$$\mathfrak{L}(s) := \int_0^1 F(\dot\sigma_s)\,\mathrm{d}t = -\mathcal{A}_1(\sigma_s).$$

Then it follows from (3.3) that

$$\frac{l(\sigma_s(0), y) + l(\sigma_{-s}(0), y) - 2\,l(x,y)}{2\,s^2} \geq \frac{\mathfrak{L}(s) + \mathfrak{L}(-s) - 2\,\mathfrak{L}(0)}{2\,s^2},$$

and the right-hand side smoothly depends on x, y, $\dot\gamma(0)$, and w. Precisely, it is written by the second variation formula from Proposition 2.7 as

$$I(W^\perp, W^\perp) - \left[g_{\dot\gamma}\left(D_V^T V(\cdot, 0), \frac{\dot\gamma}{l(x,y)} \right) \right]_0^1,$$

where $T := \partial_t \sigma$, $V := \partial_s \sigma$, and

$$W^\perp := W + g_{\dot\gamma}\left(W, \frac{\dot\gamma}{l(x,y)} \right) \frac{\dot\gamma}{l(x,y)}.$$

Observing that

$$c(x,y) = \inf_{w,\gamma} \left(I(W^\perp, W^\perp) - \left[g_{\dot\gamma}\left(D_V^T V(\cdot, 0), \frac{\dot\gamma}{l(x,y)} \right) \right]_0^1 \right)$$

shows the claim, where the infimum runs over all $w \in T_x M$ with $|w|_h = 1$ and over all maximizing geodesics $\gamma\colon [0,1] \longrightarrow M$ from x to y. □

The following characterization of timelike cut loci plays an important role in the sequel. We refer to McCann [45, Theorem 3.5] for the Lorentzian case and to Ohta [55, Lemma 3.1] for the Finsler case.

Theorem 3.10 (Semi-concavity fails on timelike cut loci)**.** *Let h be a smooth Riemannian metric on M. If $(x,y) \in \{l > 0\} \cap \operatorname{sing}(l)$, then we have*

$$\limsup_{w \to 0 \in T_x M} \frac{l(\exp_x^h(w), y) + l(\exp_x^h(-w), y) - 2\, l(x,y)}{2\, |w|_h^2} = +\infty.$$

Proof. Note that either there are two distinct maximizing geodesics from x to y, or y is the first conjugate point of x along a unique maximizing geodesic that connects x to y.

In the former case, let $\gamma, \eta\colon [0, l(x,y)] \longrightarrow M$ be distinct maximizing geodesics of unit speed from x to y (i.e. $F(\dot\gamma) = 1$). Set $v := \dot\gamma(0)$, $w := \dot\eta(0)$, and $y_\varepsilon := \gamma(l(x,y) - \varepsilon)$ for small $\varepsilon > 0$. Then, on the one hand, we deduce from the reverse triangle inequality and the first variation formula from Lemma 2.6 that

$$\begin{aligned} l\big(\eta(-t), y\big) - l(x,y) &\geq l\big(\eta(-t), y_\varepsilon\big) + \varepsilon - (l(x, y_\varepsilon) + \varepsilon) \\ &= g_v(v, -t\, w) + O(t^2) \end{aligned}$$

for $t > 0$, where we extended the domain of the geodesic η to $(-\tau, l(x,y)]$ for small $\tau > 0$. On the other hand,

$$l\big(\eta(t), y\big) - l(x,y) = -t.$$

Thus we have

$$\begin{aligned} &\frac{l(\eta(t), y) + l(\eta(-t), y) - 2\, l(x,y)}{t^2} \\ &\qquad \geq -\frac{g_v(v,w) + 1}{t} + \frac{1}{t^2}\, O(t^2) \to +\infty \qquad \text{as } t \to 0 \end{aligned}$$

since by the reverse Cauchy–Schwarz inequality (2.6),

$$g_v(v,w) = \mathscr{L}(v)(w) < -\sqrt{4\,L(v)\,L(w)} = -1.$$

In the latter case, note first that, along the argument in the proof of [45, Theorem 3.5], it is sufficient to show

$$\limsup_{w\to 0\in T_xM} \frac{l(\xi_w(1),y) + l(\xi_w(-1),y) - 2\,l(x,y)}{2\,|w|_h^2} = +\infty,$$

where $\xi_w\colon [-1,1] \longrightarrow M$ denotes the geodesic with $\dot\xi_w(0) = w$ with respect to L (where the values $\xi_w(\pm 1)$ exist provided $|w|_h$ is sufficiently small). Let J be a Jacobi field along the unique maximizing geodesic $\gamma\colon [0,1] \longrightarrow M$ from x to y vanishing only at 0 and 1. Then $w := D^{\dot\gamma}_{\dot\gamma} J(0) \in T_xM \setminus \{0\}$ and, in the same manner as in the proof of Proposition 3.9, we consider the vector field P along γ with $D^{\dot\gamma}_{\dot\gamma} P = 0$ and $P(0) = w$, and set

$$W(t) := (1-t)P(t).$$

We further define $J_\varepsilon := J + \varepsilon W$ for small $\varepsilon > 0$. Observe that $J_\varepsilon(0) = \varepsilon w$ and $J_\varepsilon(1) = 0$. We remark that J is $g_{\dot\gamma}$-orthogonal to $\dot\gamma$ since it is a Jacobi field vanishing at $t = 0$ and $t = 1$, and that

$$g_{\dot\gamma}\big(P(t),\dot\gamma(t)\big) = g_{\dot\gamma}\big(w,\dot\gamma(0)\big) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_0 g_{\dot\gamma}(J,\dot\gamma) = 0$$

for all $t \in [0,1]$. Hence, J_ε is $g_{\dot\gamma}$-orthogonal to $\dot\gamma$. Now, for a small parameter $\varepsilon > 0$ we consider the variation $\sigma\colon [0,1]\times(-\varepsilon,\varepsilon) \longrightarrow M$ given by

$$\sigma_s(t) = \sigma(t,s) := \xi_{J_\varepsilon(t)}(s).$$

Note that $\sigma_s(0) = \xi_w(\varepsilon\, s)$. Setting again

$$\mathfrak{L}(s) := \int_0^1 F(\dot\sigma_s)\,\mathrm{d}t,$$

we observe from the second variation formula from Proposition 2.7 that

$$\begin{aligned}\frac{l(\xi_w(\varepsilon\, s),y) + l(\xi_w(-\varepsilon\, s),y) - 2\,l(x,y)}{2\,(\varepsilon\, s)^2} \\ \geq \frac{\mathfrak{L}''(0)}{\varepsilon^2} + \frac{o(\varepsilon^2)}{\varepsilon^2} = \frac{I(J_\varepsilon,J_\varepsilon)}{\varepsilon^2} + \frac{g_{\dot\gamma}(D^T_V V(0,0),\dot\gamma(0))}{\varepsilon^2\,F(\dot\gamma)} + \frac{o(\varepsilon^2)}{\varepsilon^2},\end{aligned}$$

where $T = \partial_t\sigma$ and $V := \partial_s\sigma$. As J is a Jacobi field, the first summand of the latter term is rewritten with the help of (2.7) as

$$\frac{I(J,J)}{\varepsilon^2} + 2\,\frac{I(J,W)}{\varepsilon} + I(W,W) = \frac{2\,g_{\dot\gamma}(w,w)}{\varepsilon\,l(x,y)} + I(W,W).$$

Note also that, since $w \neq 0$ and $g_{\dot\gamma}\big(w, \dot\gamma(0)\big) = 0$, we have $g_{\dot\gamma}(w,w) > 0$. As for the second summand above, we directly infer from the definition of σ that $D_V^V V = 0$. Then we have

$$g_{\dot\gamma}\big(D_V^T V(0,0), \dot\gamma(0)\big) = g_{\dot\gamma}\big(D_V^T V(0,0) - D_V^V V(0,0), \dot\gamma(0)\big) = -\varepsilon^2\, \mathcal{T}_{\dot\gamma(0)}(w),$$

where

$$\mathcal{T}_v(w) := \sum_{\alpha,\beta,\delta,\lambda=1}^{n} g_{\alpha\lambda}(v)\, \big(\Gamma^\alpha_{\beta\delta}(w) - \Gamma^\alpha_{\beta\delta}(v)\big)\, w^\beta\, w^\delta\, v^\lambda$$

is called the $\mathbf{T}$-*curvature* (or the *tangent curvature*) measuring the variation of tangent spaces; see [58, §6.3.2, §8.3.3]. Therefore, we deduce that

$$\begin{aligned} &\frac{l(\xi_w(\varepsilon\, s), y) + l(\xi_w(-\varepsilon\, s), y) - 2\, l(x,y)}{2\, (\varepsilon\, s)^2} \\ &\qquad \geq \frac{2\, g_{\dot\gamma}(w,w)}{\varepsilon\, l(x,y)} + I(W,W) - \frac{\mathcal{T}_{\dot\gamma(0)}(w)}{F(\dot\gamma)} + \frac{o(\varepsilon^2)}{\varepsilon^2} \to +\infty \end{aligned}$$

as $\varepsilon \to 0$. This completes the proof. □

Finally, we address smoothness properties of l_q. We start with the usual case of $q = 1$ in Theorem 3.11, based on which we establish the case of $q \in (0,1)$ in Corollary 3.12. From Lemma 3.3, we already know that l_q is upper semi-continuous on $M \times M$ and continuous on $\{l \geq 0\}$. In addition, recall from Subsection 3.2 that the singular set $\mathrm{sing}(l)$ from Definition 3.4 is closed.

For a function $u\colon M \longrightarrow [-\infty, +\infty]$, we say that u is *subdifferentiable* at $x \in M$ with a *subgradient* $\zeta \in T_x^*M$ if $u(x) \in \mathbb{R}$ and

$$u\big(\exp_x(v)\big) \geq u(x) + \zeta(v) + o(|v|_h) \tag{3.7}$$

for $v \in T_xM$ close to 0, where h is an auxiliary Riemannian metric. The set of subgradients of u at x will be denoted by $\partial_\bullet u(x)$. The *superdifferentiability* and *supergradients* $\partial^\bullet u(x)$ are defined in the same way with the reverse inequality. As for the notions of subdifferentiability and subgradients of a function on M introduced above, the properties defined in the previous clauses do not depend on the choice of h.

Theorem 3.11 (Smoothness of l)**.** *The Lorentz–Finsler distance function l satisfies the following properties.*

(i) *It is smooth precisely on the complement of* $\mathrm{sing}(l)$*.*

(ii) *It is locally Lipschitz and locally semi-convex on the open superlevel set $\{l > 0\}$ of $M \times M$. Furthermore, for $(x,y) \in \{l > 0\}$ and any maximizing geodesic $\gamma\colon [0, l(x,y)] \longrightarrow M$ of unit speed from x to y, we have*

$$\mathscr{L}\big(\dot\gamma(0)\big) \in \partial_\bullet\big(l(\cdot, y)\big)(x), \qquad -\mathscr{L}\big(\dot\gamma\big(l(x,y)\big)\big) \in \partial_\bullet\big(l(x,\cdot)\big)(y).$$

(iii) *For $(x,y) \in \{l = 0\}$, $l(\cdot, y)$ is not superdifferentiable at x unless $x = y$, and $\partial^\bullet\big(l(\cdot,x)\big)(x) = \mathcal{H}_1^{-1}(0) \cap T_x^*M$.*

Proof. (i) and (ii) are seen in the same way as [45, Theorem 3.6] with the help of Proposition 3.9 and Theorem 3.10. (We remark that $-\mathscr{L}(\dot{\gamma}(l(x,y)))$ can be regarded as $\overline{\mathscr{L}}(\dot{\overline{\gamma}}(0))$ for the Legendre transform $\overline{\mathscr{L}}$ with respect to the *reverse Lorentz–Finsler structure* $\overline{L}(v) := L(-v)$ and the reverse curve $\overline{\gamma}(t) := \gamma(l(x,y)-t)$ which is a maximizing geodesic with respect to $\overline{L}$.) (iii) also follows from a similar argument to [45, Theorem 3.6] by virtue of Corollary 3.2(iii). □

We recall from (3.6) that $l_q(x,y) = l(x,y)^q/q$ provided $l(x,y) > -\infty$.

Corollary 3.12 (Properties of l_q)**.** *For $q \in (0,1)$, l_q satisfies the following.*

(i) *l_q is smooth precisely on $(M \times M) \setminus \mathrm{sing}(l)$ and is locally Lipschitz and locally semi-convex on $\{l > 0\}$.*

(ii) *Given $y \in M$, if $l_q(\cdot, y)$ has a supergradient $\zeta \in T_x^*M$, then $l(x,y) > 0$ necessarily holds and we have*

$$y = \exp_x \left(\sum_{\alpha=1}^{n} \frac{\partial \mathcal{H}_q}{\partial \zeta_\alpha}(\zeta) \frac{\partial}{\partial x^\alpha} \right).$$

(iii) *The following matrix is non-degenerate for $(x,y) \in (M \times M) \setminus \mathrm{sing}(l)$:*

$$\left(\frac{\partial^2 l_q}{\partial x^\alpha \partial y^\beta}(x,y) \right)_{\alpha,\beta=1}^{n}.$$

(iv) *For every $(x,y) \in \{l > 0\} \cap \mathrm{sing}(l)$ and every smooth Riemannian metric h on M, we have*

$$\sup_{0<|w|_h<1} \frac{l_q(\exp_x^h(w), y) + l_q(\exp_x^h(-w), y) - 2\, l_q(x,y)}{2\,|w|_h^2} = +\infty.$$

Proof. (i) This is shown along the same lines as [45, Corollary 3.7] thanks to Theorem 3.11(i), (ii).

(ii) It follows from Theorem 3.11(iii) that $l(x,y) > 0$. Then, together with the subdifferentiability asserted in Theorem 3.11(ii), $l(\cdot, y)$ is differentiable at x and we have the claim by noting that

$$\zeta = \mathrm{d}\big(l_q(\cdot, y)\big)_x = l(x,y)^{q-1} \cdot \mathrm{d}\big(l(\cdot, y)\big)_x$$

and, recalling Lemma 3.1,

$$\sum_{\alpha=1}^{n} \frac{\partial \mathcal{H}_q}{\partial \zeta_\alpha}(\zeta)\, \frac{\partial}{\partial x^\alpha} = F^*(\zeta)^{p-2}\, \mathscr{L}^*(\zeta) = l(x,y)\, \frac{\mathscr{L}^*(\zeta)}{F^*(\zeta)} = l(x,y)\, \dot{\gamma}(0).$$

(iii) Note first that $(x^\alpha)_{\alpha=1}^n$ and $(y^\beta)_{\beta=1}^n$ are local coordinate systems around x and y, respectively. If the matrix in question is degenerate, then there exists $v \in T_xM \setminus \{0\}$ such that

$$\sum_{\alpha=1}^{n} v^\alpha \frac{\partial^2 l_q}{\partial x^\alpha \partial y^\beta}(x,y) = 0$$

for all β. By the first variation formula from Lemma 2.6, this implies that the function $f(z) := g_{w_z}(v, w_z) = \mathscr{L}(w_z)(v)$ with $w_z := \exp_x^{-1}(z)$ satisfies $\mathrm{d}f_y = 0$, which cannot happen.

(iv) This is an immediate consequence of Theorem 3.10. □

The non-degeneracy in (iii) corresponds to a *twist condition* for l^q, cf. [65].

4 Geodesics of probability measures

Now we study the spacetime geometry of the space $\mathcal{P}(M)$ of Borel probability measures on M. This geometry is described by the concept of q-geodesics taken from McCann [45, Definition 1.1], cf. Definition 4.2. We continue assuming the global hyperbolicity of (M, L).

In Section 4.1 we recall basics of optimal transport in Lorentzian geometry (cf. e.g. the works of Cavalletti–Mondino [21], Eckstein–Miller [24], or McCann [45]), which carry over to the Lorentz–Finsler setting with no change. Then we show q-geodesics to exist in large generality in Section 4.2. Finally, based on the duality theory outlined in Section 4.3, in Section 4.4 we address the uniqueness of q-geodesics (and simultaneously the one of couplings attaining the supremum in (4.1)). Here we follow [45, Sections 2, 4, 5].

4.1 The q-Lorentz–Wasserstein distance and q-geodesics

Let $\mathrm{proj}_i \colon M \times M \longrightarrow M$ denote the projection map defined by $\mathrm{proj}_i(x) := x_i$, where $i = 1, 2$; similarly, for $i, j = 1, 2, 3$ with $i \neq j$ we will define $\mathrm{proj}_{ij} \colon M \times M \times M \longrightarrow M \times M$ by $\mathrm{proj}_{ij}(x) := (x_i, x_j)$. As said, $\mathcal{P}(M)$ is the space of Borel probability measures on M. Let $\mathcal{P}_{\mathrm{c}}(M)$ be the set of all $\mu \in \mathcal{P}(M)$ with compact support $\operatorname{spt}\mu \subset M$. Given any Borel measure $\mathfrak{m}$ on M, let $\mathcal{P}^{\mathrm{ac}}(M, \mathfrak{m})$ denote the space of all $\mathfrak{m}$-absolutely continuous elements of $\mathcal{P}(M)$, and set $\mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M, \mathfrak{m}) := \mathcal{P}^{\mathrm{ac}}(M, \mathfrak{m}) \cap \mathcal{P}_{\mathrm{c}}(M)$.

Given $\mu, \nu \in \mathcal{P}(M)$, let $\Pi(\mu, \nu)$ be the set of all couplings π of μ and ν. We call such a coupling *causal* if $\pi[\{l \geq 0\}] = 1$, and *chronological* if $\pi[\{l > 0\}] = 1$. By Lemma 3.3, π is causal if and only if $\operatorname{spt}\pi \subset \{l \geq 0\}$.

Throughout, we fix $q \in (0, 1]$, and we state explicitly when $q < 1$ is assumed in addition. Recalling (3.6) and (3.4), let $\ell_q \colon \mathcal{P}(M) \times \mathcal{P}(M) \longrightarrow [0, +\infty] \cup \{-\infty\}$ denote the *q-Lorentz–Wasserstein distance*

$$\ell_q(\mu, \nu) := \sup_{\pi \in \Pi(\mu,\nu)} \left(\int_{M \times M} l(x, y)^q \, \mathrm{d}\pi(x, y) \right)^{1/q}. \tag{4.1}$$

Our choice of the cost function l^q which can drop to $-\infty$ ensures $\ell_q(\mu, \nu) \geq 0$ if and only if μ and ν admit a causal coupling.

We term $\pi \in \Pi(\mu, \nu)$ *ℓ_q-optimal* if it is causal and it attains the supremum in (4.1). We occasionally employ the subsequent standard result, which follows from Lemma 3.3 and [65, Theorem 4.1].

Lemma 4.1 (Existence of ℓ_q-optimal couplings)**.** *Assume $\mu, \nu \in \mathcal{P}_c(M)$ to obey $\ell_q(\mu, \nu) \geq 0$. Then there exists an ℓ_q-optimal coupling of μ and ν.*

Definition 4.2 (Geodesics of probability measures)**.** Given any $q \in (0, 1]$, a family $(\mu_t)_{t\in[0,1]}$ in $\mathcal{P}(M)$ is called a *q-geodesic* if

(1) $\ell_q(\mu_0, \mu_1)$ is positive and finite, and

(2) for every $s, t \in [0, 1]$ with $s < t$,

$$\ell_q(\mu_s, \mu_t) = (t - s)\, \ell_q(\mu_0, \mu_1).$$

Akin to the corresponding property of Wasserstein distances, ℓ_q satisfies the *reverse* triangle inequality. This follows from a standard gluing argument, cf. e.g. [24, Theorem 13]. To analyze q-geodesics, we also need to study cases of equality. This is collected in Proposition 4.3 taken from [45, Proposition 2.9]. One immediate consequence of it is that the property of having endpoints with compact support propagates through the interior of q-geodesics, cf. Corollary 4.4, a proof of which we give for convenience.

Recall Lemma 3.6 for the definitions of $Z_s(E)$ and $Z(E)$.

Proposition 4.3 (Reverse triangle inequality and cases of equality)**.** *For every $\mu_1, \mu_2, \mu_3 \in \mathcal{P}(M)$, we have*

$$\ell_q(\mu_1, \mu_3) \geq \ell_q(\mu_1, \mu_2) + \ell_q(\mu_2, \mu_3). \tag{4.2}$$

Moreover, if $\ell_q(\mu_1, \mu_2)+\ell_q(\mu_2, \mu_3) \neq \pm\infty$ according to our conventions (3.4) *and there are Borel sets $X_1, X_3 \subset M$ with $\mu_1[X_1] = \mu_3[X_3] = 1$ yet $\mu_2[Z(X_1 \times X_3)] < 1$, then the inequality in* (4.2) *is strict.*

Conversely, assume that

(1) *$q < 1$,*

(2) *$\ell_q(\mu_1, \mu_3)$ is positive and finite,*

(3) *equality holds in* (4.2)*, and*

(4) *μ_1 and μ_2 as well as μ_2 and μ_3 admit ℓ_q-optimal couplings (e.g. when μ_1, μ_2, and μ_3 are compactly supported, cf. Lemma* 4.1*).*

Then there exists a measure $\omega \in \mathcal{P}(M \times M \times M)$ such that $\pi_{ij} := (\mathrm{proj}_{ij})_\sharp \pi$ is an ℓ_q-optimal coupling of its marginals μ_i and μ_j for every $i, j \in \{1, 2, 3\}$ with $i < j$, and every $(x, z, y) \in \operatorname{spt} \omega$ satisfies

$$l(x, z) = t\, l(x, y), \qquad l(z, y) = (1 - t)\, l(x, y),$$

where $t := \ell_q(\mu_1, \mu_2)/\ell_q(\mu_1, \mu_3)$. Furthermore, if π_{13} is concentrated on a Borel set $E \subset M \times M$, then μ_2 is concentrated on $Z_t(E)$. In particular, if $Z_t(x, y)$ is a singleton $\{z_t(x, y)\}$ for π_{13}-a.e. $(x, y) \in M \times M$, then

$$\omega = (z_0, z_t, z_1)_\sharp \pi_{13}, \qquad \mu_2 = (z_t)_\sharp \pi_{13}.$$

Corollary 4.4 (Interpolants inherit compact support)**.** *Let $(\mu_t)_{t\in[0,1]}$ be a q-geodesic in $\mathcal{P}(M)$ for some $q \in (0,1)$. If the endpoints μ_0 and μ_1 are compactly supported, so is the intermediate point μ_t for every $t \in (0,1)$; more precisely,* $\operatorname{spt}\mu_t \subset Z_t(\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1)$.

Proof. Lemma 3.6 yields the compactness of $Z(E)$ and $Z_t(E)$, where we set $E := \operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1$. Since $\ell_q(\mu_0,\mu_1) \in (0,+\infty)$, applying the first part of Proposition 4.3 yields $\mu_t[Z(E)] = 1$. In particular, μ_t has compact support. Then, moreover, it follows from Lemma 4.1 that μ_0 and μ_t as well as μ_t and μ_1 admit ℓ_q-optimal couplings. Hence, by the latter part of Proposition 4.3, we obtain in fact $\mu_t[Z_t(E)] = 1$. □

4.2 Existence of q-geodesics

Employing a standard measurable selection argument [3, 65], we now construct q-geodesics by taking a chronological ℓ_q-optimal coupling π between appropriate endpoints μ_0 and μ_1 under consideration, connecting π-a.e. $(x,y) \in M \times M$ by a maximizing timelike geodesic, and a lifting procedure.

We need the subsequent technical Lemma 4.5; its proof uses Lemma 3.3 and Remark 3.7 and is analogous to its Lorentzian version by McCann [45, Lemma 2.8]. Recall that a real-valued map defined on a topological space X is termed *universally measurable* if for every Borel probability measure μ on X, it is measurable with respect to the completion of the Borel σ-algebra of X with respect to μ.

Lemma 4.5 (Selecting intermediate points on the timelike cut locus)**.** *For every $t \in [0,1]$, the map $z_t\colon (M\times M)\setminus \operatorname{sing}(l) \longrightarrow M$ from Lemma 3.5 has a non-relabeled universally measurable extension $z_t\colon \{l>0\} \longrightarrow M$ such that for every $(x,y) \in \{l>0\}$, $t \longmapsto z_t(x,y)$ is a maximizing geodesic from x to y.*

Corollary 4.6 (Existence of q-geodesics)**.** *Let $\mu_0,\mu_1 \in \mathcal{P}(M)$, and let $\pi \in \Pi(\mu_0,\mu_1)$ be a chronological ℓ_q-optimal coupling with finite total ℓ_q-cost. Then the following hold.*

(i) *The assignment $\mu_t := (z_t)_\sharp\pi$, where z_t is the map provided by Lemma 4.5, defines a q-geodesic $(\mu_t)_{t\in[0,1]}$ connecting μ_0 to μ_1.*

(ii) *For every $s,t \in [0,1]$ with $s<t$, $(z_s,z_t)_\sharp\pi$ is a chronological ℓ_q-optimal coupling of the members μ_s and μ_t of that q-geodesic.*

Proof. The coupling $(z_s,z_t)_\sharp\pi$ of μ_s and μ_t is well-defined and chronological by Lemma 4.5. In addition,

$$\begin{aligned}\ell_q(\mu_s,\mu_t)^q &\geq \int_{M\times M} l\big(z_s(x,y),z_t(x,y)\big)^q \,\mathrm{d}\pi(x,y)\\ &= (t-s)^q \int_{M\times M} l(x,y)^q \,\mathrm{d}\pi(x,y)\\ &= (t-s)^q\, \ell_q(\mu_0,\mu_1)^q.\end{aligned}$$

Since s and t are arbitrary, this estimate combined with the reverse triangle inequality from Proposition 4.3 forces the inequality appearing in the above computation to become equality. Thus, $(\mu_t)_{t\in[0,1]}$ is a q-geodesic, and $(z_s, z_t)_\sharp \pi$ is an ℓ_q-optimal coupling of μ_s and μ_t. □

4.3 Kantorovich duality

Now we revisit the duality theory developed by McCann in [45, Section 4] in the Lorentzian setting (see also Cavalletti–Mondino [21, §2.4]). Therein, the additional difficulty is that l drops to $-\infty$ at points not in causal relation, hence the infimum in (4.3) below might not be attained in general (see [36] for an example). To ensure that the infimum is a minimum, cf. Theorem 4.9, the q-separation property stated in Definition 4.8 below has been introduced in [45, Definition 4.1]. Lemma 4.10 will show that this notion is not void: q-separation holds for every pair of measures for which *every* transport stays away from the light cone in a uniform way. This property is in practice easier to verify than q-separation, yet it does not suffice to consider only such measures. Indeed, the intermediate points μ_s and μ_t of a q-geodesic $(\mu_t)_{t\in[0,1]}$ will not satisfy this condition in general as their supports typically overlap for $s, t \in [0, 1]$ close to each other. On the other hand, q-separation is well-behaved along q-geodesics, cf. Proposition 4.14.

Since l is bounded from above on compact subsets of $M \times M$, the standard Kantorovich duality [65, Theorem 5.10] yields the following.

Proposition 4.7 (Basic duality)**.** *For every $\mu, \nu \in \mathcal{P}_c(M)$ we have*

$$q^{-1}\,\ell_q(\mu,\nu)^q = \inf_{(u,v)} \bigg(\int_M u \,\mathrm{d}\mu + \int_M v \,\mathrm{d}\nu \bigg), \tag{4.3}$$

where $(u, v)\colon M \longrightarrow \mathbb{R}^2$ runs over all pairs of functions in $L^1(M,\mu) \times L^1(M,\nu)$ satisfying $l^q/q \le u \oplus v$ on $M \times M$, where

$$(u \oplus v)(x, y) := u(x) + v(y).$$

Moreover, given any $X \supset \operatorname{spt}\mu$ and $Y \supset \operatorname{spt}\nu$, the infimum in (4.3) *is unchanged if we restrict it to those pairs of lower semi-continuous functions $u \in L^1(M,\mu)$ and $v \in L^1(M,\nu)$ such that $u = {}^{(l^q)}v$ on X and $v = u^{(l^q)}$ on Y, where*

$$\begin{aligned} {}^{(l^q)}v(x) &:= \sup_{y\in Y} \big(q^{-1}\, l(x,y)^q - v(y) \big), \\ u^{(l^q)}(y) &:= \sup_{x\in X} \big(q^{-1}\, l(x,y)^q - u(x) \big). \end{aligned} \tag{4.4}$$

Note that the transforms in (4.4) depend on X and Y, respectively.

Definition 4.8 (q-separation)**.** We term $(\mu, \nu) \in \mathcal{P}_c(M) \times \mathcal{P}_c(M)$ *q-separated* (*by* (π, u, v)) if there exist $\pi \in \Pi(\mu,\nu)$ and lower semi-continuous functions $u\colon \operatorname{spt}\mu \longrightarrow \mathbb{R}\cup\{+\infty\}$ and $v\colon \operatorname{spt}\nu \longrightarrow \mathbb{R}\cup\{+\infty\}$ with the following properties.

(1) We have $l^q/q \leq u \oplus v$ on $\operatorname{spt}\mu \times \operatorname{spt}\nu$.

(2) The equality set

$$E := \left\{(x,y) \in \operatorname{spt}\mu \times \operatorname{spt}\nu \mid u(x) + v(y) = q^{-1}\, l(x,y)^q\right\} \tag{4.5}$$

satisfies $\operatorname{spt}\pi \subset E \subset \{l > 0\}$.

Note that the lower semi-continuity of u and v plus the upper semi-continuity of l imply the compactness of E, cf. Theorem 4.9 below.

We omit the technical proof of the following Theorem 4.9, which can be found in [45, Theorem 4.3]. It transfers *verbatim* to the Lorentz–Finsler case by employing Theorem 3.11, Corollary 3.12, as well as an auxiliary smooth Riemannian metric on M.

To formulate it, we recall that a set $E \subset M \times M$ is *l^q-cyclically monotone* (cf. e.g. [21, Definition 2.6] and [65, Definition 5.1]) if for every $k \in \mathbb{N}$, every $(x_1, y_1), \ldots, (x_k, y_k) \in E$, and every permutation σ of length k we have

$$\sum_{i=1}^{k} l(x_i, y_i)^q \geq \sum_{i=1}^{k} l(x_i, y_{\sigma(i)})^q.$$

Every ℓ_q-optimal coupling is concentrated on an l^q-cyclically monotone set, but unlike the case of real-valued cost functions discussed in Villani's monograph [65, Theorem 5.10], the converse requires an additional hypothesis on the coupling, cf. Cavalletti–Mondino [21, Proposition 2.8].

Theorem 4.9 (Duality by q-separation)**.** *Assume $(\mu,\nu) \in \mathcal{P}_c(M) \times \mathcal{P}_c(M)$ to be q-separated by (π, u, v). Then the following hold.*

(i) *We have $u = {}^{(l^q)}v$ on $\operatorname{spt}\mu$ and $v = u^{(l^q)}$ on $\operatorname{spt}\nu$ according to* (4.4)*.*

(ii) *The equality set E from* (4.5) *is compact and l^q-cyclically monotone.*

(iii) *The coupling π is ℓ_q-optimal, while (u, v) minimizes the right-hand side of* (4.3)*.*

(iv) *The functions u and v extend to semi-convex Lipschitz functions on open neighborhoods of $\operatorname{spt}\mu$ and $\operatorname{spt}\nu$, respectively, with Lipschitz and semi-convexity constants estimated by those of $l^q|_E$.*

Lemma 4.10 (Existence of q-separation)**.** *Let $\mu, \nu \in \mathcal{P}_c(M)$ satisfy $\operatorname{spt}\mu \times \operatorname{spt}\nu \subset \{l > 0\}$. Then (μ,ν) is q-separated.*

Proof. By Lemma 3.3, the function l is continuous, bounded, and bounded away from 0 on $\operatorname{spt}\mu \times \operatorname{spt}\nu$. Thus the supremum in (4.1) and, by [65, Theorem 5.10], the infimum in (4.3) are attained by some $\pi \in \Pi(\mu,\nu)$ and uniformly continuous functions $u\colon \operatorname{spt}\mu \longrightarrow \mathbb{R}$ and $v\colon \operatorname{spt}\nu \longrightarrow \mathbb{R}$ of the form (4.4), respectively. This duality also implies π to be concentrated on the equality set E as in (4.5), and $E \subset \operatorname{spt}\mu \times \operatorname{spt}\nu \subset \{l > 0\}$. □

4.4 Characterization of q-geodesics

Now we turn to the description of q-geodesics between q-separated endpoints, ultimately stated in Corollary 4.18. To this aim, a central part is the solvability of the ℓ_q-Monge problem shown in Theorem 4.17. This is the Lorentz–Finsler analogue of the Brenier–McCann theorem in Riemannian geometry [16, 44] and of the Lorentzian result by McCann [45]. Comparable results are obtained in [63] under certain geometric conditions on the support of the target measure.

Since this subsection constitutes a central part of our work, and in order to emphasize the difference between the Lorentz–Finsler and the genuine Lorentzian settings, we provide full proofs even if they are essentially identical to those in [45, Section 5].

We start by collecting various properties linked to intermediate points of geodesics in M and $\mathcal{P}(M)$. Proposition 4.11 and Lemma 4.13 are based on the following trivial consequence of (3.5): for every triple $(x, z, y) \in M \times M \times M$ and every $t \in (0, 1)$,

$$l(x, y) \geq t\, \frac{l(x, z)}{t} + (1 - t)\, \frac{l(z, y)}{1 - t}. \tag{4.6}$$

Proposition 4.11 (Lagrangian trajectories do not cross)**.** *Let $q \in (0, 1)$, $t \in (0, 1)$, and $x, x', y, y' \in M$ with $l(x, y) > 0$ and $l(x', y') > 0$. If the set $Z_t(x, y)$ intersects $Z_t(x', y')$ yet*

$$l(x, y')^q + l(x', y)^q \leq l(x, y)^q + l(x', y')^q, \tag{4.7}$$

then we have $x = x'$ and $y = y'$.

Proof. Given any $z \in Z_t(x, y) \cap Z_t(x', y')$, the inequality (4.6) for the triple (x, z, y'), our conventions (3.4), and the concavity of $r \longmapsto r^q$ on $[0, +\infty)$ yield

$$\begin{aligned} l(x, y')^q &\geq \big(l(x, z) + l(z, y')\big)^q \\ &= \big(t\, l(x, y) + (1 - t)\, l(x', y')\big)^q \\ &\geq t\, l(x, y)^q + (1 - t)\, l(x', y')^q, \end{aligned} \tag{4.8}$$

and analogously

$$l(x', y)^q \geq t\, l(x', y')^q + (1 - t)\, l(x, y)^q. \tag{4.9}$$

This gives $l(x, y') > 0$ and $l(x', y) > 0$. Furthermore, by (4.7), summing up the previous inequalities forces equalities all over (4.6) — namely for the triples (x, z, y') and (x', z, y) —, (4.8), and (4.9). By Remark 3.7, equality in (4.6) implies z to lie in the interior of a maximizing geodesic from x to y' and in the interior of a maximizing geodesic from x' to y. Hence, all five points under consideration lie on the same maximizing geodesic $\gamma\colon [0, 1] \longrightarrow M$. Finally, identity in the second inequality of (4.8), together with *strict* concavity of $r \longmapsto r^q$ on $[0, +\infty)$, implies $l(x, y) = l(x', y')$. Since z divides the segments of γ from x to y and from x' to y' in the same ratio, we obtain $x = x'$ and $y = y'$. □

Theorem 4.12 (Continuous inverse maps)**.** *Let $q \in (0,1)$ and $t \in (0,1)$. Suppose $(\mu_0, \mu_1) \in \mathcal{P}_c(M) \times \mathcal{P}_c(M)$ to be q-separated. Then there exists a continuous map $W\colon Z_t(E) \longrightarrow E$ such that $z_t \circ W$ is the identity map on $z_t(E)$, and whenever $\mu_t \in \mathcal{P}(M)$ is a t-intermediate point of a q-geodesic from μ_0 to μ_1, $W_\sharp \mu_t$ belongs to $\Pi(\mu_0, \mu_1)$ and maximizes $\ell_q(\mu_0, \mu_1)$. Here $E \subset \operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1$ denotes the equality set from* (4.5)*, and z_t is the map given by Lemma* 4.5*.*

Proof. Recall that E is compact and l^q-cyclically monotone by Theorem 4.9. Furthermore, $E \subset \{l > 0\}$ by Definition 4.8. Given any $z \in Z_t(E)$ we define $W(z) := (x, y)$, where $(x, y) \in E$ is chosen such that $z \in Z_t(x, y)$ according to the definition of $Z_t(E)$ from Lemma 3.6.

We claim that W is well-defined and continuous. If $z \in Z_t(x, y) \cap Z_t(x', y')$ for $(x, y), (x', y') \in E$, the l^q-cyclical monotonicity of E yields (4.7), whence $x = x'$ and $y = y'$ by Proposition 4.11. Concerning continuity, let $(z_k)_{k\in\mathbb{N}}$ be a sequence in $Z_t(E)$ converging to $z \in Z_t(E)$, and consider a fixed non-relabeled subsequence of $(W(z_k))_{k\in\mathbb{N}}$. Then the compactness of E entails the existence of a limit $(x', y') \in E$ of a further non-relabeled subsequence of the latter. By the above well-definedness proof, (x', y') necessarily coincides with the unique pair $(x, y) \in E$ with $z \in Z_t(x, y)$, i.e. with $W(z)$. Since the initial subsequence was arbitrary, the continuity of W follows. The construction also ensures that W is the right-inverse of z_t on $z_t(E)$.

Finally, let μ_t be as hypothesized. Since the pair (μ_0, μ_1) is q-separated and l is bounded from above on the compact set $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1$ by Lemma 3.3, we have $\ell_q(\mu_0, \mu_1) \in (0, +\infty)$; in particular $\ell_q(\mu_0, \mu_t) > 0$ and $\ell_q(\mu_t, \mu_1) > 0$. Hence, by applying Corollary 4.4, Lemma 3.6, and Lemma 4.1, μ_0 and μ_t as well as μ_t and μ_1 admit ℓ_q-optimal couplings. Proposition 4.3 thus yields $\omega \in \mathcal{P}(M \times M \times M)$ with marginals μ_0, μ_t, and μ_1 such that $\pi_{ij} := (\mathrm{proj}_{ij})_\sharp \omega$ is ℓ_q-optimal for μ_i and μ_j for every $i, j = 0, t, 1$ with $i < j$. By duality, cf. Theorem 4.9 and Proposition 4.7, π_{01} is concentrated on E. By Proposition 4.3, $\mu_t = (\mathrm{proj}_t)_\sharp \omega$ is concentrated on $Z_t(E)$, which is compact by Lemma 3.6. Setting $W_i := \mathrm{proj}_i \circ W$ for $i = 0, 1$, the above argument gives the concentration of ω on $W_0(Z_t(E)) \times Z_t(E) \times W_1(Z_t(E))$. This implies $\omega = (W_0, \mathrm{Id}, W_1)_\sharp \mu_t$ by [1, Lemma 3.1], and hence $\pi_{01} = W_\sharp \mu_t$. □

Lemma 4.13 (Variational characterization of geodesic endpoints)**.** *Let $q \in (0,1)$ and $\gamma\colon [0,1] \longrightarrow M$ be a maximizing timelike geodesic. Then for every $t \in (0,1)$ and $x \in M$, we have*

$$l\big(x, \gamma(1)\big)^q \geq t^{1-q}\, l\big(x, \gamma(t)\big)^q + (1-t)^{1-q}\, l\big(\gamma(t), \gamma(1)\big)^q. \tag{4.10}$$

Equality holds therein if and only if $x = \gamma(0)$.

Proof. The claimed bound (4.10) follows from (4.6) for $(x, \gamma(t), \gamma(1))$ and the concavity of the function $r \longmapsto r^q$. Moreover, it is clear that $x = \gamma(0)$ achieves equality therein. Now, assume that equality holds in (4.10) for some $x \in M$. Then equality holds in (4.6) for $(x, \gamma(t), \gamma(1))$, which implies some timelike maximizing geodesic ending at $\gamma(1)$ to pass through x and $\gamma(t)$. As $\gamma(t)$ lies in the interior of γ, this geodesic necessarily coincides with γ, and equality in (4.10) and the strict concavity of $r \longmapsto r^q$ on $[0, +\infty)$ force $x = \gamma(0)$. □

Proposition 4.14 (Star-shapedness of q-separation)**.** *Let $q \in (0,1)$, $\mu_0, \mu_1 \in \mathcal{P}_c(M)$ and $(\mu_t)_{t\in[0,1]}$ be a q-geodesic from μ_0 to μ_1. If (μ_0,μ_1) is q-separated, then so is (μ_s,μ_t) for every $s,t\in[0,1]$ with $s<t$.*

Proof. It clearly suffices to prove the q-separation of (μ_0,μ_t) and (μ_t,μ_1) for every $t\in(0,1)$. We concentrate on (μ_0,μ_t); (μ_t,μ_1) is treated analogously by taking Remark 4.15 into account (or reduced to the reverse Lorentz–Finsler structure).

Assume the pair (μ_0,μ_1) to be q-separated by (π,u,v). Theorem 4.9 implies the continuity of u and v on $\operatorname{spt}\mu_0$ and $\operatorname{spt}\mu_1$, respectively. Denote by E the equality set from (4.5) relative to (μ_0,μ_1), let $W\colon Z_t(E)\longrightarrow E$ be the continuous map provided by Theorem 4.12, and define $W_i := \operatorname{proj}_i \circ W$ for $i=1,2$. We claim the q-separation of (μ_0,μ_t) by (π_t,u_t,v_t), where

$$\begin{aligned} \pi_t &:= (W_1,\mathrm{Id})_\sharp\mu_t,\\ u_t &:= t^{q-1}\,u,\\ v_t &:= t^{q-1}\,v\circ W_2 - t^{q-1}\,(1-t)^{1-q}\,q^{-1}\,l^q\circ(\mathrm{Id},W_2). \end{aligned} \tag{4.11}$$

It follows from Theorem 4.12 that $\pi_t\in\Pi(\mu_0,\mu_t)$. By Proposition 4.3 and the compactness of $Z_t(E)$ from Lemma 3.6, we find $\operatorname{spt}\mu_t\subset Z_t(E)$. Since $(\mathrm{Id},W_2)\big(Z_t(E)\big)\subset\{l>0\}$, Theorem 4.12 and Lemma 3.3 then imply the continuity of u_t and v_t on $\operatorname{spt}\mu_0$ and $\operatorname{spt}\mu_t$, respectively. Now, given any $x\in\operatorname{spt}\mu_0$, $z\in\operatorname{spt}\mu_t$, and $y:=W_2(z)$, the hypothesized q-separation of (μ_0,μ_1) entails

$$u(x)+v(y)\geq q^{-1}\,l(x,y)^q.$$

By Remark 3.7 and Lemma 4.13, this yields

$$\begin{aligned} t^{q-1}\,u(x)+t^{q-1}\,v(y) &\geq q^{-1}\,t^{q-1}\,l(x,y)^q\\ &\geq q^{-1}\,l(x,z)^q + q^{-1}\,t^{q-1}\,(1-t)^{1-q}\,l(z,y)^q. \end{aligned}$$

Thus, we have

$$u_t(x)+v_t(z)\geq q^{-1}\,l(x,z)^q. \tag{4.12}$$

Moreover, we deduce from Lemma 4.13 and $W(z)\in E$ that equality holds in (4.12) if and only if $x=W_1(z)$. This proves π_t to be concentrated on

$$E_t := \big\{(x,z)\in\operatorname{spt}\mu_0\times\operatorname{spt}\mu_t \mid u_t(x)+v_t(z)=q^{-1}\,l(x,z)^q\big\},$$

which is compact by the compactness of $\operatorname{spt}\mu_0\times\operatorname{spt}\mu_t$ (cf. Corollary 4.4) and by Lemma 3.3. Hence, we have $\operatorname{spt}\pi_t\subset E_t$. It remains to prove $E_t\subset\{l>0\}$. For any $(x,z)\in E_t$, recall that $x=W_1(z)$ necessarily holds, and $W(Z_t(E))\subset E\subset\{l>0\}$ yields $l(W_1(z),W_2(z))>0$. Since $z\in Z_t(E)$, this implies $l(x,z)=t\,l(W_1(z),W_2(z))>0$. □

Remark 4.15. To prove the q-separation of (μ_t, μ_1), the appropriate replacements of the quantities from (4.11) are

$$\begin{aligned}
\pi'_t &:= (\mathrm{Id}, W_2)_\sharp \mu_t, \\
u'_t &:= (1-t)^{q-1}\, u \circ W_1 - (1-t)^{q-1}\, t^{1-q}\, q^{-1}\, l^q \circ (W_1, \mathrm{Id}), \\
v'_t &:= (1-t)^{q-1}\, v.
\end{aligned}$$

Now we recall the definition of sub- and superdifferentiability of a function $u\colon M \longrightarrow [-\infty, +\infty]$ before Theorem 3.11. For $x \in M$ we write $x \in \mathcal{D}(\mathrm{d}u)$ if $u(x) \in \mathbb{R}$ and u is differentiable at x. The latter is equivalent to u being both sub- and superdifferentiable at x, in which case $\partial_\bullet u(x) = \partial^\bullet u(x) = \{\mathrm{d}u(x)\}$. Finally, $\mathcal{D}(\mathrm{d}^2 u)$ is the set of all $x \in \mathcal{D}(\mathrm{d}u)$ at which u is twice differentiable.

In the rest of this section, let $\mathfrak{m}$ be an arbitrary measure on M comparable with the Lebesgue measure in each local coordinate system. It is only used to distinguish $\mathfrak{m}$-negligible sets in view of differentiability properties: for example, by Alexandrov's theorem, semi-convex or semi-concave functions are twice differentiable $\mathfrak{m}$-a.e. (see [8, 54] for the Riemannian and Finsler cases, as well as [65, Chapter 14] for more details).

In the sequel, we adopt the following conventions without further notice. Given any $X \subset M$, every semi-convex Lipschitz function $u\colon X \longrightarrow \mathbb{R}$ (may and) will be treated as the restriction of an appropriate semi-convex Lipschitz function defined on a neighborhood of X, cf. Theorem 4.9. Moreover, any map defined on $\mathcal{D}(\mathrm{d}u)$ will be extended to M by an arbitrarily chosen constant. Our results will not depend on the respective choices of extensions.

We also make use of the reverse Lorentz–Finsler strucure $\overline{L}(v) := L(-v)$ (briefly mentioned in the proof of Theorem 3.11), equipped with the time orientation that $v \in TM$ is future-directed timelike for $\overline{L}$ if and only if $-v$ is future-directed timelike for L. Note that, given a timelike geodesic $\gamma\colon [0,1] \longrightarrow M$ for L, its reverse curve defined by $\overline{\gamma}(t) := \gamma(-t)$ is a timelike geodesic for $\overline{L}$.

Lemma 4.16 (Transport maps and Jacobi fields)**.** *Let $q \in (0,1)$, $X, Y \subset M$ be compact, $U \subset M$ be a given relatively compact neighborhood of X, and $u\colon U \longrightarrow \mathbb{R}$ be semi-convex and Lipschitz continuous with*

$$u(x) \geq {}^{(l^q)}(u^{(l^q)})(x) \tag{4.13}$$

for every $x \in U$ according to (4.4) defined in terms of X and Y. For $t \in [0,1]$, define $\mathcal{F}_t\colon M \longrightarrow M$ and $\mathcal{G}_t\colon M \longrightarrow M$ by

$$\begin{aligned}
\mathcal{F}_t(x) &:= \exp_x\bigl(t\, F^*(\mathrm{d}u)^{p-2}\, \mathscr{L}^*(\mathrm{d}u)\bigr), \\
\mathcal{G}_t(y) &:= \overline{\exp}_y\bigl(-t\, F^*(-\mathrm{d}u^{(l^q)})^{p-2}\, \mathscr{L}^*(-\mathrm{d}u^{(l^q)})\bigr),
\end{aligned} \tag{4.14}$$

provided $x \in \mathcal{D}(\mathrm{d}u)$ and $y \in \mathcal{D}(\mathrm{d}u^{(l^q)})$, respectively, where $\overline{\exp}_y$ denotes the exponential map for the reverse structure $\overline{L}$.

(i) *For every $(x,y) \in U \times Y$, we have*

$$u(x) + u^{(l^q)}(y) \geq q^{-1}\, l(x,y)^q. \tag{4.15}$$

Also, assuming that equality holds at $(x,y) \in (X \times Y) \cap \{l > 0\}$, we have $y = \mathcal{F}_1(x)$ if $x \in \mathcal{D}(\mathrm{d}u)$, and $(x,y) \notin \mathrm{sing}(l)$ if $x \in \mathcal{D}(\mathrm{d}^2 u)$. Analogously, $y \in \mathcal{D}(\mathrm{d}u^{(l^q)})$ implies $x = \mathcal{G}_1(y)$, and $y \in \mathcal{D}(\mathrm{d}^2 u^{(l^q)})$ gives $(x,y) \notin \mathrm{sing}(l)$.

(ii) *For $\mathfrak{m}$-a.e. $x \in X \cap \{\mathrm{d}u \neq 0\}$, the derivative $\mathrm{d}\mathcal{F}_t(x)\colon T_x M \longrightarrow T_{\mathcal{F}_t(x)} M$ exists, depends smoothly on $t \in [0,1]$, and $t \longmapsto \mathrm{d}\mathcal{F}_t(x)(w)$ is a Jacobi field along the geodesic $\gamma_x(t) := \mathcal{F}_t(x)$ for every $w \in T_x M$.*

(iii) *For $\mathfrak{m}$-a.e. $x \in X \cap \{\mathrm{d}u \neq 0\}$ and every $w \in T_x M$, we have*

$$D^{\dot\gamma_x}_{\gamma_x}\big[\mathrm{d}\mathcal{F}_t(x)(w)\big]_0 = F^*\big(\mathrm{d}u(x)\big)^{p-2}\,\nabla^2 u(w) + \mathrm{d}[F^*(\mathrm{d}u)^{p-2}](w)\,\nabla u(x),$$

where $\nabla u := \mathscr{L}^*(\mathrm{d}u)$ is the gradient vector *and $\nabla^2 u\colon T_x M \longrightarrow T_x M$ is the* Hessian *defined by*

$$\nabla^2 u(w) := D^{\nabla u}_w(\nabla u). \tag{4.16}$$

Proof. (i) The first claim (4.15) is a consequence of the definition (4.4) and (4.13). We turn to the equality case. When $x \in \mathcal{D}(\mathrm{d}u)$, the first claim and our assumption on x and y entail that, for every sufficiently small $w \in T_x M$,

$$\begin{aligned} q^{-1}\, l\big(\exp^h_x(w), y\big)^q &\le u\big(\exp^h_x(w)\big) + u^{(l^q)}(y) \\ &= u(x) + \mathrm{d}u(w) + u^{(l^q)}(y) + o(|w|_h) \\ &= q^{-1}\, l(x,y)^q + \mathrm{d}u(w) + o(|w|_h). \end{aligned}$$

In other words, $q^{-1}\, l(\cdot, y)^q$ is superdifferentiable at x with a supergradient $\mathrm{d}u \in T^*_x M$. Hence, Corollary 3.12(ii) implies $y = \mathcal{F}_1(x)$ as claimed.

Next, let $x \in \mathcal{D}(\mathrm{d}^2 u)$. Arguing similarly to the previous step and employing Taylor's theorem around x we obtain

$$\begin{aligned} q^{-1}\, l\big(\exp^h_x(w), y\big)^q &\le u\big(\exp^h_x(w)\big) + u^{(l^q)}(y) \\ &= u(x) + \mathrm{d}u(w) + \mathrm{d}^2 u(w \otimes w) + u^{(l^q)}(y) + o(|w|^2_h) \\ &= q^{-1}\, l(x,y)^q + \mathrm{d}u(w) + \mathrm{d}^2 u(w \otimes w) + o(|w|^2_h) \end{aligned}$$

for every sufficiently small $w \in T_x M$. By replacing w with $-w$ and adding up the resulting inequalities, the terms involving $\mathrm{d}u$ cancel out. Thus, we get an upper bound for the numerator in Corollary 3.12(iv), and we find $(x,y) \notin \mathrm{sing}(l)$.

The analogous statements replacing the role of x and y can be seen by the same argument for $\overline{L}$ (by noticing $\overline{F}^*(\zeta) = F^*(-\zeta)$ and $\overline{\mathscr{L}}^*(\zeta) = -\mathscr{L}^*(-\zeta)$ for $-\zeta \in \Omega^*_y$, where $\overline{F}^*$ and $\overline{\mathscr{L}}^*$ are associated with $\overline{L}$).

(ii) On the one hand, since u is semi-convex, it is twice differentiable $\mathfrak{m}$-a.e. In particular, for $\mathfrak{m}$-a.e. $x \in X$, $\mathrm{d}u(x)$ exists and, when $\mathrm{d}u(x) \neq 0$, the Hessian $\nabla^2 u$ as in (4.16) is well-defined. On the other hand, by definition, $\gamma_z(t) = \mathcal{F}_t(z)$ is a geodesic for every $z \in \mathcal{D}(\mathrm{d}u)$. Hence, $\mathrm{d}\mathcal{F}_t(x)$ exists and $t \longmapsto \mathrm{d}\mathcal{F}_t(x)(w)$ is a Jacobi field (and hence smooth).

(iii) Observe that

$$\dot\gamma_x(0) = F^*(\mathrm{d}u)^{p-2}\mathscr{L}^*(\mathrm{d}u) = F^*(\mathrm{d}u)^{p-2}\,\nabla u(x).$$

This implies, for every $w \in T_xM$,

$$\begin{aligned} D^{\dot\gamma_x}_{\gamma_x}\big[\mathrm{d}\mathcal{F}_t(x)(w)\big]_0 &= D^{\dot\gamma_x}_{w}\big[F^*(\mathrm{d}u)^{p-2}\,\nabla u\big] \\ &= F^*(\mathrm{d}u)^{p-2}\,\nabla^2 u(w) + \mathrm{d}[F^*(\mathrm{d}u)^{p-2}](w)\,\nabla u(x). \end{aligned}$$

This terminates the proof. □

We remark that, when equality holds in (4.15) as in (i), then $\mathrm{d}u(x) \in \Omega^*_x$. Therefore, $F^*(\mathrm{d}u(x))$ is well-defined and $\nabla u(x) = \mathscr{L}^*(\mathrm{d}u(x)) \in \Omega_x$. Analogously, we find $-\mathrm{d}u^{(l^q)}(y) \in \Omega^*_y$ and $\mathscr{L}^*(-\mathrm{d}u^{(l^q)}(y)) \in \Omega_y$. Thus, the vector $-\mathscr{L}^*(-\mathrm{d}u^{(l^q)}(y))$ is future-directed with respect to the reverse structure $\overline{L}$, and the map $t \longmapsto \mathcal{G}_t(y)$ constitutes a future-directed timelike geodesic for $\overline{L}$.

Theorem 4.17 (Characterizing optimal maps)**.** *Let $q \in (0,1)$ and assume $(\mu,\nu) \in \mathcal{P}_{\mathrm{c}}(M) \times \mathcal{P}_{\mathrm{c}}(M)$ to be q-separated by (π,u,v). Then the following hold.*

(i) *Suppose $\mu \in \mathcal{P}^{\mathrm{ac}}(M,\mathfrak{m})$ and define $\mathcal{F}_1 \colon M \longrightarrow M$ as in (4.14). Then we have $\pi = (\mathrm{Id},\mathcal{F}_1)_\sharp\mu$, which is a unique maximizer of $\ell_q(\mu,\nu)$, and $(x,\mathcal{F}_1(x)) \notin \mathrm{sing}(l)$ holds for μ-a.e. $x \in M$.*

(ii) *The map $\mathcal{F}_1$ in* (i) *is unique in the following sense. Assume that $\tilde u \colon U \longrightarrow \mathbb{R}$ is a Lipschitz function defined on a neighborhood U of $\operatorname{spt}\mu$ with*

$$\tilde u(x) = \sup_{y \in \operatorname{spt}\nu} \big(q^{-1}\,l(x,y)^q - \tilde u^{(l^q)}(y)\big), \tag{4.17}$$

and suppose that the map $\tilde{\mathcal{F}}_1 \colon M \longrightarrow M$ given by

$$\tilde{\mathcal{F}}_1(x) = \exp_x\big(F^*(\mathrm{d}\tilde u)^{p-2}\,\mathscr{L}^*(\mathrm{d}\tilde u)\big)$$

provided $x \in \mathcal{D}(\mathrm{d}\tilde u)$, obeys $\nu = (\tilde{\mathcal{F}}_1)_\sharp\mu$. Then $\mathrm{d}\tilde u = \mathrm{d}u$ μ-a.e.

(iii) *In addition to the hypothesis of* (i)*, assume $\nu \in \mathcal{P}^{\mathrm{ac}}(M,\mathfrak{m})$ and define $\mathcal{G}_1 \colon M \longrightarrow M$ as in (4.14). Then we have $\mathcal{F}_1 \circ \mathcal{G}_1(y) = y$ for ν-a.e. $y \in M$ as well as $\mathcal{G}_1 \circ \mathcal{F}_1(x) = x$ for μ-a.e. $x \in M$.*

Proof. (i) Theorem 4.9 yields that $u = {}^{(l^q)}v$ on $\operatorname{spt}\mu$, $v = u^{(l^q)}$ on $\operatorname{spt}\nu$, and that u and v extend to semi-convex Lipschitz functions on neighborhoods of $\operatorname{spt}\mu$ and $\operatorname{spt}\nu$, respectively. Rademacher's theorem ensures that u is differentiable $\mathfrak{m}$-a.e. and thus μ-a.e. In particular, the measure $(\mathrm{Id},\mathcal{F}_1)_\sharp\mu$ makes sense.

Again by Theorem 4.9, π and (u,v) attain the respective supremum and infimum in (4.3). Let E be the equality set from (4.5). By Lemma 4.16(i), for π-a.e. $(x,y) \in E$ with $x \in \mathcal{D}(\mathrm{d}^2u)$ we have $y = \mathcal{F}_1(x)$ and $(x,\mathcal{F}_1(x)) \notin \mathrm{sing}(l)$. Since $M \setminus \mathcal{D}(\mathrm{d}^2u)$ is $\mathfrak{m}$-negligible by Alexandrov's theorem and hence μ-negligible, we obtain $\pi = (\mathrm{Id},\mathcal{F}_1)_\sharp\mu$, therefore, $(\mathrm{Id},\mathcal{F}_1)_\sharp\mu$ is a maximizer of $\ell_q(\mu,\nu)$. Given

any such maximizer π', Proposition 4.7 gives $\pi'[E] = 1$, and arguing as above shows $\pi' = \pi$ and hence the uniqueness of π.

(ii) First, we claim that $\tilde{\pi} := (\mathrm{Id}, \tilde{\mathcal{F}}_1)_\sharp \mu$ maximizes $\ell_q(\mu, \nu)$. By Lemma 4.16(i), for every $x \in \mathcal{D}(\mathrm{d}\tilde{u})$ the supremum of (4.17) is necessarily attained at $y = \tilde{\mathcal{F}}_1(x)$. As in (i), we see that $M \setminus \mathcal{D}(\mathrm{d}\tilde{u})$ is μ-negligible. Therefore,

$$q^{-1}\, l\big(x, \tilde{\mathcal{F}}_1(x)\big) = \tilde{u}(x) + \tilde{u}^{(l^q)}\big(\tilde{\mathcal{F}}_1(x)\big)$$

for μ-a.e. $x \in M$. By integrating this inequality against μ and using $\nu = (\tilde{\mathcal{F}}_1)_\sharp \mu$ as well as Proposition 4.7, we deduce the claim.

The uniqueness part of (i) thus gives $(\mathrm{Id}, \mathcal{F}_1)_\sharp \mu = (\mathrm{Id}, \tilde{\mathcal{F}}_1)_\sharp \mu$, whence $\mathcal{F}_1 = \tilde{\mathcal{F}}_1$ μ-a.e. We claim that in fact $\mathrm{d}u = \mathrm{d}\tilde{u}$ on $\mathcal{D}(\mathrm{d}u) \cap \mathcal{D}(\mathrm{d}\tilde{u})$ (a set which has μ-full measure, as seen above). If that fails at all points $x \in \mathcal{D}(\mathrm{d}u) \cap \mathcal{D}(\mathrm{d}\tilde{u})$ with $\mathcal{F}_1(x) = \tilde{\mathcal{F}}_1(x)$ in a set of positive μ-measure, there are two different maximizing geodesics connecting x and $\mathcal{F}_1(x)$. Hence $(x, \mathcal{F}_1(x)) \in \mathrm{sing}(l)$, contradicting (i).

(iii) An analogous argument as in (i), taking into account that $v = u^{(l^q)}$ on $\operatorname{spt}\nu$, gives $\pi = (\mathcal{G}_1, \mathrm{Id})_\sharp \nu$. Therefore, π is concentrated on $\mathcal{D}(\mathrm{d}u) \times \mathcal{D}(\mathrm{d}v)$. Hence $y = \mathcal{F}_1(x)$ and $x = \mathcal{G}_1(y)$ for π-a.e. $(x, y) \in M \times M$ by Lemma 4.16(i), which is the claim. □

Corollary 4.18 (Lagrangian characterization of q-geodesics)**.** *Given any $q \in (0,1)$, assume $(\mu_0, \mu_1) \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M, \mathfrak{m}) \times \mathcal{P}_{\mathrm{c}}(M)$ to be q-separated by (π, u, v). Define $\mathcal{F}_t \colon M \longrightarrow M$ as in (4.14), where $t \in [0,1]$. Then $((\mathcal{F}_t)_\sharp \mu_0)_{t\in[0,1]}$ is a unique q-geodesic from μ_0 to μ_1.*

Proof. By Theorem 4.17, π is the unique ℓ_q-optimal coupling of μ_0 and μ_1, and has the form $\pi = (\mathrm{Id}, \mathcal{F}_1)_\sharp \mu_0$. In particular, since

$$l\big(\mathcal{F}_s(x), \mathcal{F}_t(x)\big)^q = (t-s)^q\, l\big(x, \mathcal{F}_1(x)\big)^q$$

for μ_0-a.e. $x \in M$ and every $s, t \in [0,1]$ with $s < t$, $((\mathcal{F}_t)_\sharp \mu_0)_{t\in[0,1]}$ is a q-geodesic from μ_0 to μ_1; Corollary 4.4 thus implies the compactness of $\operatorname{spt}[(\mathcal{F}_t)_\sharp \mu_0]$.

To prove the uniqueness, let $(\mu_t)_{t\in[0,1]}$ be a q-geodesic from μ_0 to μ_1, and fix $t \in (0,1)$. Again μ_t is compactly supported by Corollary 4.4. In particular, by Lemma 4.1 μ_0 and μ_t as well as μ_t and μ_1 admit ℓ_q-optimal couplings. Then we fix $\omega \in \mathcal{P}(M \times M \times M)$ as given by Proposition 4.3. The ℓ_q-optimality of its marginals and uniqueness of π ensures that $\pi_{13} = \pi$, where $\pi_{13} := (\mathrm{proj}_{13})_\sharp \omega$. Hence $\pi_{13}[\mathrm{sing}(l)] = 0$ by Theorem 4.17(i). Combining Lemma 3.5 with the last part of Proposition 4.3 and $z_t(x, \mathcal{F}_1(x)) = \mathcal{F}_t(x)$ for every $x \in \mathcal{D}(\mathrm{d}u)$, cf. Remark 3.7, finally implies

$$\mu_t = (z_t)_\sharp \pi_{13} = (z_t)_\sharp \big((\mathrm{Id}, \mathcal{F}_1)_\sharp \mu_0\big) = (\mathcal{F}_t)_\sharp \mu_0$$

and hence the desired conclusion. □

In fact, under the hypotheses of Corollary 4.18 the interpolants $(\mathcal{F}_t)_\sharp \mu_0$ are absolutely continuous with respect to $\mathfrak{m}$ provided $t < 1$. We defer the proof of the following corresponding proposition to the next section since we need the measure contraction property stated in Lemma 5.4.

Proposition 4.19 (Absolute continuity of interpolants)**.** *Given any $q \in (0,1)$, let $(\mu_t)_{t\in[0,1]}$ denote a unique q-geodesic joining a pair of q-separated endpoints $(\mu_0,\mu_1) \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m}) \times \mathcal{P}_{\mathrm{c}}(M)$. Then $\mu_t \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m})$ for all $t \in (0,1)$.*

A direct consequence is the following *Monge–Ampère type equation* (also called the *Jacobian equation*; cf. Villani [65, Theorem 11.1]). It is shown in exactly the same way as McCann's result [45, Corollary 5.11] (whose main ingredient [45, Theorem 5.10] is a local statement that can be established chart-wise in our setting) and thus the proof is omitted.

Corollary 4.20 (Monge–Ampère type equation)**.** *Retain the assumptions and the notation of Proposition* 4.19*, and set $\mu_t := (\mathcal{F}_t)_\sharp \mu_0 = \rho_t\,\mathfrak{m}$ for $t \in [0,1)$. Then, for every $t \in (0,1)$ and μ_0-a.e. $x \in M$, $\mathrm{d}\mathcal{F}_t(x)$ exists and we have*

$$\rho_0(x) = \rho_t\big(\mathcal{F}_t(x)\big)\,\det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)].$$

This conclusion also holds for $t = 1$ provided $\mu_1 \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m})$.

Here, $\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]$ is the Jacobian (determinant) of $\mathrm{d}\mathcal{F}_t(x)$ with respect to $\mathfrak{m}$, namely

$$\det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)] := \frac{\mathfrak{m}_{\mathcal{F}_t(x)}[\mathrm{d}\mathcal{F}_t(x)(A)]}{\mathfrak{m}_x[A]},$$

where $\mathfrak{m}_x$ is the Lebesgue measure on T_xM induced from $\mathfrak{m}$ via coordinates and $A \subset T_xM$ is an arbitrary nonempty, bounded open set. We remark that $\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_0(x)] = 1$ and $\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)] > 0$ for μ_0-a.e. $x \in M$.

Finally, we remark that the statements from Theorem 4.17, Corollary 4.18, and Proposition 4.19 can be extended beyond q-separated measures. Compare this with Remark 5.8 below.

Corollary 4.21 (Optimal maps and q-geodesics without duality)**.** *Let $q \in (0,1)$, and let $\mu_0 \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m})$ as well as $\mu_1 \in \mathcal{P}_{\mathrm{c}}(M)$.*

(i) *There exists at most one chronological ℓ_q-optimal coupling $\pi \in \Pi(\mu_0,\mu_1)$.*

Moreover, if π is such a coupling, then we have the following.

(ii) *It can be written as $\pi = (\mathrm{Id},\mathcal{F}_1)_\sharp\mu_0$, where $\mathcal{F}_1 \colon M \longrightarrow M$ is a certain Borel map, and satisfies $\pi[\mathrm{sing}(l)] = 0$.*

(iii) *It decomposes as $\pi = \sum_{i\in\mathbb{N}} \pi_i$, where $(\pi_i)_{i\in\mathbb{N}}$ is a sequence of mutually singular sub-probability measures on $M\times M$ with the following properties for every $i \in \mathbb{N}$: The set $\operatorname{spt}\mu^i \times \operatorname{spt}\nu^i$, where μ^i and ν^i are the marginals of $\hat{\pi}_i := \pi_i[M\times M]^{-1}\,\pi_i$, is compact and disjoint from $\{l \geq 0\}$. In particular, the pair (μ^i,ν^i) is q-separated. $\mathcal{F}_1$ agrees μ^i-a.e. with the transport map $\mathcal{F}_1^i$ pushing forward μ^i to ν^i from Theorem* 4.17*. Lastly, the above sum is finite if and only if $\operatorname{spt}\pi$ is compact and disjoint from $\{l \leq 0\}$.*

(iv) *The assignment $\mu_t := (z_t)_\sharp\pi$, where z_t designates the map from Lemma* 4.5*, defines a q-geodesic $(\mu_t)_{t\in[0,1]}$ from μ_0 to μ_1 such that $\mu_t \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m})$ provided $t < 1$. Subject to the decomposition of π from* (ii)*, $((z_t)_\sharp\pi_i)_{i\in[0,1]}$ are mutually singular sub-probability measures for $t < 1$.*

Proof. This is proven in the same way as [45, Theorem 7.1]. A decomposition of π as in (ii) is obtained by dividing $\{l > 0\} \cap \operatorname{spt}\pi$ into countably many open and relatively compact rectangles whose closures are contained in the open set $\{l > 0\}$, and restriction of π to these. We skip the details. □

In view of the last claim of (iii), we stress that the condition $\pi[\{l > 0\}] = 1$ does not imply the disjointness of $\operatorname{spt}\pi$ from $\{l \leq 0\}$. The latter is thus a strictly stronger property.

5 Curvature-dimension condition from Ricci curvature bounds

In the previous section we saw that optimal transports are done along geodesics, and hence Jacobi fields play a key role to control their behavior. Then, it is natural to expect that the Ricci curvature comes into play, and we show in this section that a lower bound of the weighted Ricci curvature provides a convexity estimate for a certain entropy functional, called the curvature-dimension condition. We refer to [22, 39, 60, 61, 62] for the Riemannian case and [55] for the Finsler case.

5.1 Weighted Ricci curvature

Here and henceforth, we fix a smooth positive measure $\mathfrak{m}$ on a Finsler spacetime (M, L). Then the associated *weight function* $\psi_{\mathfrak{m}} \colon \overline{\Omega} \setminus \{0\} \longrightarrow \mathbb{R}$ is defined by

$$\mathrm{d}\mathfrak{m} = \mathrm{e}^{-\psi_{\mathfrak{m}}(\dot{\gamma}(t))} \sqrt{-\det\bigl[g_{\alpha\beta}\bigl(\dot{\gamma}(t)\bigr)\bigr]}\, \mathrm{d}x^1\, \mathrm{d}x^2 \cdots \mathrm{d}x^n$$

along timelike geodesics γ. We remark that $\psi_{\mathfrak{m}}$ is positively 0-homogeneous in the sense that $\psi_{\mathfrak{m}}(cv) = \psi_{\mathfrak{m}}(v)$ for every $v \in \overline{\Omega} \setminus \{0\}$ and $c > 0$.

Remark 5.1. In Lu–Minguzzi–Ohta [40, 41, 42], we dealt with a more general framework by directly considering a positively 0-homogeneous weight function $\psi \colon \overline{\Omega} \setminus \{0\} \longrightarrow \mathbb{R}$, which is not necessarily associated with a measure as above. In this paper, however, we begin with a measure $\mathfrak{m}$ since we need it as a reference measure to define entropy functionals. In addition, we will not use the *ϵ-range* introduced in [40] (in other words, we put $\epsilon = 1$).

For later convenience, along a timelike geodesic $\gamma \colon [0, l) \longrightarrow M$, we denote

$$\psi_\gamma(t) := \psi_{\mathfrak{m}}\bigl(\dot{\gamma}(t)\bigr). \tag{5.1}$$

Definition 5.2 (Weighted Ricci curvature)**.** Given a causal vector $v \in \overline{\Omega} \setminus \{0\}$, let $\gamma \colon (-\varepsilon, \varepsilon) \longrightarrow M$ be the causal geodesic with $\dot{\gamma}(0) = v$. Then, for $N \in \mathbb{R} \setminus \{n\}$, we define the *weighted Ricci curvature* at v by

$$\mathrm{Ric}_N(v) := \mathrm{Ric}(v) + \psi_\gamma''(0) - \frac{\psi_\gamma'(0)^2}{N - n}, \tag{5.2}$$

where $\mathrm{Ric}(v)$ is the usual Ricci curvature defined in Definition 2.2. We also set

$$\mathrm{Ric}_\infty(v) := \mathrm{Ric}(v) + \psi''_\gamma(0), \qquad \mathrm{Ric}_n(v) := \lim_{N \downarrow n} \mathrm{Ric}_N(v),$$

and $\mathrm{Ric}_N(0) := 0$.

Remark 5.3. (a) By definition, the quantity $\mathrm{Ric}_N(v)$ is non-decreasing in N in the ranges $[n, +\infty]$ and $(-\infty, n)$. Moreover, for $N \in (n, +\infty)$ and $N' \in (-\infty, 0)$ we have

$$\mathrm{Ric}_n(v) \le \mathrm{Ric}_N(v) \le \mathrm{Ric}_\infty(v) \le \mathrm{Ric}_{N'}(v) \le \mathrm{Ric}_0(v).$$

(b) In our previous papers [40, 41, 42], we chose the notation $\dim M = n + 1$ but adopted the same definition (5.2). As a result, Ric_{N+1} in this paper (and [45]) corresponds to Ric_N in [40, 41, 42].

One can obtain several singularity theorems [40] and comparison theorems [41] under the curvature bound $\mathrm{Ric}_N \ge K$ for $K \in \mathbb{R}$, by which we mean $\mathrm{Ric}_N(v) \ge KF^2(v) = -2KL(v)$ for all $v \in \Omega$ (and hence for all $v \in \overline{\Omega}$ by continuity). For giving a deferred proof of Proposition 4.19, here we explain the *timelike measure contraction property* as a corollary to (the proof of) the Bishop–Gromov volume comparison theorem in [41, Theorem 5.10] (see [53, 62] for the Riemannian measure contraction property). To this end, we define

$$\mathfrak{s}_\kappa(r) := \begin{cases} \dfrac{1}{\sqrt{\kappa}} \sin(\sqrt{\kappa} r) & \kappa > 0, \\ r & \kappa = 0, \\ \dfrac{1}{\sqrt{-\kappa}} \sinh(\sqrt{-\kappa} r) & \kappa < 0 \end{cases}$$

for $r \in \mathbb{R}$ when $\kappa \le 0$, and for $r \in [0, \pi/\sqrt{\kappa}]$ when $\kappa > 0$, as well as

$$\tau^{(t)}_{K,N}(r) := t^{1/N} \left(\frac{\mathfrak{s}_{K/(N-1)}(tr)}{\mathfrak{s}_{K/(N-1)}(r)} \right)^{(N-1)/N} \tag{5.3}$$

for $K \in \mathbb{R}$, $N \in (-\infty, 0) \cup [n, +\infty)$, and $t \in (0, 1)$, which is defined for $r > 0$ when $K/(N-1) \le 0$, and for $r \in (0, \pi\sqrt{(N-1)/K})$ when $K/(N-1) > 0$. We also define $\tau^{(t)}_{K,N}(0) := t$. The case $N = 0$ is not included in the definition (5.3) and will be dealt with separately in Definition 5.7.

Lemma 5.4 (Timelike measure contraction property)**.** *Let $(M, L, \mathfrak{m})$ satisfy $\mathrm{Ric}_N \ge K$ for some $K \in \mathbb{R}$ and $N \in [n, +\infty)$. Then we have*

$$\mathfrak{m}[Z_t(x, B)] \ge \inf_{y \in B} \tau^{(t)}_{K,N}\big(l(x,y)\big)^N \, \mathfrak{m}[B]$$

for every $x \in M$, every compact set $B \subset I^+(x)$, and every $t \in (0, 1)$. Moreover,

$$\mathfrak{m}[Z_t(A, y)] \ge \inf_{x \in A} \tau^{(1-t)}_{K,N}\big(l(x,y)\big)^N \, \mathfrak{m}[A]$$

for every $y \in M$, every compact set $A \subset I^-(y)$, and every $t \in (0, 1)$.

Proof. First of all, if $K > 0$, then the Lorentzian counterpart to the Bonnet–Myers theorem [40, Theorem 5.17] implies $l(x,y) \le \pi\sqrt{(N-1)/K}$, and hence $l(x,y) < \pi\sqrt{(N-1)/K}$ for $\mathfrak{m}$-a.e. $y \in I^+(x)$.

We first consider the set $Z_t(x,B)$. Note that x and $\mathfrak{m}$-a.e. $y \in I^+(x)$ are connected by a unique maximizing timelike geodesic $\gamma_y \colon [0,1] \longrightarrow M$, and

$$\mathfrak{m}[Z_t(x,B)] = \int_B \det\nolimits_{\mathfrak{m}}[\mathrm{d}\Psi_t]\,\mathrm{d}\mathfrak{m},$$

where $\det_{\mathfrak{m}}[\mathrm{d}\Psi_t]$ is the Jacobian of the map $\Psi_t(y) := \gamma_y(t)$ with respect to the measure $\mathfrak{m}$. Precisely, setting $w_y := \dot\gamma_y(0) \in T_xM$ and using an orthonormal basis $(e_\alpha)_{\alpha=1}^n$ of (T_xM, g_v) with $e_1 = w_y/l(x,y)$ as well as the vector fields $J_i(t) := \mathrm{d}[\exp_{tw_y}](te_i)$ along γ_y, where $i = 2,\dots,n$, we have

$$\det\nolimits_{\mathfrak{m}}[\mathrm{d}\Psi_t(y)] = \frac{t\,\phi_y(t)}{\phi_y(1)}, \qquad \phi_y(t) := \mathrm{e}^{-\psi_{\gamma_y}(t)}\sqrt{\det\big[g_{\dot\gamma_y(t)}\big(J_i(t),J_j(t)\big)\big]}.$$

The function $s \longmapsto \phi_y(s/l(x,y))^{1/(N-1)}$ coincides with $h = h_1$ in the proof of [41, Theorem 5.10] with $\epsilon = 1$ and $c = 1/(N-1)$, and it follows from the hypothesis $\mathrm{Ric}_N \ge K$ that $h/\mathfrak{s}_{K/(N-1)}$ is non-increasing. Hence,

$$\begin{aligned}
\mathfrak{m}[Z_t(x,B)] &= \int_B \frac{t\,\phi_y(t)}{\phi_y(1)}\,\mathrm{d}\mathfrak{m}(y) \\
&\ge \int_B t\,\frac{\mathfrak{s}_{K/(N-1)}(t\,l(x,y))^{N-1}}{\mathfrak{s}_{K/(N-1)}(l(x,y))^{N-1}}\,\mathrm{d}\mathfrak{m}(y) \\
&\ge \inf_{y\in B} \tau_{K,N}^{(t)}\big(l(x,y)\big)^N\,\mathfrak{m}[B].
\end{aligned}$$

The assertion on $Z_t(A,y)$ can be reduced to the reverse structure $\overline{L}$, for which $\mathrm{Ric}_N \ge K$ (with respect to $\mathfrak{m}$) also holds since

$$\mathrm{Ric}_N^{\overline{L}}(v) = \mathrm{Ric}_N^{L}(-v) \ge -2KL(-v) = -2K\overline{L}(v)$$

for every tangent vector v which is causal with respect to $\overline{L}$. □

5.2 Absolute continuity of intermediate measures

We are ready to provide a proof of Proposition 4.19, which we restate below for convenience. The Lorentzian version of it has been shown by McCann [45, Theorem A.1] by establishing a modified Monge–Mather shortening estimate, see also [17, 65]. We propose an alternative proof based on ideas of Figalli–Juillet [26]; a byproduct of the proof is a qualitative density bound we do not state explicitly, but which should be compared to Braun [14].

Proposition 5.5. *Given any $q \in (0,1)$, let $(\mu_t)_{t\in[0,1]}$ denote a unique q-geodesic joining a pair of q-separated endpoints $(\mu_0,\mu_1) \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m}) \times \mathcal{P}_{\mathrm{c}}(M)$ according to Corollary 4.18. Then $\mu_t \in \mathcal{P}_{\mathrm{c}}^{\mathrm{ac}}(M,\mathfrak{m})$ for every $t \in (0,1)$.*

Proof. We first check the case $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1 \subset \{l > 0\}$, i.e. l is bounded and bounded away from zero on the product $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1$. Let $K \in \mathbb{R}$ and $N \in (n, +\infty)$ be fixed numbers such that $\mathrm{Ric}_N \geq K$ in all timelike directions on the compact causal diamond $J^+(\mu_0) \cap J^-(\mu_1)$. Then we claim that

$$\int_{\{\rho_0>0\}} f \circ \mathcal{F}_t \,\mathrm{d}\mathfrak{m} \leq \sup_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{-N} \int_M f \,\mathrm{d}\mathfrak{m} \tag{5.4}$$

for every nonnegative, continuous function f on M with compact support.

To this aim, we fix a sequence $(\mu_1^k)_{k\in\mathbb{N}}$ of compactly supported probability measures of the form $\mu_1^k = \lambda_1^k\,\delta_{x_1^k} + \cdots + \lambda_k^k\,\delta_{x_k^k}$ which weakly converges to μ_1; here, given any $k \in \mathbb{N}$, $x_1^k, \dots, x_k^k$ are points in $\operatorname{spt}\mu_1$, and $\lambda_1^k, \dots, \lambda_k^k$ are nonnegative real numbers which sum up to 1. Since $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1^k \subset \{l > 0\}$, (μ_0, μ_1^k) is q-separated by Lemma 4.10. Theorem 4.17 and Corollary 4.18 imply the existence of appropriate Borel maps $\mathcal{F}_t^k \colon M \longrightarrow M$, where $t \in [0,1]$, with the properties

- $\mathcal{F}_1^k$ induces the unique ℓ_q-optimal coupling of (μ_0, μ_1^k), and
- $((\mathcal{F}_t^k)_\sharp\mu_0)_{t\in[0,1]}$ constitutes the unique q-geodesic from μ_0 to μ_1^k.

Let $A_i^k \subset \{\rho_0 > 0\}$ be the set of all points $x \in \{\rho_0 > 0\}$ with $\mathcal{F}_1^k(x) = x_i^k$, where $i = 1, \dots, k$. Then $A_1^k, \dots, A_k^k$ are mutually disjoint, and their union has full μ_0-measure. Given any $i = 1, \dots, k$, Theorem 4.17 and Remark 3.7 imply the curve $(\mathcal{F}_t^k(x))_{t\in[0,1]}$ to form the unique maximizing timelike geodesic from x to x_i^k for μ_0-a.e. $x \in A_i^k$. Since μ_0 and $\mathfrak{m}$ are equivalent on $\{\rho_0 > 0\}$, the same statement holds for every $x \in B_i^k$, where $B_i^k \subset A_i^k$ is an appropriate Borel set with $\mathfrak{m}[A_i^k \setminus B_i^k] = 0$. In particular $Z_t(E', x_i^k) = \mathcal{F}_t^k(E')$ for every Borel set $E' \subset B_i^k$, for which Lemma 5.4 (applied to the Finsler spacetime built by the causal diamond $J^+(\mu_0) \cap J^-(\mu_1)$) implies

$$\mathfrak{m}\big[\mathcal{F}_t^k(E')\big] \geq \inf_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{N}\,\mathfrak{m}[E'].$$

In particular, by appropriately decomposing an arbitrary Borel set $E \subset \{\rho_0 > 0\}$ we get

$$\mathfrak{m}\big[\mathcal{F}_t^k(E)\big] \geq \inf_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{N}\,\mathfrak{m}[E].$$

This inequality can be rewritten as

$$\mathfrak{m}[F] \geq \inf_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{N}\,\mathfrak{m}\big[(\mathcal{F}_t^k)^{-1}(F) \cap \{\rho_0 > 0\}\big]$$

for every Borel set $F \subset \mathcal{F}_t^k(\{\rho_0 > 0\})$, or equivalently

$$\int_{\{\rho_0>0\}} f \circ \mathcal{F}_t^k \,\mathrm{d}\mathfrak{m} \leq \sup_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{-N} \int_M f \,\mathrm{d}\mathfrak{m} \tag{5.5}$$

for every function f as above. Now by the q-separation of (μ_0, μ_1) and the stability of optimal maps [65, Corollary 5.23], $(\mathcal{F}_t^k)_{k\in\mathbb{N}}$ converges to $\mathcal{F}_t$ in μ_0-measure, and thus $\mathcal{F}_t^k \to \mathcal{F}_t$ $\mathfrak{m}$-a.e. on $\{\rho_0 > 0\}$ as $k \to +\infty$ up to a non-relabeled subsequence. Passing to the limit in (5.5) then shows (5.4).

An approximation argument now allows us to extend (5.4) to nonnegative Borel functions f on M. For $f := 1_{\mathcal{F}_t(\{\rho_0>0\})}\, h\, \rho_0 \circ \mathcal{F}_t^{-1}$, where h is a given nonnegative and bounded Borel function on M, this translates into

$$\begin{aligned}\int_{\mathcal{F}_t(\{\rho_0>0\})} h \,\mathrm{d}\mu_t &= \int_{\{\rho_0>0\}} h \circ \mathcal{F}_t \,\mathrm{d}\mu_0 \\ &= \int_{\{\rho_0>0\}} (h \circ \mathcal{F}_t)\, \rho_0 \,\mathrm{d}\mathfrak{m} \\ &\le \sup_{x\in\operatorname{spt}\mu_0, y\in\operatorname{spt}\mu_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)^{-N} \int_M h\, \rho_0 \circ \mathcal{F}_t^{-1} \,\mathrm{d}\mathfrak{m}.\end{aligned}$$

This implies the claim under our initial assumptions on μ_0 and μ_1.

The general case now follows by sub-partition as in [13, Proposition 3.38]. The proof is similar to the one of Corollary 4.21. We skip the details. □

5.3 Timelike curvature-dimension condition

To introduce the timelike curvature-dimension conditions we will study in a moment, we need to define several entropy functionals on $\mathcal{P}(M)$. Though we give a definition for general probability measures, in fact we only deal with $\mathfrak{m}$-absolutely continuous measures.

Definition 5.6 (Entropies)**.** Let $\mu = \rho\,\mathfrak{m} + \mu_\perp$ denote the Lebesgue decomposition of a given probability measure $\mu \in \mathcal{P}(M)$ into its $\mathfrak{m}$-absolutely continuous and $\mathfrak{m}$-singular parts.

(1) The *Boltzmann–Shannon entropy* of μ is defined by

$$\operatorname{Ent}_{\mathfrak{m}}(\mu) := \int_M \rho \log \rho \,\mathrm{d}\mathfrak{m}$$

if $\mu \in \mathcal{P}^{\mathrm{ac}}(M,\mathfrak{m})$ and $\int_{\{\rho<1\}} \rho\log\rho\,\mathrm{d}\mathfrak{m} > -\infty$, otherwise $\operatorname{Ent}_{\mathfrak{m}}(\mu) := +\infty$.

(2) For $n \le N < +\infty$ we define the *N-Rényi entropy* of μ by

$$S_{\mathfrak{m}}^N(\mu) := -\int_M \rho^{(N-1)/N} \,\mathrm{d}\mathfrak{m}.$$

(3) For $N < 0$, the corresponding *N-Rényi entropy* of μ is

$$S_{\mathfrak{m}}^N(\mu) := \int_M \rho^{(N-1)/N} \,\mathrm{d}\mathfrak{m}$$

if $\mu \in \mathcal{P}^{\mathrm{ac}}(M,\mathfrak{m})$, otherwise $S_{\mathfrak{m}}^N(\mu) := +\infty$.

(4) The *0-Rényi entropy* of μ is given by

$$S_{\mathfrak{m}}^0(\mu) := \operatorname*{ess\,sup}_{x\in M} \rho(x) = \|\rho\|_{L^\infty(M,\mathfrak{m})}$$

if $\mu \in \mathcal{P}^{\mathrm{ac}}(M,\mathfrak{m})$, otherwise $S_{\mathfrak{m}}^0(\mu) := +\infty$.

We remark that the generation functions $r \longmapsto r \log r$, $r \longmapsto -r^{(N-1)/N}$ (for $n \le N < +\infty$), and $r \longmapsto r^{(N-1)/N}$ (for $N < 0$) are all convex on $[0, \infty)$. (This explains the change of sign for $S^N_{\mathfrak{m}}$ in passing from positive to negative N.) In addition, we have $\mathrm{Ent}_{\mathfrak{m}}(\mu) \ge -\log \mathfrak{m}[\mathrm{spt}\,\mu]$ for $\mu \in \mathcal{P}_{\mathrm{c}}(M)$ by Jensen's inequality, and similarly we obtain a lower bound for $S^N_{\mathfrak{m}}(\mu)$ for $N \in (-\infty, 0] \cup [n, +\infty)$, see the proof of Theorem 5.11.

Now we introduce the *timelike curvature-dimension condition* $\mathrm{TCD}_q(K, N)$ partly using the function $\tau^{(t)}_{K,N}$ in (5.3).

Definition 5.7 (Timelike curvature-dimension condition)**.** Let $q \in (0, 1)$, $K \in \mathbb{R}$, and $N \in (-\infty, 0] \cup [n, +\infty]$. We say that a weighted Finsler spacetime $(M, L, \mathfrak{m})$ satisfies $\mathrm{TCD}_q(K, N)$ if, for every q-separated $(\mu_0, \mu_1) \in \mathcal{P}^{\mathrm{ac}}_{\mathrm{c}}(M, \mathfrak{m}) \times \mathcal{P}^{\mathrm{ac}}_{\mathrm{c}}(M, \mathfrak{m})$ there exist a q-geodesic $(\mu_t)_{t\in[0,1]}$ from μ_0 and μ_1 and an ℓ_q-optimal coupling $\pi \in \Pi(\mu_0, \mu_1)$ such that the following hold.

(1) When $N = +\infty$, for every $t \in [0, 1]$ we have

$$\begin{aligned} \mathrm{Ent}_{\mathfrak{m}}(\mu_t) \le (1-t)\,\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\,\mathrm{Ent}_{\mathfrak{m}}(\mu_1) \\ - \frac{K}{2}\,t\,(1-t) \int_{M\times M} l(x,y)^2 \,\mathrm{d}\pi(x,y). \end{aligned} \tag{5.6}$$

(2) When $n \le N < +\infty$, for every $N' \ge N$ and every $t \in [0, 1]$ we have

$$\begin{aligned} S^{N'}_{\mathfrak{m}}(\mu_t) \le -\int_{M\times M} \tau^{(1-t)}_{K,N'}\big(l(x,y)\big)\,\rho_0(x)^{-1/N'}\,\mathrm{d}\pi(x,y) \\ - \int_{M\times M} \tau^{(t)}_{K,N'}\big(l(x,y)\big)\,\rho_1(y)^{-1/N'}\,\mathrm{d}\pi(x,y). \end{aligned} \tag{5.7}$$

Here ρ_0 and ρ_1 denote the $\mathfrak{m}$-densities of μ_0 and μ_1, respectively.

(3) When $N < 0$, for every $N' \in [N, 0)$ and every $t \in [0, 1]$ we have

$$\begin{aligned} S^{N'}_{\mathfrak{m}}(\mu_t) \le \int_{M\times M} \tau^{(1-t)}_{K,N'}\big(l(x,y)\big)\,\rho_0(x)^{-1/N'}\,\mathrm{d}\pi(x,y) \\ + \int_{M\times M} \tau^{(t)}_{K,N'}\big(l(x,y)\big)\,\rho_1(y)^{-1/N'}\mathrm{d}\pi(x,y), \end{aligned} \tag{5.8}$$

where we set $\tau^{(t)}_{K,N'}(r) := +\infty$ if $K < 0$ and $r \ge \pi\sqrt{(N'-1)/K}$.

(4) When $N = 0$, for every $t \in [0, 1]$ we have

$$\begin{aligned} S^0_{\mathfrak{m}}(\mu_t) \le \operatorname*{ess\,sup}_{(x,y)\in \mathrm{spt}\,\pi} \max\left\{ \frac{\mathfrak{s}_{-K}((1-t)\,l(x,y))}{(1-t)\,\mathfrak{s}_{-K}(l(x,y))}\,\rho_0(x), \right. \\ \left. \frac{\mathfrak{s}_{-K}(t\,l(x,y))}{t\,\mathfrak{s}_{-K}(l(x,y))}\,\rho_1(y) \right\}, \end{aligned} \tag{5.9}$$

where we set $\mathfrak{s}_{-K}(tl(x,y))/(t\mathfrak{s}_{-K}(l(x,y))) := 1$ if $l(x, y) = 0$.

We refer to Ohta [56, 57] for the curvature-dimension condition for Riemannian or Finsler manifolds with $N < 0$ and $N = 0$, respectively. Note that, in the case of $N < 0$ and $K < 0$, $\mathrm{TCD}_q(K, N)$ provides only a local control since the defining inequality (5.8) is trivial if

$$\pi\big[\big\{(x, y) \in M \times M \,\big|\, l(x, y) \geq \pi\sqrt{(N-1)/K}\big\}\big] > 0.$$

We also remark that π is an ℓ_q-optimal coupling of μ_0 and μ_1, thereby the last term $\int_{M\times M} l(x, y)^2 \,\mathrm{d}\pi(x, y)$ in (5.6) is not directly written as a transport cost. In fact, π is a unique ℓ_q-optimal coupling of μ_0 and μ_1 by Lemma 4.10 and Theorem 4.17, and $(\mu_t)_{t\in[0,1]}$ is a unique q-geodesic connecting its endpoints by Corollary 4.18. Here we have implicitly used that all couplings of μ_0 and μ_1 are chronological, which leads us to a further remark on the chronology assumption on μ_0 and μ_1 in Definition 5.7.

Remark 5.8 (Timelike q-dualizability)**.** The hypothesis of q-separation in Definition 5.7 lies in between strong timelike p-dualizability and timelike p-dualizability of (μ_0, μ_1) as introduced by Cavalletti–Mondino [21, Definitions 2.18, 2.27]. These are the properties in whose terms strong and weak timelike curvature-dimension conditions have been defined in Cavalletti–Mondino [21] and Braun [13]. In the young literature about TCD spaces this distinction is still necessary for stability questions, since chronology is not a closed condition.

The results in Theorem 5.9 and Theorem 6.1 still hold if Definition 4.8 was instead given for timelike p-dualizable or strongly timelike p-dualizable pairs (μ_0, μ_1). For $N \in [n, +\infty)$ the equivalence of these various TCD conditions is due to Braun [13]. The proof employs a construction similar to Corollary 4.21, convexity, and a timelike non-branching property (which is clear in our case). With some work, this can be extended to the range $N \in (-\infty, 0] \cup \{+\infty\}$. One could even merely assume (μ_0, μ_1) to obey $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1 \subset \{l > 0\}$, cf. Lemma 4.10. This property should be easier to verify in practice, but as observed in Section 4.3, unlike the other conditions (see Proposition 4.14 and [14, Lemma 3.1]) it does not propagate along q-geodesics.

Integrating the behavior of optimal transports in the previous section and the control of $\mathfrak{m}$ by the weighted Ricci curvature, we arrive at the following main result.

Theorem 5.9 ($\mathrm{Ric}_N \geq K$ implies $\mathrm{TCD}_q(K, N)$)**.** *Let $(M, L, \mathfrak{m})$ be a globally hyperbolic weighted Finsler spacetime satisfying $\mathrm{Ric}_N \geq K$ in timelike directions for some $K \in \mathbb{R}$ and $N \in (-\infty, 0] \cup [n, +\infty]$. Then it satisfies $\mathrm{TCD}_q(K, N)$ for any $q \in (0, 1)$.*

Proof. We divide the proof into five steps.

Step 1. We start with some preparations. Throughout this proof, we fix a pair $(\mu_0, \mu_1) \in \mathcal{P}_c^{\mathrm{ac}}(M, \mathfrak{m}) \times \mathcal{P}_c^{\mathrm{ac}}(M, \mathfrak{m})$ which is q-separated, say by (π, u, v). Let $(\mu_t)_{t\in[0,1]}$ be the unique q-geodesic which connects μ_0 to μ_1. Recall from Theorem 4.17 and Corollary 4.18 that $\mu_t = (\mathcal{F}_t)_\sharp\mu_0$ for every $t \in (0, 1)$, where

$$\mathcal{F}_t(x) := \exp_x\big(t\, F^*(\mathrm{d}u)^{p-2}\, \mathscr{L}^*(\mathrm{d}u)\big).$$

Moreover, $\mu_t = \rho_t\,\mathfrak{m}$ is $\mathfrak{m}$-absolutely continuous, and Corollary 4.20 yields

$$\rho_0(x) = \rho_t\big(\mathcal{F}_t(x)\big)\,\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]$$

for μ_0-a.e. $x \in M$. To be more precise, if $\mathrm{d}u(x) \neq 0$ (equivalently, $x \neq \mathcal{F}_1(x)$), one can write down

$$\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)] = \mathrm{e}^{\psi_\gamma(0)-\psi_\gamma(t)}\,\det[\mathrm{d}\mathcal{F}_t(x)]$$

for the maximizing geodesic $\gamma\colon [0,1] \longrightarrow M$ from x to $\mathcal{F}_1(x)$ and ψ_γ given by (5.1), where $\det[\mathrm{d}\mathcal{F}_t(x)]$ denotes the Jacobian associated with the measure $\mathrm{vol}_{g_{\dot\gamma}}$ along γ induced from the Lorentzian metric $g_{\dot\gamma}$ as in (2.2).

Step 2. Now, we can follow essentially the same lines as the Lorentzian case in [45], however, we need more detailed calculations to improve $\mathrm{TCD}^*(K,N)$ to $\mathrm{TCD}(K,N)$. Here we give the detailed proof for completeness. Fix $x \in \operatorname{spt}\mu_0$ with $\mathrm{d}u(x) \neq 0$ as above and take an $g_{\dot\gamma}$-orthonormal basis $(e_\alpha)_{\alpha=1}^n$ of T_xM with $e_1 = \dot\gamma(0)/F(\dot\gamma(0))$. That is, we have $g_{\dot\gamma}(e_1,e_1) = -1$, $g_{\dot\gamma}(e_i,e_i) = 1$ for $i \geq 2$, and $g_{\dot\gamma}(e_\alpha,e_\beta) = 0$ for $1 \leq \alpha < \beta \leq n$. We consider the $g_{\dot\gamma}$-parallel vector fields E_α along γ with $E_\alpha(0) = e_\alpha$ (i.e. $D^{\dot\gamma}_{\dot\gamma}E_\alpha = 0$) as well as the Jacobi fields $J_\alpha(t) := \mathrm{d}\mathcal{F}_t(x)(e_\alpha)$ along γ. We remark that $J_1(t)$ is not necessarily tangent to $\dot\gamma(t)$. For simplicity, we denote the covariant derivative along γ by $'$, namely $J'_\alpha := D^{\dot\gamma}_{\dot\gamma}J_\alpha$. Then we define the $n\times n$ matrix $J(t) = (J_{\alpha\beta}(t))_{\alpha,\beta=1}^n$ by

$$J_\alpha(t) := \sum_{\beta=1}^n J_{\alpha\beta}(t)\,E_\beta(t).$$

We have $J''_\alpha = \sum_{\beta=1}^n J''_{\alpha\beta}\,E_\beta$ and observe from the Jacobi equation that

$$\begin{aligned}
J''_\alpha &= -R_{\dot\gamma}(J_\alpha) = -\sum_{\delta=1}^n J_{\alpha\delta}\,R_{\dot\gamma}(E_\delta)\\
&= \sum_{\delta=1}^n J_{\alpha\delta}\,g_{\dot\gamma}\big(R_{\dot\gamma}(E_\delta),E_1\big)E_1 - \sum_{\delta=1}^n\sum_{j=2}^n J_{\alpha\delta}\,g_{\dot\gamma}\big(R_{\dot\gamma}(E_\delta),E_j\big)E_j\\
&= -\sum_{\beta,\delta=1}^n J_{\alpha\delta}\,g_{\dot\gamma}\big(R_{\dot\gamma}(E_\delta),E_\beta\big)E_\beta;
\end{aligned}$$

in the last line, we have used the identity $g_{\dot\gamma}(R_{\dot\gamma}(E_\delta),E_1) = g_{\dot\gamma}(R_{\dot\gamma}(E_1),E_\delta) = 0$ provided by $E_1 = \dot\gamma/F(\dot\gamma)$ and (2.5). Hence, by letting

$$B := J'\,J^{-1}, \qquad R(t) := \Big(g_{\dot\gamma}\big(R_{\dot\gamma}(E_\alpha),E_\beta\big)(t)\Big)_{\alpha,\beta=1}^n,$$

we have the *Riccati equation*

$$B' = J''\,J^{-1} - (J'\,J^{-1})^2 = -J\,R\,J^{-1} - B^2. \tag{5.10}$$

This can be rewritten as

$$D' = -J^{-1}\,J'\,J^{-1}\,B\,J - R - J^{-1}\,B^2\,J + J^{-1}\,B\,J' = -R - D^2,$$

where $D := J^{-1}\,B\,J$. Therefore, taking the trace yields

$$\mathrm{trace}[B'] + \mathrm{trace}[B^2] = \mathrm{trace}[D'] + \mathrm{trace}[D^2] = -\mathrm{Ric}(\dot\gamma). \tag{5.11}$$

Step 3. Next, we have a closer look on $\mathrm{trace}[B^2]$ at $t = 0$. This is the step where we need a more delicate analysis to achieve $\mathrm{TCD}(K, N)$. First of all, since $J_\alpha(0) = E_\alpha(0)$, $J(0)$ is the identity matrix and we have $D(0) = B(0)$. By the choice of $(e_\alpha)_{\alpha=1}^n$,

$$B_{\alpha 1}(0) = -g_{\dot\gamma}\big(J'_\alpha(0), e_1\big), \qquad B_{\alpha j}(0) = g_{\dot\gamma}\big(J'_\alpha(0), e_j\big)$$

for every $j \geq 2$. We recall from Lemma 4.16(iii) that

$$J'_\alpha(0) = F^*(\mathrm{d}u)^{p-2}\,\nabla^2 u(e_\alpha) + \mathrm{d}[F^*(\mathrm{d}u)^{p-2}](e_\alpha)\,\nabla u(x),$$

and observe from the identities $F^*(\mathrm{d}u)^2 = -g_{\nabla u}(\nabla u, \nabla u)$, [40, (3.2)], as well as $e_1 = \nabla u(x)/F^*(\mathrm{d}u)$ that

$$\begin{aligned}\mathrm{d}[F^*(\mathrm{d}u)^{p-2}](e_\alpha)\,\nabla u(x) &= -(p-2)\,F^*(\mathrm{d}u)^{p-4}\,g_{\nabla u}\big(\nabla^2 u(e_\alpha), \nabla u\big)\,\nabla u(x)\\ &= -(p-2)\,F^*(\mathrm{d}u)^{p-2}\,g_{\nabla u}\big(\nabla^2 u(e_\alpha), e_1\big)\,e_1.\end{aligned}$$

Therefore,

$$B_{\alpha 1}(0) = (1-p)\,F^*(\mathrm{d}u)^{p-2}\,g_{\nabla u}\big(\nabla^2 u(e_\alpha), e_1\big),$$

while for $j \geq 2$ we obtain

$$B_{\alpha j}(0) = F^*(\mathrm{d}u)^{p-2}\,g_{\nabla u}\big(\nabla^2 u(e_\alpha), e_j\big).$$

Thus we can decompose $B(0) = F^*(\mathrm{d}u)^{p-2}\,H^u\,Q^2$, where

$$H^u := \Big(g_{\nabla u}\big(\nabla^2 u(e_\alpha), e_\beta\big)\Big)_{\alpha,\beta=1}^n, \qquad Q := \mathrm{diag}\big(\sqrt{1-p}, 1, \dots, 1\big). \tag{5.12}$$

Note that H^u is symmetric [41, Lemma 4.13].

Since $D(0) = B(0)$, we deduce from (5.10), (2.5), and the symmetry of H^u that

$$\begin{aligned}D'_{11}(0) &= -\sum_{\alpha=1}^n D_{1\alpha}(0)\,D_{\alpha 1}(0)\\ &= -F^*(\mathrm{d}u)^{2p-4}\,\big(H^u\,Q^2\,H^u\,Q^2\big)_{11}\\ &= -F^*(\mathrm{d}u)^{2p-4}\left((1-p)^2\,(H^u_{11})^2 + (1-p)\sum_{j=2}^n (H^u_{1j})^2\right)\\ &\leq -F^*(\mathrm{d}u)^{2p-4}\,(1-p)^2\,(H^u_{11})^2\\ &= -D_{11}(0)^2.\end{aligned} \tag{5.13}$$

We set $\phi(t) := \log\det[\mathrm{d}\mathcal{F}_t(x)] = \log\det[J(t)]$ and decompose it as

$$\phi_1(t) := \int_0^t D_{11}(s)\,\mathrm{d}s, \qquad \phi_2(t) := \phi(t) - \phi_1(t).$$

Note that $\phi(0) = \phi_1(0) = \phi_2(0) = 0$ and

$$\begin{aligned}\phi'(t) &= \operatorname{trace}\big[J'(t)\,J(t)^{-1}\big] = \operatorname{trace}[B(t)] = \operatorname{trace}[D(t)],\\ \phi_2'(0) &= \sum_{i=2}^n D_{ii}(0) = F^*(\mathrm{d}u)^{p-2}\sum_{i=2}^n H^u_{ii},\end{aligned}$$

and (5.13) yields

$$\big[\mathrm{e}^{\phi_1(t)}\big]_0'' = \big(D_{11}'(0) + D_{11}(0)^2\big)\,\mathrm{e}^{\phi_1(0)} \le 0. \tag{5.14}$$

Moreover, (5.11) implies

$$\phi_2''(t) = \phi''(t) - \phi_1''(t) = -\mathrm{Ric}\big(\dot\gamma(t)\big) - \operatorname{trace}[D(t)^2] - D_{11}'(t).$$

If follows from the symmetry of H^u and the Cauchy–Schwarz inequality that

$$\begin{aligned}\operatorname{trace}[D(0)^2] + D_{11}'(0) &= \sum_{\alpha,\beta=1}^n D_{\alpha\beta}(0)\,D_{\beta\alpha}(0) - \sum_{\beta=1}^n D_{1\beta}(0)\,D_{\beta1}(0)\\ &= F^*(\mathrm{d}u)^{2p-4}\left((1-p)\sum_{j=2}^n (H^u_{1j})^2 + \sum_{i,j=2}^n (H^u_{ij})^2\right)\\ &\ge F^*(\mathrm{d}u)^{2p-4}\sum_{i=2}^n (H^u_{ii})^2\\ &\ge F^*(\mathrm{d}u)^{2p-4}\,\frac{1}{n-1}\left(\sum_{i=2}^n H^u_{ii}\right)^2\\ &= \frac{1}{n-1}\,\phi_2'(0)^2.\end{aligned}$$

Combining these we obtain

$$\frac{\big[\mathrm{e}^{\phi_2(t)/(n-1)}\big]_0''}{\mathrm{e}^{\phi_2(0)/(n-1)}} = \frac{\phi_2''(0)}{n-1} + \frac{\phi_2'(0)^2}{(n-1)^2} \le -\frac{\mathrm{Ric}(\dot\gamma(0))}{n-1}. \tag{5.15}$$

We further consider the functions, for $N \in (-\infty, 0] \cup (n, +\infty)$,

$$\begin{aligned} c(t) &:= \det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]^{1/N}, & c_1(t) &:= \exp\left(\frac{\psi_\gamma(0) - \psi_\gamma(t)}{N-n}\right),\\ c_2(t) &:= \mathrm{e}^{\phi(t)/n}, & c_3(t) &:= c_1(t)^{(N-n)/(N-1)}\,\mathrm{e}^{\phi_2(t)/(N-1)},\end{aligned}$$

where we assume $N \neq 0$ for $c(t)$ and $c_2(t)$ will be used in the proof of Theorem 6.1. Note that, since $N \in (-\infty, 0] \cup (n, +\infty)$,

$$\begin{aligned}(N-1)\,\frac{c_3''}{c_3} &= (N-n)\,\frac{c_1''}{c_1} + (n-1)\,\frac{[\mathrm{e}^{\phi_2/(n-1)}]''}{\mathrm{e}^{\phi_2/(n-1)}} \\ &\qquad - \frac{(n-1)(N-n)}{N-1}\left(\frac{c_1'}{c_1} - \frac{[\mathrm{e}^{\phi_2/(n-1)}]'}{\mathrm{e}^{\phi_2/(n-1)}}\right)^2 \\ &\leq -\psi_\gamma'' + \frac{(\psi_\gamma')^2}{N-n} + (n-1)\,\frac{[\mathrm{e}^{\phi_2/(n-1)}]''}{\mathrm{e}^{\phi_2/(n-1)}}.\end{aligned}$$

Therefore, at $t = 0$, we obtain from (5.15) and the hypothesis $\mathrm{Ric}_N \geq K$ that

$$(N-1)\,\frac{c_3''(0)}{c_3(0)} \leq -\mathrm{Ric}_N\big(\dot\gamma(0)\big) \leq -K\,F^2\big(\dot\gamma(0)\big). \tag{5.16}$$

Step 4. We shall generalize (5.14) and (5.16) to $t > 0$. When we perform the same calculation at $t_0 \in (0,1)$ with the orthonormal basis $(E_\alpha(t_0))_{\alpha=1}^n$, then $J(t)$, $B(t)$, and $D(t)$ are respectively replaced by

$$J^{t_0}(t) = J(t_0)^{-1}\,J(t), \qquad B^{t_0}(t) = J(t_0)^{-1}\,B(t)\,J(t_0), \qquad D^{t_0}(t) = D(t)$$

for $t \geq t_0$. This immediately implies (5.14) as well as (5.15) at t_0, and then (5.16) also follows.

On the one hand, we deduce from the concavity inequality (5.14) that

$$\mathrm{e}^{\phi_1(t)} \geq (1-t)\,\mathrm{e}^{\phi_1(0)} + t\,\mathrm{e}^{\phi_1(1)} = (1-t) + t\,\mathrm{e}^{\phi_1(1)}. \tag{5.17}$$

On the other hand, (5.16) implies that $-(N-1)\log c_3$ is $(KF^2(\dot\gamma), N-1)$-convex in the sense of [25, Lemma 2.2] for $N \in (n, +\infty)$ or [56, Lemma 2.1] for $N \leq 0$, respectively. Thus, we obtain

$$c_3(t) \geq \frac{\mathfrak{s}_{K/(N-1)}((1-t)\,d(x))}{\mathfrak{s}_{K/(N-1)}(d(x))} + \frac{\mathfrak{s}_{K/(N-1)}(t\,d(x))}{\mathfrak{s}_{K/(N-1)}(d(x))}\,c_3(1) \tag{5.18}$$

for $N \in (n, +\infty)$, and

$$c_3(t) \leq \frac{\mathfrak{s}_{K/(N-1)}((1-t)\,d(x))}{\mathfrak{s}_{K/(N-1)}(d(x))} + \frac{\mathfrak{s}_{K/(N-1)}(t\,d(x))}{\mathfrak{s}_{K/(N-1)}(d(x))}\,c_3(1) \tag{5.19}$$

for $N \leq 0$, where we set $d(x) := l(x, \mathcal{F}_1(x)) = F(\dot\gamma)$ for simplicity.

Step 5. In this final step, we turn to the proofs of the respective entropic displacement semi-convexity. We first consider the case of $N \in (n, +\infty)$. We remark that, thanks to the monotonicity as in Remark 5.3, it suffices to show

(5.7) for $N' = N$. Combining (5.17), (5.18), and Hölder's inequality, we obtain

$$\begin{aligned}
c(t) &= c_3(t)^{(N-1)/N}\,\mathrm{e}^{\phi_1(t)/N} \\
&\geq \left(\frac{\mathfrak{s}_{K/(N-1)}\big((1-t)\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)} + \frac{\mathfrak{s}_{K/(N-1)}\big(t\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\,c_3(1)\right)^{(N-1)/N} \\
&\qquad\qquad \times \big((1-t) + t\,\mathrm{e}^{\phi_1(1)}\big)^{1/N} \\
&\geq \left(\frac{\mathfrak{s}_{K/(N-1)}\big((1-t)\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\right)^{(N-1)/N} (1-t)^{1/N} \\
&\qquad\qquad + \left(\frac{\mathfrak{s}_{K/(N-1)}\big(t\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\,c_3(1)\right)^{(N-1)/N} \big(t\,\mathrm{e}^{\phi_1(1)}\big)^{1/N} \\
&= \tau_{K,N}^{(1-t)}\big(d(x)\big) + \tau_{K,N}^{(t)}\big(d(x)\big)\,c(1).
\end{aligned}$$

We remark that this inequality holds as equality for μ_0-a.e. $x \in \{\mathrm{d}u = 0\}$ since $d(x) = 0$. Therefore, with the help of Corollary 4.20, we obtain the desired convexity (5.7) of $S_{\mathfrak{m}}^N$ as

$$\begin{aligned}
S_{\mathfrak{m}}^N(\mu_t) &= -\int_M \rho_t\big(\mathcal{F}_t(x)\big)^{(N-1)/N}\,\mathrm{det}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]\,\mathrm{d}\mathfrak{m}(x) \\
&= -\int_M \rho_0(x)^{(N-1)/N}\,\mathrm{det}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]^{1/N}\,\mathrm{d}\mathfrak{m}(x) \\
&\leq -\int_M \rho_0(x)^{(N-1)/N}\tau_{K,N}^{(1-t)}\big(d(x)\big)\,\mathrm{d}\mathfrak{m}(x) \\
&\qquad\qquad - \int_M \rho_0(x)^{(N-1)/N}\,\tau_{K,N}^{(t)}\big(d(x)\big)\,\mathrm{det}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)]^{1/N}\,\mathrm{d}\mathfrak{m}(x) \\
&= -\int_M \tau_{K,N}^{(1-t)}\big(d(x)\big)\,\rho_0(x)^{(N-1)/N}\,\mathrm{d}\mathfrak{m}(x) \\
&\qquad\qquad - \int_M \tau_{K,N}^{(t)}\big(d(x)\big)\,\rho_1\big(\mathcal{F}_1(x)\big)^{(N-1)/N}\,\mathrm{det}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)]\,\mathrm{d}\mathfrak{m}(x) \\
&= -\int_{M\times M} \tau_{K,N}^{(1-t)}\big(l(x,y)\big)\,\rho_0(x)^{-1/N}\,\mathrm{d}\pi(x,y) \\
&\qquad\qquad - \int_{M\times M} \tau_{K,N}^{(t)}\big(l(x,y)\big)\,\rho_1(y)^{-1/N}\,\mathrm{d}\pi(x,y).
\end{aligned}$$

This completes the proof in the case of $N \in (n, +\infty)$. Then the case of $N = n$ is obtained by a limit argument, since $\mathrm{Ric}_n \geq K$ implies $\mathrm{Ric}_N \geq K$ for $N > n$.

For $N<0$, we make use of (5.19) instead of (5.18) to see

$$\begin{aligned}
c(t) &= c_3(t)^{(N-1)/N}\,\mathrm{e}^{\phi_1(t)/N} \\
&\le \left(\frac{\mathfrak{s}_{K/(N-1)}\big((1-t)\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)} + \frac{\mathfrak{s}_{K/(N-1)}\big(t\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\,c_3(1)\right)^{(N-1)/N} \\
&\qquad \times \big((1-t)+t\,\mathrm{e}^{\phi_1(1)}\big)^{1/N} \\
&\le \left(\frac{\mathfrak{s}_{K/(N-1)}\big((1-t)\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\right)^{(N-1)/N} (1-t)^{1/N} \\
&\qquad + \left(\frac{\mathfrak{s}_{K/(N-1)}\big(t\,d(x)\big)}{\mathfrak{s}_{K/(N-1)}\big(d(x)\big)}\,c_3(1)\right)^{(N-1)/N} \big(t\,\mathrm{e}^{\phi_1(1)}\big)^{1/N} \\
&= \tau_{K,N}^{(1-t)}\big(d(x)\big) + \tau_{K,N}^{(t)}\big(d(x)\big)\,c(1),
\end{aligned}$$

where the latter inequality follows from the Hölder inequality of the form

$$\begin{aligned}
(a_1+b_1)^{(N-1)/N} &= \left(\frac{a_1\cdot a_2^{1/(N-1)}}{a_2^{1/(N-1)}} + \frac{b_1\cdot b_2^{1/(N-1)}}{b_2^{1/(N-1)}}\right)^{(N-1)/N} \\
&\le \big(a_1^{(N-1)/N}a_2^{1/N} + b_1^{(N-1)/N}b_2^{1/N}\big)\,(a_2+b_2)^{-1/N}
\end{aligned}$$

for $a_1,a_2,b_1,b_2>0$. Then (5.8) is deduced along the same lines as above.

Next, in the case of $N=+\infty$, we more directly see that $\log[\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]] = \psi_\gamma(0)-\psi_\gamma(t)+\phi(t)$ satisfies

$$\begin{aligned}
[\psi_\gamma(0)-\psi_\gamma(t)+\phi(t)]'' &= -\psi_\gamma''(t) - \mathrm{trace}[D(t)^2] - \mathrm{Ric}\big(\dot\gamma(t)\big) \\
&\le -\mathrm{Ric}_\infty\big(\dot\gamma(t)\big) \le -K\,d(x)^2.
\end{aligned}$$

This implies

$$\log\det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)] \ge t\log\det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)] + \frac{K}{2}\,t\,(1-t)\,d(x)^2$$

(which again holds true also for μ_0-a.e. $x\in\{\mathrm{d}u=0\}$ since $d(x)=0$), and we deduce

$$\begin{aligned}
\mathrm{Ent}_{\mathfrak{m}}(\mu_t) &= \int_M \rho_t\big(\mathcal{F}_t(x)\big)\log\rho_t\big(\mathcal{F}_t(x)\big)\,\det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]\,\mathrm{d}\mathfrak{m}(x) \\
&= \int_M \rho_0(x)\log\frac{\rho_0(x)}{\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]}\,\mathrm{d}\mathfrak{m}(x) \\
&\le (1-t)\,\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\int_M \rho_0(x)\log\frac{\rho_0(x)}{\det_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)]}\,\mathrm{d}\mathfrak{m}(x) \\
&\qquad - \frac{K}{2}\,t\,(1-t)\int_M d(x)^2\,\mathrm{d}\mu_0(x) \\
&= (1-t)\,\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\,\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}\,t\,(1-t)\int_{M\times M} l(x,y)^2\,\mathrm{d}\pi(x,y).
\end{aligned}$$

Finally, when $N = 0$, (5.19) is still available, namely

$$c_3(t) \le \frac{\mathfrak{s}_{-K}((1-t)\,d(x))}{\mathfrak{s}_{-K}(d(x))} + \frac{\mathfrak{s}_{-K}(t\,d(x))}{\mathfrak{s}_{-K}(d(x))}\,c_3(1).$$

In this case, we have

$$\begin{aligned} c_3(t) &= c_1(t)^n\,\mathrm{e}^{-\phi_2(t)} \\ &= \mathrm{e}^{\psi_\gamma(t)-\psi_\gamma(0)}\,\mathrm{e}^{\phi_1(t)}\,\det[\mathrm{d}\mathcal{F}_t(x)]^{-1} \\ &= \mathrm{e}^{\phi_1(t)}\,\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]^{-1}. \end{aligned}$$

Combining this with (5.17) yields

$$\begin{aligned} \det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)] &= \mathrm{e}^{\phi_1(t)}\,c_3(t)^{-1} \\ &\ge \big((1-t)\ + t\,\mathrm{e}^{\phi_1(1)}\big)\left(\frac{\mathfrak{s}_{-K}((1-t)\,d(x))}{\mathfrak{s}_{-K}(d(x))} + \frac{\mathfrak{s}_{-K}(t\,d(x))}{\mathfrak{s}_{-K}(d(x))}\,c_3(1)\right)^{-1} \\ &\ge \min\left\{\frac{(1-t)\,\mathfrak{s}_{-K}(d(x))}{\mathfrak{s}_{-K}((1-t)\,d(x))}, \frac{t\,\mathfrak{s}_{-K}(d(x))}{\mathfrak{s}_{-K}(t\,d(x))}\,\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)]\right\}, \end{aligned}$$

where the latter inequality follows from

$$\frac{a_1+a_2}{b_1+b_2} = \frac{b_1}{b_1+b_2}\frac{a_1}{b_1} + \frac{b_2}{b_1+b_2}\frac{a_2}{b_2} \ge \min\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\}$$

for $a_1, a_2, b_1, b_2 > 0$. This implies

$$\begin{aligned} \rho_t\big(\mathcal{F}_t(x)\big) &= \frac{\rho_0(x)}{\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(x)]} \\ &\le \max\left\{\frac{\mathfrak{s}_{-K}((1-t)\,d)}{(1-t)\,\mathfrak{s}_{-K}(d)}\,\rho_0(x), \frac{\mathfrak{s}_{-K}(t\,d)}{t\,\mathfrak{s}_{-K}(d)}\,\frac{\rho_0(x)}{\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_1(x)]}\right\} \\ &= \max\left\{\frac{\mathfrak{s}_{-K}((1-t)\,d)}{(1-t)\,\mathfrak{s}_{-K}(d)}\,\rho_0(x), \frac{\mathfrak{s}_{-K}(t\,d)}{t\,\mathfrak{s}_{-K}(d)}\,\rho_1\big(\mathcal{F}_1(x)\big)\right\}, \end{aligned}$$

while $\rho_t(\mathcal{F}_t(x)) = \rho_0(x) = \rho_1(\mathcal{F}_1(x))$ for μ_0-a.e. $x \in \{\mathrm{d}u = 0\}$. Therefore, we obtain (5.9), and the proof is finished. □

We remark that, in the Riccati equation in Ohta [55] concerning the Finsler case, as $R_{\alpha\beta}$ we employed $g_{\dot\gamma}(R_{\dot\gamma}(J_\alpha), J_\beta)$ instead of $g_{\dot\gamma}(R_{\dot\gamma}(E_\alpha), E_\beta)$. They both work well in the positive definite case, however, we found that the latter is more convenient in the current setting (due to the fact that $g_{\dot\gamma}(E_\alpha, E_\beta)$ is not the identity matrix).

Remark 5.10. For negative N, one can define the entropic timelike curvature-dimension condition $\mathrm{TCD}^e_q(K,N)$ akin to the approach of Cavalletti–Mondino [21]. However, in this dimensional range it seems that $\mathrm{TCD}^e_q(K,N)$ is stronger than $\mathrm{Ric}_N \ge K$ [56, Remark 4.16]. On the other hand, if $n \le N < +\infty$ these properties imply their entropic counterpart by Braun [13].

5.4 Timelike Brunn–Minkowski inequality

An important consequence of Theorem 5.9 is the timelike Brunn–Minkowski inequality. In either dimensional case, it follows quite directly from $\mathrm{TCD}_q(K,N)$ by Jensen's inequality. In the range $N \in [1,\infty)$, it has been derived first in [21] from the *entropic* timelike curvature-dimension condition after [25], and then in [13] in sharp form. See e.g. [55, 56, 62, 65] for the positive signature case.

Theorem 5.11 (Timelike Brunn–Minkowski inequality)**.** *Assume that $(M,L,\mathfrak{m})$ is a globally hyperbolic weighted Finsler spacetime obeying $\mathrm{Ric}_N \geq K$ for some $K \in \mathbb{R}$ and $N \in (-\infty,0] \cup [n,+\infty]$. Let $A_0, A_1 \subset M$ be two relatively compact Borel sets with positive $\mathfrak{m}$-measure with $\overline{A}_0 \times \overline{A}_1 \subset \{l > 0\}$. Then the following hold for every $t \in (0,1)$.*

(i) *When $N = +\infty$, we have*

$$\log \mathfrak{m}\big[Z_t(A_0,A_1)\big] \geq (1-t)\log \mathfrak{m}[A_0] + t \log \mathfrak{m}[A_1] + \inf_{x\in A_0, y\in A_1} \frac{K}{2}\, t\,(1-t)\, l(x,y)^2.$$

(ii) *When $n \leq N < +\infty$, we have*

$$\mathfrak{m}\big[Z_t(A_0,A_1)\big]^{1/N} \geq \inf_{x\in A_0, y\in A_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big) \cdot \mathfrak{m}[A_0]^{1/N} + \inf_{x\in A_0, y\in A_1} \tau_{K,N}^{(t)}\big(l(x,y)\big) \cdot \mathfrak{m}[A_1]^{1/N}.$$

(iii) *When $N < 0$, we have*

$$\mathfrak{m}\big[Z_t(A_0,A_1)\big]^{1/N} \leq \sup_{x\in A_0, y\in A_1} \tau_{K,N}^{(1-t)}\big(l(x,y)\big) \cdot \mathfrak{m}[A_0]^{1/N} + \sup_{x\in A_0, y\in A_1} \tau_{K,N}^{(t)}\big(l(x,y)\big) \cdot \mathfrak{m}[A_1]^{1/N}.$$

(iv) *When $N = 0$, we have*

$$\mathfrak{m}\big[Z_t(A_0,A_1)\big] \geq \min\Big\{ \inf_{x\in A_0, y\in A_1} \frac{(1-t)\,\mathfrak{s}_{-K}(l(x,y))}{\mathfrak{s}_{-K}((1-t)\,l(x,y))}\, \mathfrak{m}[A_0], \inf_{x\in A_0, y\in A_1} \frac{t\,\mathfrak{s}_{-K}(l(x,y))}{\mathfrak{s}_{-K}(t\,l(x,y))}\, \mathfrak{m}[A_1] \Big\}.$$

Proof. Let $\mu_0, \mu_1 \in \mathcal{P}_c^{\mathrm{ac}}(M,\mathfrak{m})$ be the uniform distributions of A_0 and A_1 with respect to $\mathfrak{m}$, respectively, i.e. $\mu_0 := \mathfrak{m}[A_0]^{-1}\,\mathfrak{m}|_{A_0}$ and $\mu_1 := \mathfrak{m}[A_1]^{-1}\,\mathfrak{m}|_{A_1}$. By our assumption on A_0 and A_1, Lemma 4.10, and Theorem 5.9 there exists a q-geodesic $(\mu_t)_{t\in[0,1]}$ from μ_0 to μ_1 and an ℓ_q-optimal coupling $\pi \in \Pi(\mu_0,\mu_1)$ witnessing the entropic displacement semi-convexity given by $\mathrm{TCD}_q(K,N)$. Depending on the range of N, from the precise form of the latter we will derive the claims in each respective case.

Some small preparations are in order. By Lemma 4.10, the pair (μ_0, μ_1) is q-separated. Therefore $(\mu_t)_{t\in[0,1]}$ is in fact the unique q-geodesic connecting μ_0 to μ_1 by Corollary 4.18. By Proposition 4.19, it only consists of $\mathfrak{m}$-absolutely continuous measures, and we write $\mu_t = \rho_t\,\mathfrak{m}$. Moreover, recall that $\operatorname{spt}\mu_t \subset Z_t(A_0, A_1)$ by Corollary 4.4 for given $t \in (0,1)$.

(i) In the case $N = +\infty$, the hypothesis implies

$$\begin{aligned}\operatorname{Ent}_{\mathfrak{m}}(\mu_t) \le (1-t)\operatorname{Ent}_{\mathfrak{m}}(\mu_0) + t\operatorname{Ent}_{\mathfrak{m}}(\mu_1)\\ - \frac{K}{2}\,t\,(1-t)\int_{M\times M} l(x,y)^2\,\mathrm{d}\pi(x,y).\end{aligned} \tag{5.20}$$

It is easy to see that

$$\operatorname{Ent}_{\mathfrak{m}}(\mu_0) = -\log\mathfrak{m}[A_0], \qquad \operatorname{Ent}_{\mathfrak{m}}(\mu_1) = -\log\mathfrak{m}[A_1],$$

and thus the right-hand side of (5.20) is bounded from above by

$$-(1-t)\log\mathfrak{m}[A_0] - t\log\mathfrak{m}[A_1] - \inf_{x\in A_0, y\in A_1}\frac{K}{2}\,t\,(1-t)\,l(x,y)^2.$$

To bound the left-hand side of (5.20) from below, by Jensen's inequality we find

$$\operatorname{Ent}_{\mathfrak{m}}(\mu_t) \ge -\log\mathfrak{m}[\operatorname{spt}\mu_t] \ge -\log\mathfrak{m}\big[Z_t(A_0,A_1)\big].$$

(ii) In the case $n \le N < +\infty$, the hypothesis implies

$$\begin{aligned}S^N_{\mathfrak{m}}(\mu_t) \le -\int_{M\times M}\tau^{(1-t)}_{K,N}\big(l(x,y)\big)\,\rho_0(x)^{-1/N}\,\mathrm{d}\pi(x,y)\\ -\int_{M\times M}\tau^{(t)}_{K,N}\big(l(x,y)\big)\,\rho_1(y)^{-1/N}\,\mathrm{d}\pi(x,y).\end{aligned}$$

The right-hand side of this inequality is bounded from above by

$$-\inf_{x\in A_0, y\in A_1}\tau^{(1-t)}_{K,N}\big(l(x,y)\big)\,\mathfrak{m}[A_0]^{1/N} - \inf_{x\in A_0, y\in A_1}\tau^{(t)}_{K,N}\big(l(x,y)\big)\,\mathfrak{m}[A_1]^{1/N}.$$

The claim follows again by applying Jensen's inequality to the convex function $r \longmapsto -r^{(N-1)/N}$ on $[0,+\infty)$, thereby obtaining

$$S^N_{\mathfrak{m}}(\mu_t) \ge -\mathfrak{m}[\operatorname{spt}\mu_t]^{1/N} \ge -\mathfrak{m}\big[Z_t(A_0,A_1)\big]^{1/N}.$$

(iii) This is argued analogously to the previous item and thus omitted.

(iv) In the case $N = 0$, the hypothesis implies

$$S^0_{\mathfrak{m}}(\mu_t) \le \operatorname*{ess\,sup}_{(x,y)\in\operatorname{spt}\pi}\max\left\{\frac{\mathfrak{s}_{-K}\big((1-t)\,l(x,y)\big)}{(1-t)\,\mathfrak{s}_{-K}\big(l(x,y)\big)}\,\rho_0(x), \frac{\mathfrak{s}_{-K}\big(t\,l(x,y)\big)}{t\,\mathfrak{s}_{-K}(x,y)}\,\rho_1(y)\right\}.$$

The right-hand side of this inequality is bounded from above by

$$\max\Bigg\{ \sup_{x\in A_0, y\in A_1} \frac{\mathfrak{s}_{-K}((1-t)\,l(x,y))}{(1-t)\,\mathfrak{s}_{-K}(l(x,y))}\,\mathfrak{m}[A_0]^{-1}, \\ \sup_{x\in A_0, y\in A_1} \frac{\mathfrak{s}_{-K}(t\,l(x,y))}{t\,\mathfrak{s}_{-K}(l(x,y))}\,\mathfrak{m}[A_1]^{-1} \Bigg\}.$$

The claim follows by estimating

$$\mathfrak{m}\big[Z_t(A_0,A_1)\big]^{-1} \leq \mathfrak{m}[\operatorname{spt}\mu_t]^{-1} \leq \operatorname*{ess\,sup}_{z\in\operatorname{spt}\mu_t} \rho_t(z) = S^0_{\mathfrak{m}}(\mu_t).$$

This terminates the proof. □

6 Ricci curvature bounds from curvature-dimension condition

Now we turn to the converse of Theorem 5.9, i.e. the derivation of the lower bound on the weighted Ricci curvature given the timelike curvature-dimension condition. Our strategy follows the proofs of Lott–Villani [39, Theorem 7.3] and Mondino–Suhr [52, Theorem 4.3].

Theorem 6.1 ($\mathrm{TCD}_q(K,N)$ implies $\mathrm{Ric}_N \geq K$)**.** *Let $(M,L,\mathfrak{m})$ be a globally hyperbolic weighted Finsler spacetime satisfying $\mathrm{TCD}_q(K,N)$ for some $q\in(0,1)$, $K\in\mathbb{R}$, and $N\in(-\infty,0]\cup[n,+\infty]$. Then we have $\mathrm{Ric}_N \geq K$ in timelike directions.*

Proof. We first consider the case $n < N < +\infty$. Let $p<0$ denote the dual exponent of q, i.e. $p^{-1}+q^{-1}=1$. Suppose to the contrary that for some $\varepsilon>0$, $z\in M$, and a timelike tangent vector $w\in\Omega_z$,

$$\mathrm{Ric}_N(w) < (K-2\,\varepsilon)\,F^2(w).$$

For every sufficiently small $s>0$, we find $\delta>0$ as well as a smooth function $u\colon B_\delta(z)\longrightarrow\mathbb{R}$ defined on a ball $B_\delta(z)$ (with respect to a fixed Riemannian metric) around z such that its first and second derivatives at z are prescribed as

$$\begin{aligned} F^*\big(\mathrm{d}u(z)\big)^{p-2}\,\mathscr{L}^*\big(\mathrm{d}u(z)\big) &= s\,w, \\ F^*\big(\mathrm{d}u(z)\big)^{p-2}\,Q\,H^u\,Q &= -\frac{\psi'_\gamma(0)}{N-n}\,I_n, \end{aligned} \tag{6.1}$$

where we define $\gamma(t) := \exp_z(t\,s\,w)$ for $t\in[0,1]$, Q and H^u are the matrices defined in (5.12), and I_n is the identity matrix. We choose s and δ such that $\mathcal{F}_t\colon B_\delta(z)\longrightarrow M$, where

$$\mathcal{F}_t(x) := \exp_x\big(t\,F^*(\mathrm{d}u)^{p-2}\,\mathscr{L}^*(\mathrm{d}u)\big),$$

becomes a diffeomorphism onto its image for every $t \in [0,1]$. By (6.1) and [45, Lemma 8.3], up to reducing s and δ we may and will assume u to be (l^q/q)-convex, so that the above map will eventually give an optimal transport. Lastly, by the continuity of u and possibly reducing δ further, we may and will also assume that the time derivative of the curve $t \longmapsto \mathcal{F}_t(x)$ is timelike and lies uniformly away from the light cone in $x \in B_\delta(z)$ and $t \in [0,1]$.

Define the measures $\mu_0 := \mathfrak{m}[B_\delta(z)]^{-1}\,\mathfrak{m}|_{B_\delta(z)} = \rho_0\,\mathfrak{m}$ and $\mu_t := (\mathcal{F}_t)_\sharp \mu_0$ for $t \in [0,1]$. By further reducing δ if necessary, we achieve $\operatorname{spt}\mu_0 \times \operatorname{spt}\mu_1 \subset \{l > 0\}$. In particular, (μ_0, μ_1) is q-separated by Lemma 4.10 and the hypothesized (l^q/q)-convexity of u, and $(\mu_t)_{t\in[0,1]}$ constitutes the unique q-geodesic connecting μ_0 to μ_1 by Corollary 4.18. After rescaling, μ_t is $\mathfrak{m}$-absolutely continuous for every $t \in [0,1]$ by Proposition 4.19.

Along $\gamma(t) = \exp_z(t\,s\,w)$ as above, we consider the decomposition

$$c(t) := \det\nolimits_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(z)]^{1/N} = c_1(t)^{(N-n)/N}\,c_2(t)^{n/N}$$

as in the proof of Theorem 5.9, where

$$c_1(t) := \exp\left(\frac{\psi_\gamma(0) - \psi_\gamma(t)}{N-n}\right), \qquad c_2(t) := \mathrm{e}^{\phi(t)/n} = \det[\mathrm{d}\mathcal{F}_t(z)]^{1/n}.$$

Then we have, by $\phi' = \operatorname{trace}[B]$ and the Riccati equation (5.11),

$$\begin{aligned} N\,\frac{c''}{c} &= (N-n)\,\frac{c_1''}{c_1} + n\,\frac{c_2''}{c_2} - \frac{n(N-n)}{N}\left(\frac{c_1'}{c_1} - \frac{c_2'}{c_2}\right)^2 \\ &= -\mathrm{Ric}_N(\dot\gamma) - \operatorname{trace}[B^2] + \frac{(\operatorname{trace}[B])^2}{n} - \frac{n(N-n)}{N}\left(\frac{c_1'}{c_1} - \frac{c_2'}{c_2}\right)^2. \end{aligned}$$

At $t = 0$, recalling $B(0) = F^*(\mathrm{d}u(z))^{p-2}\,H^u\,Q^2$, we deduce from (6.1) that

$$\frac{c_1'(0)}{c_1(0)} = -\frac{\psi_\gamma'(0)}{N-n} = \frac{1}{n}\operatorname{trace}[B(0)] = \frac{c_2'(0)}{c_2(0)}$$

and, since H^u is symmetric,

$$\operatorname{trace}[B(0)^2] = \frac{(\operatorname{trace}[B(0)])^2}{n} = \frac{n\,\psi_\gamma'(0)^2}{(N-n)^2}.$$

Hence, we obtain

$$N\,\frac{c''(0)}{c(0)} = -\mathrm{Ric}_N(s\,w) > -(K - 2\,\varepsilon)\,F^2(s\,w).$$

Thus, by continuity and a further rescaling, we may and will assume

$$N\,\frac{c''(t)}{c(t)} > -(K-\varepsilon)\,F^2\big(\dot{\mathcal{F}}_t(x)\big)$$

for every $x \in B_\delta(z)$ and every $t \in [0,1]$. This inequality implies, by putting $d(x) := F(\dot{\mathcal{F}}_t(x))$,

$$c(t) \le \frac{\mathfrak{s}_{(K-\varepsilon)/N}((1-t)\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,c(0) + \frac{\mathfrak{s}_{(K-\varepsilon)/N}(t\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,c(1) \tag{6.2}$$

thanks to (an analogue of) [25, Lemma 2.2]. Then it follows from the Monge–Ampère equation (Corollary 4.20) that

$$\begin{aligned}\rho_t\big(\mathcal{F}_t(x)\big)^{-1/N} \le{}& \frac{\mathfrak{s}_{(K-\varepsilon)/N}((1-t)\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,\rho_0(x)^{-1/N} \\ &+ \frac{\mathfrak{s}_{(K-\varepsilon)/N}(t\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,\rho_1\big(\mathcal{F}_1(x)\big)^{-1/N}.\end{aligned}$$

Therefore, we find from [13, Theorem 1.5(v)] the failure of the *reduced* timelike curvature-dimension condition $\mathrm{TCD}^*_q(K,N)$, which is weaker than $\mathrm{TCD}_q(K,N)$ (see [13, Proposition 3.6]).

The case $N = n$ can be reduced to the above argument by employing the smoothness of $(M, L, \mathfrak{m})$. Precisely, if $\mathrm{Ric}_n(w) < (K - 2\,\varepsilon)F^2(w)$, then we have $\mathrm{Ric}_N(w) < (K-\varepsilon)F^2(w)$ for sufficiently small $N > n$. This implies that $\mathrm{TCD}^*_q(K,N)$ fails, and hence the stronger condition $\mathrm{TCD}^*_q(K,n)$ does not hold as well (see [13, Proposition 3.7]).

The case $N < 0$ is also argued as in (6.2). We have

$$c(t) \ge \frac{\mathfrak{s}_{(K-\varepsilon)/N}((1-t)\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,c(0) + \frac{\mathfrak{s}_{(K-\varepsilon)/N}(t\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,c(1),$$

and hence

$$\begin{aligned}\rho_t\big(\mathcal{F}_t(x)\big)^{-1/N} \ge{}& \frac{\mathfrak{s}_{(K-\varepsilon)/N}((1-t)\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,\rho_0(x)^{-1/N} \\ &+ \frac{\mathfrak{s}_{(K-\varepsilon)/N}(t\,d(x))}{\mathfrak{s}_{(K-\varepsilon)/N}(d(x))}\,\rho_1\big(\mathcal{F}_1(x)\big)^{-1/N}.\end{aligned}$$

Recalling that the N-Rényi entropy has no minus sign for $N < 0$, we deduce that $\mathrm{TCD}^*_q(K,N)$ fails, which is again weaker than $\mathrm{TCD}_q(K,N)$ (which is argued as for [56, Proposition 4.7]).

In the case $N = +\infty$, we replace the second condition of (6.1) with $H^u = 0$. Then we have, as in the proof of Theorem 5.9,

$$\begin{aligned}\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_0 \log\det{}_{\mathfrak{m}}[\mathrm{d}\mathcal{F}_t(z)] &= \left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}\right|_0 \big[\psi_\gamma(0) - \psi_\gamma(t) + \phi(t)\big] \\ &= -\mathrm{Ric}_\infty(s\,w) > -(K-2\,\varepsilon)F^2(s\,w),\end{aligned}$$

with which one can show the failure of $\mathrm{TCD}_q(K,\infty)$ in the same way.

The case $N = 0$ uses the hierarchy from Remark 5.3. Upper bounds on Ric_0 yield the same upper bounds on Ric_N for every $N < 0$, and following the above arguments we have

$$\rho_t\big(\mathcal{F}_t(x)\big) \geq \bigg(\frac{\mathfrak{s}_{(K-\varepsilon)/N}\big((1-t)\,d(x)\big)}{\mathfrak{s}_{(K-\varepsilon)/N}\big(d(x)\big)}\,\rho_0(x)^{-1/N} + \frac{\mathfrak{s}_{(K-\varepsilon)/N}\big(t\,d(x)\big)}{\mathfrak{s}_{(K-\varepsilon)/N}\big(d(x)\big)}\,\rho_1\big(\mathcal{F}_1(x)\big)^{-1/N} \bigg)^{-N}.$$

By virtue of [56, Proposition 4.7], this implies

$$\rho_t\big(\mathcal{F}_t(x)\big) \geq \Big(\tau^{(1-t)}_{K-\varepsilon,N}\big(d(x)\big)\,\rho_0(x)^{-1/N} + \tau^{(t)}_{K-\varepsilon,N}\big(d(x)\big)\,\rho_1\big(\mathcal{F}_1(x)\big)^{-1/N} \Big)^{-N}.$$

Therefore, letting $N \to 0$ (and therefore $-1/N \to +\infty$), we obtain

$$\rho_t\big(\mathcal{F}_t(x)\big) \geq \max\bigg\{ \frac{\mathfrak{s}_{-K+\varepsilon}\big((1-t)\,d(x)\big)}{(1-t)\,\mathfrak{s}_{-K+\varepsilon}\big(d(x)\big)}\,\rho_0(x), \frac{\mathfrak{s}_{-K+\varepsilon}\big(t\,d(x)\big)}{t\,\mathfrak{s}_{-K+\varepsilon}\big(d(x)\big)}\,\rho_1\big(\mathcal{F}_1(x)\big) \bigg\},$$

which shows the failure of $\mathrm{TCD}_q(K,0)$. □

References

[1] N. Ahmad, H. K. Kim and R. J. McCann, Optimal transportation, topology and uniqueness. Bull. Math. Sci. **1** (2011), no. 1, 13–32.

[2] S. B. Alexander and R. L. Bishop, Lorentz and semi-Riemannian spaces with Alexandrov curvature bounds. Comm. Anal. Geom. **16** (2008), 251–282.

[3] L. Ambrosio and N. Gigli, A user's guide to optimal transport. Modelling and optimisation of flows on networks, 1–155, Lecture Notes in Math., 2062, Fond. CIME/CIME Found. Subser., Springer, Heidelberg, 2013.

[4] L. Ambrosio, N. Gigli and G. Savaré, Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163** (2014), 1405–1490.

[5] L. Ambrosio, N. Gigli and G. Savaré, Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. Ann. Probab. **43** (2015), no. 1, 339–404.

[6] K. Bacher and K.-T. Sturm, Localization and tensorization properties of the curvature-dimension condition for metric measure spaces. J. Funct. Anal. **259** (2010), no. 1, 28–56.

[7] D. Bakry and M. Émery, Diffusions hypercontractives. Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

[8] V. Bangert, Analytische Eigenschaften konvexer Funktionen auf Riemannschen Mannigfaltigkeiten. J. Reine Angew. Math. **307**(**308**) (1979), 309–324.

[9] D. Bao, S.-S. Chern and Z. Shen, An Introduction to Riemann–Finsler Geometry. Graduate Texts in Mathematics, **200**. Springer-Verlag, New York, 2000.

[10] J. K. Beem, Indefinite Finsler spaces and timelike spaces. Can. J. Math. **22** (1970), 1035–1039.

[11] J. K. Beem, P. E. Ehrlich and K. L. Easley, Global Lorentzian geometry. Second edition. Monographs and Textbooks in Pure and Applied Mathematics, 202. Marcel Dekker, Inc., New York, 1996. xiv+635 pp.

[12] T. Beran, A. Ohanyan, F. Rott and D. A. Solis, The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature. Lett. Math. Phys. **113** (2023), no. 2, Paper No. 48.

[13] M. Braun, Rényi's entropy on Lorentzian spaces. Timelike curvature-dimension conditions. J. Math. Pures Appl. (9) **177** (2023), 46–128.

[14] M. Braun, Good geodesics satisfying the timelike curvature-dimension condition. Nonlinear Anal. **229** (2023), Paper No. 113205, 30 pp.

[15] M. Braun and M. Calisti, Timelike Ricci bounds for low regularity spacetimes by optimal transport. Commun. Contemp. Math, to appear.

[16] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.

[17] P. Bernard and B. Buffoni, Optimal mass transportation and Mather theory. J. Eur. Math. Soc. (JEMS) **9** (2007), no. 1, 85–121.

[18] D. Burago, Y. Burago and S. Ivanov, A Course in Metric Geometry. Graduate Studies in Mathematics, vol. **33**. American Mathematical Society, Providence (2001).

[19] S. Carroll. Spacetime and geometry. An introduction to general relativity. Addison Wesley, San Francisco, CA, 2004. xiv+513 pp.

[20] J. S. Case, Singularity theorems and the Lorentzian splitting theorem for the Bakry–Emery–Ricci tensor. J. Geom. Phys. **60** (2010), 477–490.

[21] F. Cavalletti and A. Mondino, Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications. Camb. J. Math., to appear.

[22] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. Invent. Math. **146** (2001), no. 2, 219–257.

[23] G. B. De Luca, N. De Ponti, A. Mondino and A. Tomasiello, Gravity from thermodynamics: optimal transport and negative effective dimensions. SciPost Phys. **15** (2023), no. 2, Paper No. 039, 55 pp.

[24] M. Eckstein and T. Miller, Causality for nonlocal phenomena. Ann. Henri Poincaré **18** (2017), no. 9, 3049–3096.

[25] M. Erbar, K. Kuwada and K. T. Sturm, On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201** (2015), no. 3, 993–1071.

[26] A. Figalli and N. Juillet, Absolute continuity of Wasserstein geodesics in the Heisenberg group. J. Funct. Anal. **255** (2008), no. 1, 133–141.

[27] G. J. Galloway and E. Woolgar, Cosmological singularities in Bakry–Émery spacetimes. J. Geom. Phys. **86** (2014), 359–369.

[28] N. Gigli, The splitting theorem in non-smooth context. arXiv:1302.5555

[29] N. Gigli, On the differential structure of metric measure spaces and applications. Mem. Am. Math. Soc. **236** (2015), no. 1113.

[30] M. Graf. Singularity theorems for C^1-Lorentzian metrics. Comm. Math. Phys. **378** (2020), no. 2, 1417–1450.

[31] S. G. Harris, A triangle comparison theorem for Lorentz manifolds. Indiana Univ. Math. J. **31** (1982), no. 3, 289–308.

[32] S. W. Hawking, The occurrence of singularities in cosmology. I. Proc. Roy. Soc. London Ser. A **294** (1966), 511–521.

[33] S. W. Hawking and G. F. R. Ellis, The large scale structure of space-time. Cambridge Monographs on Mathematical Physics, No. 1. Cambridge University Press, London-New York, 1973. xi+391 pp.

[34] S. W. Hawking and R. Penrose, The singularities of gravitational collapse and cosmology. Proc. Roy. Soc. London Ser. A **314** (1970), 529–548.

[35] M. Hohmann, C. Pfeifer and N. Voicu, Mathematical foundations for field theories on Finsler spacetimes. J. Math. Phys. **63** (2022), no. 3, Paper No. 032503, 33 pp.

[36] M. Kell and S. Suhr, On the existence of dual solutions for Lorentzian cost functions. Ann. Inst. H. Poincaré Anal. Non Linéaire **37** (2020), 343–372.

[37] M. Kunzinger and C. Sämann, Lorentzian length spaces. Ann. Global Anal. Geom. **54** (2018), 399–447.

[38] M. Kunzinger, R. Steinbauer, M. Stojković and J. A. Vickers. Hawking's singularity theorem for $C^{1,1}$-metrics. Classical Quantum Gravity **32** (2015), no. 7, 075012, 19 pp.

[39] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport. Ann. of Math. (2) **169** (2009), no. 3, 903–991.

[40] Y. Lu, E. Minguzzi and S. Ohta, Geometry of weighted Lorentz–Finsler manifolds I: Singularity theorems. J. Lond. Math. Soc. **104** (2021), 362–393.

[41] Y. Lu, E. Minguzzi and S. Ohta, Comparison theorems on weighted Finsler manifolds and spacetimes with ϵ-range. Anal. Geom. Metr. Spaces **10** (2022), 1–30.

[42] Y. Lu, E. Minguzzi and S. Ohta, Geometry of weighted Lorentz–Finsler manifolds II: A splitting theorem. Internat. J. Math. **34** (2023), no. 1, 2350002.

[43] M. Matsumoto, Foundations of Finsler geometry and special Finsler spaces. Kaiseisha Press, Shigaken, 1986. vi+344 pp.

[44] R. J. McCann, Polar factorization of maps on Riemannian manifolds. Geom. Funct. Anal. **11** (2001), no. 3, 589–608.

[45] R. J. McCann, Displacement convexity of Boltzmann's entropy characterizes the strong energy condition from general relativity. Camb. J. Math. **8** (2020), 609–681.

[46] E. Minguzzi, Convex neighborhoods for Lipschitz connections and sprays. Monatsh. Math. **177** (2015), 569–625.

[47] E. Minguzzi, Light cones in Finsler spacetimes. Comm. Math. Phys. **334** (2015), 1529–1551.

[48] E. Minguzzi, Raychaudhuri equation and singularity theorems in Finsler spacetimes. Class. Quantum Grav. **32**, 185008 (2015).

[49] E. Minguzzi, Causality theory for closed cone structures with applications. Rev. Math. Phys. **31** (2019), 1930001, 139pp.

[50] E. Minguzzi, Lorentzian causality theory. Living Reviews in Relativity **22**, 3 (2019). https://doi.org/10.1007/s41114-019-0019-x.

[51] E. Minguzzi and S. Suhr, Lorentzian metric spaces and their Gromov–Hausdorff convergence. arXiv:2209.14384

[52] A. Mondino and S. Suhr, An optimal transport formulation of the Einstein equations of general relativity. J. Eur. Math. Soc. **25** (2023), no. 3, 933–994.

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

[54] S. Ohta, Uniform convexity and smoothness, and their applications in Finsler geometry. Math. Ann. **343** (2009), 669–699.

[55] S. Ohta, Finsler interpolation inequalities. Calc. Var. Partial Differential Equations **36** (2009), 211–249.

[56] S. Ohta, (K, N)-convexity and the curvature-dimension condition for negative N. J. Geom. Anal. **26** (2016), 2067–2096.

[57] S. Ohta, Needle decompositions and isoperimetric inequalities in Finsler geometry. J. Math. Soc. Japan **70** (2018), 651–693.

[58] S. Ohta, Comparison Finsler geometry. Springer Monographs in Mathematics, Springer, Cham, 2021.

[59] B. O'Neill, Semi-Riemannian geometry. With applications to relativity. Pure and Applied Mathematics, 103. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983. xiii+468 pp.

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

[61] K.-T. Sturm, On the geometry of metric measure spaces. I. Acta Math. **196** (2006), no. 1, 65–131.

[62] K.-T. Sturm, On the geometry of metric measure spaces. II. Acta Math. **196** (2006), no. 1, 133–177.

[63] S. Suhr, Theory of optimal transport for Lorentzian cost functions. Münster J. Math. **11** (2018), 13–47.

[64] J. Szilasi, R. L. Lovas and D. C. Kertész, Connections, sprays and Finsler structures. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014. xxii+709 pp.

[65] C. Villani, Optimal transport. Old and new. Grundlehren der mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009. xxii+973 pp.

[66] E. Woolgar and W. Wylie, Cosmological singularity theorems and splitting theorems for N-Bakry–Émery spacetimes. J. Math. Phys. **57** (2016), 022504, 1–12.

[67] E. Woolgar and W. Wylie, Curvature-dimension bounds for Lorentzian splitting theorems. J. Geom. Phys. **132** (2018), 131–145.