On the existence of dual solutions for Lorentzian cost functions

Martin Kell [a,*], Stefan Suhr [b]

[a] *Fachbereich Mathematik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany*
[b] *Ruhr-Universität Bochum, Fakultät für Mathematik, Gebäude NA 4/33, D-44801 Bochum, Germany*

Abstract

The dual problem of optimal transportation in Lorentz-Finsler geometry is studied. It is shown that in general no solution exists even in the presence of an optimal coupling. Under natural assumptions dual solutions are established. It is further shown that the existence of a dual solution implies that the optimal transport is timelike on a set of full measure. In the second part the persistence of absolute continuity along an optimal transportation under obvious assumptions is proven and a solution to the relativistic Monge problem is provided.
© 2019 L'Association Publications de l'Institut Henri Poincaré. Published by Elsevier B.V. All rights reserved.

Keywords: Optimal transport; Lorentz cost; Dual solution; Relativistic Monge problem

1. Introduction

The theory of optimal transportation on Riemannian manifolds has revolutionized Riemannian geometry during the last decade with its characterization of lower bounds on the Ricci curvature in terms of optimal transport and the formulation of synthetic Ricci curvature for metric measure spaces. Einstein's field equations, the central equations of general relativity, are equations for the Ricci curvature of a Lorentzian metric. Thus the prospect of developing generalized notions of spacetimes and solutions to the Einstein field equations readily motivates a theory of optimal transportation in Lorentzian geometry. The works [15,21,23,27] are first steps in this direction, with the work by McCann giving a first characterization of lower Ricci curvature bounds for globally hyperbolic spacetimes. Optimal transportation in the context of special relativity was proposed in [8] and studied in [6,7,19,22].

The present theory is formulated for globally hyperbolic Lorentz-Finsler spacetimes. See Section 2 and [5] for definitions and properties. The cost function in Lorentz-Finsler geometry is the negative of the time separation, or Lorentzian distance function, for future causally related points and extended by ∞ to non-future causally related points. Because of the distance-like character of this cost function the problem is a relativistic version of the original Monge problem. This ensures that optimal couplings of finite cost transport along future pointing causal geodesics.

---

* Corresponding author.
 E-mail addresses: martin.kell@math.uni-tuebingen.de (M. Kell), stefan.suhr@ruhr-uni-bochum.de (S. Suhr).

The non-finiteness, non-Lipschitzity and the discontinuity of the cost function at the boundary of J^+ cause additional difficulties, though.

This article is a continuation of the work [27] by the second author on optimal transportation in Lorentz-Finsler geometry. The first major result, Theorem 2.8, focuses on the existence of solutions to the dual problem of Lorentzian optimal transportation, also known as the *dual Kantorovich problem*, see [28]. In [27] the second author gave natural conditions on the marginals to obtain optimal couplings and the weak Kantorovich duality. Here the weak Kantorovich duality says that the "inf-sup-equality" holds. No statement on the existence of solutions was made, though. The dual problem does in general not admit a solution as Section 3.1 shows. The problem lies on the lightcones. In general only a negligible part of mass is allowed to be transported along lightlike geodesics if a solution is to exist. To circumvent the underlying phenomenon the condition of *strict timelikeness* is introduced in Definition 2.6. Theorem 2.8 then shows that dual solutions exists if the marginals satisfy the strict timelikeness condition and at least one marginal has connected support. The conditions are met on a weakly dense subset of pairs of measures by Corollary 2.9. Theorem 2.8 generalizes results in [6,7,21]. The condition of strict timelikeness is related to parts of the definition of *q-separated measures* in [21]. Conversely the existence of a dual solution necessitates that only a negligible part of mass is transported along lightlike geodesics, see Theorem 2.12 the second major result of the present paper. A related result in special relativity is [7, Theorem C].

It should be noted that both the existence and non-existence of dual solutions adapt to the Lorentzian cost with $q \in (0, 1)$ as studied in [15,21]. Indeed, the proof of Theorem 2.8 makes only use of the causality structure. In Example 3.1 and Theorem 2.12 the adjusted upper bound is $-\frac{C}{2} n^{1-\frac{q}{2}}$ and one may replace Lemma 4.1 by McCann's result on starshapedness of q-separated measures [21, Proposition 5.5].

An important question for Lorentzian optimal transportation is whether the interpolation measures of an optimal transport are absolutely continuous with respect to the volume measure if at least one marginal is absolutely continuous. Theorem 2.14, the third major result, shows that intermediate measures are absolutely continuous if one marginal is absolutely continuous and the other marginal is concentrated on an achronal set. This result seems optimal also in the non-relativistic setting, as the non-uniqueness of optimal couplings and interpolation measures usually prevent intermediate measures to be absolutely continuous. However, the optimal couplings constructed from the solution of the relativistic Monge problem can be used to show that there are indeed absolutely continuous intermediate measure though they are non-unique even assuming that transport is along time-affinely parametrized geodesics. We emphasize that the proof of Theorem 2.14 does not rely on the Lipschitz regularity of the transport directions as e.g. in [4], since Lipschitz regularity is not available, see [27].

We remark that the synthetic proof of existence of optimal transport maps adapts easily to Lorentzian cost functions with $q \in (0, 1)$. Indeed, excluding lightlike geodesics one may parametrized geodesics with respect to arc length. Then the non-branching property (Lemma 4.8) and a weak measure contraction property (Lemma 4.6) or alternatively the (K, N)-convexity of the entropy as obtained in [21] are for example sufficient to follow mutatis mutandis the proof of Cavalletti-Huesmann [11].

The importance of the non-branching property – or equivalently the non-overlapping property of interpolation measures for costs that are strictly convex functions of the distance – was already observed by McCann [20] and appeared in the setting of general metric spaces in [25]. Non-branching turned out to be essential to solve the Monge problem in $\mathbb{R}^n$ [10] via Sudakov's needle decomposition [26] and via density estimates using the measure contraction property in $\mathbb{R}^n$ [13] and in the Heisenberg group [14].

The last major result, Theorem 2.16, provides a solution to the relativistic Monge problem. It is shown that there exists an optimal transport map between any two causally related measures whenever the first measure is absolutely continuous with respect to any volume form of the differentiable manifold. If the second measure is concentrated on an achronal set this was already proven by the second author, see [27, Theorem 2.12]. The theorem therein also contains a uniqueness statement. Note, however, the existence proof in this article is independent of [27] and only relies on a non-branching property of time-affinely parametrized geodesics, see Lemma 4.8 below. Uniqueness then follows using [27, Proposition 3.21]. This can be seen as a stronger non-branching property that is related to the volume form.

The article is organized as follows: In Section 2 the setting is introduced and the main results are formulated. In Section 3 two examples are given. One example shows that not all pairs of measures with an optimal coupling admit a solution to the dual problem. The second example shows that the dual solution does not need to be Lipschitz, i.e. the optimal transport is not bounded away from the lightcones, despite the existence of a strictly timelike coupling. Finally Section 4 contains the proofs of all results.

2. Results

Let M be a smooth manifold. Throughout the article one fixes a complete Riemannian metric h on M, though local changes to the metric will be allowed. Consider a continuous function $\mathbb{L} \colon TM \to \mathbb{R}$, smooth on $TM \setminus T^0 M$ (here $T^0 M$ denotes the zero section in TM) and positive homogenous of degree 2 such that the second fiber derivative is non-degenerate with index $\dim M - 1$. A set $\mathcal{C} \subset TM$ is a *closed cone field* if $\mathcal{C}_p := \mathcal{C} \cap TM_p$ is a closed convex cone for all $p \in M$ and $\mathcal{C} \cup T^0 M$ is a closed subset of TM. Choose a *casual structure* $\mathcal{C}$ of $(M, \mathbb{L})$, see [27], i.e. select a closed cone field $\mathcal{C}$ with $\pi_{TM}(\mathcal{C}) = M$ such that the interior of $\mathcal{C}$ is a connected component of $\{\mathbb{L} > 0\}$. Define the Lagrangian L on TM by setting

$$
L(v) := \begin{cases} -\sqrt{\mathbb{L}(v)}, & \text{for } v \in \mathcal{C}, \\ \infty, & \text{otherwise.} \end{cases}
$$

The pair (M, L) is referred to as a *Lorentz-Finsler manifold.*

One calls an absolutely continuous curve $\gamma \colon I \to M$ *($\mathcal{C}$-)causal* if $\dot{\gamma} \in \mathcal{C}$ for almost all $t \in I$. Note that this condition already implies that the tangent vector is contained in $\mathcal{C}$ whenever it exists.

Denote with $J^+(x)$ the set of points $y \in M$ such that there exists a causal curve $\gamma \colon [a, b] \to M$ with $\gamma(a) = x$ and $\gamma(b) = y$. Two points x and y will be called *causally related* if $y \in J^+(x)$. Note that this relation is in general asymmetric. Define the set

$$
J^+ := \{(x, y) \in M \times M \mid y \in J^+(x)\}
$$

and $J^-(y) \colon \{x \in M \mid y \in J^+(x)\}$.

A Lorentz-Finsler manifold is said to be *causal* if it does not admit a closed causal curve, i.e. $J^+ \cap \triangle = \emptyset$ for $\triangle := \{(x, x) \mid x \in M\}$.

Definition 2.1. A causal Lorentz-Finsler manifold (M, L) is *globally hyperbolic* if the sets $J^+(x) \cap J^-(y)$ are compact for all $x, y \in M$.

By [5] every globally hyperbolic Lorentz-Finsler spacetime admits a diffeomorphism (called a *splitting*) $M \cong \mathbb{R} \times N$ such that the projection $\tau \colon \mathbb{R} \times N \to \mathbb{R}$, $(\theta, p) \mapsto \theta$ satisfies

$$
-d\tau(v) \le \min\{L(v), -|v|\}
$$

for all $v \in \mathcal{C}$. In the following one fixes a splitting τ and refers to it as a *time function*. Note that though the proofs below use a particular highly non-unique time function, the existence and uniqueness results do not depend on the choice of such a function.

Define the *Lagrangian action (relative to L)* of an absolutely continuous curve $\gamma \colon [a, b] \to M$:

$$
\mathcal{A}(\gamma) := \int_a^b L(\dot{\gamma}) dt \in \mathbb{R} \cup \{\infty\}
$$

Note that $\mathcal{A}(\gamma) \in \mathbb{R}$ if and only if γ is causal. The following result is proven in the same fashion as in the Lorentzian case, see [27].

Proposition 2.2. *Let (M, L) be globally hyperbolic. Then for every pair $(x, y) \in J^+$ there exists a minimizer of $\mathcal{A}$ with finite action connecting the two points. The minimizer γ solves the Euler-Lagrange equation of $\mathbb{L}$ up to monotone reparametrization and satisfies $\dot{\gamma} \in \mathcal{C}$ everywhere.*

The Euler-Lagrange equation of $\mathbb{L}$ defines a maximal local flow

$$
\Phi^{\mathbb{L}} \colon \mathbb{U} \to TM,
$$

where $\mathbb{U} \subset \mathbb{R} \times TM$ is an open neighborhood of $\{0\} \times (TM \setminus T^0 M)$. Note that $\mathcal{C}$ and $\partial \mathcal{C}$ are invariant under $\Phi^{\mathbb{L}}$. A curve $\gamma : I \to M$ is a $\Phi^{\mathbb{L}}$-*orbit* if it solves the Euler-Lagrange equation of $\mathbb{L}$, see [27].

For a globally hyperbolic Lorentz-Finsler manifold (M, L) define the *Lorentzian cost function*

$$c_L : M \times M \to \mathbb{R} \cup \{\infty\}$$
$$x, y \mapsto \min \{\mathcal{A}(\gamma) | \ \gamma \text{ connects } x \text{ and } y\}.$$

It is immediate that c_L is non-positive for causally related points and infinite otherwise.

Define $\overline{\tau} : M \times M \to \mathbb{R}, (x, y) \mapsto \tau(y) - \tau(x)$. Let $\mathcal{P}_\tau(M \times M)$ be the space of Borel probability measures π on $M \times M$ such that $\overline{\tau} \in L^1(\pi)$. The *Lorentzian cost* is the functional

$$\mathcal{P}_\tau(M \times M) \to \mathbb{R} \cup \{\infty\}, \ \pi \mapsto \int c_L d\pi.$$

The minimization problem for the Lorentzian cost is called the *Relativistic Monge-Kantorovich problem*: Given two Borel probability measures μ_0 and μ_1 on M find a minimizer of the Lorentzian cost among all Borel probability measures on $M \times M$ with first marginal μ_0 and second marginal μ_1. Any minimizer will be called an *optimal coupling* between μ_0 and μ_1.

Let $\mathcal{P}(M)$ denote the space of Borel probability measures on M. For a splitting $\tau : M \to \mathbb{R}$ define

$$\mathcal{P}_\tau^+(M) := \{(\mu_0, \mu_1) | \ \tau \in L^1(\mu_0) \cap L^1(\mu_1) \text{ and } \mu_0, \mu_1 \text{ are } J^+\text{-related}\}$$

where two probability measures are J^+-*related* (or just *causally related*) if there exists a coupling π with $\pi(J^+) = 1$. Note that if π is a coupling of two J^+-related probability measures μ_0 and μ_1, then $(\mu_0, \mu_1) \in \mathcal{P}_\tau^+(M)$ if and only if $\pi \in \mathcal{P}_\tau(M \times M)$.

A function $\psi : M \to \mathbb{R} \cup \{\infty\}$ with $\psi \not\equiv \infty$ is c_L-*convex* if there exists a function $\zeta : M \to \mathbb{R} \cup \{\pm\infty\}$ such that

$$\psi(x) = \sup \{\zeta(y) - c_L(x, y) | \ y \in M\}$$

for all $x \in M$. The function

$$\psi^{c_L} : M \to \mathbb{R} \cup \{-\infty\}$$
$$y \mapsto \inf \{\psi(x) + c_L(x, y) | \ x \in M\}$$

is called the c_L-*transform* of ψ. A pair $(x, y) \in M \times M$ belongs to the c_L-*subdifferential* $\partial_{c_L} \psi$ if

$$\psi^{c_L}(y) - \psi(x) = c_L(x, y).$$

Proposition 2.3 *([27]). Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$. One has*

$$\inf \left\{ \int c_L d\pi \ \middle| \ \pi \text{ is a coupling of } \mu \text{ and } \nu \right\} = \sup \left\{ \int_M \varphi(y) d\nu(y) - \int_M \psi(x) d\mu(x) \right\},$$

where the supremum is taken over the functions $\psi \in L^1(\mu), \varphi \in L^1(\nu)$ with $\varphi(y) - \psi(x) \le c_L(x, y)$.

Proposition 2.3 shows that the *weak Kantorovich duality* holds.

Definition 2.4. Let $(\mu, \nu) \in P_\tau^+(M)$. A c_L-convex function

$$\psi : M \to \mathbb{R} \cup \{\infty\}$$

is a *solution to the dual Kantorovich problem (DKP) for (μ, ν)* if ψ is μ-almost surely finite and

$$\int_M \psi^{c_L}(y) d\nu(y) - \int_M \psi(x) d\mu(x) = \inf \left\{ \int c_L d\pi' \ \middle| \ \pi' \text{ is a coupling of } \mu \text{ and } \nu \right\}.$$

Since $\psi^{c_L}(y) - \psi(x) \le c_L(x, y)$ every coupling π of μ and ν with $\psi^{c_L}(y) - \psi(x) = c_L(x, y)$ π-almost surely will necessarily be optimal.

Definition 2.5 *([28]).* A *dynamical coupling* of two probability measures μ_0 and μ_1 is a probability measure Π on the space of continuous curves $\eta \colon [0, 1] \to M$ such that $(\mathrm{ev}_0)_\sharp \Pi = \mu_0$ and $(\mathrm{ev}_1)_\sharp \Pi = \mu_1$.

Dynamical couplings in Lorentzian geometry have been studied in [15,23].

Definition 2.6. A pair (μ, ν) of probability measures is *strictly timelike* if there exists a dynamical coupling Π supported in the subspace of causal curves such that $(\partial_t \mathrm{ev})_\sharp (\Pi \times \mathcal{L}|_{[0,1)})$ is locally bounded away from $\partial \mathcal{C}$ where $\partial_t \mathrm{ev}(\gamma, t) := \dot{\gamma}(t)$.

Remark 2.7.

(1) Recall that every causal curve admits a Lipschitz parameterization. Further the condition of strict timelikeness is convex in the sense that the set of strictly timelike pairs of measures is convex.
(2) The condition of *strict timelikeness* generalizes the *supercritical speed* for *relativistic cost functions* in [6,7,19]. It is further related to the condition of *q-separatedness* in [21].

Theorem 2.8 *(Existence of dual solutions).* *Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ be strictly timelike and assume that $\mathrm{supp}\,\mu$ is connected. Then the DKP for (μ, ν) has a solution, i.e. there exists a c_L-convex function $\psi \colon M \to \mathbb{R} \cup \{\infty\}$ real-valued on $\mathrm{supp}\,\mu$ with*

$$\int_M \psi^{c_L}(y) d\nu(y) - \int_M \psi(x) d\mu(x) = \inf \left\{ \int c_L d\pi' \,\middle|\, \pi' \text{ is a coupling of } \mu \text{ and } \nu \right\}.$$

For any optimal coupling π of μ and ν one has $\psi^{c_L}(y) - \psi(x) = c_L(x, y)$ π-almost everywhere, i.e. $\mathrm{supp}\,\pi \subset \partial_{c_L} \psi$.

The theorem generalizes [6, Theorem 5.13] and [7, Theorem B]. Note that Theorem 2.8 is most likely optimal, as Theorem 2.12 shows that dual solutions cannot exist whenever there are optimal couplings transporting a set of positive measure along $\partial \mathcal{C}$, the boundary of the lightcone. A $1 + 1$-dimensional example of the non-existence of dual solution is provided in Section 3.1 below.

Corollary 2.9. *Every pair $(\mu_0, \mu_1) \in \mathcal{P}_\tau^+(M)$ can be approximated in the weak topology by a sequence $\{(\mu_0^n, \mu_1^n)\}_{n \in \mathbb{N}} \subset \mathcal{P}_\tau^+(M)$ such that every pair (μ_0^n, μ_1^n) admits a solution to the DKP.*

Proof of Corollary 2.9. Choose a vector field $X \in \Gamma(TM)$ with $L(X) < 0$ and $d\tau(X) = 1$. Denote with φ_t^X the flow of X. Then

$$(\nu_0^n, \nu_1^n) := ((\varphi_{-t}^X)_\sharp \mu_0, \mu_1) \in \mathcal{P}_\tau^+(M)$$

is strictly timelike for all $n \in \mathbb{N}$. The measure ν_0^n can approximated by a measure o^n with connected support and

$$\mathrm{supp}\,\nu_0^n \subset \mathrm{supp}\,o^n.$$

Now the coupling

$$(\mu_0^n, \mu_1^n) := \left(\frac{n-1}{n} \nu_0^n + \frac{1}{n} o^n, \frac{n-1}{n} \nu_1^n + \frac{1}{n} (\varphi_1^X)_\sharp o^n \right)$$

satisfies the assumptions of Theorem 2.8. $\square$

If $c_L(x, y) = 0$ then by Proposition 2.2 there exists a lightlike $\Phi^{\mathbb{L}}$-orbit η, i.e. $L(\dot{\eta}) \equiv 0$, which cannot be parametrized by "arc length", i.e. $L(\dot{\eta}) \equiv 1$. In particular, such lightlike $\Phi^{\mathbb{L}}$-orbits do not admit a "preferred affine parametrization" in any sense. However, using the time function τ one may reparametrize every causal $\Phi^{\mathbb{L}}$-orbit as

follows: Denote with Γ the set of causal minimizers $\gamma : [0, 1] \to M$ of $\mathcal{A}$ such that $d\tau(\dot{\gamma}) \equiv \mathrm{const}(\gamma)$. Elements of Γ are called *time-affinely* parametrized geodesics. For $(x, y) \in J^+$ consider the subspace

$$\Gamma_{x \to y} := \{\gamma \in \Gamma |\ \mathrm{ev}_0(\gamma) = x,\ \mathrm{ev}_1(\gamma) = y\}$$

where

$$\mathrm{ev} \colon \Gamma \times [0, 1] \to M,\ (\gamma, t) \mapsto \gamma(t) \text{ and } \mathrm{ev}_t := \mathrm{ev}(., t).$$

Since (M, L) is assumed to be globally hyperbolic one has $\Gamma_{x \to y} \neq \varnothing$.

Definition 2.10. A Borel measure Π on Γ is a *dynamical optimal coupling* of $\mu_0 := (\mathrm{ev}_0)_\sharp \Pi$ and $\mu_1 := (\mathrm{ev}_1)_\sharp \Pi$ if $\pi := (\mathrm{ev}_0, \mathrm{ev}_1)_\sharp \Pi$ is an optimal coupling between μ_0 and μ_1.

Proposition 2.11 *([27]). Let $(\mu_0, \mu_1) \in \mathcal{P}_\tau^+(M)$. Then there exists a dynamical optimal coupling Π for μ_0 and μ_1 with $\mathrm{supp}\,\Pi \subseteq \Gamma$.*

In the following all Lebesgue measures are understood to be induced by the Riemannian metric h. The assumptions (i) and (ii) of the next theorem are similar and yield the same conclusion. It is not necessarily obvious that Theorem 2.12(i) is the analogue of the classical solution to the dual Kantorovich problem for real-valued cost functions, see [28]. The conclusion of Theorem 2.12 under assumption (ii) on the other hand has no counterpart there.

Theorem 2.12 *(Non-existence of dual solutions). Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$. Assume that $\mathrm{supp}\,\mu \cap \mathrm{supp}\,\nu = \emptyset$ and that the DKP for (μ, ν) admits a solution $\psi \colon M \to \mathbb{R} \cup \{\infty\}$. Further assume μ to be either*

 (i) *supported on a spacelike hypersurface H and that it is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}_H$ on H or*
(ii) *absolutely continuous with respect to the Lebesgue measure $\mathcal{L}$ on M.*

Denote with Π a dynamical optimal coupling of μ and ν and with Γ_0 the set of lightlike minimizers in Γ. Then $\Pi(\Gamma_0) = 0$, i.e. only a μ-negligible set of points is transported along lightlike minimizers.

The theorem generalizes [6, Corollary 3.6] and [7, Theorem C]. Note that Theorem 2.12 is proven indirectly and relies on a very similar construction as the $1 + 1$-dimensional example in Section 3.1.

The following two theorems are the second main result of this article. The first theorem has a counterpart in the work of the second author, see [27, Theorem 2.13] and the second theorem is a solution to the *relativistic Monge problem*: For $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ find a Borel-measurable map $F \colon M \to M$ such that $\pi := (\mathrm{id}, F)_\sharp \mu$ is an optimal coupling of μ and ν.

Note that the proofs are independent of [27, Theorem 2.13] and rely only on a straightforward geometric argument, see [27, Proposition 3.22].

Definition 2.13. A set $A \subset M$ is *achronal* if $c_L|_{A \times A} \geq 0$. In case $c_L|_{A \times A} \equiv \infty$ one says that A is *acausal*.

It is not difficult to see that any time slice $\{\tau = \tau_0\}$ is acausal. The definition is in accordance with the classical definitions of acausal and achronal sets in Lorentzian geometry.

Theorem 2.14 *(Existence and uniqueness for achronal targets). Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ such that μ is absolutely continuous with respect to the Lebesgue measure on M and ν is concentrated on an achronal set. Then there exists a unique dynamical optimal coupling Π such that $(\mathrm{ev}_t)_\sharp \Pi$ is absolutely continuous for $t \in [0, 1)$ and the optimal couplings $(\mathrm{ev}_t, \mathrm{ev}_1)_\sharp \Pi$ are induced by transport maps.*

Remark 2.15. Note that the Monge problem is in general highly non-unique even in the non-relativistic setting. In the non-relativistic setting the equivalent to being supported on a time slice would be to assume the second measure is concentrated in a level set of a dual solution to the Monge problem. However, such a condition depends on the first measure.

Theorem 2.16 *(Solution to the relativistic Monge problem). Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ such that μ is absolutely continuous with respect to the Lebesgue measure on M. Then there exists a Borel-measurable map $F : M \to M$ such that $\pi :=$ $(\mathrm{id}, F)_\sharp \mu$ is an optimal coupling of μ and ν.*

3. Two examples

3.1. An example with no dual solution

Let $M = \mathbb{R}^2$ and

$$\mathbb{L} : TM \to \mathbb{R}, \ (x, v) \mapsto v_1^2 - v_2^2$$

with

$$\mathcal{C} := \{(x, v) \in TM \mid v_2 \geq |v_1|\},$$

where $v = (v_1, v_2)$. It follows that c_L is the negative Lorentzian distance on the 2-dimensional Minkowski space. Fix the splitting

$$\tau : M \to \mathbb{R}, \ (s, t) \mapsto t.$$

Denote with $i_0, i_1 : \mathbb{R} \to M$, the maps $i_0(s) := (s, 0)$ and $i_1(s) := (s, 1)$, respectively and with $\mathcal{L}_1$ the Lebesgue measure on the real line $\mathbb{R}$. Consider the transport problem between

$$\overline{\mu} := (i_0)_\sharp(\mathcal{L}_1|_{[0,1]}) \text{ and } \overline{\nu} := (i_1)_\sharp(\mathcal{L}_1|_{[1,2]}).$$

The map

$$\overline{T} : M \to M, \ (s, t) \mapsto (s + 1, t + 1)$$

induces a causal coupling $(\mathrm{id}, \overline{T})_\sharp \overline{\mu}$ of $\overline{\mu}$ and $\overline{\nu}$, i.e. $(\overline{\mu}, \overline{\nu}) \in \mathcal{P}_\tau^+(M)$.

Proposition 3.1. *The DKP for $(\overline{\mu}, \overline{\nu})$ does not have a solution.*

The transport problem for $(\overline{\mu}, \overline{\nu})$ is equivalent to the following transport problem on the real line: The restriction of c_L to

$$\{((s, 0), (t, 1)) \mid s, t \in \mathbb{R}\} \subset \mathbb{R}^2 \times \mathbb{R}^2$$

and the identification

$$\{((s, 0), (t, 1)) \mid s, t \in \mathbb{R}\} \cong \mathbb{R} \times \mathbb{R}, \ ((s, 0), (t, 1)) \cong (s, t)$$

yield the cost function

$$c : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \ (s, t) \mapsto \begin{cases} -\sqrt{1 - (s - t)^2}, & \text{for } |s - t| \leq 1 \\ \infty, & \text{for } |s - t| \geq 1 \end{cases}$$

and the probability measures $\overline{\mu}$ and $\overline{\nu}$ are identified with

$$\mu = \mathcal{L}_1|_{[0,1]} \text{ and } \nu = \mathcal{L}_1|_{[1,2]},$$

respectively.

Lemma 3.2. *If π is a coupling of μ and ν with finite c-cost, then*

$$\pi = (\mathrm{id}, T)_\sharp \mu,$$

where $T : \mathbb{R} \to \mathbb{R}, \ s \mapsto s + 1$.

Proof. Let π be a coupling of μ and ν with finite cost. For every $\varepsilon > 0$ one has

$$\mu([0, \varepsilon]) = \nu([1, 1 + \varepsilon]) = \nu([-1, 1 + \varepsilon]).$$

The support of π is contained in $\{(s, t) \mid |s - t| \leq 1\}$ since it has finite cost. Therefore

$$\mu([0, \varepsilon]) = \pi([0, \varepsilon] \times \mathbb{R}) = \pi([0, \varepsilon] \times [1, 1 + \varepsilon]).$$

By complementary reasoning one concludes that

$$\pi([\varepsilon, 1] \times \mathbb{R}) = \pi([\varepsilon, 1] \times [1 + \varepsilon, 2]).$$

An induction over n then implies that the support of π is contained in

$$\bigcup_{k=0}^{2^n} \left([k \cdot 2^{-n}, (k + 1) \cdot 2^{-n}] \times [1 + k \cdot 2^{-n}, 1 + (k + 1) \cdot 2^{-n}]\right)$$

for every $n \in \mathbb{N}$. The claim follows in the limit for $n \to \infty$. $\square$

Lemma 3.3. *Let $[a, b] \subset \mathbb{R}$, $\varepsilon > 0$ and a Borel measurable set $B \subset [a, b]$ be given with $\mathcal{L}_1(B) \geq \varepsilon(b - a)$. Then for all $n \in \mathbb{N}$ there exists $\{t_i\}_{1 \leq i \leq n} \subset B$ with $t_1 < \ldots < t_n$ and $t_{i+1} - t_i \geq \frac{\varepsilon}{2n}(b - a)$.*

Proof. Let $n \in \mathbb{N}$ be given. Consider the function $t \mapsto \mathcal{L}_1(B \cap [a, t])$ and choose $t_1 \in B$ such that

$$\mathcal{L}_1(B \cap [a, t_1]) \in \left(0, \frac{\varepsilon}{2n}(b - a)\right).$$

Then one has

$$\mathcal{L}_1\left(B \cap \left[a, t_1 + \frac{\varepsilon}{2n}(b - a)\right]\right) \leq \frac{\varepsilon}{n}(b - a).$$

Next consider the function

$$t \mapsto \mathcal{L}_1\left(B \cap \left[t_1 + \frac{\varepsilon}{2n}(b - a), t\right]\right), \quad \text{for } t > t_1 + \frac{\varepsilon}{2n}(b - a).$$

Choose $t_2 \in B$ such that

$$\mathcal{L}_1\left(B \cap \left[t_1 + \frac{\varepsilon}{2n}(b - a), t_2\right]\right) \in \left(0, \frac{\varepsilon}{2n}(b - a)\right).$$

Then one has

$$\mathcal{L}_1\left(B \cap \left[a, t_2 + \frac{\varepsilon}{2n}(b - a)\right]\right) \leq 2\frac{\varepsilon}{n}(b - a) = 2\frac{\varepsilon}{n}(b - a).$$

Continue inductively. For $k < n$ one has

$$\mathcal{L}_1\left(\left[a, t_k + \frac{\varepsilon}{2n}(b - a)\right]\right) \leq k\frac{\varepsilon}{n}(b - a) = \frac{k}{n}\varepsilon(b - a) \leq \frac{n - 1}{n}\varepsilon(b - a).$$

Thus one concludes $t_k + \frac{\varepsilon}{2n}(b - a) < b$ for all $k < n$. This shows that the construction does not terminate before n points have been chosen. The claimed properties are clear from the construction. $\square$

Lemma 3.4. *There does not exist a c-convex function $\psi : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ with $\psi|_{[0,1]} \not\equiv \infty$ and $\psi^c(y) - \psi(x) = c(x, y)$ for π-almost all $(x, y) \in \mathbb{R} \times \mathbb{R}$, where π is the coupling in Lemma 3.2.*

The existence of a solution to the dual problem is independent of an additive constant in the definition of the cost function, i.e. for $c' := c + 1$ the pair (φ', ψ') solves the DKP for c' iff $(\varphi', \psi - 1)$ solves the DKP for c.

Proof. Let (φ, ψ) be a solution to the DKP for (μ, ν), i.e. $c(x, y) \geq \varphi(x) + \psi(y)$ and

$$\int c \, d\pi' = \int \varphi \, d\mu + \int \psi \, d\nu = \int (\varphi + \psi) \, d\pi'.$$

Thus one has $c = \varphi + \psi \ \pi'$-almost surely.

Let $s < t \in [0, 1]$. By Lemma 3.3 for every $n \in \mathbb{N}$ there exist $s \leq t_1 < \dots t_n \leq t$ with $t_{k+1} - t_k \geq \frac{t-s}{2n}$ and $\varphi(t_k) + \psi(t_k + 1) = 0$. Thus one has

$$c \, (t_{k+1}, 1 + t_k) \leq c \left(0, 1 - \frac{1}{2n}(t - s) \right) \leq -\sqrt{\frac{1}{2n}(t - s)}.$$

As in [28, page 61] it follows that

$$\psi(s) \leq \psi(t) + \sum_{k=1}^{n} c \, (t_{k+1}, 1 + t_k)$$

$$\leq \psi(t) - n\sqrt{\frac{1}{2n}(t - s)} = \psi(t) - \sqrt{\frac{n(t - s)}{2}}$$

for all n. Therefore $\psi(s) = -\infty$ for all $s < 1$. But this contradicts the definition of ψ. $\square$

Proof of Proposition 3.1. The claim follows directly from Lemma 3.4 by reversing the identification

$$\{((s, 0), (t, 1)) | \, s, t \in \mathbb{R}\} \cong \mathbb{R} \times \mathbb{R}. \quad \square$$

3.2. An example with non-Lipschitz dual solution

An example is given of a strictly timelike pair (μ, ν) for which the optimal coupling is not bounded away from ∂J^+. This counters the intuition that the optimal coupling of strictly timelike pairs is supported away from ∂J^+.

Let $0 < \varepsilon < \frac{1}{2}$. Choose a function $\bar{f} \in C^{0, \frac{1}{2}}([0, 5]) \cap C^\infty([0, 5] \setminus \{2\})$ with

(1) $\bar{f} \equiv 1 + \varepsilon$ on $[0, 1]$,
(2) $\bar{f}(x) > -c_L((x, 0), (1, 1))$ for $x \in [1, 2)$,
(3) $\bar{f}(2) = 0$
(4) $\bar{f}(x) > c_L((3, 0), (x, 1))$ for $x \in (2, 3]$,
(5) $\bar{f}' < 0$ near 2.
(6) $\bar{f} \equiv \varepsilon - 1$ on $[3, 4]$

that induces a $\frac{1}{2}$-Hölder continuous function f on $\mathbb{R}/5\mathbb{Z}$, smooth except at $[2] \in \mathbb{R}/5\mathbb{Z}$.

Now consider $\mathbb{R}/5\mathbb{Z} \times \mathbb{R}$ with the inner product

$$(\mathbb{L} =)g := d\theta^2 - dt^2$$

for $(\theta, t) \in \mathbb{R}/5\mathbb{Z} \times \mathbb{R}$ where

$$\mathcal{C} := \{v | \, g(v, v) \leq 0, \ dt(v) \geq 0\}.$$

The cost function c_L for the pair $(g, \mathcal{C})$ is

$$c_L((\eta, s), (\theta, t)) = \begin{cases} -\sqrt{(t - s)^2 - (\theta - \eta)^2}, & s \leq t, \ t - s \geq \theta - \eta \\ \infty, & \text{else.} \end{cases}$$

Define

$$\varphi \colon \mathbb{R}/5\mathbb{Z} \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}, \quad \varphi(y) := \inf\{f(\theta) + c_L((\theta, 0), y) | \, \theta \in \mathbb{R}/5\mathbb{Z}\}.$$

Lemma 3.5. *One has*

$$\varphi^{c_L}(y) = \inf_x \{\varphi(x) + c_L(x, y)\} \equiv \varphi(y)$$

for all $y \in \mathbb{R}/5\mathbb{Z} \times \mathbb{R}$. It follows that φ is c_L-concave.

Proof. Indeed first since $c_L(x, x) = 0$ one has

$$\varphi^{c_L}(y) = \inf_x \{\varphi(x) + c_L(x, y)\} \leq \varphi(y)$$

for all y.

Fix $y \in \mathbb{R}/5\mathbb{Z} \times \mathbb{R}$ and choose $z \in \mathbb{R}/5\mathbb{Z} \times \mathbb{R}$ with $\varphi^{c_L}(y) = \varphi(z) + c_L(z, y)$. For z choose $\theta \in \mathbb{R}/5\mathbb{Z}$ with $\varphi(z) = f(\theta) + c_L((\theta, 0), z)$. Then one has

$$\varphi(y) \leq f(\theta) + c_L((\theta, 0), y) \leq f(\theta) + c_L((\theta, 0), z) + c_L(z, y) = \varphi(z) + c_L(z, y) = \varphi^{c_L}(y)$$

by the triangle inequality for c_L. Thus one has

$$\varphi(y) = \inf_x \{\varphi(x) + c_L(x, y)\} = \varphi^{c_L}(y)$$

for all y. $\quad\square$

As usual define

$$\partial_c \varphi := \{(x, y) \mid \varphi(y) - \varphi(x) = c_L(x, y)\} \subseteq (\mathbb{R}/5\mathbb{Z} \times \mathbb{R}) \times (\mathbb{R}/5\mathbb{Z} \times \mathbb{R})$$

and $\partial_c \varphi_x := p_2(\partial_c \varphi \cap (\{x\} \times (\mathbb{R}/5\mathbb{Z} \times \mathbb{R})))$. Note that for all (θ', t) with $t \geq 0$ there exists $\theta \in \mathbb{R}/5\mathbb{Z}$ with $(\theta', t) \in \partial_c \varphi_{(\theta, 0)}$ since c_L is continuous on its domain.

Lemma 3.6. *For $\theta \neq [2]$ and $y \in \partial_c \varphi_{(\theta, 0)}$ with $t(y) > 0$ one has $y \in I^+(\theta, 0)$. Further for every $(\theta, t) \in \mathbb{R}/5\mathbb{Z} \times [0, 1]$ with $\theta \neq [2]$ the set $\partial_c \varphi_{(\theta, 0)} \cap \mathbb{R}/5\mathbb{Z} \times \{t\}$ has exactly one element.*

Proof. Indeed $y \in \partial_c \varphi_{(\theta, 0)}$ implies

$$\inf_\eta \{f(\eta) + c_L((\eta, 0), y)\} = \varphi(y) = f(\theta) + c_L((\theta, 0), y),$$

i.e. the function $\eta \mapsto f(\eta) + c_L((\eta, 0), y)$ has a minimum in θ. If $c_L((\theta, 0), y) = 0$ then $\eta \mapsto c_L((\eta, 0), y)$ falls off to one side of $\theta \neq [2]$ faster than f can rise by construction. Therefore in this case θ cannot be a minimum. Thus it follows that $c_L((\theta, 0), y) < 0$, i.e. $y \in I^+((\theta, 0))$.

Now fix $\theta \neq [2]$ and $t \in [0, 1]$. Then the equation

$$\frac{\partial}{\partial \theta} f(\theta) + \frac{\partial}{\partial \theta} c_L((\theta, 0), (\theta', t)) = 0$$

has exactly one solution θ'. Since by the previous paragraph the points in $\partial_c \varphi_{(\theta, 0)}$ are characterized as solutions to this equation, the second part of the claim follows. $\quad\square$

Lemma 3.7. *For every neighborhood U of ∂J^+ and every $t \in (0, 1]$, the 1-dimensional Lebesgue measure of*

$$\{\theta \in \mathbb{R}/5\mathbb{Z} \setminus \{[2]\} \mid ((\theta, 0), y_{(\theta, t)}) \in U\}$$

is positive, where $y_{(\theta, t)}$ denotes the unique point in $\partial_c \varphi_{(\theta, 0)} \cap \mathbb{R}/5\mathbb{Z} \times \{t\}$.

Proof. The Lebesgue measure of points θ such that $-f'(\theta) \geq C$ is bounded from below by the Lebesgue measure of the set of points with $\frac{\partial}{\partial \theta} c_L((\theta, 0), (1, 1)) \geq C$ for C sufficiently large by the assumptions (2) and (3) above. The last set has positive Lebesgue measure for every $C < \infty$. For every neighborhood U of ∂J^+ there exists $C_U < \infty$ such that $\left| \frac{\partial}{\partial \theta} c_L((\theta, 0), y) \right| \geq C_U$ for all $((\theta, 0), y) \in U$. Now $y = y_{(\theta, t)}$ is the unique solution to the equation $\frac{\partial}{\partial \theta} f(\theta) + \frac{\partial}{\partial \theta} c_L((\theta, 0), y) = 0$ with $t(y) = t$, and the claim follows. $\quad\square$

Lemma 3.8. *There exists $\delta = \delta(t) > 0$ such that $\mathrm{dist}((\theta - t, t), \partial_c \varphi_{(\theta,0)}) \geq \delta$ for all $\theta \neq [2]$.*

Proof. For $\theta \neq [2]$ let $(\theta_t, t) \in \partial_c \varphi_{(\theta,0)}$. By the Lemma 3.7 one has that

$$\eta \mapsto c_L((\eta, 0), (\theta_t, t))$$

is smooth at θ and

$$\frac{\partial}{\partial \theta} f(\theta) + \frac{\partial}{\partial \theta} c_L((\theta, 0), (\theta_t, t)) = 0 \Leftrightarrow \frac{\partial}{\partial \theta} c_L((\theta, 0), (\theta_t, t)) = -\frac{\partial}{\partial \theta} f(\theta).$$

Since $\frac{\partial}{\partial \theta} f(\theta)$ is bounded outside every neighborhood of $[2] \in \mathbb{R}/5\mathbb{Z}$, the existence of δ follows there. By the assumption $\frac{\partial}{\partial \theta} f < 0$ on a neighborhood of $[2]$ the bound follows in fact for all $\theta \neq [2]$ since $\frac{\partial}{\partial \theta} c_L((\theta, 0), (\theta', t)) \to -\infty$ for $\theta' \downarrow \theta - t$. $\square$

Now consider the probability measure $\mu := I_\sharp \mathcal{L}'$ where

$$I : \mathbb{R}/5\mathbb{Z} \hookrightarrow \mathbb{R}/5\mathbb{Z} \times \mathbb{R}, \ \theta \mapsto (\theta, 0),$$

and $\mathcal{L}'$ is the normalized Lebesgue measure on $\mathbb{R}/5\mathbb{Z}$. The following result is a reformulation of [3, Proposition 3] adapted to the present situation.

Proposition *([3]). Let φ be a c_L-convex function, and let μ be a probability measure on M. Then there exists a probability measure ν on M such that φ solves the DKP for (μ, ν).*

By the proposition there exists a probability measure ν supported on $\mathbb{R}/5\mathbb{Z} \times \{1\}$ such that φ is optimal for the pair (μ, ν). By Lemma 3.7 the transport is not bounded away from ∂J^+.

Now a rotation of $\mathbb{R}/5\mathbb{Z}$ in the negative direction leaves μ unchanged, but the coupling induced by φ is twisted into a coupling whose support has positive distance from ∂J^+ by Lemma 3.8. Thus the pair (μ, ν) is strictly timelike.

4. Proofs

4.1. Proof of Theorem 2.8

Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ be strictly timelike. Further let ρ be an optimal coupling of μ and ν. Note that ρ is causal since its cost is finite. Fix $(x_0, y_0) \in \mathrm{supp}\,\rho$. Define

$$\psi : J^-(\mathrm{supp}\,\nu) \to \mathbb{R} \cup \{\pm\infty\},$$

$$x \mapsto \sup \left\{ \sum_{i=0}^k [c_L(x_i', y_i') - c_L(x_{i+1}', y_i')] \right\}$$

where the supremum is taken over all $k \in \mathbb{N}$ and all sequences

$$\{(x_i', y_i')\}_{0 \leq i \leq k+1} \subset \mathrm{supp}\,\rho \text{ with } x_{k+1}' = x \text{ and } (x_0', y_0') = (x_0, y_0).$$

One has $\psi(x_0) \geq c_L(x_0, y_0) - c_L(x_0, y_0) = 0$. At the same time the right hand side of the above definition is nonpositive for $x = x_0$ by cyclic monotonicity, see [27, Proposition 2.7]. Therefore one has $\psi(x_0) = 0$.

Next one shows that ψ is real-valued and measurable on $\mathrm{supp}\,\mu$. Fix $(u, w) \in \mathrm{supp}\,\rho$. Consider for $k \in \mathbb{N}$ chains $\{(u_i, w_i)\}_{0 \leq i \leq k} \subseteq \mathrm{supp}\,\rho$ with $u_{i+1} \in J^-(w_i) \cap \mathrm{supp}\,\mu$ for $0 \leq i \leq k - 1$ where $(u_0, w_0) = (u, w)$. Define

$$A := \{w_k | \{(u_i, w_i)\}_{0 \leq i \leq k} \text{ as just defined}\}.$$

The claim is that

$$p_2(p_1^{-1}(J^-(A)) \cap \mathrm{supp}\,\rho) = A,$$

where $p_1, p_2 : M \times M \to M$ are the canonical projections onto the first and second factor respectively. It is easy to see that A is contained in $p_2(p_1^{-1}(J^-(A)) \cap \mathrm{supp}\,\rho)$. More precisely let $y \in A$. Choose $(x, y) \in \mathrm{supp}\,\rho$. Then

$x \in J^-(y) \subset J^-(A)$ since ρ is causal, i.e. $y \in p_2(p_1^{-1}(J^-(A)) \cap \operatorname{supp} \rho)$. For the opposite inclusion consider $y \in p_2(p_1^{-1}(J^-(A)) \cap \operatorname{supp} \rho)$. Then there exists $x \in M$ with $(x, y) \in p_1^{-1}(J^-(A)) \cap \operatorname{supp} \rho$, i.e. $(x, y) \in \operatorname{supp} \rho$ and $x \in J^-(A)$. So there exists a chain

$$\{(u_i, w_i)\}_{0 \le i \le k} \subset \operatorname{supp} \rho$$

with $u_{i+1} \in J^-(w_i) \cap \operatorname{supp} \mu$ for $0 \le i \le k-1$, $(u_0, w_0) = (u, w)$ and $x \in J^-(w_k)$. Now define a new chain

$$\{(u_i, w_i)\}_{0 \le i \le k+1} \subset \operatorname{supp} \rho$$

identical with the original chain for $i \le k$ and

$$(u_{k+1}, w_{k+1}) := (x, y).$$

Since $(x, y) \in \operatorname{supp} \rho$ this shows that $y \in A$.

By the marginal property one has

$$\mu(J^-(A)) = \rho(p_1^{-1}(J^-(A)) \cap \operatorname{supp} \rho)$$

and

$$\nu(A) = \rho(p_2^{-1}(A)) = \rho(p_2^{-1}(p_2(p_1^{-1}(J^-(A)) \cap \operatorname{supp} \rho)))$$

by the above characterization of A. Since $B \subset p_2^{-1}(p_2(B))$ for any set $B \subset M \times M$ one knows that $\nu(A) \ge \mu(J^-(A))$. With the inclusion $\operatorname{supp} \rho \subseteq J^+$ one has on the other hand $\nu(A) \le \mu(J^-(A))$, i.e. $\nu(A) = \mu(J^-(A))$.

Consequently every causal coupling ρ' of (μ, ν) has to couple $J^-(A)$ with A, especially the coupling guaranteed by the definition of strict timelikeness. But that means $J^-(A) \cap \operatorname{supp} \mu$ is locally uniformly bounded away from $\partial J^-(A)$. Since $J^-(A) \cap \operatorname{supp} \mu$ is nonempty and open it has to be equal to $\operatorname{supp} \mu$ since $\operatorname{supp} \mu$ is connected. This implies $A = \operatorname{supp} \nu$ by the construction of A.

Now let $x \in \operatorname{supp} \mu$ be given. Choose $y \in \operatorname{supp} \nu$ with $(x, y) \in \operatorname{supp} \rho$. The above argument for the set A with $(u_0, w_0) = (x, y)$ yields that there exists a finite chain $\{(u_i, w_i)\}_{0 \le i \le k+1} \subset \operatorname{supp} \rho$ with $c_L(u_{i+1}, w_i) < \infty$ and $(u_{k+1}, w_{k+1}) = (x_0, y_0)$. By definition of ψ one has

$$\psi(x) + \sum_{i=0}^{k} [c_L(u_i, w_i) - c_L(u_{i+1}, w_i)] \le \psi(x_0) = 0.$$

Since

$$\sum_{i=0}^{k} [c_L(u_i, w_i) - c_L(u_{i+1}, w_i)] > -\infty$$

one obtains $\psi(x) < \infty$.

Next consider the construction of A with $(u_0, w_0) = (x_0, y_0)$. Then there exists a finite chain $\{(u_i', w_i')\}_{0 \le i \le k+1} \subset \operatorname{supp} \rho$ with $c_L(u_{i+1}', w_i') < \infty$ and $(u_{k+1}', w_{k+1}') = (x, y)$. This time one has

$$\psi(x_0) + \sum_{i=0}^{k} [c_L(u_i', w_i') - c_L(u_{i+1}', w_i')] \le \psi(x)$$

and it follows $\psi(x) > -\infty$. Since c_L is continuous and real-valued on J^+, one concludes that ψ is measurable.

Define $\zeta \colon \operatorname{supp} \nu \to \mathbb{R} \cup \{-\infty\}$,

$$\zeta(y) := \sup \left\{ \sum_{i=0}^{k} [c_L(x_i', y_i') - c_L(x_{i+1}', y_i')] + c_L(x_{k+1}', y) \right\},$$

where the supremum is taken over all $k \in \mathbb{N}$ and sequences

$$\{(x_i', y_i')\}_{0 \le i \le k+1} \in \operatorname{supp} \rho \text{ with } y_{k+1}' = y \text{ and } (x_0', y_0') = (x_0, y_0).$$

Then one has $\psi(x) = \sup_y\{\zeta(y) - c_L(x, y)\}$, i.e. ψ is c_L-convex. It follows, like in step 3 of the proof of [28, Theorem 5.10 (i)] that

$$\psi^{c_L}(y) - \psi(x) = c_L(x, y)$$

for ρ-almost all $(x, y) \in M \times M$. Since $\psi^{c_L}(y) - \psi(x) \leq c_L(x, y)$ everywhere one obtains equality π-almost everywhere for any optimal coupling π of μ and ν. This finishes the proof of Theorem 2.8.

4.2. Proof of Theorem 2.12(i)

Let Π be a dynamical optimal coupling between μ and ν and $\psi \colon M \to \mathbb{R} \cup \{\infty\}$ be a solution to the DKP for (μ, ν). The proof is carried out via contradiction, i.e. one assumes that

$$\Pi(\Gamma_0) > 0$$

or equivalently $\Pi|_{\Gamma_0} \neq 0$.

One has $(\mathrm{ev}_0)_\sharp \Pi|_{\Gamma_0} \leq \mu$ and therefore

$$(\mathrm{ev}_0)_\sharp \Pi|_{\Gamma_0} \ll \mathcal{L}_H.$$

The goal is to find a set V with positive measure relative to μ such that $\psi|_V \equiv -\infty$, i.e. contradicting the definition of a solution to the DKP for (μ, ν).

Since the problem is local one can assume that $\mathrm{supp}\,\mu$ and $\mathrm{supp}\,\nu$ are compact. The Borel measurable map $\Gamma \to TM$, $\gamma \mapsto \dot\gamma(0)$ induces a measure on $\partial\mathcal{C}$ via the push forward of $\Pi|_{\Gamma_0}$.

Since μ and ν have disjoint compact support there exists a lower bound $\varepsilon_0 > 0$ on the distance between points in the supports. In order to illuminate the construction one can, by diminishing ε_0 and considering an intermediate transport, assume that

(1) there exists a submanifold chart $U \to \mathbb{R}^m$ of $H \cong \mathbb{R}^{m-1} \times \{0\}$ such that $\partial_m \in \mathrm{int}\,\mathcal{C}$ everywhere and

$$\mathrm{supp}\,\mu \cup \mathrm{supp}\,\nu \subset B_{3\varepsilon_0}(0) \subset \mathbb{R}^m \cong U,$$

(2) the Riemannian metric is induced by the euclidian metric on a convex neighborhood of $\mathrm{supp}\,\mu \cup \mathrm{supp}\,\nu$ and
(3) $\exp^{\mathbb{L}}$ is a diffeomorphism from an open set in $T\mathbb{R}^m$ onto $B_1(0) \times B_1(0)$.

In order to justify these assumptions one has to show that the intermediate transport has a solution to the DKP for the transported measures.

Lemma 4.1. *Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ such that the DKP for (μ, ν) has a solution and let Π be a dynamical optimal coupling of μ and ν. Further let $\sigma_1, \sigma_2 \colon \Gamma \to [0, 1]$ be measurable with $\sigma_1 \leq \sigma_2$. Then there exists a solution of the DKP for the intermediate transport between $(\mathrm{ev}\circ\sigma_1)_\sharp\Pi$ and $(\mathrm{ev}\circ\sigma_2)_\sharp\Pi$.*

Proof. (i) First consider the special case $\sigma_1 \equiv 0$. The assertion then claims that for $\sigma \colon \Gamma \to [0, 1]$ the DKP has a solution for the martingales μ and $\nu_\sigma := (\mathrm{ev}\circ\sigma)_\sharp\Pi$.

Set $\pi := (\mathrm{ev}_0, \mathrm{ev}_1)_\sharp\Pi$ and let $\psi \colon M \to \mathbb{R} \cup \{\infty\}$ be a solution of the DKP, i.e. $\psi|_{\mathrm{supp}\,\mu} \not\equiv \infty$ and

$$\psi^{c_L}(y) - \psi(x) = c_L(x, y), \ \pi\text{-almost everywhere.} \tag{1}$$

Choose a set $\Sigma \subset \mathrm{supp}\,\pi$ of full π-measure where (1) is satisfied. Then Π is concentrated on $\Sigma_\Gamma := (\mathrm{ev}_0, \mathrm{ev}_1)^{-1}(\Sigma)$. By definition of ψ^{c_L} one has

$$\psi^{c_L}(y) - \psi(x) \leq c_L(x, y), \ \text{for all } x, y \in M.$$

Assume that there exists $\gamma \in \Sigma_\Gamma$ and $t \in [0, 1]$ such that

$$\psi^{c_L}(\gamma(t)) - \psi(\gamma(0)) < c_L(\gamma(0), \gamma(t)).$$

Then there exists $x \in M$ with

$$\psi(x) + c_L(x, \gamma(t)) < \psi(\gamma(0)) + c_L(\gamma(0), \gamma(t))$$

by the definition of the c_L-transform ψ^{c_L}. Adding $c_L(\gamma(t), \gamma(1))$ to both sides and applying the triangle inequality, which is an equality on the right hand side, one obtains

$$\psi^{c_L}(\gamma(1)) \le \psi(x) + c_L(x, \gamma(1)) < \psi(\gamma(0)) + c_L(\gamma(0), \gamma(1)),$$

which implies

$$\psi^{c_L}(\gamma(1)) - \psi(\gamma(0)) < c_L(\gamma(0), \gamma(1)),$$

a contradiction. Therefore one has

$$\psi^{c_L}(\gamma(t)) - \psi(\gamma(0)) = c_L(\gamma(0), \gamma(t))$$

for all $\gamma \in \Sigma_\Gamma$ and the lemma in the special case $\sigma_1 \equiv 0$ is proved.

(ii) Second consider the case $\sigma_1 \equiv 0$. By definition ψ^{c_L} is a c_L-concave function and since ψ is c_L-convex one has, cf. [28],

$$\psi(x) = \sup\{\psi^{c_L}(y) - c_L(x, y) \mid y \in M\} = (\psi^{c_L})^{c_L}(x)$$

for all $x \in M$. This implies that the transport problem between μ and ν has a solution of the DKP if and only if there exists a c_L-concave function $\varphi \colon M \to \mathbb{R} \cup \{-\infty\}$ with

$$\varphi(y) - \varphi^{c_L}(x) = c_L(x, y), \quad \pi\text{-almost everywhere}$$

for all optimal coupling π of μ and ν. With an analogous argument as in case (i) one obtains

$$\varphi(\gamma(1)) - \varphi^{c_L}(\gamma(t)) = c_L(\gamma(t), \gamma(1))$$

for all $\gamma \in \Sigma_\Gamma$ and $t \in [0, 1]$ with the notation of the special case. Setting $\varphi = \psi^{c_L}$ yields the assertion.

(iii) To complete the proof consider the succession of first the intermediate transport between $\sigma_1' \equiv 0$ and $\sigma_2' \equiv \sigma$ and second the intermediate transport between $\sigma_1'' \equiv \sigma_1$ and $\sigma_2'' \equiv 1$. $\quad\square$

For $v \in T\mathbb{R}^m \cong TU$ denote with v_H the projection of v along ∂_m onto $\text{span}\{\partial_1, \ldots, \partial_{m-1}\}$. Further let γ_v be the unique Φ-orbit with $\dot{\gamma}_v(0) = v$ for $v \in \mathcal{C}$.

Lemma 4.2. *For $\varepsilon_0 > 0$ sufficiently small there exists $C_0 < \infty$ and $\varphi_0 > 0$ such that for all $v \in \partial \mathcal{C} \cap T^1 B_{3\varepsilon_0}(0)$ and $t > 0$ such that*

$$\text{dist}(\gamma_v(0), \gamma_v(t)) \in [\varepsilon_0, 6\varepsilon_0]$$

the intersection

$$P_{v,t} := \partial J^-(\gamma_v(t)) \cap (\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\})$$

is a smooth hypersurface in $\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\}$ with

(a) *the norm of the second fundamental form bounded by C_0 and*
(b) *$\angle(v_H, T P_{v,t}) > \varphi_0$.*

Proof. Consider for $\varepsilon > 0$ the map

$$r_\varepsilon \colon B_1(0) \to B_\varepsilon(0), \quad x \mapsto \varepsilon x.$$

The Lagrangians $\mathbb{L}_\varepsilon := \frac{1}{\varepsilon^2} r_\varepsilon^* \mathbb{L}$ converge uniformly to $\mathbb{L}_0 := \mathbb{L}(0)$ on every compact subset of $T B_1(0)$ for $\varepsilon \to 0$ in every C^k-topology. The Euler-Lagrange flow of $\mathbb{L}_\varepsilon$ then converges uniformly on compact subsets in every C^k-topology to the Euler-Lagrange flow of $\mathbb{L}_0$. Note that the Euler-Lagrange equation for $\mathbb{L}_0$ is $\ddot{x} = 0$. Thus one has $\exp_x^{\mathbb{L}_0}(\mathcal{C}_0) = x + \mathcal{C}_0$ and therefore the assertion (a) and (b) hold for $\mathbb{L}_0$. Since the Riemannian metric h is euclidian on $B_1(0)$ one has $\frac{1}{\varepsilon^2} r_\varepsilon^* h = h$. This together with the fact that

$$\exp_x^{\mathbb{L}_\varepsilon}(V_{\varepsilon,x} \cap \partial\mathcal{C}_{\varepsilon,x}) = \partial J_\varepsilon^+(x) \subset B_1(0)$$

for some neighborhood $V_{\varepsilon,x}$ of 0_x in $T_x B_1(0)$ one obtains the lemma for $\varepsilon > 0$ sufficiently small. $\quad\square$

Statement (b) is equivalent to requiring

$$\angle(v_H, N_{v,t}) \le \frac{\pi}{2} - \varphi_0$$

where $\angle(v, w)$ denotes the euclidian angle between v and w and $N_{v,t}$ denotes the inward pointing unit normal to $P_{v,t}$. Property (a) implies that for $\varepsilon_1 := \frac{1}{\sqrt{C_0}} > 0$ one has

$$B_{\varepsilon_1}(\gamma_v(0) + \varepsilon_1 N_{v,t}) \cap (\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\}) \subset J^-(\gamma_v(t)) \cap (\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\}).$$

This and (b) then imply that

$$\{\gamma_v(0)\} + \mathrm{Cone}\left(v, \frac{\varphi_0}{2}, \varepsilon_2\right) \subset J^-(\gamma_v(t)) \cap (\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\})$$

where

$$\mathrm{Cone}(v, \varphi, \varepsilon) := \{w \in \mathbb{R}^{m-1} \times \{0\} \mid \angle(v_H, w) < \varphi, |w| < \varepsilon\}$$

and

$$\varepsilon_2 := 2\varepsilon_1 \sin\frac{\varphi_0}{2}.$$

Now for $p \in B_{3\varepsilon_0}(0)$, $v \in \partial\mathcal{C}_p$, $q \in \{p\} - \mathrm{Cone}(v, \frac{\varphi_0}{4}, \varepsilon_2)$ and $w \in \partial\mathcal{C}_q$ with $\angle(w_H, v_H) < \frac{\varphi_0}{4}$ one has

$$B_{\mathrm{dist}(p,q) \sin\frac{\varphi_0}{4}}(p) \cap (\{\gamma_v(0)\} + \mathbb{R}^{m-1} \times \{0\}) \subset J^-(\gamma_w(t))$$

for $\mathrm{dist}(q, \gamma_w(t)) \in [\varepsilon_0, 3\varepsilon_0]$.

Abbreviate

$$A := \mathrm{supp}[(\mathrm{ev}_0)_\sharp(\Pi|_{\Gamma_0})].$$

Define a map

$$T: \mathrm{supp}(\Pi|_{\Gamma_0}) \to S^{m-2}, \quad \gamma \mapsto \dot\gamma(0)_H.$$

Since $(\pi_{TM} \circ T)_\sharp \Pi|_{\Gamma_0} = \mu|_A$ and $\mu(A) \ne 0$ one has

$$T_\sharp(\Pi|_{\Gamma_0}) \ne 0$$

where $\pi_{TM}: TM \to M$ denotes the canonical projection. Choose $v_0 \in S^{m-2}$ such that

$$T_\sharp(\Pi|_{\Gamma_0})(B_{\frac{\varphi_0}{4}}(v_0)) > 0.$$

Then $A' := \pi_{TM}(B_{\frac{\varphi_0}{4}}(v_0))$ has positive measure with respect to $\mathcal{L}_H$, since $\mu|_A \ll \mathcal{L}_H$. Let p be a Lebesgue point of A', i.e.

$$\lim_{\delta\downarrow 0} \frac{\mathcal{L}_H(A' \cap B_\delta(p))}{\mathcal{L}_H(B_\delta(p))} = 1.$$

Then there exist $\varepsilon_3, \varepsilon_4 > 0$ such that for the unique $v \in \partial\mathcal{C}$ with $v_H = v_0$

$$\mathcal{L}_H\left(A' \cap B_{2\varepsilon_3}(p) \setminus B_{\varepsilon_3}(p) \cap \left(\{p\} - \mathrm{Cone}\left(v, \frac{\varphi_0}{4}, \varepsilon_2\right)\right)\right) \ge \varepsilon_4.$$

Choose polar coordinates $(r, \theta_1, \ldots, \theta_{m-2})$ on $B_{2\varepsilon_3}(p)) \subset \mathbb{R}^{m-1} \times \{0\}$ centered at p. By Fubini's Theorem there exists $(\eta_1, \ldots, \eta_{m-2})$ such that

$$\int\limits_{\varepsilon_3}^{2\varepsilon_3} \chi_{A'}(r, \eta_1, \ldots, \eta_{m-2}) r\,dr \ge \varepsilon_4.$$

Thus the 1-dimensional Lebesgue measure

$$\mathcal{L}_1(A' \cap \{r \in (\varepsilon_3, 2\varepsilon_3), \theta_1 = \eta_1, \ldots, \theta_{m-2} = \eta_{m-2}\}) \geq \frac{\varepsilon_4}{2\varepsilon_3}.$$

Recall Lemma 3.3:

Lemma. *Let $[a, b] \subset \mathbb{R}$, $\varepsilon > 0$ and $B \subset [a, b]$ Borel measurable be given with $\mathcal{L}_1(B) \geq \varepsilon(b - a)$. Then for all $n \in \mathbb{N}$ there exists $\{t_i\}_{1 \leq i \leq n} \subset B$ with $t_1 < \ldots < t_n$ and $t_{i+1} - t_i \geq \frac{\varepsilon}{2n}$.*

Applying Lemma 3.3 to

$$B := A' \cap \{r \in (\varepsilon_3, 2\varepsilon_3), \theta_1 = \eta_1, \ldots, \theta_{m-2} = \eta_{m-2}\}$$

yields for every $n \in \mathbb{N}$ points

$$x_1, \ldots x_n \in A' \cap \{r \in (\varepsilon_3, 2\varepsilon_3), \theta_1 = \eta_1, \ldots, \theta_{m-2} = \eta_{m-2}\}$$

and $y_1, \ldots, y_n \in M$ such that $(x_i, y_i) \in \text{supp}\,\pi$ for the optimal coupling π induced by Π, $x_{i+1} \in J^-(y_i)$, $\text{dist}(x_i, x_{i+1}) \geq \frac{\varepsilon_4}{4n\varepsilon_3}$ and $B_{\varepsilon_3 \sin \frac{\varphi_0}{4}}(p) \subset J^-(y_n)$. Using the exponential map of $\mathbb{L}$ one sees that there exists $C > 0$ such that $c_L(x_{i+1}, y_i) \leq -\frac{C}{\sqrt{n}}$. Then one has for every point $x' \in B_{\varepsilon_3 \sin \frac{\varphi_0}{4}}(p)$

$$\psi(x') \leq \psi(x_1) + \sum c_L(x_{i+1}, y_i) \leq \psi(x_1) - (n - 1)\frac{C}{\sqrt{n}} \leq \psi(x_1) - \frac{C}{2}\sqrt{n}$$

for all $n \in \mathbb{N}$ where ψ denotes the solution to the DKP for (μ, ν). Thus

$$\psi|_{B_{\varepsilon_3 \sin \frac{\varphi_0}{4}}(p)} \equiv -\infty$$

therefore contradicting that ψ is c_L-convex.

4.3. Proof of Theorem 2.12(ii)

One can prove Theorem 2.12(ii) in the same fashion as Theorem 2.12(i), but for the sake of avoiding repetition Theorem 2.12(ii) is reduced to Theorem 2.12(i).

Let $f_\mu \in L^1(\mathcal{L})$ be the density of μ with respect to the Lebesgue measure $\mathcal{L}$ on M. If $\Pi(\Gamma_0) > 0$ there exists a chart $U \to \mathbb{R}^m$ of M such that U is foliated by spacelike hypersurfaces $\{H_s = \mathbb{R}^{m-1} \times \{s\}\}_{s \in \mathbb{R}}$ and $(\text{ev}_0)_\sharp \Pi|_{\Gamma_0}(U) > 0$. By [27, Corollary 3.5] one can then assume that $\text{supp}\,\mu$ is contained in U. By Fubini's Theorem the restriction of f_μ to H_s is integrable with respect to the Lebesgue measure $\mathcal{L}_{m-1}$ on $\mathbb{R}^{m-1}$ for almost all $s \in \mathbb{R}$. For $(x_m)_\sharp \mu$-almost all $s \in \mathbb{R}$ one has

$$m_s := \int_{H_s} f_\mu \, d\mathcal{L}_{m-1} > 0$$

where x_m denotes the m-th coordinate function on $\mathbb{R}^m$. For s with $m_s \in (0, \infty)$ define μ_s by

$$\mu_s(A) := \frac{1}{m_s} \int_{A \cap H_s} f_\mu \, d\mathcal{L}_{m-1}.$$

Note that μ_s is a probability measure on H_s. Consider the disintegration $\{\Pi_s\}_{s \in \mathbb{R}}$ of Π along $X_m := x_m \circ \text{ev}_0 \colon \Gamma \to \mathbb{R}$. Then one has $(\text{ev}_0)_\sharp \Pi_s = \mu_s$ for μ_m-almost all s. This can be seen as follows: Obviously one has $(X_m)_\sharp \Pi = \mu_m$ since $(\text{ev}_0)_\sharp \Pi = \mu$. It follows that for all Borel measurable $B \subset M$ one has:

$$\int_{\mathbb{R}} (\text{ev}_0)_\sharp \Pi_s(B) \, d\mu_m = \int_{\mathbb{R}} \Pi_s(\text{ev}_0^{-1}(B)) \, d\mu_m$$

$$= \int_{\mathbb{R}} \Pi_s(\text{ev}_0^{-1}(B)) \, d(X_m)_\sharp \Pi = \Pi(\text{ev}_0^{-1}(B)) = \mu(B)$$

Thus one has

$$\mu(.) = \int\limits_{\mathbb{R}} (\mathrm{ev}_0)_\sharp \Pi_s(.) d\mu_m$$

and the claim follows from the uniqueness part of the Disintegration Theorem. Finally define $\nu_s := (\mathrm{ev}_1)_\sharp \Pi_s$.

Lemma 4.3. Π_s *is an optimal dynamical coupling of μ_s and ν_s for $\mu_m := (x_m)_\sharp \mu$-almost all $s \in \mathbb{R}$. Further ψ solves the DKP for (μ_s, ν_s) for μ_m-almost all s.*

Proof. The first assertion follows from the second since the pair (ψ, ψ^{c_L}) is admissible (see Section 2). Define $\pi_s := (\mathrm{ev}_0, \mathrm{ev}_1)_\sharp \Pi_s$. Assume that ψ does not solve the DKP for (μ_s, ν_s) for s in a set $B \subset \mathbb{R}$ of positive μ_m-measure, i.e.

$$\int\limits_{M \times M} c_L(x, y) d\pi_s > \int\limits_M \psi^{c_L}(y) d\nu_s - \int\limits_M \psi(x) d\mu_s$$

for all $s \in B$. Then for some $\delta > 0$ one has

$$\int\limits_{M \times M} c_L(x, y) d\pi_s - \delta > \int\limits_M \psi^{c_L}(y) d\nu_s - \int\limits_M \psi(x) d\mu_s$$

for s in a smaller but μ_m-non-negligible set $B_\delta \subset \mathbb{R}$. This implies that

$$\int\limits_{M \times M} c_L(x, y) d\pi - \delta \int\limits_{B_\delta} m_s d\mu_m > \int\limits_M \psi^{c_L}(y) d\nu - \int\limits_M \psi(x) d\mu$$

since $c_L(x, y) \geq \psi^{c_L}(y) - \psi(x)$ for all (x, y). Note that $\int_{A_\varepsilon} m_s d\mu_m > 0$ since B_δ is μ_m-non-negligible. This contradicts the assumption that ψ is a solution to the dual problem for (μ, ν) since π is minimal. $\quad\square$

Theorem 2.12(i) then yields that for almost all $s \in \mathbb{R}$ one has $\Pi_s(\Gamma_0) = 0$. Since

$$\Pi(\Gamma_0) = \int\limits_{\mathbb{R}} m_s \Pi_s(\Gamma_0) ds$$

the claim follows. The last equation follows from the respective statement about μ_s, i.e.

$$\mu(A) = \int\limits_{\mathbb{R}} m_s \mu_s(A) ds$$

for all measurable $A \subset \mathbb{R}^m$.

4.4. Proof of Theorem 2.14

The proof of Theorem 2.14 relies on ideas of [11], see also [17,12,18]. Note, however, there is no unique equivalent to the assumption of achronality resp. acausality of the support of the second measure. Indeed, the Monge problem is in general highly non-unique. However, the proof in the non-relativistic as well as the relativistic setting relies essentially on the following two properties: Geodesics with endpoints in a given set are non-branching (Lemma 4.8) and that there is a weak form of the measure contraction property (Lemma 4.6). The latter holds due to differentiability of the time function and the exponential map.

For the proof it suffices to consider the case that both $\mathrm{supp}\,\mu$ and $\mathrm{supp}\,\nu$ are compact and disjoint. This follows from the observation that absolute continuity is equivalent to absolute continuity on every compact subset.

Recall that

$$\Phi^{\mathbb{L}} : \mathbb{U} \subset \mathbb{R} \times TM \to TM$$

denotes the Euler-Lagrange flow of $\mathbb{L}$. Set

$$\{1\} \times \mathbb{V} := (\{1\} \times TM) \cap \mathbb{U}.$$

By [27, Proposition 3.14] the map

$$\exp^{\mathbb{L}} : \mathbb{V} \to M \times M, \ v \mapsto (\pi_{TM}(v), \pi_{TM} \circ \Phi^{\mathbb{L}}(1, v))$$

is a C^1-diffeomorphism of a neighborhood of the zero section onto its image and smooth outside $T^0 M$. Here

$$\pi_{TM} : TM \to M$$

denotes the canonical projection. Set

$$\exp_x : \mathbb{V} \cap TM_x \to M, \ v \mapsto \pi_{TM} \circ \Phi^{\mathbb{L}}(1, v).$$

Further since $\partial_t (\pi_{TM} \circ \Phi^{\mathbb{L}}(t, v)) = \Phi^{\mathbb{L}}(t, v)$ and $\Phi^{\mathbb{L}}(1, tv) = \Phi^{\mathbb{L}}(t, v)$ one has

$$d(\exp_x)_0 = \partial_v (\pi_{TM} \circ \Phi^{\mathbb{L}}(1, v))|_{v=0} = \mathrm{id}_{TM} \tag{2}$$

via the canonical identification $TM_x \cong T(TM_x)_0$.

Proposition 4.4. *Let $v \in \mathcal{C}_p := \mathcal{C} \cap TM_p$ such that*

$$d(\exp_p)_v : T(TM_p)_v \to TM_{\exp_p^{\mathbb{L}}(v)}$$

is singular. Then for every $T > 1$ the geodesic

$$\eta_v : [0, T] \to M, \ t \mapsto \exp_p(tv)$$

is not $\mathcal{A}$-minimizing between its endpoints.

Proof. In the case of $v \in \mathrm{int}\,\mathcal{C}$ the claim follows mutatis mutandis as in [2, Proposition 7.4.1.], since C^1-small variations of timelike curves remain timelike.

The case $v \in \partial\mathcal{C}$ is the subject of [1, Proposition 6.8]. Note that the definition of Lorentz-Finsler metrics therein is equivalent to the presently used by virtue of [24]. □

Corollary 4.5. *Let $K \subset M$ be compact. Consider the set $\mathcal{K}$ of $\mathcal{A}$-minimal causal $\mathbb{L}$-geodesics $\eta : [0, 1] \to M$, i.e. $\eta(t) = \exp_{\eta(0)}(t\dot\eta(0))$ and $\mathcal{A}(\eta) = c_L(\eta(0), \eta(1))$, with $\eta(0), \eta(1) \in K$. Then there exists a continuous function $e : (0, 1) \to \mathbb{R}_{>0}$ with*

$$\|d(\exp_{\eta(0)})_{t\dot\eta(0)}^{-1}\| \leq e(t)$$

for all $\eta \in \mathcal{K}$ and $t \in [0, 1)$. Further one has $\lim_{t \to 0} e(t) = 1$.

Proof. For a single $\mathcal{A}$-minimal geodesic $\eta : [0, 1] \to M$ the claim follows directly from Proposition 4.4. Further by (2) one has $e(t) \to 1$ for $t \to 0$.

Since K is compact the corollary follows from the continuity of the differential of the exponential map. □

For $v \in \mathcal{C}$ consider the unique time affinely parameterized local $\mathcal{A}$-minimizer $\gamma_v : \mathbb{R} \to M$ with $\dot\gamma_v(0) = v$. According to [27, Section 3.5] γ_v is an orbit of a flow on $\mathcal{C} \setminus T^0 M$. Define the map

$$\exp^\tau : \mathcal{C} \to M, \ v \mapsto \gamma_v(1).$$

Denote with $\exp_p^\tau$ the restriction of $\exp^\tau$ to $\mathcal{C} \cap TM_p$. The map satisfies $\exp^\tau(tv) = \gamma_v(t)$ which implies

$$d(\exp_p^\tau)_v \to \mathrm{id}_{TM_p} \tag{3}$$

for $v \to 0$ and

$$d(\exp_p^\tau)_v(v) = \dot\gamma_v(1). \tag{4}$$

Further by [27, Proposition 3.5] there exists a smooth map

$$s : \mathcal{C} \setminus T^0 M \to [0, \infty)$$

such that $\exp^\tau(v) = \exp(s(v)v)$ for all $v \in \mathcal{C} \setminus T^0 M$. One has

$$d(\exp_p^\tau)_v = d(\exp_p)_{s(v)v}(ds_v \otimes v + s(v) \cdot \mathrm{id}_{T M_p}). \tag{5}$$

The following lemma is at the heart of the proof of Theorem 2.14. It is an easy consequence of differentiability of the time function τ and the exponential map.

Lemma 4.6 (*weak measure contraction property*). *Let $K \subset M$ be compact. Then there exists a function $f : (0, 1) \to (0, 1)$ with $\lim_{t \to 0} f(t) = 1$ such that for all $y \in K$ and all measurable $A \subset K \cap J^-(y)$ it holds*

$$\mathcal{L}(A_{t,y}) \geq f(t) \mathcal{L}(A)$$

where $A_{t,y} = \{ev_t \gamma \mid \gamma \in (ev_0, ev_1)^{-1}(A \times \{y\})\}$.

Proof. From (4) and (5) follows that $d(\exp_p^\tau)_v$ is singular if and only if $d(\exp_p)_v$ is singular. By Corollary 4.5 and continuity of both $d \exp_p^\tau$ and $d \exp_p$ there exists a function $e^\tau : (0, 1) \to \mathbb{R}_{>0}$ with

$$\|d(\exp_{\gamma(0)}^\tau)_{t\dot\gamma(0)}^{-1}\| \leq e^\tau(t) \tag{6}$$

for any $\gamma \in (ev_0, ev_1)^{-1}(K \times \{y\})$. Now consider

$$T_{K,y} := \{v \in \mathcal{C} \mid \exists \gamma \in (ev_0, ev_1)^{-1}(K \times \{y\}) \text{ with } \dot\gamma(0) = v\}.$$

It follows that

$$A_{t,y} = \exp^\tau(t(T_{K,y} \cap (\pi_{TM})^{-1}(A))).$$

By equation (6) there exists a function $f : (0, 1) \to (0, 1)$ independent of A with

$$\mathcal{L}(A_{t,y}) \geq f(t) \mathcal{L}(A)$$

for all $t \in (0, 1)$. With property (3) one concludes $f(t) \to 1$ for $t \to 0$. □

For the following proposition observe that due to the fact that geodesics γ with $(\tau \circ \gamma)' \equiv \mathrm{const}$ are uniquely defined by their initial velocity one knows that the image of such a geodesic is a one-dimensional rectifiable curve. In particular, it has zero measure with respect to the Lebesgue measure $\mathcal{L}$ on M. This implies that for distinct points $x, y \in M$ the set

$$B_{x,y} := \{z \in J^-(y) \cap J^-(x) \mid c_L(z, y) = c_L(z, x) + c_L(x, y)$$
$$\text{or } c_L(z, x) = c_L(z, y) + c_L(y, x)\}$$

has vanishing $\mathcal{L}$-measure. In particular,

$$\mathcal{L}(A_{t,x} \cap A_{t,y}) = 0$$

for $x \neq y$. Note that if x and y are not causally related then $A_{t,x} \cap A_{t,y} = \emptyset$.

A more general statement of this form was obtained by the second author in [27].

Lemma 4.7. *Let B be a closed achronal set. Then the set*

$$\bigcup_{x \neq y \in B} B_{x,y}$$

has vanishing $\mathcal{L}$-measure.

Proof. For $z \in \cup_{x \neq y \in B} B_{x,y}$ there exist $x, y \in B$ with $z \in B_{x,y}$. Especially one has $c_L(z, x), c_L(z, y) < \infty$. By the definition of $B_{x,y}$ one concludes that $c_L(x, y) < \infty$ or $c_L(y, x) < \infty$.

Now let $\Gamma_B \subset \Gamma$ be the subspace of minimizers which intersect B at least twice. By [27, Proposition 3.22] the set

$$\bigcup_{\gamma \in \Gamma_B} \gamma(\mathbb{R}) \cap \{\tau = r\}$$

has vanishing Lebesgue measure in $\{\tau = r\}$ for all $r \in \mathbb{R}$. Thus

$$\bigcup_{\gamma \in \Gamma_B} \gamma(\mathbb{R})$$

has vanishing Lebesgue measure in M. Since

$$\bigcup_{x \neq y \in B} B_{x,y} \subset \bigcup_{\gamma \in \Gamma_B} \gamma(\mathbb{R})$$

the claim follows.   □

For a map $\sigma : \Gamma \to [0, 1]$ and a geodesic $\gamma \in \Gamma$ one writes $\gamma_\sigma = \gamma(\sigma(\gamma))$. Let $S \subset \Gamma$ be a subset of the space of minimizing geodesics. Then for $s, t \in [0, 1]$ one defines

$$S_{t,s} := (\mathrm{ev}_s, \mathrm{ev}_t)(S) \subset M \times M,$$

$S_t := \mathrm{ev}_t(S)$ and $S_\sigma := \{\gamma_\sigma\}_{\gamma \in S} \subset M$.

Lemma 4.8. *Let $S \subset \Gamma$ be such that $S_{0,1}$ is c_L-cyclically monotone. If S_1 is achronal then for all Borel-measurable maps $\sigma : S \to (0, 1)$ one has*

$$\mathcal{L}(S_\sigma^{(1)} \cap S_\sigma^{(2)}) = 0$$

for all Borel-measurable $S^{(1)}, S^{(2)} \subset S$ with $S_1^{(1)} \cap S_1^{(2)} = \varnothing$.

Remark 4.9. If S_1 is additionally acausal, e.g. S_1 is contained in a time-slice $\{\tau = \tau_0\}$, then one even has $S_\sigma^1 \cap S_\sigma^2 = \varnothing$.

Proof. Let $z \in S_\sigma^1 \cap S_\sigma^2$. Then there exist $\gamma_1 \in S^{(1)}$ and $\gamma_2 \in S^{(2)}$ such that

$$c_L(x_i, y_i) = c_L(x_i, z) + c_L(z, y_i)$$

for $i = 1, 2$, $x_i := \gamma_i(0)$ and $y_i := \gamma_i(1)$.

Since $S_{0,1}$ is c_L-cyclically monotone one has

$$
\begin{aligned}
c_L(x_1, y_1) + c_L(x_2, y_2) &\leq c_L(x_2, y_1) + c_L(x_1, y_2) \\
&\leq (c_L(x_2, z) + c_L(z, y_1)) + (c_L(x_1, z) + c_L(z, y_2)) \\
&= c_L(x_1, y_1) + c_L(x_2, y_2).
\end{aligned}
$$

This shows that there is a minimizing geodesic connecting x_1 with y_2 and passing through z. As geodesics are locally unique, either y_1 is on the geodesic connecting z and y_2 or y_2 is on the geodesic connecting z and y_1. Thus $z \in B_{y_1, y_2}$ with $y_1 \neq y_2$.

In particular,

$$S_\sigma^1 \cap S_\sigma^2 \subset \bigcup_{y \neq y' \in S_1} B_{y,y'}.$$

Since S_1 is achronal it follows that

$$\mathcal{L}(S_\sigma^1 \cap S_\sigma^2) = 0$$

by Lemma 4.7.   □

Now one combines Lemma 4.8 with the weak measure contraction property to obtain the following.

Proposition 4.10. *Assume that*

$$A \times \{y, z\} \subset \operatorname{supp} \mu \times \operatorname{supp} \nu$$

is c_L-cyclically monotone for an achronal two-point set $\{y, z\}$ and some measurable set A. Then A has vanishing Lebesgue measure.

Proof. By inner regularity of $\mathcal{L}$ one may assume A is compact so that for a fixed $\epsilon > 0$ and t sufficiently close to 0 it holds

$$A_{t,y} \cup A_{t,z} \subset A_\epsilon$$

where A_ϵ is the ϵ-neighborhood of A with respect to the distance dist. Lemma 4.8 implies that

$$\mathcal{L}(A_{t,y} \cap A_{t,z}) = 0 \quad \text{for all } t \in [0, 1).$$

Then the weak measure contraction property yields

$$
\begin{aligned}
\mathcal{L}(A) &= \lim_{\epsilon \to 0} \mathcal{L}(A_\epsilon) \\
&\geq \limsup_{t \to 0} \mathcal{L}(A_{t,y} \cup A_{t,z}) \\
&= \limsup_{t \to 0} \mathcal{L}(A_{t,y}) + \mathcal{L}(A_{t,z}) \\
&\geq 2 \limsup_{t \to 0} f(t) \mathcal{L}(A) = 2\mathcal{L}(A)
\end{aligned}
$$

which can hold only if A has zero measure. $\quad\square$

Lemma 4.8 can be used to prove an interpolation inequality in form of the weak measure contraction property between any absolutely continuous measure and a causally related achronal discrete measure. In order to prove such an interpolation inequality for general achronal target measures one needs to approximate the target measures via finite measures which satisfy the achronality assumption. As such an approximation seems difficult, one proceeds in two steps: As measures supported in a time slice can be easily approximated one first proves the interpolation inequality for those measures. In a second step one uses this fact together with the strong non-branching property implied by Lemma 4.8 to approximate general achronal target measures.

Given a subset $C \subset M \times M$ and $s, t \in (0, 1)$ define

$$C_{s,t} = \{(\operatorname{ev}_s \gamma, \operatorname{ev}_t \gamma) \mid \gamma \in (ev_0, ev_1)^{-1}(C)\}$$

and $C_t = p_1(C_{t,t})$.

Lemma 4.11. *Assume π is an optimal coupling with compact support between an absolutely continuous measure μ and a measure ν with support in a time-slice $\{\tau = \tau_0\}$. Then there is a sequence $\nu_n = \sum_{i=1}^{N_n} \lambda_i^n \delta_{x_i^n}$ such that $\operatorname{supp} \nu_n \subset \{\tau = \tau_0\}$ and the optimal couplings π_n of (μ, ν_n) converge weakly to an optimal coupling π' of (μ, ν).*

Proof. Let $X \in \Gamma(TM)$ be a vector field with $L(X) < 0$ and $d\tau(X) = 1$ everywhere and denote with φ_t the flow of X. The coupling

$$\pi'_\varepsilon := (\mathrm{id}, \varphi_\varepsilon)_\sharp \pi$$

is supported in $\{c_L < 0\}$. Recall that by [27, Proposition 3.10] there exists a Borel map $S \colon J^+ \to C^0([0, 1], M)$ with $S(x, y) \in \Gamma_{x \to y}$. The push forward $\Pi'_\varepsilon := S_\sharp \pi'_\varepsilon$ is then a dynamical coupling. Consider the subset $\Gamma_{t_0} \subset \Gamma$ of minimizers γ with $\tau \circ \gamma(0) \leq t_0$ and $\tau \circ \gamma(1) \geq t_0 + \varepsilon$. The maps

$$T_0 \colon \Gamma_{t_0} \to [0, 1], \quad \gamma \mapsto T_0(\gamma)$$

such that $\tau(\gamma(T_0(\gamma))) = t_0$ and

$$R \colon \Gamma_{t_0} \to \Gamma_{t_0}, \ \gamma \mapsto [t \mapsto \gamma(T_0(\gamma)t)]$$

are continuous. Define

$$\Pi_\varepsilon := R_\sharp(\Pi'_\varepsilon) \text{ and } \pi_\varepsilon := (\mathrm{ev}_0, \mathrm{ev}_1)_\sharp \Pi_\varepsilon.$$

It then follows that $\operatorname{supp} \pi_\varepsilon \subset \{c_L < 0\}$ and $\operatorname{supp}(p_2)_\sharp \pi_\varepsilon \subset \{\tau = \tau_0\}$. Furthermore, the Prokhorov distance between $\nu_\epsilon := (p_2)_\sharp \pi_\epsilon$ and ν tends to zero for $\varepsilon \to 0$.

Observe that for any approximation by finite measures $(\nu_{n,\epsilon})$ of ν_ϵ the C_L-cost between μ and $\nu_{n,\epsilon}$ is eventually finite and the distance between ν and ν_n is eventually less than 2ϵ. One may also assume that $(\nu_{n,\epsilon})$ has support in $\{\tau = \tau_0\}$.

Denote by $\pi_{n,\epsilon}$ the c_L-optimal coupling of $(\mu, \nu_{n,\epsilon})$. Then

$$\liminf_{n \to \infty} \int c_L d\pi_{n,\epsilon} \le \int c_L d\pi_\epsilon.$$

To conclude just observe that for a diagonal sequence $\pi_{(k)} = \pi_{n_k, \frac{1}{k}}$ one has $\pi_{(k)} \rightharpoonup \tilde{\pi}$ satisfying

$$\int c_L d\tilde{\pi} \le \liminf_{k \to \infty} \int c_L d\pi_{(k)} \le \int c_L d\pi.$$

Since π is optimal the $\tilde{\pi}$ must be optimal as well. $\square$

Remark 4.12. If an optimal coupling π is supported in the interior of J^+ then it is possible to obtain an approximation π_n with finite target measures which have support in $\operatorname{supp}((p_2)_\sharp \pi)$. Thus it follows that it is possible to keep the target approximation ν_n in a fixed achronal set B. Note, however, such an approximation for purely lightlike optimal couplings is not always possible. It even seems difficult to prove Lemma 4.11 under the assumption that $\operatorname{supp} \nu$ is achronal.

Proposition 4.13. *Let $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ with μ being absolutely continuous and $\operatorname{supp} \nu \subset \{\tau = \tau_0\}$. Then there is an optimal coupling π of (μ, ν) such that for all c_L-cyclically monotone sets $C \subset \operatorname{supp} \pi$ with $\pi(C) = 1$ it holds*

$$\mathcal{L}(C_t) \ge f(t)\mathcal{L}(C_0).$$

Proof. First note if ν is a finite measure then the support $C = \operatorname{supp} \pi$ of any c_L-cyclically monotone coupling π satisfies the assumption. Indeed, the set of points $x \in C_0$ such that $(x, y), (x, y') \in C$ for distinct $y, y' \in C_1$ has zero $\mathcal{L}$-measure, i.e. $\mathcal{L}(A^i_{t, y_i} \cap A^j_{t, y_j}) = 0$ for $i \ne j$ and $t \in [0, 1)$ where $\{y_i\}_{i=1}^n = C_1$ and $A^i = p_1((p_2)^{-1}(y_i))$. Observe now

$$\mathcal{L}(C_t) = \sum_{i=1}^n \mathcal{L}(A^i_{t, y_i}) \ge \sum_{i=1}^n f(t)\mathcal{L}(A^i) = \mathcal{L}(C_0)$$

For more general ν let π be the weak limit of a sequence π_n with $(p_2)_\sharp \pi_n$ finite as given by Lemma 4.11. Note by restricting the first marginal of π_n slightly one can assume that the support of π_n converges in the Hausdorff metric to the support of π. Note that since π_n converges weakly to π one must have $\mu(C_0^n) \to 1$ where $C^n = \operatorname{supp} \pi_n$.

If $C = \operatorname{supp} \pi$ is c_L-cyclically monotone then for all $\epsilon > 0$ and for sufficiently large $n \in \mathbb{N}$ it holds $(C_t^{(n)}) \subset (C_t)_\epsilon$. Since C_t is compact and $C_0^{(n)} = C_0$ one obtains

$$\mathcal{L}(C_t) = \lim_{\epsilon \to 0} \mathcal{L}((C_t)_\epsilon)$$
$$= \limsup_{n \to \infty} \mathcal{L}(C_t^{(n)})$$
$$\ge \limsup_{n \to \infty} f(t)\mathcal{L}(C^{(n)_t})$$
$$= f(t)\mathcal{L}(C_0).$$

If the support of π is not c_L-cyclically monotone, one may find a c_L-cyclically monotone subset $C \subset \operatorname{supp} \pi$ of full π-measure and compact sets $C^k \subset C$ such that $\pi(C^k) \to \pi(C)$ and $\mathcal{L}(C_0^k) \to \mathcal{L}(C_0)$.

Denote by π_k the coupling obtained by restricting π to C^k and renormalizing. Note that each π_k is supported in C^k and is given as a weak limit of an appropriate restriction of the approximating sequence π_n. In particular, one sees that the claim of the proposition holds for C^k so that one concludes with the following chain of inequalities

$$
\begin{aligned}
\mathcal{L}(C_t) &\geq \limsup_{k \to \infty} \mathcal{L}(C_t^k) \\
&\geq \limsup_{k \to \infty} f(t)\mathcal{L}(C_0^k) = f(t)\mathcal{L}(C_0). \quad \square
\end{aligned}
$$

Combining the results above one obtains the existence and uniqueness of optimal transport maps if the target is supported in a time-slice.

Proposition 4.14. *Between any absolutely continuous probability measure μ and any probability measure ν supported in a time-slice $\{\tau = \tau_0\}$ such that $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$ there exists a unique c_L-optimal coupling π and this coupling is induced by a transport map.*

Proof. Let π be an optimal coupling for (μ, ν) and choose a c_L-cyclically monotone measurable set $C \subset \operatorname{supp} \pi$ of full π-measure.

We claim π is induced by a transport map. Note that this implies that π is unique.

Suppose the statement was wrong. Then the Selection Dichotomy in [18, Theorem 2.3] gives couplings $\pi_1, \pi_2 \ll \pi$ which are supported on disjoint sets $K \times A_1$ and $K \times A_2$ and their first marginals are equal to $\mu_K = \frac{1}{\mu(K)}\mu\big|_K$, where $K \subset M$ is compact. Since μ is absolutely continuous one can additionally assume μ_K and $\mathcal{L}\big|_K$ are mutually absolutely continuous.

It is easy to see that all three measures π_1, π_2 and $\frac{1}{2}(\pi_1 + \pi_2)$ are optimal. Thus by Proposition 4.13 there are optimal couplings $\tilde{\pi}_i$ between $(\mu_K, (p_2)_\sharp \pi_i)$ such that the couplings $\tilde{\pi}_i$ are concentrated on disjoint c_L-cyclically monotone sets C^i satisfying

$$
\mathcal{L}(C_t^i) \geq f(t)\mathcal{L}(C^i).
$$

Let $\epsilon > 1$. Then $C_t^i \subset K_\epsilon$ for sufficiently small t where K_ϵ denotes the ϵ-neighborhood of K. Since the sets C_0^1 and C_0^2 are disjoint by Lemma 4.8 the sets C_t^1 and C_t^2 are disjoint as well so that one obtains

$$
\begin{aligned}
\mathcal{L}(K) &= \lim_{\epsilon \to 0} \mathcal{L}(K_\epsilon) \\
&\geq \limsup_{t \to 0} \mathcal{L}(C_t^1 \dot{\cup} C_t^2) \\
&= \limsup_{t \to 0} \mathcal{L}(C_t^1) + \mathcal{L}(C_t^2) \\
&\geq 2\limsup_{t \to 0} f(t)\mathcal{L}(C_0),
\end{aligned}
$$

which is a contradiction as $\mu_K(K) = \mu_K(C_0) = 1$ and μ_K and $\mathcal{L}\big|_K$ are mutually absolutely continuous. $\square$

Using Lemma 4.8 one can extend the proposition to general achronal target measures.

Proposition 4.15. *The previous preposition also holds for probability measures ν supported in an achronal set. Furthermore, for the unique dynamical optimal coupling Π and any c_L-cyclically monotone set $C \subset \operatorname{supp} \pi$ with $\pi(C) = 1$ where $\pi = (\mathrm{ev}_0, \mathrm{ev}_1)_\sharp \Pi$ it holds*

$$
\mathcal{L}(C_t) \geq f(t)\mathcal{L}(C_0).
$$

Proof. Assume π is a c_L-optimal coupling for (μ, ν) and choose a c_L-cyclically monotone measurable set $C \subset \operatorname{supp} \pi$ of full π-measure.

Set $C^0 = \Delta \cap C$ and $C^{>0} = C \backslash \Delta$ where Δ is the diagonal in $M \times M$. The intersection $C_0^0 \cap C_0^{>0}$ is μ-negligible.

Indeed, all points in the intersection $C_0^0 \cap C_0^{>0}$ would have a minimizer passing through that point which intersects two (necessarily distinct) points in C_0^0 and $C_1^{>0}$. Hence the set must be $\mathcal{L}$-negligible which also shows $\mu(C_0^0 \cap C_0^{>0}) = 0$.

Observe that the claim implies that π is induced by a transport map if and only if π restricted to $C^{>0}$ is induced by a transport map. If either of the cases holds then π must be unique.

So without loss of generality one can assume that π is concentrated away from the diagonal Δ. In this case π must be concentrated on $\bigcup_{\tau_0 \in \mathbb{Q}, n \in \mathbb{N}} \Omega_{\tau_0, n}$ where

$$\Omega_{\tau_0, n} = \left\{ \tau \leq \tau_0 - \frac{1}{n} \right\} \times \left\{ \tau \geq \tau_0 + \frac{1}{n} \right\}.$$

Furthermore, π is induced by a transport map if and only if for each $\tau_0 \in \mathbb{Q}$ and $n \in \mathbb{N}$, $\pi|_{\Omega_{\tau_0, n}}$ is either the zero measure or induced by a transport map. Thus one may assume that π is supported in $\Omega_{\tau_0, n}$ for some $\tau_0 \in \mathbb{Q}$ and $n \in \mathbb{N}$.

Let $\sigma : \Gamma \to (0, 1)$ be measurable with $\tau(\gamma_\sigma) = \tau_0$ whenever $\gamma(0) \leq \tau_0 \leq \gamma(1)$. Then given an optimal dynamical coupling Π one obtains an intermediate measure μ_σ which is supported in the time-slice $\{\tau = \tau_0\}$. By Proposition 4.14 for any Π there is a unique optimal coupling π_σ between μ and μ_σ and a measurable map T_σ such that $\pi_\sigma = (\mathrm{id} \otimes T_\sigma)_\sharp \mu$.

We claim that Π is unique among the dynamical couplings representing π. Assume μ_σ', π_σ' and T_σ' are obtained from a distinct optimal dynamical coupling Π'. In this case the maps T_σ and T_σ' do not agree on a set of positive μ-measure. By construction the measure $\frac{1}{2}(\pi_\sigma + \pi_\sigma')$ is the unique optimal coupling between μ and $\frac{1}{2}(\mu_\sigma + \mu_\sigma')$ which is induced by a transport map. However, this is only possible if T_σ and T_σ' agree μ-almost everywhere. This is a contradiction and shows that the dynamical coupling Π representing π is unique.

Note that for π-almost all $(x, y) \in M \times M$ the point $T_\sigma(x)$ is on a geodesic connecting x and y. Since the value of T_σ is unique almost everywhere and geodesics are non-branching, for μ-almost all $x \in M$ there can be at most one geodesics γ with $\gamma_0 = x$ and $\gamma_\sigma = T_\sigma(x)$. In particular, for μ-almost all $x \in M$ there is a unique $(x, y) \in \mathrm{supp}\,\pi$. But then π is induced by a transport map and hence the unique optimal coupling between μ and ν.

It remains to show that the interpolation inequality holds as well: Let Π be the unique dynamical optimal coupling and π be the unique induced optimal coupling.

Let $\chi : \Gamma \to (1 - \epsilon, 1]$ be a measurable map such that for a set Γ' of full Π measure the set $\tau \circ \chi(\Gamma')$ is countable and whenever $\gamma(1) = \eta(1)$ then $\tau(\gamma_\chi) = \tau(\eta_\chi)$.

Let μ_χ be the intermediate measures obtained from χ. Then μ_χ is concentrated in countably many time-slices $\{\tau = \tau_k\}_{k \in \mathbb{N}}$. Observe that the interpolation property holds when we restrict the coupling to $M \times \{\tau_k\}$. Since the endpoints for two different time-slices are disjoint Lemma 4.8 implies that the interpolated points never intersect. Thus if C is a c_L-cyclically monotone subset of $\mathrm{supp}\,\pi$ of full π-measure then the set

$$C^\chi = \{(\gamma_0, \gamma_\chi) \mid \gamma \in (\mathrm{ev}_0, \mathrm{ev}_1)^{-1}(C)\}$$

is c_L-cyclically monotone and has full $(\mathrm{ev}_0, \mathrm{ev}_\chi)_\sharp \Pi$-measure and it holds

$$\mathcal{L}(C_t^\chi) \geq f(t) \mathcal{L}(C_0).$$

Via approximation it suffices to show the interpolation property assuming C is compact. Observe now that for compact C and all $\delta > 0$ it holds

$$C_t^\chi \subset (C_t)_\delta$$

for $\epsilon > 0$ sufficiently small. Thus

$$\begin{aligned}
\mathcal{L}(C_t) &= \lim_{\delta \to 0} \mathcal{L}((C_t)_\delta) \\
&\geq \limsup_{\epsilon \to 0} \mathcal{L}(C_t^\chi) \geq f(t) \mathcal{L}(C_0). \quad \square
\end{aligned}$$

Proof of Theorem 2.14. The only thing that is left is to show that the intermediate measures $\mu_t = (\mathrm{ev}_t)_\sharp \Pi$ are absolutely continuous. For this let $C = \mathrm{supp}\,\pi$ and assume μ_t was not absolutely continuous. Then there is a compact set $\tilde{C} \subset C$ such that $\mu(\tilde{C}_0) = \mu_t(\tilde{C}_t) > 0$ and $\mathcal{L}(\tilde{C}_t) = 0$. In particular, $\mathcal{L}(\tilde{C}_0) > 0$. However, the interpolation property shows $0 = \mathcal{L}(\tilde{C}_t) \geq f(t)\mathcal{L}(\tilde{C}_0)$ which is clearly a contradiction and thus proving that μ_t is absolutely continuous. $\square$

The following corollary turns out to be useful in the next section.

Corollary 4.16 (*Self-intersection lemma*). *If μ and ν are causally related, μ is absolutely continuous and ν is supported on an achronal set then for all sets A of full μ-measure there is a $t_0 \ll 1$ such that the intermediate measures μ_t, $t \in (0, t_0)$ satisfy $\mu_t(A) > 0$. In particular, μ and μ_t cannot be mutually singular.*

Proof. By restricting μ we may assume μ has density by M. Then uniqueness of μ_t implies that the density of μ_t is bounded by $M \cdot f(t)^{-1}$, see [18, 5.15]. Since $f(t) \to 1$ as $t \to 0$ we see that the densities of μ_t for sufficiently small t can be uniformly bounded. Now the claim follows directly from the Self-Intersection Lemma in [18, Lemma 6.4]. $\square$

Remark 4.17. The argument shows that for $\mu = g\mathcal{L}$ and $\mu_t = g_t\mathcal{L}$ one has the estimate

$$g_t(\gamma_t) \leq \frac{1}{f(t)} g(\gamma_0)$$

for Π-almost all $\gamma \in \Gamma$ where Π is the unique optimal dynamical coupling between μ and ν.

4.5. Proof of Theorem 2.16

The goal is to reduce the problem to the 1-dimensional case and then construct a map from that solution. The proof is very similar to the proof of Bianchini–Cavalletti [9] for general non-branching geodesic spaces, see also [10]. However, the lack of a natural parametrization of lightlike geodesics prevents a direct application of their proof. One of the features of the proof will be to show how the time function τ and time-affinely parametrized geodesics can be used to overcome this obstacle and give a complete solution to the Monge problem in the relativistic setting.

Note that the proof shows that the optimal coupling is in general non-unique without assuming some relative form of achronality. Indeed, in order to prove uniqueness using the reduction to the 1-dimensional setting on a set of full measure there must be an almost everywhere defined injective map from the set of transport rays to M which corresponds to the target of the transport.

By [27, Proposition 2.7] one knows that the any optimal coupling is concentrated on a measurable c_L-cyclically monotone set C.

Definition 4.18 (*Maximal c_L-cyclically monotone set*). A set $A \subset M \times M$ is *maximal c_L-cyclically monotone in a set* $\Sigma \subset \{c_L \leq 0\}$ if it is c_L-cyclically monotone and is maximal with respect to inclusion among subsets of Σ.

It is not difficult to see that a maximal c_L-cyclically monotone set A must be closed if Σ is closed. One calls any maximal element A_{max} of a c_L-cyclically monotone set A a *maximal hull*. Note that the maximal hull is in general not unique.

Lemma 4.19. *Every c_L-cyclically monotone set $A \subset \{c_L \leq 0\}$ is contained in a maximal c_L-cyclically monotone set $A_{max} \subset \{c_L \leq 0\}$. In particular, if $(\mu, \nu) \in \mathcal{P}_\tau^+(M)$, then any optimal coupling is supported in a maximal c_L-cyclically monotone set $A_{max} \subset \{c_L \leq 0\}$.*

Proof. Just observe that if $\{A_i\}_{i \in I}$ is a chain of c_L-cyclically monotone sets in the closed set $\{c_L \leq 0\}$ then

$$\tilde{A} = \bigcup_{i \in I} A_i$$

is maximal in $\{c_L \leq 0\}$ and c_L-cyclically monotone. Thus Zorn's Lemma gives the existence of a maximal element A_{max} with

$$A \subset A_{max} \subset \{c_L \leq 0\}.$$

The last statement follows by observing that a coupling with finite cost must have support in $\{c_L \leq 0\}$. $\quad\square$

Let A_{max} be a maximal c_L-cyclically monotone hull of the support of an optimal coupling π of μ and ν in J^+. Further let Π be a dynamical optimal coupling of (μ, ν).

Lemma 4.20. *For any point $(x, y) \in A_{max}$ and any point $z \in \gamma \in \Gamma_{x \to y}$ one has $(x, z), (z, y) \in A_{max}$.*

Proof. Let $\{(x_i, y_i)\}_{1 \leq i \leq N} \subset A_{max}$. Then one has

$$
\begin{aligned}
c_L(x, z) + c_L(z, y) + \sum_{i=1}^{N} c_L(x_i, y_i) &= c_L(x, y) + \sum_{i=1}^{N} c_L(x_i, y_i) \\
&\leq c_L(x, y_1) + \sum_{i=1}^{n-1} c_L(x_i, y_{i+1}) + c_L(x_N, y) \\
&\leq c_L(x, y_1) + \sum_{i=1}^{n-1} c_L(x_i, y_{i+1}) + c_L(x_N, z) + c_L(z, y)
\end{aligned}
$$

where the next to last inequality follows from the cyclic monotonicity and the last inequality is the triangle inequality for c_L. This implies that $(x, z) \in A_{max}$. The other case is analogous. Note that $c_L(x, y) < \infty$ since $(x, y) \in J^+$ and thus $c_L(x, z), c_L(z, y) < \infty$. $\quad\square$

Consider the relation $R \subset M \times M$ with

$$(x, y) \in R :\Leftrightarrow (x, y) \in A_{max} \text{ or } (y, x) \in A_{max}.$$

Set $R_{>1} := \{(x, y) \in R \mid \exists z \neq x : (x, z) \in R\}$. Then one can assume without loss of generality that $R_{>1}$ has full measure relative to any optimal coupling. This follows from the observation that on $R \setminus R_{>1}$ all optimal transports are constant. It is assumed from here on that $R = R_{>1}$.

Next define the following two sets:

$$A^+ := \{x \in M \mid \exists z \neq w \in M : (x, z), (x, w) \in A_{max} \text{ and } (z, w) \notin R\}$$

and

$$A^- := \{y \in M \mid \exists x \neq z \in M : (x, y), (z, y) \in A_{max} \text{ and } (x, z) \notin R\}.$$

Assume the disintegration of π with respect to the first projection is given by

$$\pi = \mu \otimes \pi_x.$$

Lemma 4.21. *For μ-almost all $x \in A^+$ the measures π_x are supported in $\{(x, x)\}$.*

Proof. If A^+ is μ-negligible there is nothing to prove. Therefore one can assume by [27, Corollary 3.12] that $\mu(A^+) = 1$. After possibly further restricting the transport problem one can suppose that

$$\operatorname{supp} \pi \subset \{\tau \leq \tau_0 - \varepsilon\} \times \{\tau \geq \tau_0 + \varepsilon\}$$

for a sufficiently small $\varepsilon > 0$. Let now

$$\Gamma_{\tau_0} := \{\gamma \in \Gamma \mid \tau(\gamma(0)) \leq \tau_0 \leq \tau(\gamma(1))\}$$

and $\sigma : \Gamma_{\tau_0} \to [0, 1]$ be the map defined by $\tau(\operatorname{ev}(\gamma, \sigma(\gamma))) := \tau_0$. Note that Π is supported in Γ_{τ_0} under the above assumptions. Then

$$\mu_\sigma := (\operatorname{ev} \circ (\operatorname{id}, \sigma))_\sharp \Pi$$

is an intermediate measure of μ and ν which is supported in the time-slice $\{\tau = \tau_0\}$. Thus by Proposition 4.14 there is a unique coupling which, in addition, is induced by a transport map T_σ. The assumption shows that for μ-almost all $x \in A^+$ the (unique) geodesics connecting x and $T_\sigma(x)$ never intersects A^+. Thus $\mu_{\sigma,t}(A^+) = 0$ for the any intermediate measure $\mu_{\sigma,t}$ between μ and μ_σ. This, however, violates Corollary 4.16. $\quad\square$

Remark 4.22. Under the assumptions of Theorem 2.16 it is even possible to show $\mathcal{L}(A^+) = 0$. For this one needs to know whether any coupling $\tilde{\pi}$ concentrated in $A_{\max} \cap (A^+ \times \{\tau = \tau_0\})$ is induced by a transport map. By Theorem 2.14 a sufficient condition would be that $\tilde{\pi}$ is optimal.

If $\pi_x = \delta_x \otimes \delta_x$ for all x in a measurable set $A \subset M$ then $\pi|_{A \times M} = (\mathrm{id} \times \mathrm{id})_\sharp \mu|_A$. Thus in the following one will always assume that $\pi_x \neq \delta_x \otimes \delta_x$ for μ-almost all $x \in M$. In particular, the measures μ and $(p_1)_\sharp(\pi|_{M \times M \setminus \Delta})$ are mutually absolutely continuous. In combination with Lemma 4.21 one concludes that $\mu(A^+) = 0$.

Given a symmetric relation $R \subset M \times M$ let the *domain of R* be defined by

$$\mathrm{dom}(R) := p_1(R).$$

In the following one will use the following short hand notation to define a new symmetric relations $R' \subset R$: R' is *given by* $\mathrm{dom}(R') = A$ for a subset $A \subset \mathrm{dom}\, R$ if

$$R' = A \times A \cap R.$$

One easily verifies that $\mathrm{dom}\, R'$ is indeed equal to A. Furthermore, if A is (Borel) measurable then R' is (Borel) measurable or analytic if R is (Borel) measurable or analytic, respectively.

Let R_{red} be obtained by requiring $\mathrm{dom}(R_{red}) = \mathrm{dom}(R) \setminus A^+ \cup A^-$. Then one may verify that R_{red} is an equivalence relation.

Decompose μ into two measures μ_1 and μ_2 such that μ_1 is concentrated $\mathrm{dom}(R_{red})$ and μ_2 on A^-. Choose an optimal coupling π along the same decomposition $\mu = \mu_1 + \mu_2$. Denote the second marginals by ν_1 and ν_2, respectively. By the definition of A^- one sees that π_2 is concentrated on the diagonal. Thus if one finds an optimal coupling $\tilde{\pi}_1$ between μ_1 and ν_1 which is induced by a transport map then $\tilde{\pi} = \tilde{\pi}_1 + \pi_2$ is an optimal coupling between μ_1 and ν_1 which is induced by transport maps. Thus one may assume $\mu(A^\pm) = 0$.

Lemma 4.23. *There exists a measurable projection* $T : dom(R_{red}) \to dom(R_{red})$ *with* $(x, T(x)) \in R_{red}$.

Proof. Choose an enumeration $\{q_n\}_{n \in \mathbb{N}}$ of $\mathbb{Q}$. Define inductively disjoint relations $\{R_n\}_{n \in \mathbb{N}}$ as follows: Set $\tilde{R}_0 := R_{red}$. Assume that R_k for $k \leq n$ has been constructed. Define R_{n+1} by

$$(x, z) \in R_{n+1} :\Leftrightarrow (x, z) \in \tilde{R}_n \wedge \exists y, y' \in \{\tau = q_{n+1}\} : (x, y), (z, y') \in \tilde{R}_n.$$

R_{n+1} is an equivalence relation since it is the intersection of two equivalence relations. Thus $\tilde{R}_{n+1} := \tilde{R}_n \setminus R_{n+1}$ is an equivalence relation. Continuing one obtains a measurable partition $\{R_n\}_{n \in \mathbb{N}}$ of R_{red}. This follows from the initial assumption that all minimizer are non-constant.

For all $r \in \mathbb{R}$ there exists a measurable selection

$$S_r : p_1(R_{red} \cap (M \times \{\tau = r\})) \to \mathrm{dom}(R_{red}) \cap \{\tau = r\}.$$

Define the map

$$T : \mathrm{dom}\, R_{red} \to \mathrm{dom}\, R_{red}, \quad x \mapsto S_{q_n}(x) \text{ for } x \in R_n. \quad\square$$

Disintegrate μ along T, i.e. for $\mu_{red} := (T)_\sharp \mu$ let $\{t_x\}_{x \in \mathrm{dom}(R_{red})}$ be the almost everywhere defined family of probability measures on $T(\mathrm{dom}(R_{red}))$ such that

$$\mu = \mu_{red} \otimes t_x.$$

Lemma 4.24. *For* μ_{red}-*almost all* $x \in M$ *is the measure* $\bar{t}_x = \tau_\#(t_x)$ *is non-atomic.*

Proof. By the assumptions the statement holds for π if it holds for π restricted to $M \times M \setminus \Delta$. In particular, one can assume that for π-almost all $(x, y) \in M \times M$ one has $\tau(x) < \tau(y)$.

Assume now for a set A of positive μ_{red}-measure the measure $\bar{t}_x$ has atoms for all $x \in A$. Then there is a compact set $K \subset M$ of positive μ-measure such that the map

$$K \to \mathcal{P}(\mathbb{R}), \ x \mapsto \bar{t}_x$$

is weakly continuous. Thus the function

$$F \colon K \times \mathbb{R} \to [0, 1], \ F(x, r) = \bar{t}_x(\{r\})$$

is upper semi-continuous. In particular, the set

$$C = F^{-1}((0, 1])$$

is a Borel set and for each $(x, r) \in C$ the point r is an atom of $\bar{t}_x$. Applying the Selection Theorem [16, Section 423] to C yields a measurable selection $T \colon p_1(C) \to C$ such that $(x, T(x)) \in C$ for all $x \in p_1(C)$. In particular, $\mu_{red}\big|_K \otimes \delta_{T(x)}$ is non-trivial. Since t_x is atomic for all $x \in K$ one also has

$$\mu_{red}\big|_K \otimes \delta_{T(x)} \ll \mu_{red} \otimes \bar{t}_x.$$

Translating back to the coupling π one sees that there exists a measurable map $S \colon M \to M$ and a set $\tilde{K}$ of positive μ-measure such that

$$\pi' := \frac{1}{\mu(K)}(\mathrm{id} \otimes S)_\sharp \mu\big|_K \ll \pi$$

and for all $x \neq y \in \tilde{K}$ one has $(x, T(y)) \notin R_{red}$ and $\tau(x) < \tau(T(x))$.

This implies that for any $\sigma \colon \Gamma \to (0, 1)$ the intermediate measures μ'_σ of $(p_1)_\sharp \pi'$ and $(p_2)_\sharp \pi'$ would be mutually singular with respect to $(p_1)_\sharp \pi'$. However, as in the proof of Lemma 4.21, this yields a contradiction. $\quad\square$

Remark 4.25. A more involved proof shows that $\bar{t}_x$ is absolutely continuous. As this strengthened statement is not needed the details are left to the interested reader.

Disintegrating an optimal coupling π along $T \circ p_1$ yields a family of probability measures $\{s_x\}$ such that

$$\pi = \mu_{red} \otimes s_x,$$

where s_x is a probability measure on

$$(R_x \cap \mathrm{dom}(R_{red})) \times R_x \text{ with } R_x := p_2((\{x\} \times M) \cap R).$$

Lemma 4.26. *For all $x \in T(\mathrm{dom}(R_{red}))$ the set R_x is diffeomorphic to an interval and the time function τ is injective on R_x.*

Proof. From Lemma 4.20 and $x \in R_x$ one sees that R_x is formed by the image of geodesics which contain x and meet at most at their endpoints. As $x \in R_x$ is not in A^+ or A^-, it must be in the interior of R_x. Thus, because geodesics are non-branching and the time function τ is strictly increasing along causal curves one sees that R_x is the image of precisely one geodesic. $\quad\square$

Define for $x \in T(\mathrm{dom}(R_{red}))$ the measures $r_x := (p_2)_\sharp(s_x)$, i.e.

$$\nu = \mu_{red} \otimes r_x.$$

Next one constructs a transport map for the optimal couplings between t_x and r_x. By the previous Lemma the measures t_x and r_x are concentrated on a single geodesic such that the time function τ give a uniquely defined parametrization. Thus it suffices to solve the one-dimensional optimal transport problem between t_x and r_x. First observe the following.

Lemma 4.27. *Let $\gamma \in \Gamma$ and μ, ν be causally related probability measures on γ. Then any causal coupling π_γ is optimal.*

Proof. Choose a monotone reparameterization $[0, 1] \to [0, 1]$ of γ to an affine parameter. The cost function for $s, t \in [0, 1]$ then is

$$c_L(s, t) = \begin{cases} c(t - s), & \text{if } s \le t, \\ \infty & \text{else} \end{cases}$$

for some constant $c = c(\gamma) \le 0$. It is now easy to see that any causal coupling is cyclically monotone, i.e. optimal. $\quad\square$

By Lusin's Theorem one can assume that $x \mapsto (t_x, r_x)$ is continuous. Set

$$m(x, a) := t_x(\tau^{-1}(-\infty, a])$$

and

$$n(x, b) := r_x(\tau^{-1}(-\infty, b]).$$

With this define $\varphi(x, a) = b$ if $b = \text{argmin}\{m(x, a) \le n(x, b)\}$. Observe that φ is measurable and (T, τ) is injective on $\text{dom}(R_{red})$ so that there is a measurable map $\psi : \text{dom}(R_{red}) \to M$ such that $\psi(y) = \varphi(T(y), \tau(y))$ for μ_{red}-almost all $y \in \text{dom}(R_{red})$.

Again by Lusin's Theorem one may assume ψ is continuous. Define a set $\mathcal{T} \subset M \times M$ as follows

$$\mathcal{T} = \{(y, z) \mid (y, z) \in R, \psi(y) = \tau(z)\}.$$

Note that $\mathcal{T}$ is analytic and for each $x \in \text{dom}(R_{red})$ there is exactly one $(x, y) \in \mathcal{T}$. Thus $\mathcal{T}$ agrees on $\text{dom}(R_{red}) \times M$ with the graph of a measurable function $\Psi : \text{dom}(R_{red}) \to M$.

The choice of φ implies $\Psi_\sharp t_x = r_x$. Thus $(\text{id} \times \Psi)_\sharp \mu$ is a coupling of μ and ν. Since Ψ transports monotonously along each R_x one sees that Ψ is an optimal transport map between t_x and r_x. As the initial coupling was optimal, we see that along each transport ray the cost is not change In particular, the coupling $(\text{id} \times \Psi)_\sharp \mu$ is optimal between μ and ν. This finishes the proof of Theorem 2.16.

Declaration of competing interest

There are no competing interests.

References

[1] A.B. Aazami, M.A. Javaloyes, Penrose's singularity theorem in a Finsler spacetime, Class. Quantum Gravity 33 (2) (2016) 025003.
[2] D. Bao, S.-S. Chern, Z. Shen, An Introduction to Riemann-Finsler Geometry, Graduate Texts in Mathematics, vol. 200, Springer-Verlag, New York, 2000.
[3] P. Bernard, B. Buffoni, Weak KAM pairs and Monge-Kantorovich duality, in: Asymptotic Analysis and Singularities, Adv. Stud. Pure Math. 47 (2) (2007) 397–420.
[4] P. Bernard, B. Buffoni, Optimal mass transportation and Mather theory, J. Eur. Math. Soc. (JEMS) 9 (1) (2007) 85–121.
[5] P. Bernard, S. Suhr, Lyapounov functions of closed cone fields: from Conley theory to time functions, Commun. Math. Phys. 359 (2) (2018) 467–498.
[6] J. Bertrand, M. Puel, The optimal transport problem for relativistic costs, Calc. Var. Partial Differ. Equ. 46 (1–2) (2013) 353–374.
[7] J. Bertrand, A. Pratelli, M. Puel, Kantorovich potentials and continuity of total cost for relativistic cost functions, J. Math. Pures Appl. (9) 110 (2018) 93–122.
[8] Y. Brenier, Extended Monge-Kantorovich theory, in: Optimal Transportation and Applications, Martina Franca, 2001, in: Lecture Notes in Math., vol. 1813, Springer-Verlag, Berlin, 2003, pp. 91–121.
[9] S. Bianchini, F. Cavalletti, The Monge problem for distance cost in geodesic spaces, Commun. Math. Phys. 318 (2013) 615–673.
[10] L. Caravenna, A proof of Sudakov theorem with strictly convex norms, Math. Z. 286 (1–2) (2011) 371–407.
[11] F. Cavalletti, M. Huesmann, Existence and uniqueness of optimal transport maps, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 32 (6) (2015) 1367–1377.
[12] F. Cavalletti, A. Mondino, Optimal maps in essentially non-branching spaces, Commun. Contemp. Math. 19 (6) (2017) 1–27.
[13] Th. Champion, L. De Pascale, The Monge problem for strictly convex norms in $\mathbb{R}^n$, J. Eur. Math. Soc. 12 (2010) 1355–1369.

[14] L. De Pascale, S. Rigot, Monge's transport problem in the Heisenberg group, Adv. Calc. Var. 4 (2) (2011) 195–227.

[15] M. Eckstein, T. Miller, Causal evolution of wave packets, Phys. Rev. A 95 (2017) 032106.

[16] D.H. Fremlin, Measure Theory, vol. 4, Torres Fremlin, Colchester, 2006.

[17] N. Gigli, Optimal maps in non-branching spaces with Ricci curvature bounded from below, Geom. Funct. Anal. 22 (4) (2012) 990–999.

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

[19] J. Louet, A. Pratelli, F. Zeisler, On the continuity of the total cost in the mass transport problem with relativistic cost functions, arXiv:1612.06229.

[20] R. McCann, A convexity principle for interacting gases, Adv. Math. 128 (1997) 153–179.

[21] R. McCann, Displacement concavity of Boltzmann's entropy characterizes positive energy in general relativity, arXiv:1808.01536.

[22] R. McCann, M. Puel, Constructing a relativistic heat flow by transport time steps, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 26 (6) (2009) 2539–2580.

[23] T. Miller, Polish spaces of causal curves, J. Geom. Phys. 116 (2017) 295–315.

[24] E. Minguzzi, An equivalence of Finslerian relativistic theories, Rep. Math. Phys. 77 (1) (2016) 45–55.

[25] K.-Th. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (1) (2006) 65–131.

[26] V.N. Sudakov, Geometric problems in the theory of infinite-dimensional probability distributions, in: Proceedings of Steklov Inst. Math., vol. 2, 1979, pp. 1–178, Number in Russian Series Statements, vol. 141, 1976.

[27] S. Suhr, Optimal transportation for Lorentzian cost functions, Münster J. Math. 11 (2018) 13–47.

[28] C. Villani, Optimal Transport, Old and New, Springer-Verlag, Berlin, Heidelberg, 2009.