JEMS

Andrea Mondino · Stefan Suhr

An optimal transport formulation of the Einstein equations of general relativity

Abstract. The goal of the paper is to give an optimal transport formulation of the full Einstein equations of general relativity, linking the (Ricci) curvature of a space-time with the cosmological constant and the energy-momentum tensor. Such an optimal transport formulation is in terms of convexity/concavity properties of the Boltzmann–Shannon entropy along curves of probability measures extremizing suitable optimal transport costs. The result gives a new connection between general relativity and optimal transport; moreover, it gives a mathematical reinforcement of the strong link between general relativity and thermodynamics/information theory that emerged in the physics literature of the last years.

1. Introduction

In recent years, optimal transport revealed to be a very effective and innovative tool in several fields of mathematics and applications. By way of example, let us mention fluid mechanics (e.g. Brenier [16] and Benamou–Brenier [10]), partial differential equations (e.g. Jordan–Kinderlehrer–Otto [48] and Otto [67]), random matrices (e.g. Figalli–Guionnet [32]), optimization (e.g. Bouchitté–Buttazzo [14]), non-linear σ-models (e.g. Carfora [21]), geometric and functional inequalities (e.g. Cordero-Erausquin–Nazaret–Villani [27], Figalli–Maggi–Pratelli [33], Klartag [53], Cavalletti–Mondino [23]) Ricci curvature in Riemannian geometry (e.g. Otto–Villani [68], Cordero-Erausquin–McCann–Schmuckenschläger [25], Sturm–von Renesse [81]) and in metric measure spaces (e.g.

Andrea Mondino (corresponding author): Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford OX2 6GG, United Kingdom; andrea.mondino@maths.ox.ac.uk

Stefan Suhr: Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstrasse 150, 44780 Bochum, Germany; stefan.suhr@ruhr-uni-bochum.de

Lott–Villani [56], Sturm [73, 74], Ambrosio–Gigli–Savaré [5]). For more details about optimal transport and its applications in both pure and applied mathematics, we refer the reader to the many books on the topic, e.g. [1, 3, 72, 79, 80].

Here let us just quote two of the many applications to partial differential equations. In the pioneering work of Jordan–Kinderlehrer–Otto [48] a new optimal transport formulation of the Fokker–Planck equation (and in particular of the heat equation) was discovered as the gradient flow of a suitable functional (roughly, the Boltzmann–Shannon entropy defined below in (1.5) plus a potential) in the Wasserstein space (i.e. the space of probability measures with finite second moments endowed with the quadratic Kantorovich–Wasserstein distance); later, Otto [67] found a related optimal transport formulation of the porous medium equation. The impact of these works in the optimal transport community has been huge, and opened the way to a more general theory of gradient flows (see for instance the monograph by Ambrosio–Gigli–Savaré [3]).

The goal of the present work is to give a new optimal transport formulation of another fundamental class of partial differential equations: the Einstein equations of general relativity. First published by Einstein in 1915, the Einstein equations describe gravitation as a result of space-time being curved by mass and energy; more precisely, the space-time (Ricci) curvature is related to the local energy and momentum expressed by the energy-momentum tensor. Before entering into the topic, let us first recall that the Einstein equations are *hyperbolic* evolution equations (for a comprehensive treatment see the recent monograph by Klainerman–Nicolò [52]). Instead of a gradient flow/PDE approach, we will see the evolution from a geometric/thermodynamic/information point of view.

Next we briefly recall the formulation of the Einstein equations. Let M^n be an n-dimensional manifold ($n \geq 3$, the physical dimension being $n = 4$) endowed with a Lorentzian metric g, i.e. g is a non-degenerate symmetric bilinear form of signature $(-, +, \ldots, +)$. Denote by Ric and Scal the Ricci and the scalar curvatures of (M^n, g). The Einstein equations read

$$\text{Ric} - \tfrac{1}{2} \text{Scal} \, g + \Lambda g = 8\pi T, \tag{1.1}$$

where $\Lambda \in \mathbb{R}$ is the cosmological constant, and T is the energy-momentum tensor. Physically, the cosmological constant Λ corresponds to the energy density of the vacuum; the energy-momentum tensor is a symmetric bilinear form on M representing the density of energy and momentum, acting as the source of the gravitational field.

1.1. Statement of the main results

Recall that in a Lorentzian manifold (M^n, g), a non-zero tangent vector $v \in T_x M$ is called *time-like* if $g(v, v) < 0$. If M admits a continuous nowhere vanishing time-like vector field X, then (M, g) is said to be *time-oriented* and it is called a *space-time*. The vector field X induces a partition of the set of time-like vectors into two classes: the *future pointing* tangent vectors v for which $g(X, v) < 0$ and the *past pointing* tangent vectors v for which $g(X, v) > 0$. The closure of the set of future pointing time-like vectors

is denoted

$$\mathcal{C} = \text{Cl}(\{v \in TM \mid g(v, v) < 0 \text{ and } g(X, v) < 0\}) \subset TM.$$

A physical particle moving in the space-time $(M, g, \mathcal{C})$ is represented by a *causal curve* which is an absolutely continuous curve, γ, satisfying

$$\dot{\gamma}_t \in \mathcal{C} \quad \text{a.e. } t \in [0, 1].$$

If the particle cannot reach the speed of light (e.g. a massive particle), then it is represented by a *chronological* curve which is an absolutely continuous curve, γ, satisfying

$$\dot{\gamma}_t \in \text{Int}(\mathcal{C}) \quad \text{a.e. } t \in [0, 1],$$

where $\text{Int}(\mathcal{C})$ is the interior of the cone $\mathcal{C}$ made up of future pointing time-like vectors. The *Lorentz length* of a causal curve is

$$L_g(\gamma) = \int_0^1 \sqrt{-g(\dot{\gamma}_t, \dot{\gamma}_t)} \, dt.$$

A point y is in the future of x if there is a future oriented chronological curve from x to y; in this case, the *Lorentz distance* or *proper time* between x and y is defined by

$$\sup \{L_g(\gamma) \mid \gamma_0 = x \text{ and } \gamma_1 = y, \, \gamma \text{ chronological}\} > 0,$$

which is achieved by a geodesic which is called a *maximal geodesic*. See for example [43, 66, 82].

In this paper, we consider the following Lorentzian Lagrangian on TM for $p \in (0, 1)$:

$$\mathcal{L}_p(v) := \begin{cases} -\frac{1}{p}(-g(v, v))^{p/2} & \text{if } v \in \mathcal{C}, \\ +\infty & \text{otherwise.} \end{cases} \tag{1.2}$$

Note that if p were 1 this would be the negative of the integrand for the Lorentz length given above. Here we study $p \in (0, 1)$ because the Lorentzian Lagrangian $\mathcal{L}_p$ has good convexity properties for such p (see Lemma 2.1).

Let $AC([0, 1], M)$ denote the space of absolutely continuous curves from $[0, 1]$ to M. The Lagrangian action $\mathcal{A}_p$, corresponding to the Lagrangian $\mathcal{L}_p$ and defined for any $\gamma \in AC([0, 1], M)$, is given by

$$\mathcal{A}_p(\gamma) := \int_0^1 \mathcal{L}_p(\dot{\gamma}_t) \, dt \in (-\infty, 0] \cup \{+\infty\}. \tag{1.3}$$

Observe that $\mathcal{A}_p(\gamma) \in (-\infty, 0]$ if and only if γ is a causal curve. Note that if p were 1, this would be the negative of the Lorentz length of γ or the proper time along γ. Thus $-\mathcal{A}_p(\dot{\gamma})$ can be seen as a kind of non-linear p-proper time along γ, enjoying better convexity properties. The reader may note the parallel with the theory of Riemannian geodesics, where one often studies the energy functional $\int |\dot{\gamma}|^2$ in place of the length functional $\int |\dot{\gamma}|$, due to the analogous advantages.

The choice of the minus sign in (1.2) is motivated by optimal transport theory, in order to have a minimization problem instead of a maximization one (as in the sup defining the Lorentz distance between points above). It is readily checked that the critical points of $\mathcal{A}_p$ with negative action are time-like geodesics (see Lemma 2.2). The advantage of $\mathcal{A}_p$ with $p \in (0, 1)$ is that it automatically selects an affine parametrization for its critical points with negative action.

The *cost function*, $c_p : M \times M \to (-\infty, 0] \cup \{+\infty\}$, relative to the p-action $\mathcal{A}_p$ is defined by

$$c_p(x, y) = \inf \{\mathcal{A}_p(\gamma) \mid \gamma \in AC([0, 1], M), \, \gamma_0 = x, \, \gamma_1 = y\}.$$

Note that if p were 1, then this would be the negative of the Lorentz distance between x and y.

Consider a relatively compact open subset

$$E \subset\subset \text{Int}(\mathcal{C}) \subset TM \quad \text{and} \quad r \in (0, \text{inj}_g(E)),$$

where $\text{inj}_g(E) > 0$ is the injectivity radius of the exponential map of g restricted to E. If we take $p_{TM \to M} : TM \to M$ to be the canonical projection map then for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -r^2$, we have a maximal geodesic $\gamma_x : [0, 1] \to M$ defined by

$$\gamma_x(t) = \exp_x((t - 1/2)v) \quad \text{such that} \quad \gamma_x(1/2) = x.$$

The Ricci curvature $\text{Ric}_x(v, v)$ at a point $x \in M$ in the direction v is the trace of the curvature tensor so that intuitively it measures the average way in which geodesics near γ_x bend towards or away from it. See Section 1.3. In Riemannian geometry, the Ricci curvature influences the volumes of balls. Here, instead of balls we define, for any $x \in p_{TM \to M}(E)$,

$$B_r^{g,E}(x) := \{\exp_x^g(tw) \mid w \in T_x M \cap E, \, g(w, w) = -1, \, t \in [0, r]\}$$

precisely to avoid the null directions.

Rather than considering individual paths between a given pair of points, we will consider distributions of paths between a given pair of distributions of points using the optimal transport approach.

We denote by $\mathcal{P}(M)$ the set of Borel probability measures on M. For any $\mu_1, \mu_2 \in \mathcal{P}(M)$, we say that a Borel probability measure $\pi \in \mathcal{P}(M \times M)$ is a *coupling* of μ_1 and μ_2 if

$$(p_i)_\sharp \pi = \mu_i, \quad i = 1, 2,$$

where $p_1, p_2 : M \times M \to M$ are the projections onto the first and second coordinate. Recall that the push-forward $(p_1)_\sharp \pi$ is defined by

$$(p_1)_\sharp \pi(A) := \pi(p_1^{-1}(A))$$

for any Borel subset $A \subset M$. The set of couplings of μ_1, μ_2 is denoted by $\mathrm{Cpl}(\mu_1, \mu_2)$. The c_p-*cost* of a coupling π is given by

$$\int_{M \times M} c_p(x, y) \, d\pi(x, y) \in [-\infty, 0] \cup \{+\infty\}.$$

Denote by $C_p(\mu_1, \mu_2)$ the *minimal cost relative to c_p* among all couplings from μ_1 to μ_2, i.e.

$$C_p(\mu_1, \mu_2) := \inf \left\{ \int c_p \, d\pi \,\middle|\, \pi \in \mathrm{Cpl}(\mu_1, \mu_2) \right\} \in [-\infty, 0] \cup \{+\infty\}.$$

If $C_p(\mu_1, \mu_2) \in \mathbb{R}$, a coupling achieving the infimum is said to be c_p-*optimal*.

For $t \in [0, 1]$ denote by $\mathrm{e}_t : \mathrm{AC}([0, 1], M) \to M$ the evaluation map

$$\mathrm{e}_t(\gamma) := \gamma_t.$$

A c_p-*optimal dynamical plan* is a probability measure Π on $\mathrm{AC}([0, 1], M)$ such that $(\mathrm{e}_0, \mathrm{e}_1)_\sharp \Pi$ is a c_p-optimal coupling from $\mu_0 := (\mathrm{e}_0)_\sharp \Pi$ to $\mu_1 := (\mathrm{e}_1)_\sharp \Pi$. One can naturally associate to Π a curve

$$(\mu_t := (\mathrm{e}_t)_\sharp \Pi)_{t \in [0, 1]} \subset \mathcal{P}(M)$$

of probability measures. The condition that Π is a c_p-optimal dynamical plan corresponds to saying that the curve $(\mu_t)_{t \in [0, 1]} \subset \mathcal{P}(M)$ is a length minimizing geodesic with respect to C_p, i.e.

$$C_p(\mu_s, \mu_t) = |t - s| C_p(\mu_0, \mu_1) \quad \forall s, t \in [0, 1].$$

We will mainly consider a special class of c_p-optimal dynamical plans, which we call *regular*: roughly, a c_p-optimal dynamical plan is said to be regular if it is obtained by exponentiating the gradient (which is assumed to be time-like) of a smooth Kantorovich potential ϕ:

$$\mu_t = (\Psi_{1/2}^t)_\sharp \mu_{1/2}, \quad \Psi_{1/2}^t(x) := \exp_x^g\big(-(t - 1/2)|\nabla_g \phi|_g^{q-2} \nabla_g \phi(x)\big), \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{1.4}$$

and moreover $\mu_t \ll \mathrm{vol}_g$ for all $t \in (0, 1)$, where vol_g denotes the standard volume measure of (M, g). For the precise notions, the reader is referred to Section 2.5.

A key role in our optimal transport formulation of the Einstein equations will be played by the (relative) Boltzmann–Shannon entropy. Denote by vol_g the standard volume measure on (M, g). Given an absolutely continuous probability measure $\mu = \varrho \, \mathrm{vol}_g$ with density $\varrho \in C_c(M)$, its *Boltzmann–Shannon entropy* (relative to vol_g) is defined as

$$\mathrm{Ent}(\mu | \mathrm{vol}_g) := \int_M \varrho \log \varrho \, d\mathrm{vol}_g. \tag{1.5}$$

In Theorem 4.10 we will prove that the Einstein equations are equivalent to an equation involving (at least at a formal level) the second order derivative of this entropy along a c_p-optimal dynamical plan. See Figure 1.
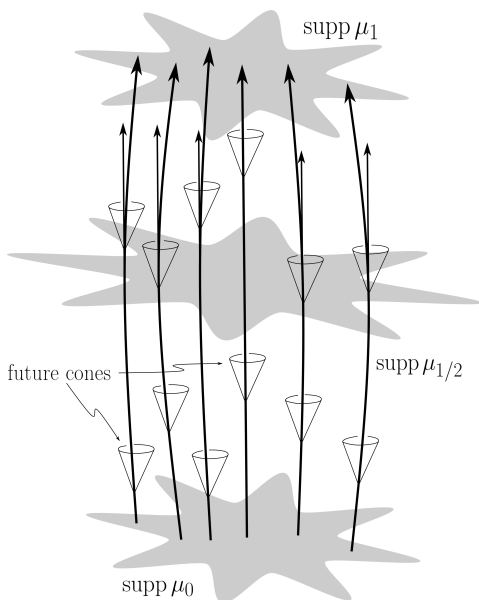
Fig. 1. The transport in Theorem 1.1.

Throughout the paper we will assume the cosmological constant Λ and the energy momentum tensor T to be given, say from physics and/or mathematical general relativity. Given g, Λ and T it is convenient to set

$$\tilde{T} := \frac{2\Lambda}{n-2} g + 8\pi T - \frac{8\pi}{n-2} \mathrm{Tr}_g(T)g \tag{1.6}$$

so that the Einstein equations can be written as $\mathrm{Ric} = \tilde{T}$ (see Lemma 4.1).

Theorem 1.1 (Theorem 4.10). *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 3$. Then the following assertions are equivalent:*

(1) *$(M, g, \mathcal{C})$ satisfies the Einstein equations (1.1), which can be rewritten in terms of $\tilde{T}$ of (1.6) as $\mathrm{Ric} = \tilde{T}$.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \mathrm{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function*

$$\epsilon = \epsilon_E : (0, \infty) \to (0, \infty) \quad with \quad \lim_{r \downarrow 0} \epsilon(r) = 0$$

such that for all

$$x \in p_{TM \to M}(E) \quad and \quad v \in T_x M \cap E \text{ with } g(v, v) = -R^2$$

the following assertion holds.

For every $r \in (0, R)$, setting $y = \exp_x^g(rv)$, there exists a regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ with associated curve of probability measures

$$(\mu_t := (e_t)_\sharp \Pi)_{t \in [0,1]} \subset \mathcal{P}(M)$$

such that

- $\mu_{1/2} = \text{vol}_g(B_{r^4}^{g,E}(x))^{-1} \text{vol}_g \llcorner B_{r^4}^{g,E}(x)$,
- $\text{supp}\, \mu_1 \subset \{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$

and which has convex/concave entropy in the following sense:

$$\left| \frac{4}{r^2} [\text{Ent}(\mu_1|\text{vol}_g) - 2\,\text{Ent}(\mu_{1/2}|\text{vol}_g) + \text{Ent}(\mu_0|\text{vol}_g)] - \tilde{T}(v, v) \right| \leq \epsilon(r). \quad (1.7)$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in (2) holds true.*

Remark 1.2 (On the regularity of space-time). For simplicity, in the paper we work with a smooth space-time. However, all the statements and proofs would be valid assuming that M is a differentiable manifold endowed with a C^3-atlas and that g is a C^2-Lorentzian metric on M.

Remark 1.3 (A heuristic thermodynamic interpretation of Theorem 1.1). A curve $(\mu_t)_{t \in [0,1]} \subset \mathcal{P}(M)$ associated to a c_p-optimal dynamical plan can be interpreted as the evolution[1] of a distribution of gas passing through a given gas distribution $\mu_{1/2}$ (which in Theorem 1.1 is assumed to be concentrated in the space-time near x). Theorem 1.1 says that the Einstein equations can be equivalently formulated in terms of the convexity properties of the Boltzmann–Shannon entropy along such evolutions $(\mu_t)_{t \in [0,1]} \subset \mathcal{P}(M)$. Extrapolating a bit more, we can say that the second law of thermodynamics (i.e. in a natural thermodynamic process, the sum of the entropies of the interacting thermodynamic systems decreases, due to our sign convention) concerns the *first* derivative of the Boltzmann–Shannon entropy; gravitation (under the form of Ricci curvature) is instead related to the *second* order derivative of the Boltzmann–Shannon entropy along a natural thermodynamic process.

Remark 1.4 (Disclaimer). In Theorem 1.1 we are *not claiming to solve* the general Einstein equations via optimal transport; we are instead *proposing a novel formulation/ characterization of the solutions of the Einstein equations based on optimal transport*, assuming the cosmological constant Λ and the energy-momentum tensor T are already given (this can be a bit controversial for a general T; however, the characterization is already new and interesting in the vacuum case $T \equiv 0$ where there is no controversy). The aim is indeed to bridge optimal transport and general relativity, with the goal of stimulating fruitful connections between these two fascinating fields. In particular, optimal transport tools have been very successful in studying Ricci curvature bounds in a (low

---

[1] Strictly speaking, t is not the proper time, but only a variable parametrizing the evolution.

regularity) *Riemannian and metric-measure* framework (see later in the introduction for the related literature) and it is thus natural to expect that optimal transport can also be useful in a low-regularity Lorentzian framework, where singularities play an important part in the theory; for example it is expected that, at least generically, singularities occur in black-hole interiors.

For equivalent formulations of Theorem 1.1, see Remarks 4.11 and 4.12.

In the vacuum case $T \equiv 0$ with zero cosmological constant $\Lambda = 0$, the Einstein equations read

$$\text{Ric} \equiv 0 \tag{1.8}$$

for an n-dimensional space-time $(M, g, \mathcal{C})$. Specializing Theorem 1.1 to the choice $\tilde{T} = 0$ (plus a small extra observation to sharpen the lower bound in (1.9) from $-\epsilon(r)$ to 0; moreover, the same proof extends to $n = 2$) gives the following optimal transport formulation of Einstein vacuum equations with zero cosmological constant.

Corollary 1.5. *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$. Then the following assertions are equivalent:*

(1) *$(M, g, \mathcal{C})$ satisfies the Einstein vacuum equations with zero cosmological constant, i.e. $\text{Ric} \equiv 0$.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r)/r^2 = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.*

    For every $r \in (0, R)$, setting $y = \exp_x^g(rv)$, there exists a regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ with associated curve of probability measures

$$(\mu_t := (e_t)_\sharp \Pi)_{t \in [0,1]} \subset \mathcal{P}(M)$$

such that

- *$\mu_{1/2} = \text{vol}_g(B_{r^4}^{g,E}(x))^{-1} \text{vol}_g \llcorner B_{r^4}^{g,E}(x)$,*
- *$\text{supp}\,\mu_1 \subset \{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$*

and which has almost affine entropy in the sense that

$$0 \leq \text{Ent}(\mu_1|\text{vol}_g) - 2\,\text{Ent}(\mu_{1/2}|\text{vol}_g) + \text{Ent}(\mu_0|\text{vol}_g) \leq \epsilon(r). \tag{1.9}$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in (2) holds true.*

1.2. Outline of the argument

As already mentioned, the Einstein equations can be written as $\text{Ric} = \tilde{T}$ where $\tilde{T}$ was defined in (1.6) (see Lemma 4.1). The optimal transport formulation of the Einstein equations will consist separately of an optimal transport characterization of the two inequalities

$$\text{Ric} \geq \tilde{T} \tag{1.10}$$

and

$$\text{Ric} \leq \tilde{T}. \tag{1.11}$$

The optimal transport characterization of the lower bound (1.10) will be achieved in Theorem 4.3 and consists in showing that (1.10) is equivalent to a *convexity* property of the Boltzmann–Shannon entropy along *every* regular c_p-optimal dynamical plan. The optimal transport characterization of the upper bound (1.11) will be achieved in Theorem 4.7 and consists in showing that (1.11) is equivalent to the *existence* of a large family of regular c_p-optimal dynamical plans (roughly the ones given by exponentiating the gradient of a smooth Kantorovich potential with Hessian vanishing at a given point) along which the Boltzmann–Shannon entropy satisfies the corresponding *concavity* condition.

Important ingredients in the proofs will be the following. In Theorem 4.3, to prove that Ricci lower bounds imply convexity properties of the entropy, we will perform Jacobi fields computations relating the Ricci curvature to the Jacobian of the change of coordinates of the optimal transport map (see Proposition 3.4); in order to establish the converse implication we will argue by contradiction via constructing c_p-optimal dynamical plans very localized in space-time (Lemma 3.2).

In Theorem 4.3 we will consider the special class of regular c_p-optimal dynamical plans constructed in Lemma 3.2, roughly the ones given by exponentiating the gradient of a smooth Kantorovich potential with Hessian vanishing at a given point $x \in M$. To prove that Ricci upper bounds imply concavity properties of the entropy, we will need to establish the Hamilton–Jacobi equation satisfied by the evolved Kantorovich potentials (Proposition 3.1) and a non-linear Bochner formula involving the p-Box operator (Proposition A.1), the Lorentzian counterpart of the p-Laplacian. In order to show the converse we will argue by contradiction using Theorem 4.3.

1.3. An example: FLRW space-times

We illustrate Theorem 1.1 for the class of Friedmann–Lemaître–Robertson–Walker space-times (FLRW space-times for short), a group of cosmological models well known in general relativity. See [66, Chapter 12] for a discussion of the geometry in the case $n = 4$.

FLRW space-times are of the form

$$(M, g) = (I \times \Sigma, -ds^2 + a(s)^2 \sigma),$$

where $I \subset \mathbb{R}$ is an interval, $a : I \to (0, \infty)$ is smooth, and (Σ, σ) is a Riemannian manifold with constant sectional curvature $k \in \{-1, 0, 1\}$. The Ricci and scalar curvatures are given by

$$\text{Ric} = -(n-1)\frac{\ddot{a}}{a}ds^2 + \left[\frac{\ddot{a}}{a} + (n-2)\left(\frac{\dot{a}^2 + k}{a^2}\right)\right]a^2\sigma$$

and

$$\text{Scal} = 2(n-1)\frac{\ddot{a}}{a} + (n-1)(n-2)\frac{\dot{a}^2 + k}{a^2},$$

respectively, where $\dot{a} := \frac{da}{ds}$. The stress-energy tensor is thus (assuming $\Lambda = 0$)

$$8\pi T = \mathrm{Ric} - \tfrac{1}{2} \mathrm{Scal}\, g$$
$$= \frac{(n-1)(n-2)}{2} \frac{\dot{a}^2 + k}{a^2} ds^2 - \left[(n-2) \frac{\ddot{a}}{a} + \frac{(n-2)(n-3)}{2} \frac{\dot{a}^2 + k}{a^2} \right] a^2 \sigma.$$

The foliation

$$\mathfrak{D} := \{ s \mapsto (s, \underline{x}) \}_{\underline{x} \in \Sigma}$$

is a geodesic foliation by c_p-minimal geodesics. The orthogonal complement $\partial_s^\perp = T\Sigma$ with respect to g is integrable. Consider the projection

$$S : M = I \times \Sigma \to I$$

and for $r > 0$ the function

$$\phi : M \to \mathbb{R}, \quad x \mapsto r^{p-1} S(x).$$

It is easy to see that

$$\nabla_g^q \phi(x) := -|\nabla_g \phi|_g^{q-2} \nabla_g \phi(x) = r \partial_s.$$

For the c_p-transform we have (see Section 2.5)

$$\phi^{c_p}(s, \underline{y}) = \inf_{x \in M} [c_p(x, (s, \underline{y})) - \phi(x)]$$
$$= \inf_{s' < s} [c_p((s', \underline{y}), (s, \underline{y})) - \phi(s', \underline{y})]$$
$$= \inf_{s' < s} \left[-\frac{1}{p}(s - s')^p - r^{p-1} s' \right]$$
$$= \frac{p-1}{p} r^p - r^{p-1} s,$$

where the second equality follows from the fact that the geodesics in $\mathfrak{D}$ minimize c_p to the level sets of ϕ. It follows that

$$\phi^{c_p}(s, \underline{y}) + \phi(s - r, \underline{y}) = -\frac{1}{p} r^p = c_p((s - r, \underline{y}), (s, \underline{y})),$$

i.e. $\partial^{c_p} \phi(x) = \{\exp_x(\nabla_g^q \phi(x))\}$ for all $x \in M$ whenever the right hand side is well defined (see Section 2.5 for the definition).

It now follows by standard transportation theory (see for instance [1, Theorem 1.13]) that for a Borel probability measure $\mu_{1/2}$ on $I \times \Sigma$ the family $(\mu_t := (\Psi_{1/2}^t)_\sharp \mu_{1/2})_{t \in [0,1]}$, where

$$\Psi_{1/2}^t : I \times \Sigma \to I \times \Sigma, \quad (s, \underline{x}) \mapsto (s + r(t - 1/2), \underline{x}),$$

defines a c_p-optimal dynamical plan as long as it is defined in accordance with the notation in (1.4) and ϕ is a smooth Kantorovich potential for $(\mu_t)_{t \in [0,1]}$.

If $\mu_{1/2} \ll \mathrm{vol}_g$ with density $\rho_{1/2} \in C_c(M)$, we get (compare with the proof of Theorem 4.3)

$$
\begin{aligned}
\mathrm{Ent}(\mu_t|\mathrm{vol}_g) &= \int_M \log \rho_t(y)\, d\mu_t(y) = \int_M \log \rho_t(\Psi^t_{1/2}(x))\, d\mu_{1/2}(x) \\
&= \int_M \log[\rho_{1/2}(x)(\mathrm{Det}_g(D\Psi^t_{1/2})(x))^{-1}]\, d\mu_{1/2}(x) \\
&= \mathrm{Ent}(\mu_{1/2}|\mathrm{vol}_g) - \int_M \log[\mathrm{Det}_g(D\Psi^t_{1/2})(x)]\, d\mu_{1/2}(x).
\end{aligned}
$$

We have $\mathrm{Det}_g(D\Psi^t_{1/2})((s,\underline{x}))) = \frac{a(s+r(t-1/2))^{n-1}}{a(s)^{n-1}}$ and thus

$$
-\frac{d^2}{dt^2}\log[\mathrm{Det}_g(D\Psi^t_{1/2})(x))] = -(n-1)\frac{\ddot{a}a - \dot{a}^2}{a^2}r^2 = \mathrm{Ric}(r\partial_s, r\partial_s) + (n-1)\frac{\dot{a}^2}{a^2}r^2. \tag{1.12}
$$

Neglecting the term $\frac{\dot{a}^2}{a^2} \geq 0$ we conclude (compare with the proof of Proposition 3.4)

$$
\begin{aligned}
\frac{d^2}{dt^2}\mathrm{Ent}(\mu_t|\mathrm{vol}_g) &= -\frac{d^2}{dt^2}\int_M \log[\mathrm{Det}_g(D\Psi^t_{1/2})(x))]\, d\mu_{1/2}(x) \\
&\geq \int_M \mathrm{Ric}(r\partial_s, r\partial_s)_{\Psi^t_{1/2}(x)}\, d\mu_{1/2}(x),
\end{aligned} \tag{1.13}
$$

which implies

$$
\frac{4}{r^2}[\mathrm{Ent}(\mu_1|\mathrm{vol}_g) - 2\,\mathrm{Ent}(\mu_{1/2}|\mathrm{vol}_g) + \mathrm{Ent}(\mu_0|\mathrm{vol}_g)] \geq \int_M \mathrm{Ric}(\partial_s, \partial_s)\, d\mu_{1/2}
$$

for $r \to 0$, i.e. one side of (1.7).

Note that the gap in (1.13) is

$$
(n-1)r^2 \int_M \frac{\dot{a}^2}{a^2}\, d\mu_{1/2}.
$$

From this we see that the bound from above

$$
\frac{4}{r^2}[\mathrm{Ent}(\mu_1|\mathrm{vol}_g) - 2\,\mathrm{Ent}(\mu_{1/2}|\mathrm{vol}_g) + \mathrm{Ent}(\mu_0|\mathrm{vol}_g)]
$$

$$
\leq \int_M \mathrm{Ric}(\partial_s, \partial_s)\, d\mu_{1/2} + \epsilon(r) \tag{1.14}
$$

in (1.7) for the aforementioned transports holds for $r \to 0$ if $\mu_{1/2}$ is concentrated on $\{(s,\underline{x}) \mid \dot{a}(s) = 0\}$ or, more generally, if $\mu_{1/2} = \mu^r_{1/2}$ satisfies $\lim_{r\to 0}\int_M \frac{\dot{a}^2}{a^2}\, d\mu^r_{1/2} = 0$.

The Levi-Civita connection ∇ of g satisfies

$$
\nabla_X \partial_s = \nabla_{\partial_s} X = \frac{\dot{a}}{a}X
$$

for all vector fields X tangent to Σ. Thus the Hessian of ϕ is given by

$$
\mathrm{Hess}_\phi = r^{p-1}\mathrm{Hess}_S = r^{p-1}\nabla.\nabla_g S = -r^{p-1}\nabla.\partial_s = -r^{p-1}\frac{\dot{a}}{a}(\mathrm{Id} - \partial_s \otimes ds). \tag{1.15}
$$

It vanishes at $(s, \underline{x})$ if and only if $\dot{a}(s) = 0$. Thus the inequality (1.14) follows for these transports if the Hessian of ϕ vanishes on supp $\mu_{1/2}$ or, more generally, if

$$\lim_{r \to 0} \int_M \|r^{1-p} \mathrm{Hess}_\phi\|^2 \, d\mu^r_{1/2} = 0.$$

Finally, we illustrate the theory laid out in Appendix B via warping functions $a : I \to (0, \infty)$ of regularity below C^2. Compare with [39] for related results on metrics of low regularity.

Consider a probability measure $\nu_{1/2} \ll \mathrm{vol}^\sigma$ on Σ and set

$$\mu_{1/2} := \frac{1}{s_1 - s_0} \mathcal{L}^1|_{[s_0, s_1]} \otimes \nu_{1/2}$$

for $s_0, s_1 \in I$ with $s_0 < s_1$ and $\mathcal{L}^1$ the 1-dimensional Lebesgue measure. For $\mu_t := (\Psi^t_{1/2})_\# \mu_{1/2}$ we have

$$\frac{4}{r^2} [\mathrm{Ent}(\mu_1 | \mathrm{vol}_g) - 2 \mathrm{Ent}(\mu_{1/2} | \mathrm{vol}_g) + \mathrm{Ent}(\mu_0 | \mathrm{vol}_g)]$$
$$= -\frac{4}{r^2(s_1 - s_0)} \int_{s_0}^{s_1} \log \left(\frac{a^{n-1}(s + r/2) a^{n-1}(s - r/2)}{a^{2n-2}(s)} \right) ds.$$

Assuming $s_1 - s_0 \ll r$ and $r > 0$ sufficiently small (i.e. the setting of Theorem 1.1), since a is continuous, we obtain

$$\frac{4}{r^2} [\mathrm{Ent}(\mu_1 | \mathrm{vol}_g) - 2 \mathrm{Ent}(\mu_{1/2} | \mathrm{vol}_g) + \mathrm{Ent}(\mu_0 | \mathrm{vol}_g)]$$
$$= -\frac{4}{r^2} \log \left(\frac{a^{n-1}(s + r/2) a^{n-1}(s - r/2)}{a^{2n-2}(s)} \right) + \frac{\varepsilon(r)}{r^2}$$

for some function $\varepsilon(r) \to 0$ as $r \to 0$. As an example, we discuss the case of the following $C^{1,1}$-warping function:

$$a : I \to \mathbb{R}, \quad a(s) := \begin{cases} \lambda_+ s^2 + 1, & s \geq 0, \\ \lambda_- s^2 + 1, & s < 0, \end{cases}$$

for $\lambda_-, \lambda_+ \in \mathbb{R}$.

Ignoring terms of higher order near $s = 0$ we get

$$\frac{4}{r^2} [\mathrm{Ent}(\mu_1 | \mathrm{vol}_g) - 2 \mathrm{Ent}(\mu_{1/2} | \mathrm{vol}_g) + \mathrm{Ent}(\mu_0 | \mathrm{vol}_g)]$$
$$= \begin{cases} -(n-1)[\lambda_+ (2s/r + 1)^2 + \lambda_- (2s/r - 1)^2 - 2\lambda_+ (2s/r)^2], & s \geq 0, \\ -(n-1)[\lambda_+ (2s/r + 1)^2 + \lambda_- (2s/r - 1)^2 - 2\lambda_- (2s/r)^2], & s \leq 0, \end{cases}$$

for $r \to 0$. It is now easy to see that the possible accumulation points of the right hand side for $r \to 0$ and $s \in [-r/2, r/2]$ lie between $-2(n-1)\lambda_+ = \lim_{s \downarrow 0} \mathrm{Ric}(\partial_s, \partial_s)$ and

$-2(n-1)\lambda_- = \lim_{s\uparrow 0} \text{Ric}(\partial_s, \partial_s)$. Furthermore, every value in that interval is an accumulation point for the right hand side, as $r \to 0$. In case the warping function lies in C^2, we get from (1.12) the asymptotic formula

$$\frac{4}{r^2}[\text{Ent}(\mu_1|\text{vol}_g) - 2\text{Ent}(\mu_{1/2}|\text{vol}_g) + \text{Ent}(\mu_0|\text{vol}_g)] - (n-1)\int_M \frac{\dot{a}^2}{a^2}\,d\mu_{1/2} \quad (1.16)$$

for $\text{Ric}(\partial_s, \partial_s)$ when $r \to 0$. The second integral is well defined for $a \in W^{1,2}(I)$. Therefore we can use (1.16) as a definition for $\text{Ric}(\partial_s, \partial_s)$ in this case. This is indeed the spirit of the synthetic bounds on the time-like Ricci curvature given in Definitions B.5 and B.8.

1.4. Related literature

1.4.1. Ricci curvature via optimal transport in Riemannian setting. In the Riemannian framework, a line of research pioneered by McCann [60], Cordero-Erausquin–McCann–Schmuckenschläger [25, 26], Otto–Villani [68] and von Renesse–Sturm [81] has culminated in a characterization of Ricci curvature lower bounds (by a constant $K \in \mathbb{R}$) involving only the displacement convexity of certain information-theoretic entropies. This in turn led Sturm [73, 74] and independently Lott–Villani [56] to develop a theory for Ricci curvature lower bounds in a non-smooth metric-measure space setting. The theory of such spaces has seen a very fast development in the last years; see e.g. [2, 4–7, 18, 22, 23, 30, 37, 38, 64]. An approach to the complementary upper bounds on the Ricci tensor (again by a constant $K' \in \mathbb{R}$) has been recently proposed by Naber [65] (see also Haslhofer–Naber [41]) in terms of functional inequalities on path spaces and martingales, and by Sturm [75] (see also Erbar–Sturm [31]) in terms of contraction/expansion rate estimates of the heat flow and in terms of displacement *concavity* of the Boltzmann–Shannon entropy. The Lorentzian time-like Ricci upper bounds of this paper have been inspired in particular by the work of Sturm [75].

1.4.2. Optimal transport in Lorentzian setting. The optimal transport problem in Lorentzian geometry was first proposed by Brenier [15] and further investigated in [12, 50, 76]. An intriguing physical motivation for studying the optimal transport problem in Lorentzian setting is called the "early universe reconstruction problem" [17, 36]. The Lorentzian cost C_p, for $p \in (0, 1)$, was proposed by Eckstein–Miller [28] and thoroughly studied by McCann [61] very recently. In the same paper [61], McCann gave an optimal transport formulation of the strong energy condition $\text{Ric} \geq 0$ of Penrose–Hawking [42, 44, 70] in terms of displacement convexity of the Boltzmann–Shannon entropy under the assumption that the space-time is globally hyperbolic.

We learned of the work of McCann [61] when we were already in the final stages of writing the present paper. Though both papers (inspired by the aforementioned Riemannian setting) are based on the idea of analyzing convexity properties of entropy functionals on the space of probability measures endowed with the cost C_p, $p \in (0, 1)$, the two approaches are largely independent: while McCann develops a general theory of optimal transportation in globally hyperbolic space-times focusing on the strong energy condition

Ric ≥ 0, in this paper we decided to take the quickest path in order to reach our goal of giving an optimal transport formulation of the full Einstein equations. Compared to [61], in the present paper we remove the assumption of global hyperbolicity on the space-time, we extend the optimal transport formulation to any lower bound of the type Ric $\geq \tilde{T}$ for any symmetric bilinear form $\tilde{T}$, and we also characterize general upper bounds Ric $\leq \tilde{T}$.

1.4.3. Physics literature. The existence of strong connections between thermodynamics and general relativity is not new in the physics literature; it has its origins at least in the work of Bekenstein [9] and Hawking with collaborators [8] in the mid-1970s about the black hole thermodynamics. These works inspired a new research field in theoretical physics, called entropic gravity (also known as emergent gravity), asserting that gravity is an entropic force rather than a fundamental interaction. Let us give a brief account. In 1995 Jacobson [47] derived the Einstein equations from the proportionality of entropy and horizon area of a black hole, exploiting the fundamental relation $\delta Q = T \delta S$ linking heat Q, temperature T and entropy S. Subsequently, other physicists, most notably Padmanabhan (see for instance the recent survey [69]), have been exploring links between gravity and entropy.

More recently, in 2011 Verlinde [78] proposed a heuristic argument suggesting that (Newtonian) gravity can be identified with an entropic force caused by changes in the information associated with the positions of material bodies. A relativistic generalization of those arguments leads to the Einstein equations.

The optimal transport formulation of Einstein equations obtained in the present paper involving the Boltzmann–Shannon entropy can be seen as an additional strong connection between general relativity and thermodynamics/information theory. It would be interesting to explore this relationship further.

2. Preliminaries

2.1. Some basics of Lorentzian geometry

Let M be a smooth manifold of dimension $n \geq 2$. It is convenient to fix a complete Riemannian metric h on M. The norm $|\cdot|$ on $T_x M$ and the distance $\text{dist}(\cdot, \cdot)$: $M \times M \to \mathbb{R}^+$ are understood to be induced by h, unless otherwise specified. Recall that h induces a Riemannian metric on TM. Distances on TM are understood to be induced by that metric. The metric ball around $x \in M$ with radius r, with respect to h, is denoted by $B_r^h(x)$ or simply by $B_r(x)$.

A *Lorentzian metric g* on M is a smooth $(0, 2)$-tensor field such that

$$g|_x : T_x M \times T_x M \to \mathbb{R}$$

is symmetric and non-degenerate with signature $(-, +, \ldots, +)$ for all $x \in M$. It is well known that, if M is compact, the vanishing of the Euler characteristic of M is equivalent to the existence of a Lorentzian metric; on the other hand, any *non-compact* manifold admits a Lorentzian metric. A non-zero tangent vector $v \in T_x M$ is called

- *time-like* if $g(v, v) < 0$,

- *light-like* (or *null*): if $g(v, v) = 0$ as well as $v \neq 0$,

- *space-like* if $g(v, v) > 0$ or $v = 0$.

A non-zero tangent vector $v \in T_x M$ which is either time-like or light-like, i.e. $g(v, v) \leq 0$ and $v \neq 0$, is called *causal* (or *non-space-like*). A Lorentzian manifold (M, g) is said to be *time-oriented* if M admits a continuous nowhere vanishing time-like vector field X. The vector field X induces a partition of the set of causal vectors into two classes:

- the *future pointing* tangent vectors v for which $g(X, v) < 0$,

- the *past pointing* tangent vectors v for which $g(X, v) > 0$.

The closure of the set of future pointing time-like vectors is denoted

$$\mathcal{C} = \mathrm{Cl}(\{v \in TM \mid g(v, v) < 0 \text{ and } g(X, v) < 0\}) \subset TM.$$

Note that the fiber $\mathcal{C}_x := \mathcal{C} \cap T_x M$ is a closed convex cone and its interior $\mathrm{Int}(\mathcal{C})$ is a connected component of $\{v \mid g(v, v) < 0\}$. A time-oriented Lorentzian manifold $(M, g, \mathcal{C})$ is called a *space-time*.

An absolutely continuous curve $\gamma : I \to M$ is called $(\mathcal{C})$-*causal* if $\dot{\gamma}_t \in \mathcal{C}$ for every differentiability point $t \in I$. A causal curve $\gamma : I \to M$ is called *time-like* if for every $s \in I$ there exist $\varepsilon, \delta > 0$ such that $\mathrm{dist}(\dot{\gamma}_t, \partial \mathcal{C}) \geq \varepsilon |\dot{\gamma}_t|$ for every $t \in I$ for which $\dot{\gamma}_t$ exists and $|s - t| < \delta$. In [11, Section 2.2] time-like curves are defined in terms of the Clarke differential of a Lipschitz curve. Whereas the definition via the Clarke differential is probably more satisfying from a conceptual point of view, the definition given here is easier to state. All relevant sets and curves used below are independent of the definition used (see [11, Lemma 2.11] and Proposition 2.4), though.

We denote by $J^+(x)$ (resp. $J^-(x)$) the set of points $y \in M$ such that there exists a causal curve with initial point x (resp. y) and final point y (resp. x), i.e. the causal future (resp. past) of x. The sets $I^{\pm}(x)$ are defined analogously by replacing causal curves by time-like ones. The sets $I^{\pm}(p)$ are always open in any space-time, but the sets $J^{\pm}(p)$ are in general neither closed nor open.

For a subset $A \subset M$, define $J^{\pm}(A) := \bigcup_{x \in A} J^{\pm}(x)$, and set

$$J^+ := \{(x, y) \in M \times M \mid y \in J^+(x)\}. \tag{2.1}$$

2.2. The Lagrangian $\mathcal{L}_p$, the action $\mathcal{A}_p$ and the cost c_p

On a space-time $(M, g, \mathcal{C})$ consider, for any $p \in (0, 1)$, the following Lagrangian on TM:

$$\mathcal{L}_p(v) := \begin{cases} -\frac{1}{p}(-g(v, v))^{p/2} & \text{if } v \in \mathcal{C}, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.2}$$

The following fact appears in [61, Lemma 3.1]. We provide a proof for the readers' convenience.

Lemma 2.1. *The function $\mathcal{L}_p$ is fiberwise convex, finite (and non-positive) on its domain and positive homogeneous of degree p. Moreover, $\mathcal{L}_p$ is smooth and fiberwise strictly convex on $\text{Int}(\mathcal{C})$.*

Proof. It is clear from its very definition that the restriction of $\mathcal{L}_p$ to $\text{Int}(\mathcal{C})$ is smooth. A direct computation gives

$$\frac{\partial \mathcal{L}_p}{\partial v^i} = (-g(v,v))^{\frac{p-2}{2}} g_{ik} v^k, \quad i = 1, \ldots, n, \tag{2.3}$$

$$\frac{\partial^2 \mathcal{L}_p}{\partial v^i \partial v^j} = (-g(v,v))^{\frac{p-4}{2}} \left(-g(v,v) g_{ij} + (2-p) g_{ik} v^k g_{jl} v^l \right), \quad i, j = 1, \ldots, n. \tag{2.4}$$

Fix $v \in \text{Int}(\mathcal{C})$. Decompose $w \in T_x M$ into $w^\parallel$, the part parallel to v, and $w^\perp$, the part orthogonal to v, all with respect to g. Then we have

$$D^2_{vv} \mathcal{L}_p(w, w) = (-g(v,v))^{\frac{p-4}{2}} \left(-g(w^\perp, w^\perp) g(v,v) - g(w^\parallel, w^\parallel) g(v,v) \right.$$
$$\left. + (2-p) g(v, w^\parallel)^2 \right)$$
$$= (-g(v,v))^{\frac{p-4}{2}} \left(-g(w^\perp, w^\perp) g(v,v) + (1-p) g(v, w^\parallel)^2 \right). \tag{2.5}$$

Since $g(w^\perp, w^\perp) \geq 0$ and $p < 1$ we have $D^2_{vv} \mathcal{L}_p(w, w) > 0$ for $w \neq 0$. ∎

We define the Lagrangian action $\mathcal{A}_p$ associated to $\mathcal{L}_p$ as follows:

$$\mathcal{A}_p(\gamma) := \int_0^1 \mathcal{L}_p(\dot{\gamma}_t) \, dt \in (-\infty, 0] \cup \{+\infty\}. \tag{2.6}$$

Note that if $\mathcal{A}_p(\gamma) \in \mathbb{R}$, then γ is causal. A causal curve $\gamma : [0, 1] \to M$ is an $\mathcal{A}_p$-*minimizer* between its endpoints $x, y \in M$ if

$$\mathcal{A}_p(\gamma) = \inf \{ \mathcal{A}_p(\eta) \mid \eta \in \text{AC}([0, 1], M), \eta_0 = x, \eta_1 = y \}.$$

Lemma 2.2. *Any $\mathcal{A}_p$-minimizer with finite action is either a future pointing time-like geodesic of (M, g) or a future pointing light-like pregeodesic of (M, g), i.e. an orientation preserving reparametrization is a future pointing light-like geodesic of (M, g).*

Proof. Let $\gamma : [0, 1] \to M$ be a $\mathcal{A}_p$-minimizer with finite action. Then $\dot{\gamma}(t) \in \mathcal{C}$ for a.e. t. By Jensen's inequality we have

$$\int_0^1 -\frac{1}{p} (-g(\dot{\eta}, \dot{\eta}))^{p/2} \, dt \geq -\frac{1}{p} \left(\int_0^1 \sqrt{-g(\dot{\eta}, \dot{\eta})} \, dt \right)^p$$

for any causal curve $\eta : [0, 1] \to M$ with equality if and only if η is parametrized proportionally to arclength.

Recall that the restriction of a minimizer to any subinterval of $[0, 1]$ is a minimizer of the restricted action. Since any point in a space-time admits a globally hyperbolic neighbourhood [63, Theorem 2.14], the Avez–Seifert Theorem [66, Proposition 14.19] implies that every minimizer of $\mathcal{A}_1$ with finite action is a causal pregeodesic.

Combining both points we see that if the action of γ is negative, the curve is a time-like pregeodesic parametrized with constant arclength, i.e. a time-like geodesic. If the action of γ vanishes, the curve is a light-like pregeodesic. ∎

Consider the cost function relative to the p-action $\mathcal{A}_p$,

$$c_p : M \times M \to \mathbb{R} \cup \{+\infty\},$$
$$(x, y) \mapsto \inf\{\mathcal{A}_p(\eta) \mid \eta \in \mathrm{AC}([0, 1], M), \, \eta_0 = x, \, \eta_1 = y\}.$$

Remark 2.3. We will always assume that:

(i) The cost function is bounded from below on bounded subsets of $M \times M$. By transitivity of the causal relation this follows from the assumption that $c_p(x, y) > -\infty$ for all $x, y \in M$.

(ii) The cost function is *localizable*, i.e. every point $x \in M$ has a neighbourhood $U \subset M$ such that the cost function of the space-time $(U, g|_U, \mathcal{C}|_U)$ coincides with the global cost function.

Since the main results of this paper are local in nature, these assumptions can always be satisfied by restricting the space-time to a suitable open subset.

Proposition 2.4. *Fix $p \in (0, 1)$ and let $(M, g, \mathcal{C})$ be a space-time. Then every point has a neighbourhood U such that the following holds for the space-time $(U, g|_U, \mathcal{C}|_U)$. For $x, y \in U$ with (x, y) in J_U^+, the causal relation of $(U, g|_U, \mathcal{C}|_U)$, there exists a curve $\gamma :$ $[0, 1] \to U$ with $\gamma_0 = x$, $\gamma_1 = y$, and minimizing $\mathcal{A}_p$ among all curves $\eta \in \mathrm{AC}([0, 1], M)$ with $\eta_0 = x$ and $\eta_1 = y$. Moreover, γ is a constant speed geodesic for the metric g, $\dot{\gamma} \in \mathcal{C}$ whenever the tangent vector exists, and $\mathcal{A}_p(\gamma) \in \mathbb{R}$.*

Proof. It is well known that in a space-time every point has a globally hyperbolic neighbourhood U. If $(x, y) \in J_U^+$ there exists a curve with finite action $\mathcal{A}_p$ between x and y. At the same time the action is bounded from below, e.g. by a steep Lyapunov function [11]. Therefore any minimizer $\gamma : [0, 1] \to U$ has finite action, i.e. $\dot{\gamma}(t) \in \mathcal{C}$ for almost all t. By Jensen's inequality we have

$$\int_0^1 -\frac{1}{p}(-g(\dot{\eta}, \dot{\eta}))^{p/2} \, dt \geq -\frac{1}{p}\left(\int_0^1 \sqrt{-g(\dot{\eta}, \dot{\eta})} \, dt\right)^p$$

for any causal curve $\eta : [0, 1] \to U$ with equality if and only if η is parametrized proportionally to arclength. By the Avez–Seifert Theorem [66, Proposition 14.19] every minimizer of the right hand side is a causal pregeodesic. Combining both it follows that every $\mathcal{A}_p$-minimizer is a causal geodesic. ∎

2.3. Ricci curvature and Jacobi equation

We now fix the notation regarding curvature for a Lorentzian manifold (M, g) of dimension $n \geq 2$. Let ∇ be the Levi-Civita connection of (M, g). The Riemann curvature tensor is defined by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X X - \nabla_{[X,Y]}Z, \tag{2.7}$$

where X, Y, Z are smooth vector fields on M and $[X, Y]$ is the Lie bracket of X and Y. For each $x \in M$, the Ricci curvature is the symmetric bilinear form $\mathrm{Ric}_x : T_x M \times T_x M \to \mathbb{R}$ defined by

$$\mathrm{Ric}_x(v, w) := \sum_{i=1}^{n} g(e_i, e_i) g(R(e_i, w)v, e_i), \tag{2.8}$$

where $(e_i)_{i=1}^{n}$ is an orthonormal basis of $T_x M$, i.e. $|g(e_i, e_j)| = \delta_{ij}$ for all $i, j = 1, \ldots, n$.

Given an endomorphism $\mathcal{U} : T_x M \to T_x M$ and a g-orthonormal basis $(e_i)_{i=1}^{n}$ of $T_x M$, we associate to $\mathcal{U}$ the matrix

$$(\mathcal{U}_{ij})_{i,j=1}^{n}, \quad \mathcal{U}_{ij} := g(e_i, e_j) g(\mathcal{U} e_i, e_j). \tag{2.9}$$

The trace $\mathrm{Tr}_g(\mathcal{U})$ and the determinant $\mathrm{Det}_g(\mathcal{U})$ of the endomorphism $\mathcal{U}$ with respect to the Lorentzian metric g are by definition the trace $\mathrm{tr}(\mathcal{U}_{ij})$ and the determinant $\det(\mathcal{U}_{ij})$ of the matrix $(\mathcal{U}_{ij})_{i,j=1}^{n}$. It is standard to check that this definition is independent of the chosen orthonormal basis of $T_x M$. Note that $\mathrm{Ric}_x(v, w)$ is the trace of the curvature endomorphism $R(\cdot, w)v : T_x M \to T_x M$.

A smooth curve $\gamma : I \to M$ is called a *geodesic* if $\nabla_{\dot{\gamma}} \dot{\gamma} = 0$. A vector field J along a geodesic γ is said to be a *Jacobi field* if it satisfies the *Jacobi equation*

$$\nabla_{\dot{\gamma}}(\nabla_{\dot{\gamma}} J) + R(J, \dot{\gamma})\dot{\gamma} = 0. \tag{2.10}$$

2.4. The q-gradient of a function

Finally, let us recall the definition of gradient and Hessian. Given a smooth function $f : M \to \mathbb{R}$, the *gradient* of f denoted by $\nabla_g f$ is defined by the identity

$$g(\nabla_g f, Y) = df(Y), \quad \forall Y \in TM,$$

where df is the differential of f. The *Hessian* of f, denoted by Hess_f, is defined to be the covariant derivative of df:

$$\mathrm{Hess}_f := \nabla(df).$$

It is related to the gradient through the formula

$$\mathrm{Hess}_f(X, Y) = g(\nabla_X \nabla_g f, Y), \quad \forall X, Y \in TM,$$

and satisfies the symmetry

$$\mathrm{Hess}_f(X, Y) = \mathrm{Hess}_f(Y, X), \quad \forall X, Y \in TM. \tag{2.11}$$

Next we recall some notions related to the causal character of functions.

- A function $f : M \to \mathbb{R} \cup \{\pm\infty\}$ is a *causal function* if $f(x) \leq f(y)$ for all $(x, y) \in J^+$.
- A function $f : M \to \mathbb{R} \cup \{\pm\infty\}$ is a *time function* if $f(x) < f(y)$ for all $(x, y) \in J^+ \setminus \Delta$, where Δ denotes the diagonal in $M \times M$.

- Following [11] we call a (C^k-) differentiable function $f : M \to \mathbb{R}$ ($k \in \mathbb{N} \cup \{\infty\}$) a ($C^k$-) *Lyapunov* or ($C^k$-) *temporal* function if $df_x|_{\mathcal{C}_x \setminus \{0\}} > 0$ for all $x \in M$.

Let q be the conjugate exponent to p, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, or equivalently $(p-1)(q-1) = 1$. Notice that since p ranges in $(0, 1)$, q ranges in $(-\infty, 0)$. In order to describe the optimal transport maps later, it is useful to introduce the q-*gradient* (cf. [49])

$$\nabla_g^q \phi := -|g(\nabla_g \phi, \nabla_g \phi)|^{\frac{q-2}{2}} \nabla_g \phi \tag{2.12}$$

for differentiable Lyapunov functions $\phi : M \to \mathbb{R}$; in particular, $\nabla_g^q \phi(x) \in \mathcal{C}_x \setminus \{0\}$. Notice that

$$\text{for } v \in \mathcal{C}_x \setminus \{0\}, \quad \nabla_g \phi(x) = -|g(v, v)|^{\frac{p-2}{2}} v \text{ if and only if } \nabla_g^q \phi(x) = v.$$

Moreover,

$$x \mapsto \nabla_g^q \phi(x) \text{ is continuous (resp. } C^k, k \geq 1) \text{ on } U \subset \{|g(\nabla_g^q \phi, \nabla_g^q \phi)| > 0\}$$

if and only if

$$x \mapsto \nabla_g \phi(x) \text{ is continuous (resp. } C^k, k \geq 1) \text{ on } U \subset \{|g(\nabla_g \phi, \nabla_g \phi)| > 0\}.$$

The motivation for the use of the q-gradient comes from the Hamiltonian formulation of the dynamics; let us briefly mention a few key facts that will play a role later. For $\alpha \in T_x^* M$, let

$$\mathcal{H}_p(\alpha) = \sup_{v \in T_x M} [\alpha(v) - \mathcal{L}_p(v)] \tag{2.13}$$

be the Legendre transform of $\mathcal{L}_p$. Denote by g^* the dual Lorentzian metric on $T^* M$ and by $\mathcal{C}^* \subset T^* M$ the dual cone field to $\mathcal{C}$. Then

$$\mathcal{H}_p(\alpha) = \begin{cases} -\frac{1}{q}(-g^*(\alpha, \alpha))^{q/2} & \text{if } \alpha \in \mathcal{C}^* \setminus T^{*,0} M, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.14}$$

for $(p-1)(q-1) = 1$. By analogous computations to those in the proof of Lemma 2.1, one can check that

$$\nabla_g^q \phi(x) = D\mathcal{H}_p(-d\phi(x)). \tag{2.15}$$

By well known properties of the Legendre transform (see for instance [20, Theorem A.2.5]) it follows that $D\mathcal{H}_p$ is invertible on $\text{Int}(\mathcal{C}^*)$ with inverse $D\mathcal{L}_p$. Thus (2.15) is equivalent to

$$D\mathcal{L}_p(\nabla_g^q \phi(x)) = -d\phi(x). \tag{2.16}$$

2.5. c_p-concave functions and regular c_p-optimal dynamical plans

We denote by $\mathcal{P}(M)$ the set of Borel probability measures on M. For any $\mu_1, \mu_2 \in \mathcal{P}(M)$, we say that $\pi \in \mathcal{P}(M \times M)$ is a *coupling* of μ_1 and μ_2 if

$$(p_i)_\sharp \pi = \mu_i, \quad i = 1, 2,$$

where $p_1, p_2 : M \times M \to M$ are the projections onto the first and second coordinate. Recall that the push-forward $(p_1)_\sharp \pi$ is defined by

$$(p_1)_\sharp \pi(A) := \pi(p_1^{-1}(A))$$

for any Borel subset $A \subset M$. The set of couplings of μ_1, μ_2 is denoted by $\mathrm{Cpl}(\mu_1, \mu_2)$. The c_p-*cost* of a coupling π is given by

$$\int_{M \times M} c_p(x, y) \, d\pi(x, y) \in [-\infty, 0] \cup \{+\infty\}.$$

Denote by $C_p(\mu_1, \mu_2)$ the *minimal cost relative to* c_p among all couplings from μ_1 to μ_2, i.e.

$$C_p(\mu_1, \mu_2) := \inf \left\{ \int c_p \, d\pi \,\middle|\, \pi \in \mathrm{Cpl}(\mu_1, \mu_2) \right\} \in [-\infty, 0] \cup \{+\infty\}.$$

If $C_p(\mu_1, \mu_2) \in \mathbb{R}$, a coupling achieving the infimum is said to be c_p-*optimal*.

We next define the notion of c_p-*optimal dynamical plan*. To this end, it is convenient to consider the set of $\mathcal{A}_p$-minimizing curves, denoted by Γ_p. The set Γ_p is endowed with the sup metric induced by the auxiliary Riemannian metric h. It will be useful to consider the following maps for $t \in [0, 1]$:

$$\mathrm{e}_t : \Gamma_p \to M, \qquad \mathrm{e}_t(\gamma) := \gamma_t,$$
$$\partial \mathrm{e}_t : \Gamma_p \to TM, \quad \partial \mathrm{e}_t(\gamma) := \dot{\gamma}_t \in T_{\gamma_t} M.$$

A c_p-*optimal dynamical plan* is a probability measure Π on Γ_p such that $(\mathrm{e}_0, \mathrm{e}_1)_\sharp \Pi$ is a c_p-optimal coupling from $\mu_0 := (\mathrm{e}_0)_\sharp \Pi$ to $\mu_1 := (\mathrm{e}_1)_\sharp \Pi$.

We will be mostly interested in c_p-optimal dynamical plans obtained by "exponentiating the q-gradient of a c_p-concave function", which we will call *regular c_p-optimal dynamical plans*. In order to define them precisely, let us first recall some basics of Kantorovich duality (we adopt the convention of [1]).

Fix two subsets $X, Y \subset M$. A function $\phi : X \to \mathbb{R} \cup \{-\infty\}$ is c_p-*concave* (with respect to (X, Y)) if it is not identically $-\infty$ and there exists $u : Y \to \mathbb{R} \cup \{-\infty\}$ such that

$$\phi(x) = \inf_{y \in Y} [c_p(x, y) - u(y)] \quad \text{for every } x \in X.$$

Then its c_p-*transform* is the function $\phi^{c_p} : Y \to \mathbb{R} \cup \{-\infty\}$ defined by

$$\phi^{c_p}(y) := \inf_{x \in X} [c_p(x, y) - \phi(x)], \tag{2.17}$$

and its c_p-*superdifferential* $\partial^{c_p} \phi(x)$ at a point $x \in X$ is defined by

$$\partial^{c_p} \phi(x) := \{ y \in Y \mid \phi(x) + \phi^{c_p}(y) = c_p(x, y) \}. \tag{2.18}$$

Note that

$$\begin{cases} \phi(x) = c_p(x, y) - \phi^{c_p}(y) & \text{for all } x \in X, \, y \in \partial^{c_p} \phi(x), \\ \phi(x) \leq c_p(x, y) - \phi^{c_p}(y) & \text{for all } x \in X, \, y \in Y. \end{cases} \tag{2.19}$$

From the definition it follows readily that if ϕ is c_p-concave, then for $(x, z) \in J^+ \cap (X \times X)$ we have

$$\phi(z) = \inf_{y \in Y} [c_p(z, y) - u(y)] \geq \inf_{y \in Y} [c_p(x, y) - u(y)] = \phi(x),$$

i.e. ϕ is a causal function. The same argument shows that $-\phi^{c_p}$ is a causal function as well.

Definition 2.5 (Regular c_p-optimal dynamical plan). A c_p-optimal dynamical plan $\Pi \in \mathscr{P}(\Gamma_p)$ is *regular* if the following holds. There exist relatively compact open subsets $U, V \subset M$ and a smooth c_p-concave (with respect to (U, V)) function $\phi_{1/2} : U \to \mathbb{R}$ such that

(1) $\nabla_g^q \phi_{1/2}(x) \in \mathcal{K} \subset\subset \mathrm{Int}(\mathcal{C})$ for every $x \in U$ and

$$(-g(\nabla_g^q \phi_{1/2}, \nabla_g^q \phi_{1/2}))^{1/2} < \mathrm{inj}_g(U),$$

where $\mathrm{inj}_g(U)$ is the injectivity radius of g on U;

(2) setting

$$\Psi_{1/2}^t(x) = \exp_x^g((t - 1/2)\nabla_g^q \phi_{1/2}(x))$$

and $\mu_t := (\mathrm{e}_t)_\sharp \Pi$ for every $t \in [0, 1]$, we have

$$\mathrm{supp}\,\mu_{1/2} \subset U, \quad \mu_t = (\Psi_{1/2}^t)_\sharp \mu_{1/2} \quad \text{and} \quad \mu_t \ll \mathrm{vol}_g \quad \forall t \in [0, 1].$$

Roughly, the above notion of regularity asks that the $\mathcal{A}_p$-minimizing curves realizing the optimal transport from $\mu_0 := (\mathrm{e}_0)_\sharp \Pi$ to $\mu_1 := (\mathrm{e}_1)_\sharp \Pi$ have velocities contained in $\mathcal{K}$, i.e. they are all "uniformly" time-like future pointing. Moreover, it also implies that $\bigcup_{t \in [0,1]} \mathrm{supp}\,\mu_t \subset M$ is compact; in addition, the optimal transport is assumed to be driven by a smooth potential $\phi_{1/2}$. Even if these conditions may appear a bit strong, we will prove in Lemma 3.2 that there are a lot of such regular plans; moreover, we will show that it is enough to consider such particular optimal transports in order to characterize upper and lower bounds on the (causal-)Ricci curvature and thus characterize the solutions of the Einstein equations.

3. Existence, regularity and evolution of Kantorovich potentials

In order to characterize Lorentzian Ricci curvature upper bounds, we will use the next proposition concerning the evolution of Kantorovich potentials along a regular $\mathcal{A}_p$-minimizing curve $(\mu_t)_{t \in [0,1]}$ of probability measures given by exponentiating the q-gradient of a smooth c_p-concave function with time-like gradient. To this end it is convenient to consider, for $0 \leq s < t \leq 1$, the restricted minimal action

$$c_p^{s,t}(x, y) := \inf\left\{ \int_s^t \mathcal{L}_p(\dot\gamma(\tau))\,d\tau \;\middle|\; \gamma \in \mathrm{AC}([s, t], M), \gamma(s) = x, \gamma(t) = y \right\}.$$

Proposition 3.1. *Let $(M, g, \mathcal{C})$ be a space-time, fix $p \in (0, 1)$ and let $q \in (-\infty, 0)$ be the Hölder conjugate exponent. Let $U, V \subset M$ be relatively compact open subsets and $\phi_{1/2}$ be a smooth c_p-concave function relative to (U, V) such that*

- *$\phi_{1/2}$ is a smooth Lyapunov function on U,*
- *$(-g(\nabla_g^q \phi_{1/2}, \nabla_g^q \phi_{1/2}))^{1/2} < \mathrm{inj}_g(U)$.*

For $t \in [0, 1]$, let

$$\Psi_{1/2}^t : U \to M, \quad \Psi_{1/2}^t(x) := \exp_x((t - 1/2)\nabla_g^q \phi_{1/2}(x)).$$

For every $x \in U$, define

$$\phi_t(\Psi_{1/2}^t(x)) = \phi(t, \Psi_{1/2}^t(x)) := \begin{cases} \phi_{1/2}(x) - c_p^{1/2,t}(x, \Psi_{1/2}^t(x)) & \text{for } t \in [1/2, 1], \\ \phi_{1/2}(x) + c_p^{t,1/2}(\Psi_{1/2}^t(x), x) & \text{for } t \in [0, 1/2). \end{cases}$$
(3.1)

Then the map $(t, y) \mapsto \phi(t, y)$ defined on $\bigcup_{t \in [0,1]} \{t\} \times \Psi_{1/2}^t(U)$ is C^∞ and satisfies the Hamilton–Jacobi equation

$$\partial_t \phi_t(t, y) + \frac{1}{q}\big(-g(\nabla_g \phi_t(y), \nabla_g \phi_t(y))\big)^{q/2} = 0, \quad \forall (t, y) \in \bigcup_{t \in [0,1]} \{t\} \times \Psi_{1/2}^t(U),$$
(3.2)

with

$$\frac{d}{dt}\Psi_{1/2}^t(x) = \nabla_g^q \phi_t(\Psi_{1/2}^t(x)), \quad \forall (t, x) \in [0, 1] \times U.$$
(3.3)

Proof. **Step 1:** *Smoothness of ϕ.* The fact that $t \mapsto \Psi_{1/2}^t$ is a smooth 1-parameter family of maps realizing c_p-optimal transport shows that ϕ defined in (3.1) satisfies (cf. [20, Theorem 6.4.6])

$$\phi_t(\Psi_{1/2}^t(x)) = \phi_s(\Psi_{1/2}^s(x)) - c_p^{s,t}(\Psi_{1/2}^s(x), \Psi_{1/2}^t(x)), \quad \forall x \in U, 0 \le s < t \le 1.$$
(3.4)

In particular,

$$\phi_t(\Psi_{1/2}^t(x)) = \phi_0(\Psi_{1/2}^0(x)) - c_p^{0,t}(\Psi_{1/2}^0(x), \Psi_{1/2}^t(x)), \quad \forall x \in U, t \in [0, 1], \quad (3.5)$$

$$\phi_t(\Psi_{1/2}^t(x)) = \phi_1(\Psi_{1/2}^1(x)) + c_p^{t,1}(\Psi_{1/2}^t(x), \Psi_{1/2}^1(x)), \quad \forall x \in U, t \in [0, 1]. \quad (3.6)$$

Since by construction everything is defined inside the injectivity radius and all the transport rays are non-constant, from (3.5) (respectively (3.6)) it is manifest that the map $(t, y) \mapsto \phi(t, y)$ is C^∞ on $\bigcup_{t \in (0,1]}\{t\} \times \Psi_{1/2}^t(U)$ (resp. $\bigcup_{t \in [0,1)}\{t\} \times \Psi_{1/2}^t(U)$). The smoothness of ϕ on $\bigcup_{t \in [0,1]}\{t\} \times \Psi_{1/2}^t(U)$ follows.

Step 2: *Validity of the Hamilton–Jacobi equation* (3.2). We consider $t \in (1/2, 1]$, the case $t \in [0, 1/2]$ being analogous. Fix $y = \Psi_{1/2}^t(x)$ for some $x \in U$ and $t \in (1/2, 1]$, and let $\gamma : [0, s] \to M$ be a smooth curve with $\dot{\gamma}(0) = v \in T_y M$. From (3.1) we have

$$\phi(t + s, \gamma_s) \ge -\int_0^s \mathcal{L}_p(\dot{\gamma}_\tau) \, d\tau + \phi(t, \gamma_0),$$

with equality for $\gamma(\tau) = \Psi_{t+\tau}(x)$ for all $\tau \in [0, s]$. Dividing by s and taking the limit as $s \to 0$, we obtain

$$\lim_{s \to 0} \frac{\phi(t + s, \gamma_s) - \phi(t, \gamma_0)}{s} \geq -\mathcal{L}_p(v),$$

which in turn implies

$$\partial_t \phi_t(y) \geq -d\phi_t(v) - \mathcal{L}_p(v) \quad \text{for every } v \in T_y M.$$

Note that equality holds for $v = \nabla_g^q \phi_{1/2}(x)$. For $\alpha \in T_y^* M$, let

$$\mathcal{H}_p(\alpha) = \sup_{v \in T_y M} [\alpha(v) - \mathcal{L}_p(v)]$$

denote the Legendre transform of $\mathcal{L}_p$. Thus we get

$$\partial_t \phi_t(y) = \mathcal{H}_p(-d(\phi_t)_y). \tag{3.7}$$

Recalling that $\mathcal{H}_p$ has the representation (2.14), we have

$$\mathcal{H}_p(-d(\phi_t)_y) = -\frac{1}{q}\left(-g(\nabla_g \phi_t(y), \nabla_g \phi_t(y))\right)^{q/2},$$

which, together with (3.7), implies (3.2).

Step 3: *Validity of* (3.3). Since $\Psi_{1/2}^t$ is a smooth 1-parameter family of maps realizing c_p-optimal transport and the function ϕ defined in (3.1) is smooth, it coincides with the viscosity solution (resp. backward solution)

$$\phi_t(y) = \begin{cases} \sup_{z \in \Psi_{1/2}^s(U)} [\phi_s(z) - c_p^{s,t}(z, y)] & \text{for } t \in [s, 1], \\ \inf_{z \in \Psi_{1/2}^s(U)} [\phi_s(z) + c_p^{t,s}(y, z)] & \text{for } t \in [0, s), \end{cases} \tag{3.8}$$

for every $s \in (0, 1)$ and $y \in \Psi_{1/2}^t(U)$.

Let us discuss the case $t \in (s, 1]$, the other is analogous. From (3.4) it follows that $\Psi_{1/2}^s(x)$ is a maximum point on the right hand side of (3.8) corresponding to $y = \Psi_{1/2}^t(x)$. Thus

$$d\phi_s(\Psi_{1/2}^s(x)) = d[c_p^{s,t}(\cdot, \Psi_{1/2}^t(x))](\Psi_{1/2}^s(x)) = -D\mathcal{L}_p\left(\frac{d}{ds}\Psi_{1/2}^s(x)\right).$$

By construction $\frac{d}{ds}\Psi_{1/2}^s(x) \in \text{Int}(\mathcal{C})$ and, as already observed, $D\mathcal{L}_p$ is invertible on $\text{Int}(\mathcal{C})$ with inverse $D\mathcal{H}_p$. We conclude that

$$\frac{d}{ds}\Psi_{1/2}^s(x) = D\mathcal{H}_p(-d\phi_s(\Psi_{1/2}^s(x))) = \nabla_g^q \phi_s(\Psi_{1/2}^s(x)). \qquad \blacksquare$$

We next show that for every point $\bar{x} \in M$ and every $v \in \mathcal{C}_{\bar{x}}$ "small enough" we can find a smooth c_p-concave function ϕ defined on a neighbourhood of $\bar{x}$ such that $\nabla_g^q \phi = v$ and the Hessian of ϕ vanishes at $\bar{x}$. This is well known in the Riemannian setting (e.g. [80, Theorem 13.5]) and should be compared with the recent paper by McCann [61] in the Lorentzian framework. The second part of the next lemma shows that the class of regular c_p-optimal dynamical plans is non-empty, and actually rather rich.

Lemma 3.2. *Let $(M, g, \mathcal{C})$ be a space-time, and fix $\bar{x} \in M$ and $v \in \mathcal{C}_{\bar{x}}$ with $g(v, v) < 0$. Then there exists $\varepsilon = \varepsilon(\bar{x}, v) > 0$ with the following property:*

(1) *For every $s \in (0, \varepsilon)$ and every C^2 function $\phi : M \to \mathbb{R}$ satisfying*

$$\nabla_g^q \phi(\bar{x}) = sv, \quad \mathrm{Hess}_\phi(\bar{x}) = 0, \tag{3.9}$$

there exists a neighbourhood $U_{\bar{x}}$ of $\bar{x}$ and a neighbourhood $U_{\bar{y}}$ of $\bar{y} := \exp_{\bar{x}}^g(sv)$ such that ϕ is c_p-concave relative to $(U_{\bar{x}}, U_{\bar{y}})$.

(2) *Let*

$$\Psi_{1/2}^t(x) := \exp_x((t - 1/2)\nabla_g^q \phi(x)), \quad \forall x \in U_{\bar{x}},$$
$$\tilde{\Psi} : U_{\bar{x}} \to \mathrm{AC}([0, 1], M), \quad x \mapsto \Psi_{1/2}^{(\cdot)}(x).$$

Then, for every $\mu_{1/2} \in \mathcal{P}(M)$ with $\mathrm{supp}\, \mu_{1/2} \subset U_{\bar{x}}$, the measure $\Pi := \tilde{\Psi}_\sharp \mu_{1/2}$ is a c_p-optimal dynamical plan.

Proof. (1) Setting $\bar{y} = \bar{y}(sv) := \exp_{\bar{x}}^g(sv)$, notice that $\nabla_g^q \phi(\bar{x}) = sv$ is equivalent to

$$d\phi(\bar{x}) = D_x c_p(\bar{x}, \bar{y}), \tag{3.10}$$

where $D_x c_p(\bar{x}, \bar{y})$ denotes the differential at $\bar{x}$ of the function $x \mapsto c_p(x, \bar{y})$. Indeed, a computation shows that $D_x c_p(\bar{x}, \bar{y}) = -D\mathcal{L}_p(sv)$ and thus the claim follows from (2.16).

Let $\phi : M \to \mathbb{R}$ be any smooth function satisfying

$$\nabla_g^q \phi(\bar{x}) = sv, \quad \mathrm{Hess}_\phi(\bar{x}) = 0. \tag{3.11}$$

In what follows we denote by $\mathrm{Hess}_{x, c_p}(\bar{x}, \bar{y})$ (resp. $\mathrm{Hess}_{v, \mathcal{L}_p}(sv)$) the Hessian of the function $x \mapsto c_p(x, \bar{y})$ evaluated at $x = \bar{x}$ (resp. the Hessian of the function $T_{\bar{x}} M \ni w \mapsto \mathcal{L}_p(sv + w)$). By taking normal coordinates centred at $\bar{x}$ one can check that the operator norm satisfies

$$\|\mathrm{Hess}_{x, c_p}(\bar{x}, \bar{y}) - \mathrm{Hess}_{v, \mathcal{L}_p}(sv)\| \to 0 \quad \text{as } t \to 0.$$

Recalling from (2.5) that there exists $C_{p,v} > 0$ such that $\mathrm{Hess}_{v, \mathcal{L}_p}(sv) \geq C_{p,v} s^{-2+p}$ as quadratic forms, we infer

$$\mathrm{Hess}_{x, c_p}(\bar{x}, \bar{y}) - \mathrm{Hess}_\phi(\bar{x}) > 0 \quad \text{as quadratic forms, for every } s \in (0, \varepsilon), \tag{3.12}$$

for some $\varepsilon = \varepsilon(\bar{x}, v) > 0$ small enough. Since by construction we have $D_x c_p(\bar{x}, \bar{y}) - d\phi(\bar{x}) = 0$, by the Implicit Function Theorem there exists a neighbourhood $U_{\bar{x}} \times U_{\bar{y}}$ of $(\bar{x}, \bar{y}) \in M \times M$ and a smooth function $F : U_{\bar{y}} \to U_{\bar{x}}$ such that $F(\bar{y}) = \bar{x}$ and

$$D_x c_p(F(y), y) - d\phi(F(y)) = 0 \quad \text{for every } y \in U_{\bar{y}}.$$

Differentiating the last equation in y at $\bar{y}$ and using $\mathrm{Hess}_\phi(\bar{x}) = 0$, we obtain

$$D_{yx}^2 c_p(\bar{x}, \bar{y}) + \mathrm{Hess}_{x, c_p}(\bar{x}, \bar{y}) DF(\bar{y}) = 0. \tag{3.13}$$

Using normal coordinates centred at $\bar{x}$ and (2.4) one can check that the operator norm

$$\left\| D^2_{yx} c_p(\bar{x}, \bar{y}) - (-g(sv, sv))^{\frac{p-4}{2}} (-g(sv, sv)g + (2-p)(sv)^* \otimes (sv)^*) \right\| \to 0$$

as $s \to 0$, where $(sv)^* = g(sv, \cdot)$ is the covector associated to sv.

Since by assumption $g(v, v) < 0$ and $p \in (0, 1)$, it follows from the reverse Cauchy–Schwarz inequality that $\det[D^2_{yx} c_p(\bar{x}, \bar{y})] > 0$ for $s \in (0, \varepsilon)$. Recalling that $\det[\mathrm{Hess}_{x, c_p}(\bar{x}, \bar{y})] \neq 0$, from (3.13) we infer that $\det(DF(\bar{y})) \neq 0$. By the Inverse Function Theorem, up to reducing the neighbourhoods, we see that $F : U_{\bar{y}} \to U_{\bar{x}}$ is a smooth diffeomorphism. Define now

$$u : U_{\bar{y}} \to \mathbb{R}, \quad u(y) := c_p(F(y), y) - \phi(F(y)).$$

For every fixed $y \in U_{\bar{y}}$, the function $U_{\bar{x}} \ni x \mapsto c_p(x, y) - \phi(x) - u(y)$ vanishes at $x = F(y)$; moreover, from (3.12), it follows that $x = F(y)$ is the strict global minimum of that function on $U_{\bar{x}}$, up to further reducing $U_{\bar{x}}$ and $U_{\bar{y}}$, possibly. In other words, the function

$$U_{\bar{x}} \times U_{\bar{y}} \ni (x, y) \mapsto c_p(x, y) - \phi(x) - u(y)$$

is always non-negative and vanishes exactly on the graph of F. It follows that

$$\phi(x) = \inf_{y \in U_{\bar{y}}} [c_p(x, y) - u(y)] \quad \text{for every } x \in U_{\bar{x}}, \tag{3.14}$$

i.e. $\phi : U_{\bar{x}} \to \mathbb{R}$ is a smooth c_p-concave function relative to $(U_{\bar{x}}, U_{\bar{y}})$ satisfying (3.9).

(2) We divide the argument into two steps.

Step 1. We first show that $(e_{1/2}, e_t)_\sharp \Pi$ is a c_p-optimal coupling for $(\mu_{1/2}, \mu_t)$ and every $t \in [1/2, 1]$. To keep notation short, it is convenient to define

$$\Psi'_t(x) := \exp_x(t\nabla^q_g \phi(x)), \quad \forall x \in U_{\bar{x}},$$
$$\tilde{\Psi}' : U_{\bar{x}} \to \mathrm{AC}([0, 1], M), \quad x \mapsto \Psi'_{(\cdot)}(x).$$

Setting $\Pi' := (\tilde{\Psi}')_\sharp \mu_{1/2}$, we have $(e_t)_\sharp \Pi' = (e_{t+1/2})_\sharp \Pi$ for every $t \in [0, 1/2]$. If we show that

$$(e_0, e_1)_\sharp \Pi' \text{ is a } c_p\text{-optimal coupling for } ((e_0)_\sharp \Pi', (e_1)_\sharp \Pi'), \tag{3.15}$$

then, by the triangle inequality, it will follow that $(e_0, e_t)_\sharp \Pi'$ is a c_p-optimal coupling for $((e_0)_\sharp \Pi', (e_t)_\sharp \Pi')$ for every $t \in [0, 1]$; in particular, our claim that $(e_{1/2}, e_t)_\sharp \Pi$ is a c_p-optimal coupling for $(\mu_{1/2}, \mu_t)$, $t \in [1/2, 1]$, will be proved. Thus, the rest of Step 1 will be devoted to establishing (3.15).

Since by construction $c_p : U_{\bar{x}} \times U_{\bar{y}} \to \mathbb{R}$ is smooth, by classical optimal transport theory it is well known that the c_p-superdifferential $\partial^{c_p} \phi \subset U_{\bar{y}}$ is c_p-cyclically monotone (see for instance [1, Theorem 1.13]). Therefore, in order to have (3.15), it is enough to prove that

$$\partial^{c_p} \phi(x) = \{\exp_x(\nabla^q_g \phi(x))\} \quad \text{for every } x \in U_{\bar{x}}. \tag{3.16}$$

Let us first show that $\partial^{c_p}\phi(x) \neq \emptyset$ for every $x \in U_{\bar{x}}$. From the proof of part (1), there exists a smooth diffeomorphism $F : U_{\bar{y}} \to U_{\bar{x}}$ such that

$$\begin{cases} \phi(F(y)) = c_p(F(y), y) - u(y) & \text{for all } y \in U_{\bar{y}}, \\ \phi(x) \leq c_p(x, y) - u(y) & \text{for all } x \in U_{\bar{x}}, \ y \in U_{\bar{y}}. \end{cases} \tag{3.17}$$

From the definition of ϕ^{c_p} in (2.17), it is readily seen that $\phi^{c_p} = u$ on $U_{\bar{y}}$. Thus (3.17) combined with (2.19) implies that $y \in \partial^{c_p}\phi(F(y))$ for every $y \in U_{\bar{y}}$ or, equivalently, $F^{-1}(x) \in \partial^{c_p}\phi(x)$ for every $x \in U_{\bar{x}}$. In particular, $\partial^{c_p}\phi(x) \neq \emptyset$ for every $x \in U_{\bar{x}}$.

Now fix $x \in U_{\bar{x}}$ and pick $y \in \partial^{c_p}\phi(x) \subset U_{\bar{y}}$. Since $z \mapsto c_p(z, y)$ is differentiable on $U_{\bar{x}}$, we get

$$c_p(z, y) = c_p(x, y) + (D_x c_p(x, y))[(\exp_x^h)^{-1}(z)] + o(\mathrm{d}_h(z, x)) \quad \text{for every } z \in U_{\bar{x}}. \tag{3.18}$$

From $y \in \partial^{c_p}\phi(x)$, we have

$$\phi(z) - \phi(x) \overset{(2.19)}{\leq} c_p(z, y) - c_p(x, y)$$
$$\overset{(3.18)}{=} (D_x c_p(x, y))[(\exp_x^h)^{-1}(z)] + o(\mathrm{d}_h(z, x)) \quad \text{for every } z \in U_{\bar{x}}.$$

Since ϕ is differentiable at $x \in U_{\bar{x}}$, it follows that

$$d\phi(x) = D_x c_p(x, y) = -D\mathcal{L}_p(w),$$

where $w \in \mathrm{Int}(\mathcal{C}_x)$ is such that $y = \exp_x^g(w)$, which by (2.15) is equivalent to

$$w = D\mathcal{H}_p(-d\phi(x)) = \nabla_g^q \phi(x),$$

which yields $y = \exp_x^g(w) = \exp_x^g(\nabla_g^q \phi(x))$, concluding the proof of (3.16).

Step 2. Up to further reducing the open set $U_{\bar{x}}$ and the scale parameter $s > 0$ in the definition of ϕ, we can assume that ϕ satisfies the assumptions of Proposition 3.1 and that $\bar{\phi} : \Psi'_{1/2}(U_{\bar{x}}) \to \mathbb{R}$ defined by

$$\bar{\phi}(\Psi'_{1/2}(x)) := \phi(x) - c_p^{0,1/2}(x, \Psi'_{1/2}(x)), \quad \forall x \in \Psi'_{1/2}(U_{\bar{x}}), \tag{3.19}$$

is still a Lyapunov function satisfying

$$(-g(\nabla_g^q \bar{\phi}, \nabla_g^q \bar{\phi}))^{1/2} < \mathrm{inj}_g(\Psi'_{1/2}(U_{\bar{x}})). \tag{3.20}$$

It is easily seen that

$$x = \exp_z\left(-\tfrac{1}{2}\nabla_g^q \bar{\phi}(z)\right), \quad \forall x \in U_{\bar{x}}, \ z := \Psi'_{1/2}(x).$$

Moreover, thanks to (3.20), the curve

$$[0, 1] \ni t \mapsto \exp_z((t - 1)\nabla_g^q \bar{\phi}(z)), \quad \forall x \in U_{\bar{x}}, \ z := \Psi'_{1/2}(x),$$

is still a g-geodesic, solving the method of characteristics associated to the optimal transport problem (see for instance [20, Ch. 5.1] and [80, Ch. 7]; this is actually a variation of Step 1 and of the proof of Proposition 3.1). It follows that the map

$$\Xi : \Psi'_{1/2}(U_{\bar{x}}) \to M, \quad \Xi(z) := \exp_z(-\nabla^q_g \bar{\phi}(z)), \quad \forall x \in U_{\bar{x}}, z := \Psi'_{1/2}(x),$$

is a c_p-optimal transport map. In other words, for every $\bar{\mu} \in \mathcal{P}(M)$ with supp $\bar{\mu} \subset \Psi'_{1/2}(U_{\bar{x}})$,

$$\pi := (\Xi, \mathrm{Id})_\sharp \bar{\mu} \in \mathcal{P}(M \times M)$$

is a c_p-optimal coupling for its marginals. We conclude that

$$(e_0, e_1)_\sharp \Pi = (\Xi, \mathrm{Id})_\sharp ((e_1)_\sharp \Pi)$$

is a c_p-optimal coupling for its marginals and thus Π is a c_p-optimal dynamical plan. ∎

We next establish some basic properties of c_p-optimal dynamical plans which will turn out to be useful for the OT-characterization of Lorentzian Ricci curvature upper and lower bounds.

Lemma 3.3. *Let $(M, g, \mathcal{C})$ be a space-time and let Π be a regular c_p-optimal dynamical plan with*

$$(e_t)_\sharp \Pi = \mu_t = (\Psi^t_{1/2})_\sharp \mu_{1/2} \ll \mathrm{vol}_g, \quad \Psi^t_{1/2}(x) = \exp^g_x((t - 1/2)\nabla^q_g \phi(x)).$$

Then

(1) $\nabla \nabla^q_g \phi(x) : T_x M \to T_x M$ *is a symmetric endomorphism, i.e.*

$$g(\nabla_X \nabla^q_g \phi, Y) = g(\nabla_Y \nabla^q_g \phi, X), \quad \forall X, Y \in T_x M, \forall x \in \mathrm{supp}\, \mu_s.$$

(2) $\Psi^t_{1/2} : \mathrm{supp}\, \mu_{1/2} \to M$ *is a diffeomorphism onto its image for all $t \in [0, 1]$.*

(3) *Setting $\mu_t = \rho_t \, \mathrm{vol}_g$, the following Monge–Ampère equation holds true:*

$$\rho_{1/2}(x) = \mathrm{Det}_g[D\Psi^t_{1/2}(x)]\rho_t(\Psi^t_{1/2}(x)), \quad \mu_{1/2}\text{-a.e. } x, \forall t \in [0, 1]. \quad (3.21)$$

Proof. (1) By construction, ϕ is smooth on U and $g(\nabla_g \phi, \nabla_g \phi) < 0$. Thus also $\nabla^q_g \phi : M \to TM$ is a smooth section of the tangent bundle and the symmetry of the endomorphism $\nabla \nabla^q_g \phi(x) : T_x M \to T_x M$ follows by Schwarz's Lemma.

(2) is a straightforward consequence of assumption (1) in Definition 2.5.

(3) is a straightforward consequence of the change of variable formula. ∎

It will be convenient to consider the matrix of Jacobi fields

$$\mathcal{B}_t(x) := D\Psi^t_{1/2}(x) : T_x M \to T_{\Psi^t_{1/2}(x)} M \quad \text{for all } x \in \mathrm{supp}\, \mu_{1/2}, \quad (3.22)$$

along the geodesic $t \mapsto \gamma_t := \Psi^t_{1/2}(x)$; recalling (2.10), $\mathcal{B}_t(x)$ satisfies the Jacobi equation

$$\nabla_t \nabla_t \mathcal{B}_t(x) + R(\mathcal{B}_t(x), \dot{\gamma}_t)\dot{\gamma}_t = 0, \quad (3.23)$$

where we denote $\nabla_t := \nabla_{\dot{\gamma}_t}$ for short.

Since by Lemma 3.3 we know that $\mathcal{B}_t$ is non-singular for all $x \in \operatorname{supp} \mu_{1/2}$, we can define

$$\mathcal{U}_t(x) := \nabla_t \mathcal{B}_t \circ \mathcal{B}_t^{-1} : T_{\gamma_t} M \to T_{\gamma_t} M \quad \text{for all } x \in \operatorname{supp} \mu_{1/2}. \tag{3.24}$$

The next proposition will be key in the proof of the lower bounds for causal Ricci curvature. It is well known in Riemannian and Lorentzian geometry; see for instance [26, Lemma 3.1] and [29]; in any case we give the proof for the reader's convenience.

Proposition 3.4. *Let $\mathcal{U}_t$ be defined in* (3.24). *Then $\mathcal{U}_t$ is a symmetric endomorphism of $T_{\gamma_t} M$ (i.e. the matrix $(\mathcal{U}_t)_{ij}$ with respect to an orthonormal basis is symmetric) and*

$$\nabla_t \mathcal{U}_t + \mathcal{U}_t^2 + R(\cdot, \dot{\gamma}_t)\dot{\gamma}_t = 0.$$

Taking the trace with respect to g yields

$$\operatorname{Tr}_g(\nabla_t \mathcal{U}_t) + \operatorname{Tr}_g(\mathcal{U}_t^2) + \operatorname{Ric}(\dot{\gamma}_t, \dot{\gamma}_t) = 0. \tag{3.25}$$

Setting $y(t) := \log \operatorname{Det}_g \mathcal{B}_t$, we get

$$y''(t) + \frac{1}{n}(y'(t))^2 + \operatorname{Ric}(\dot{\gamma}_t, \dot{\gamma}_t) \leq 0. \tag{3.26}$$

Proof. Using (3.23) we get

$$\nabla_t \mathcal{U}_t = (\nabla_t \nabla_t \mathcal{B}_t)\mathcal{B}_t^{-1} + \nabla_t \mathcal{B}_t \nabla_t(\mathcal{B}_t^{-1}) = -R(\cdot, \dot{\gamma}_t)\dot{\gamma}_t - (\nabla_t \mathcal{B}_t)\mathcal{B}_t^{-1}(\nabla_t \mathcal{B}_t)\mathcal{B}_t^{-1}$$
$$= -R(\cdot, \dot{\gamma}_t)\dot{\gamma}_t - \mathcal{U}_t^2.$$

Taking the trace with respect to g yields the second identity.

The rest of the proof is devoted to showing (3.26). Let $(e_i(t))_{i=1}^n$ be an orthonormal basis of $T_{\gamma_t} M$ parallel along γ. Setting $y(t) = \log \det \mathcal{B}_t$, we find that

$$y'(t_0) = \frac{d}{dt}\bigg|_{t=t_0} \log \operatorname{Det}_g(\mathcal{B}_t \mathcal{B}_{t_0}^{-1})$$
$$= \frac{d}{dt}\bigg|_{t=t_0} \log \det[(g(e_i(t), e_j(t))g(\mathcal{B}_t \mathcal{B}_{t_0}^{-1} e_i(t), e_j(t)))_{i,j}]$$
$$= \operatorname{Tr}_g[(\nabla_t \mathcal{B}_t)\mathcal{B}_{t_0}^{-1}]|_{t=t_0} = \operatorname{Tr}_g(\mathcal{U}_{t_0}). \tag{3.27}$$

We next show that $\mathcal{U}_t$ is a symmetric endomorphism of $T_{\gamma_t} M$, i.e. the matrix $(\mathcal{U}_t)_{ij}$ is symmetric. To this end, calling the adjoint $\mathcal{U}_t^*$, we observe that

$$\mathcal{U}_t^* - \mathcal{U}_t = (\mathcal{B}_t^*)^{-1}[(\nabla_t \mathcal{B}_t^*)\mathcal{B}_t - \mathcal{B}_t^*(\nabla_t \mathcal{B}_t)]\mathcal{B}_t^{-1}, \tag{3.28}$$

and

$$\nabla_t[(\nabla_t \mathcal{B}_t^*)\mathcal{B}_t - \mathcal{B}_t^*(\nabla_t \mathcal{B}_t)] = (\nabla_t \nabla_t \mathcal{B}_t^*)\mathcal{B}_t - \mathcal{B}_t^*(\nabla_t \nabla_t \mathcal{B}_t). \tag{3.29}$$

Now the Jacobi equation (3.23) reads

$$\nabla_t \nabla_t \mathcal{B}_t = -R(\mathcal{B}_t, \dot{\gamma}_t)\dot{\gamma}_t = -\mathcal{R}(t)\mathcal{B}_t, \tag{3.30}$$

where

$$\mathcal{R}(t) : T_{\gamma_t} M \to T_{\gamma_t} M, \quad \mathcal{R}(t)[v] := R(v, \dot{\gamma}_t)\dot{\gamma}_t,$$

is symmetric; indeed, in the orthonormal basis $(e_i(t))_{i=1}^n$, it is represented by the symmetric matrix

$$(g(e_i(t), e_j(t))g(R(e_i(t), \dot{\gamma}_t)\dot{\gamma}_t, e_j(t)))_{i,j=1}^n.$$

Plugging (3.30) into (3.29), we find that

$$(\nabla_t \mathcal{B}_t^*)\mathcal{B}_t - \mathcal{B}_t^*(\nabla_t \mathcal{B}_t)$$

is constant in t. But $\mathcal{B}_{1/2} = \mathrm{Id}_{T_{\gamma_{1/2}} M}$ and

$$\nabla_t \mathcal{B}_t|_{t=1/2} = -\nabla \nabla_g^q \phi$$

is symmetric by Lemma 3.3 (1). Taking into account (3.28), we conclude that $\mathcal{U}_t$ is symmetric for every $t \in [0, 1]$. Using the fact that $\mathcal{U}_t$ is symmetric, by the Cauchy–Schwarz inequality, we obtain

$$\mathrm{Tr}_g[\mathcal{U}_t^2] = \mathrm{tr}[(\mathcal{U}^2)_{ij}] \geq \frac{1}{n}(\mathrm{tr}[\mathcal{U}_{ij}])^2 = \frac{1}{n}(\mathrm{Tr}_g[\mathcal{U}_{ij}])^2. \tag{3.31}$$

The desired estimate (3.26) then follows from the combination of (3.25), (3.27) and (3.31). ∎

4. Optimal transport formulation of the Einstein equations

The Einstein equations of general relativity for an n-dimensional space-time $(M^n, g, \mathcal{C})$, $n \geq 3$, read

$$\mathrm{Ric} - \tfrac{1}{2}\,\mathrm{Scal}\,g + \Lambda g = 8\pi T, \tag{4.1}$$

where Scal is the scalar curvature, $\Lambda \in \mathbb{R}$ is the cosmological constant, and T is the energy-momentum tensor.

Lemma 4.1. *The space-time $(M^n, g, \mathcal{C})$, $n \geq 3$, satisfies the Einstein equations (4.1) if and only if*

$$\mathrm{Ric} = \frac{2\Lambda}{n-2}g + 8\pi T - \frac{8\pi}{n-2}\,\mathrm{Tr}_g(T)g. \tag{4.2}$$

Proof. Taking the trace of (4.1), one can express the scalar curvature as

$$\mathrm{Scal} = \frac{2n\Lambda}{n-2} - \frac{16\pi}{n-2}\,\mathrm{Tr}_g(T). \tag{4.3}$$

Plugging (4.3) into (4.1) gives the equivalent formulation (4.2) of the Einstein equations just in terms of the metric, the Ricci tensor and the energy-momentum tensor. ∎

The optimal transport formulation of the Einstein equations will consist separately of an optimal transport characterization of the two inequalities

$$\mathrm{Ric} \geq \tilde{T} \quad \text{and} \quad \mathrm{Ric} \leq \tilde{T},$$

where

$$\tilde{T} := \frac{2\Lambda}{n-2}g + 8\pi T - \frac{8\pi}{n-2}\operatorname{Tr}_g(T)g. \tag{4.4}$$

Subsection 4.1 will be devoted to the lower bound and Subsection 4.2 to the upper bound on the Ricci tensor.

A key role in this optimal transport formulation will be played by the (relative) Boltzmann–Shannon entropy defined below. Denote by vol_g the standard volume measure on (M, g). Given an absolutely continuous probability measure $\mu = \varrho \operatorname{vol}_g$ with density $\varrho \in C_c(M)$, define its *Boltzmann–Shannon entropy* (relative to vol_g) as

$$\operatorname{Ent}(\mu|\operatorname{vol}_g) := \int_M \varrho \log \varrho \, d\operatorname{vol}_g. \tag{4.5}$$

4.1. OT-characterization of $\operatorname{Ric} \geq \tilde{T}$

In establishing the Ricci curvature lower bounds, the next elementary lemma will be key (for the proof see for instance [80, Chapter 16] or [6, Lemma 9.1]).

Lemma 4.2. *Define the function* $G : [0, 1] \times [0, 1] \to [0, 1]$ *by*

$$G(s, t) := \begin{cases} (1 - t)s & \text{if } s \in [0, t], \\ t(1 - s) & \text{if } s \in [t, 1], \end{cases} \tag{4.6}$$

so that for all $t \in (0, 1)$ *one has*

$$-\frac{\partial^2}{\partial s^2}G(s, t) = \delta_t \quad \text{in } \mathscr{D}'(0, 1), \quad G(0, t) = G(1, t) = 0. \tag{4.7}$$

If $u \in C([0, 1], \mathbb{R})$ *satisfies* $u'' \geq f$ *in* $\mathscr{D}'(0, 1)$ *for some* $f \in L^1(0, 1)$ *then*

$$u(t) \leq (1 - t)u(0) + tu(1) - \int_0^1 G(s, t)f(s) \, ds, \quad \forall t \in [0, 1]. \tag{4.8}$$

In particular, if $f \equiv c \in \mathbb{R}$ *then*

$$u(t) \leq (1 - t)u(0) + tu(1) - c\frac{t(1 - t)}{2}, \quad \forall t \in [0, 1]. \tag{4.9}$$

The characterization of Ricci curvature lower bounds (i.e. $\operatorname{Ric} \geq Kg$ for some constant $K \in \mathbb{R}$) via displacement convexity of the entropy is by now classical in the Riemannian setting; let us briefly recall the key contributions. Otto–Villani [68] gave a nice heuristic argument for the implication "$\operatorname{Ric} \geq Kg \Rightarrow K$-convexity of the entropy"; this implication was proved for $K = 0$ by Cordero-Erausquin–McCann–Schmuckenschläger [25]; the equivalence for every $K \in \mathbb{R}$ was then established by Sturm–von Renesse [81]. Our optimal transport characterization of $\operatorname{Ric} \geq \frac{2\Lambda}{n-2}g + 8\pi T - \frac{8\pi}{n-2}\operatorname{Tr}_g(T)g$ is inspired by those fundamental papers (cf. also with [51] for the implication (3)$\Rightarrow$(1)). Let us also mention that the characterization of $\operatorname{Ric} \geq Kg$ for $K \geq 0$ via displacement convexity in

the globally hyperbolic Lorentzian setting has recently been obtained independently by McCann [61]. Note that Corollary 4.4 extends such a result to any lower bounds $K \in \mathbb{R}$ and to the case of general (possibly non-globally-hyperbolic) space-times.

The next general result will be applied with $n \geq 3$ and $\tilde{T}$ as in (4.4).

Theorem 4.3 (OT-characterization of Ric $\geq \tilde{T}$). *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$ and let $\tilde{T}$ be a quadratic form on M. Then the following are equivalent:*

(1) $\text{Ric}(v, v) \geq \tilde{T}(v, v)$ *for every causal vector $v \in \mathcal{C}$.*

(2) *For every $p \in (0, 1)$ and every regular dynamical c_p-optimal plan Π,*

$$\text{Ent}(\mu_t|\text{vol}_g) \leq (1 - t)\,\text{Ent}(\mu_0|\text{vol}_g) + t\,\text{Ent}(\mu_1|\text{vol}_g)$$

$$- \int \int_0^1 G(s, t)\tilde{T}(\dot{\gamma}_s, \dot{\gamma}_s)\,ds\,d\Pi(\gamma), \tag{4.10}$$

where we denote $\mu_t := (e_t)_\sharp \Pi$, $t \in [0, 1]$, the curve of probability measures associated to Π.

(3) *There exists $p \in (0, 1)$ such that for every regular dynamical c_p-optimal plan Π the convexity property (4.10) holds.*

Proof. (1)$\Rightarrow$(2). Fix $p \in (0, 1)$. Let Π be a regular dynamical c_p-optimal coupling and let $(\mu_t := (e_t)_\sharp \Pi)_{t \in [0,1]}$ be the corresponding curve of probability measures with $\mu_t = \rho_t \text{vol}_g \ll \text{vol}_g$ compactly supported. By definition of regular dynamical c_p-optimal coupling there exists a smooth function $\phi_{1/2}$ such that, with

$$\Psi^t_{1/2}(x) = \exp^g_x((t - 1/2)\nabla^q_g \phi_{1/2}(x)),$$

we have $\mu_t = (\Psi^t_{1/2})_\sharp \mu_{1/2}$ for every $t \in [0, 1]$. Moreover, the Jacobian $D\Psi^t_{1/2}$ is nonsingular for every $t \in [0, 1]$ on $\text{supp}\,\mu_{1/2}$. Recall the definition of $\mathcal{B}_t$ and $\mathcal{U}_t$ along the geodesic $t \mapsto \gamma_t := \Psi^t_{1/2}(x)$:

$$\mathcal{B}_t(x) := D\Psi^t_{1/2}(x) : T_x M \to T_{\Psi^t_{1/2}(x)} M \quad \text{for } \mu_{1/2}\text{-a.e. } x,$$

$$\mathcal{U}_t(x) := \nabla_t \mathcal{B}_t \circ \mathcal{B}_t^{-1} : T_{\gamma_t} M \to T_{\gamma_t} M \quad \text{for } \mu_{1/2}\text{-a.e. } x.$$

Setting $y_x(t) := \log \text{Det}_g \mathcal{B}_t(x)$ and $\gamma^x_t := \Psi^t_{1/2}(x)$, from Proposition 3.4 we get

$$y''_x(t) + \tilde{T}(\dot{\gamma}^x_t, \dot{\gamma}^x_t) \leq y''_x(t) + \text{Ric}(\dot{\gamma}^x_t, \dot{\gamma}^x_t) \leq 0 \quad \mu_{1/2}\text{-a.e. } x. \tag{4.11}$$

Now, for $t \in [0, 1]$ we have

$$\text{Ent}(\mu_t|\text{vol}_g) = \int \log \rho_t(y)\,d\mu_t(y) = \int_M \log \rho_t(\Psi^t_{1/2}(x))\,d\mu_{1/2}(x)$$

$$= \int \log[\rho_{1/2}(x)(\text{Det}_g(D\Psi^t_{1/2})(x))^{-1}]\,d\mu_{1/2}(x)$$

$$= \text{Ent}(\mu_{1/2}|\text{vol}_g) - \int y_x(t)\,d\mu_{1/2}(x), \tag{4.12}$$

where the second to last equality follows from Lemma 3.3 (3). Using (4.11) we obtain

$$\frac{d^2}{dt^2} \operatorname{Ent}(\mu_t | \operatorname{vol}_g) \geq \int \tilde{T}(\dot{\gamma}_t^x, \dot{\gamma}_t^x) \, d\mu_{1/2}(x) = \int \tilde{T}(\dot{\gamma}_t, \dot{\gamma}_t) \, d\Pi(\gamma), \quad \forall t \in (0, 1).$$
(4.13)

Using Lemma 4.2, we get (4.10).

(2)⇒(3) is trivial.

(3)⇒(1). We argue by contradiction. Assume there exist $x_0 \in M$ and $v \in T_{x_0}M \cap \mathcal{C}$ with $g(v, v) < 0$ such that the Ricci curvature at x_0 in the direction of $v \in \mathcal{C}_x$ satisfies

$$\operatorname{Ric}(v, v) \leq (\tilde{T} + 3\epsilon g)(v, v)$$
(4.14)

for some $\epsilon > 0$. Thanks to Lemma 3.2, for $\eta \in (0, \bar{\eta}(x_0, v)]$ small enough, there exist $\bar{\delta} > 0$ and a c_p-convex function $\phi_{1/2}$, smooth on $B_{\bar{\delta}}(x_0)$ and satisfying

$$\nabla_g^q \phi_{1/2}(x_0) = \eta v \neq 0 \quad \text{and} \quad \operatorname{Hess}_{\phi_{1/2}}(x_0) = 0.$$
(4.15)

From now on we fix $\eta \in (0, \min(\bar{\eta}(x_0, v), \operatorname{inj}_g(B_{\bar{\delta}}(x_0))))]$, where $\operatorname{inj}_g(B_{\bar{\delta}}(x_0))$ is the injectivity radius of $B_{\bar{\delta}}(x_0)$ with respect to the metric g. It is easily checked that for $\delta \in (0, \bar{\delta})$ small enough the map

$$x \mapsto \Psi_{1/2}^t(x) = \exp_x^g((t - 1/2)\nabla_g^q \phi_{1/2}(x))$$

is a diffeomorphism from $B_\delta(x_0)$ onto its image for any $t \in [0, 1]$. Moreover, since $\nabla_g^q \phi_{1/2}(x_0) \in \operatorname{Int}(\mathcal{C})$ and arguing by continuity and by parallel transport along the geodesics $t \mapsto \Psi_{1/2}^t(x), x \in B_\delta(x_0)$, for $\delta > 0$ small enough we have

$$\bigcup_{t \in [0,1]} \bigcup_{x \in B_\delta(x_0)} \frac{d}{dt} \Psi_{1/2}^t(x) \subset\subset \operatorname{Int}(\mathcal{C}).$$
(4.16)

Define $\mu_{1/2} := \operatorname{vol}_g(B_\delta(x_0))^{-1} \operatorname{vol}_g \llcorner B_\delta(x_0)$. Let Π be the c_p-optimal dynamical plan representing the curve of probability measures $(\mu_t := (\Psi_{1/2}^t)_\sharp \mu_{1/2})_{t \in [0,1]}$. Note that (4.16) together with (4.15) ensures that Π is regular for $\delta > 0$ small enough.

Letting $\gamma_t^x := \Psi_{1/2}^t(x) = \exp_x^g((t - 1/2)\nabla_g^q \phi_{1/2}(x))$ for $x \in B_\delta(x_0)$ be the geodesic realizing the transport, note that by continuity there exists $\delta > 0$ small enough such that

$$\operatorname{Ric}(\dot{\gamma}_t^x, \dot{\gamma}_t^x) < (\tilde{T} + 2\epsilon g)(\dot{\gamma}_t^x, \dot{\gamma}_t^x), \quad \forall x \in B_\delta(x_0), \forall t \in [0, 1].$$
(4.17)

The identity (3.25) proved in Proposition 3.4 reads

$$[\operatorname{Tr}_g(\mathcal{U}_t^x)]' + \operatorname{Tr}_g[(\mathcal{U}_t^x)^2] + \operatorname{Ric}(\dot{\gamma}_t^x, \dot{\gamma}_t^x) = 0, \quad \forall x \in B_\delta(x_0), \forall t \in [0, 1].$$
(4.18)

Since by construction

$$\mathcal{U}_{1/2}^{x_0} := \nabla_t \mathcal{B}_{1/2}^{x_0}(\mathcal{B}_{1/2}^{x_0})^{-1} = \nabla \nabla_g^q \phi_{1/2}(x_0) = 0 \quad \text{and} \quad g(\nabla_g^q \phi_{1/2}(x_0), \nabla_g^q \phi_{1/2}(x_0)) < 0,$$

again by continuity we can choose $\delta > 0$ even smaller so that

$$\operatorname{Tr}_g[(\mathcal{U}_t^x)^2] < -\epsilon g(\dot{\gamma}_t^x, \dot{\gamma}_t^x), \quad \forall x \in B_\delta(x_0), \forall t \in [0, 1].$$
(4.19)

The combination of (4.17)–(4.19) yields

$$[\mathrm{Tr}_g(\mathcal{U}_t^x)]' + (\tilde{T} + \epsilon g)(\dot{\gamma}_t^x, \dot{\gamma}_t^x) > 0, \quad \forall x \in B_\delta(x_0), \ \forall t \in [0, 1].$$

Recalling (3.27), the last inequality can be rewritten as

$$y_x(t)'' + (\tilde{T} + \epsilon g)(\dot{\gamma}_t^x, \dot{\gamma}_t^x) > 0, \quad \forall x \in B_\delta(x_0), \ \forall t \in [0, 1].$$

The combination of the last inequality with (4.12) gives

$$\frac{d^2}{dt^2} \mathrm{Ent}(\mu_t|\mathrm{vol}_g) < \int (\tilde{T} + \epsilon g)(\dot{\gamma}_t^x, \dot{\gamma}_t^x) \, d\mu_{1/2}(x), \quad \forall t \in (0, 1). \tag{4.20}$$

By applying Lemma 4.2 we get

$$\mathrm{Ent}(\mu_t|\mathrm{vol}_g) \geq (1 - t) \, \mathrm{Ent}(\mu_0|\mathrm{vol}_g) + t \, \mathrm{Ent}(\mu_1|\mathrm{vol}_g)$$
$$- \int \int_0^1 G(s, t)(\tilde{T} + \epsilon g)(\dot{\gamma}_s, \dot{\gamma}_s) \, ds \, d\Pi(\gamma),$$
$$= (1 - t) \, \mathrm{Ent}(\mu_0|\mathrm{vol}_g) + t \, \mathrm{Ent}(\mu_1|\mathrm{vol}_g)$$
$$- \int \int_0^1 G(s, t)\tilde{T}(\dot{\gamma}_s, \dot{\gamma}_s) \, ds \, d\Pi(\gamma) - \epsilon \frac{t(1 - t)}{2} \int g(\dot{\gamma}, \dot{\gamma}) \, d\Pi(\gamma),$$

where in the equality we have used the fact that for every fixed $x \in B_\delta(x_0)$ the function $t \mapsto g(\dot{\gamma}_t^x, \dot{\gamma}_t^x)$ is constant (as $t \mapsto \gamma_t^x$ is a g-geodesic by construction).

This clearly contradicts (4.10), as $\int g(\dot{\gamma}, \dot{\gamma}) \, d\Pi(\gamma) < 0$. ∎

In the vacuum case, i.e. $T \equiv 0$, the inequality $\mathrm{Ric} \geq \tilde{T}$ with $\tilde{T}$ as in (4.4) reads $\mathrm{Ric} \geq Kg$ with $K = \frac{2\Lambda}{n-2} \in \mathbb{R}$. Note that for $v \in \mathcal{C}$ we have $g(v, v) \leq 0$, so when comparing the next result with its Riemannian counterparts [25, 68, 81], the sign of the lower bound K is reversed.

Corollary 4.4 (The vacuum case $T \equiv 0$). *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$ and let $K \in \mathbb{R}$. Then the following are equivalent:*

(1) $\mathrm{Ric}(v, v) \geq Kg(v, v)$ *for every causal vector $v \in \mathcal{C}$.*

(2) *For every $p \in (0, 1)$ and every regular dynamical c_p-optimal plan Π,*

$$\mathrm{Ent}(\mu_t|\mathrm{vol}_g)$$
$$\leq (1 - t) \, \mathrm{Ent}(\mu_0|\mathrm{vol}_g) + t \, \mathrm{Ent}(\mu_1|\mathrm{vol}_g) - K\frac{t(1 - t)}{2} \int g(\dot{\gamma}, \dot{\gamma}) \, d\Pi(\gamma)$$
$$= (1 - t) \, \mathrm{Ent}(\mu_0|\mathrm{vol}_g) + t \, \mathrm{Ent}(\mu_1|\mathrm{vol}_g) - Kt(1 - t) \int \mathcal{A}_2(\gamma) \, d\Pi(\gamma), \tag{4.21}$$

where we denote by $\mu_i := (\mathrm{e}_i)_\sharp \Pi$, $i = 0, 1$, the endpoints of the curve of probability measures associated to Π.

(3) *There exists $p \in (0, 1)$ such that for every regular dynamical c_p-optimal plan Π the convexity property (4.21) holds.*

Remark 4.5 (The strong energy condition). The strong energy condition asserts that, with T being the energy-momentum tensor, $T(v, v) \geq \frac{1}{2} \operatorname{Tr}_g(T)$ for every time-like vector $v \in TM$ satisfying $g(v, v) = -1$. Assuming that the space-time $(M, g, \mathcal{C})$ satisfies the Einstein equations (4.1) with zero cosmological constant $\Lambda = 0$, the strong energy condition is equivalent to $\operatorname{Ric}(v, v) \geq 0$ for every time-like vector $v \in TM$. This corresponds to the case $K = 0$ in Corollary 4.4.

The strong energy condition, proposed by Hawking and Penrose [42, 44, 70], plays a key role in general relativity. For instance, in the presence of trapped surfaces, it implies that the space-time has singularities (e.g. black holes) [43, 82].

4.2. OT-characterization of $\operatorname{Ric} \leq \tilde{T}$

The goal of the present section is to provide an optimal transport formulation of upper bounds on *time-like* Ricci curvature in the *Lorentzian setting*. More precisely, given a quadratic form $\tilde{T}$ (which will later be chosen to be equal to the right hand side of the Einstein equations, i.e. as in (4.4)), we aim to find an optimal transport formulation of the condition

$$\text{``}\operatorname{Ric}(v, v) \leq \tilde{T}(v, v) \text{ for every time-like vector } v \in \mathcal{C}\text{''}.$$

The Riemannian counterpart, in the special case of $\operatorname{Ric} \leq Kg$ for some constant $K \in \mathbb{R}$, has recently been established by Sturm [75].

In order to state the result, let us fix some notation. Given a relatively compact open subset $E \subset\subset \operatorname{Int}(\mathcal{C})$ let $p_{TM \to M} : TM \to M$ be the canonical projection map and $\operatorname{inj}_g(E) > 0$ be the injectivity radius of the exponential map of g restricted to E. For $x \in p_{TM \to M}(E)$ and $r \in (0, \operatorname{inj}_g(E))$ we denote

$$B_r^{g,E}(x) := \{\exp_x^g(tw) \mid w \in T_x M \cap E, g(w, w) = -1, t \in [0, r]\}.$$

Definition 4.6 (r-concentrated regular c_p-optimal dynamical plan). Fix a relatively compact open subset $E \subset\subset \operatorname{Int}(\mathcal{C})$, $x \in p_{TM \to M}(E)$, $v \in T_x M \cap E$, and $r > 0$. A regular c_p-optimal dynamical plan Π is called *r-concentrated* in the direction of v (with respect to E) if $\mu_t := (e_t)_\sharp \Pi$ for $t \in [0, 1]$ satisfies

(1) $\mu_{1/2} = \operatorname{vol}_g(B_{r^4}^{g,E}(x))^{-1} \operatorname{vol}_g \llcorner B_{r^4}^{g,E}(x)$;

(2) $\operatorname{supp} \mu_1 \subset \{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$, where $y := \exp_x^g(rv)$.

The next general result will be applied with $n \geq 3$ and $\tilde{T}$ as in (4.4).

Theorem 4.7 (OT-characterization of $\operatorname{Ric} \leq \tilde{T}$). *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$ and let $\tilde{T}$ be a quadratic form on M. Then the following assertions are equivalent:*

(1) *$\operatorname{Ric}(v, v) \leq \tilde{T}(v, v)$ for every causal vector $v \in \mathcal{C}$.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \operatorname{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$ such*

that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.

For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) which has $\tilde{T}(v, v)$-concave entropy in the sense that

$$\frac{4}{r^2}\left[\operatorname{Ent}(\mu_1|\operatorname{vol}_g) - 2\operatorname{Ent}(\mu_{1/2}|\operatorname{vol}_g) + \operatorname{Ent}(\mu_0|\operatorname{vol}_g)\right] \leq \tilde{T}(v, v) + \epsilon(r). \quad (4.22)$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in* (2) *holds true.*

Remark 4.8. Given an auxiliary Riemannian metric h on M, with arguments as in the proof below, condition (2) in Theorem 4.7 can be replaced by

(2)′ For every $p \in (0, 1)$ and for every $x \in M$ there exist $R = R(x) > 0$ and a function $\epsilon = \epsilon_x : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$ such that for all $v \in \operatorname{Int}(\mathcal{C}_x)$ with $h(v, v) \leq R^2$ the following assertion holds.

For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) satisfying (4.22).

Moreover, both in (2) and (2)′ one can replace $B_{r^4}^{g,E}(x)$ (resp. $\{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$) by $B_{r^4}^h(x)$ (resp. $B_{r^2}^h(y)$).

Proof of Theorem 4.7. (1)⇒(2). Let $(M, g, \mathcal{C})$ be a space-time and let h be an auxiliary Riemannian metric on M such that

$$\tfrac{1}{4}h(w, w) \leq |g(w, w)| \leq 4h(w, w), \quad \forall w \in E.$$

We denote by d_h^{TM} the distance on TM induced by h. Once the compact subset $E \subset\subset \operatorname{Int}(\mathcal{C})$ is fixed, thanks to Lemma 3.2 there exist a constant

$$R = R(E) \in (0, \min(1, \operatorname{inj}_g(E)))$$

and a function

$$\epsilon = \epsilon_E : (0, \infty) \to (0, \infty) \quad \text{with} \quad \lim_{r \downarrow 0} \epsilon(r) = 0$$

such that for all $r \in (0, R/10)$, $x \in p_{TM \to M}(E)$, and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ we can find a c_p-convex function $\phi : M \to \mathbb{R}$ with the following properties:

- ϕ is smooth on $B_{100r}^h(x)$, $\nabla_g^q \phi(x) = v$, $\nabla_g^q \phi \in E$ on $B_{10r}^h(x)$, $\mathrm{d}_h^{TM}(\nabla_g^q \phi, v) \leq \epsilon(r)$ on $B_{10r}^h(x)$;
- $|\operatorname{Hess}_\phi|_h \leq \epsilon(r)$ on $B_{10r}^h(x)$.

For $t \in [0, 1]$, consider the map

$$\Psi_{1/2}^t : z \mapsto \exp_z(r(t - 1/2)\nabla_g^q \phi(z)).$$

Notice that

$$\Psi_{1/2}^t(B_{r^4}^{g,E}(x)) \subset B_{10r}^h(x), \quad \forall t \in [0, 1].$$

Let

$$\mu_{1/2} = \mathrm{vol}_g \left(B_{r^4}^{g,E}(x) \right)^{-1} \mathrm{vol}_g \llcorner B_{r^4}^{g,E}(x)$$

and define

$$\mu_t := (\Psi_{1/2}^t)_\sharp (\mu_{1/2}), \quad \forall t \in [0, 1].$$

By the properties of ϕ, the plan Π representing the curve $(\mu_t)_{t\in[0,1]}$ of probability measures is a regular c_p-optimal dynamical plan and

$$\mathrm{supp}\,\mu_1 \subset \{\exp_y^g (r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}.$$

By Proposition 3.1 we can find a smooth family $(\phi_t)_{t\in[0,1]}$ of functions defined on $\bigcup_{t\in[0,1]}\{t\} \times \mathrm{supp}\,\mu_t$ with $\phi_{1/2} = \phi$ satisfying

$$(\partial_t \phi_t)(\gamma_t) + \frac{r}{q}(-g(\nabla_g \phi_t(\gamma_t), \nabla_g \phi_t(\gamma_t)))^{q/2} = 0 \quad \text{for } \Pi\text{-a.e. } \gamma \text{ and } t \in [0, 1], \tag{4.23}$$

$$r\nabla_g^q \phi_t(\gamma_t) - \dot{\gamma}_t = 0 \quad \text{for } \Pi\text{-a.e. } \gamma \text{ and } t \in [0, 1]. \tag{4.24}$$

Moreover, using the properties of $\phi_{1/2} = \phi$ and the smoothness of the family $(\phi_t)_{t\in[0,1]}$, we have

$$\begin{aligned} \mathrm{d}_h^{TM}(\nabla_g^q \phi_t(\gamma_t), v) &\leq \epsilon(r), \\ |\mathrm{Hess}_{\phi_t}(\gamma_t)|_h &\leq \epsilon(r) \quad \text{for } \Pi\text{-a.e. } \gamma \text{ and } t \in [0, 1], \end{aligned} \tag{4.25}$$

up to replacing $\epsilon(r)$ with a suitable function

$$\epsilon = \epsilon_E : (0, \infty) \to (0, \infty) \quad \text{with} \quad \lim_{r\downarrow 0} \epsilon(r) = 0.$$

The curve $[0, 1] \ni t \mapsto \mathrm{Ent}(\mu_t|\mathrm{vol}_g) \in \mathbb{R}$ is smooth and, in virtue of (4.24), it satisfies

$$\frac{d}{dt} \mathrm{Ent}(\mu_t|\mathrm{vol}_g) = r \int_M g(\nabla_g^q \phi_t, \nabla_g \rho_t)\, d\,\mathrm{vol}_g = r \int_M \Box_g^q \phi_t\, d\mu_t \quad \text{for all } t \in [0, 1], \tag{4.26}$$

where $\rho_t := \frac{d\mu_t}{d\,\mathrm{vol}_g}$ is the density of μ_t,

$$\Box_g^q \phi_t := \mathrm{div}(-\nabla_g^q \phi_t)$$

is the q-Box of ϕ_t (the Lorentzian analog of the q-Laplacian), and where we have used the continuity equation

$$\frac{d}{dt}\rho_t + r\,\mathrm{div}(\rho_t \nabla_g^q \phi_t) = 0.$$

For what follows it is useful to consider the linearization of the q-Box at a smooth function f, denoted by L_f^q and defined by the following relation:

$$\left.\frac{d}{dt}\right|_{t=0} \Box_g^q(f + tu) = L_f^q u, \quad \forall u \in C_c^\infty(M). \tag{4.27}$$

The map $[0, 1] \ni t \mapsto \int_M \square_g^q \phi_t \, d\mu_t \in \mathbb{R}$ is smooth and, in virtue of (4.23) and (4.24),

$$
\frac{d}{dt} r \int_M \square_g^q \phi_t \, d\mu_t
$$
$$
= -r^2 \int_M \left(L_\phi^q \left(\frac{1}{q} (-g(\nabla_g \phi_t, \nabla_g \phi_t))^{q/2} \right) + g(\nabla_g \square_g^q \phi_t, -\nabla_g^q \phi_t) \right) d\mu_t
$$

for every $t \in [0, 1]$. Using the q-Bochner identity (A.2) together with the assumption $\mathrm{Ric}(w, w) \leq \tilde{T}(w, w)$ for any $w \in \mathcal{C}$ and estimates (4.25) on ϕ_t, we can rewrite the last formula as

$$
\frac{d}{dt} \frac{1}{r} \int_M \square_g^q \phi_t \, d\mu_t
$$
$$
= \int_M |g(\nabla_g \phi_t, \nabla_g \phi_t)|^{q-2} \Bigg[\mathrm{Ric}(\nabla_g \phi_t, \nabla_g \phi_t) + g(\mathrm{Hess}_{\phi_t}, \mathrm{Hess}_{\phi_t})
$$
$$
+ \left((q-2) \frac{\mathrm{Hess}_{\phi_t}(\nabla_g \phi_t, \nabla_g \phi_t)}{|g(\nabla_g \phi_t, \nabla_g \phi_t)|} \right)^2
$$
$$
- 2(q-2) \frac{\mathrm{Hess}_{\phi_t}((\nabla_g \phi_t, \mathrm{Hess}_{\phi_t}(\nabla_g \phi_t)))}{|g(\nabla_g \phi_t, \nabla_g \phi_t)|} \Bigg] d\mu_t
$$
$$
\leq \int_M (\tilde{T}(v, v) + \epsilon(r)) \, d\mu_t = \tilde{T}(v, v) + \epsilon(r) \quad \text{for all } t \in [0, 1],
$$

up to replacing $\epsilon(r)$ with a suitable function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$. Thus

$$
\mathrm{Ent}(\mu_1 | \mathrm{vol}_g) - 2\,\mathrm{Ent}(\mu_{1/2} | \mathrm{vol}_g) + \mathrm{Ent}(\mu_0 | \mathrm{vol}_g)
$$
$$
= \int_0^{1/2} \left(\frac{d}{dt} \mathrm{Ent}(\mu_t | \mathrm{vol}_g) \Big|_{t=s+1/2} - \frac{d}{dt} \mathrm{Ent}(\mu_t | \mathrm{vol}_g) \Big|_{t=s} \right) ds
$$
$$
= \int_0^{1/2} \int_s^{s+1/2} \frac{d^2}{dt^2} \mathrm{Ent}(\mu_t | \mathrm{vol}_g) \, dt \, ds \leq \frac{\tilde{T}(v, v) + \epsilon(r)}{4} r^2.
$$

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

(3)$\Rightarrow$(1). Fix $p \in (0, 1)$ given by (3) and assume for contradiction that there exist $x \in M$, $\epsilon > 0$ and $v \in T_x M \cap \mathcal{C}$ with $-g(v, v) = 1$ such that

$$
\mathrm{Ric}(v, v) \geq (\tilde{T} - 2\epsilon g)(v, v).
$$

Then, by continuity, we can find a relatively compact neighbourhood $E \subset\subset \mathrm{Int}(\mathcal{C})$ of v in TM such that

$$
\mathrm{Ric}(w, w) \geq (\tilde{T} - \epsilon g)(w, w), \quad \forall w \in E. \tag{4.28}
$$

By Lemma 3.2 we can construct a c_p-convex function $\phi : M \to \mathbb{R}$ such that ϕ is smooth on a neighbourhood of x and

$$
\nabla_g^q \phi(x) = Rv.
$$

For $t \in [0, 1]$, define

$$
\Psi_{1/2}^t(z) := \exp_z^g(2r(t - 1/2) \nabla_g^q \phi(z)).
$$

By continuity, for $r \in (0, R)$ small enough, we have

$$\frac{1}{r} \frac{d}{dt} \Psi^t_{1/2}(z) \in E, \quad d^{TM}_h \left(\frac{d}{dt} \Psi^t_{1/2}(z), rv \right) \le \epsilon r, \quad \forall z \in B^{g,E}_{r^4}(x), \forall t \in [0, 1]. \quad (4.29)$$

Moreover $\Psi^1_{1/2}(B^{g,E}_{r^4}(x)) \subset B^{g,E}_{r^2}(y)$.

Set $\mu_{1/2} := \text{vol}_g(B^{g,E}_{r^4}(x))^{-1} \text{vol}_g \llcorner B^{g,E}_{r^4}(x)$ and consider $\mu_t := (\Psi^t_{1/2})_\sharp \mu_{1/2}$. Notice that

$$\text{supp}\, \mu_1 \subset B^{g,E}_{r^2}(y) \subset \{\exp^g_y(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}.$$

By the above construction, $(\mu_t)_{t \in [0,1]}$ can be represented by a regular c_p-optimal dynamical plan Π such that $\text{supp}((\partial e)_\sharp \Pi) \subset E$. Therefore (4.28) together with Theorem 4.3 yields

$$\text{Ent}(\mu_{1/2}|\text{vol}_g) \le \tfrac{1}{2} \text{Ent}(\mu_0|\text{vol}_g) + \tfrac{1}{2} \text{Ent}(\mu_1|\text{vol}_g)$$

$$- \int \int_0^1 G(s, 1/2)(\tilde{T} - \epsilon g)(\dot{\gamma}_s, \dot{\gamma}_s)\, ds\, d\Pi(\gamma)$$

$$\le \tfrac{1}{2} \text{Ent}(\mu_0|\text{vol}_g) + \tfrac{1}{2} \text{Ent}(\mu_1|\text{vol}_g) - \frac{r^2 (\tilde{T}(v, v) - \epsilon g(v, v) + C\epsilon r)}{8}, \quad (4.30)$$

where in the second inequality we have used (4.29) and that $C > 0$ is a constant independent of r and ϵ. Note that $\epsilon > 0$ in (4.30) is fixed independently of $r > 0$. Clearly (4.30) contradicts the existence of $\epsilon_E(r) \to 0$ as $r \to 0$ such that (4.22) holds. ∎

In the vacuum case when $T \equiv 0$, the inequality $\text{Ric} \le \tilde{T}$ with $\tilde{T}$ as in (4.4) reads as

$$\text{Ric} \le Kg \quad \text{with} \quad K = \frac{2\Lambda}{n - 2} \in \mathbb{R}.$$

Note that for $v \in \mathcal{C}$ we have $g(v, v) \le 0$, so when comparing the next result with its Riemannian counterpart [75], the sign of the lower bound K is reversed.

Corollary 4.9. *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \ge 2$ and let $K \in \mathbb{R}$. Then the following assertions are equivalent:*

(1) $\text{Ric}(v, v) \le Kg(v, v)$ *for every causal vector $v \in \mathcal{C}$.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.*

   For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) which has $-K$-concave entropy in the sense that

$$\frac{4}{r^2} [\text{Ent}(\mu_1|\text{vol}_g) - 2\,\text{Ent}(\mu_{1/2}|\text{vol}_g) + \text{Ent}(\mu_0|\text{vol}_g)] \le -K + \epsilon(r). \quad (4.31)$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in (2) holds true.*

4.3. Optimal transport formulation of the Einstein equations

Recall that the Einstein equations of general relativity, with cosmological constant equal to $\Lambda \in \mathbb{R}$ and energy-momentum tensor T, read

$$\text{Ric} = \frac{2\Lambda}{n-2} g + 8\pi T - \frac{8\pi}{n-2} \text{Tr}_g(T)g \tag{4.32}$$

for an n-dimensional space-time $(M, g, \mathcal{C})$. Combining Theorem 4.3 with Theorem 4.7, both with the choice

$$\tilde{T} = \frac{2\Lambda}{n-2} g + 8\pi T - \frac{8\pi}{n-2} \text{Tr}_g(T)g,$$

we obtain the following optimal transport formulation of (4.32).

Theorem 4.10. *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$ and let $\tilde{T}$ be a quadratic form on M. Then the following assertions are equivalent:*

(1) $\text{Ric}(v, v) = \tilde{T}(v, v)$ *for every $v \in T_x M$.*

(2) $\text{Ric}(v, v) = \tilde{T}(v, v)$ *for every causal vector $v \in \mathcal{C}$.*

(3) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r\downarrow 0} \epsilon(r) = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.*

   For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) satisfying

$$\tilde{T}(v, v) - \epsilon(r) \leq \frac{4}{r^2}[\text{Ent}(\mu_1|\text{vol}_g) - 2\text{Ent}(\mu_{1/2}|\text{vol}_g) + \text{Ent}(\mu_0|\text{vol}_g)]$$
$$\leq \tilde{T}(v, v) + \epsilon(r). \tag{4.33}$$

(4) *There exists $p \in (0, 1)$ such that the assertion as in (3) holds true.*

Remark 4.11 ($\mu_{1/2}$ can be chosen more general). From the proof of Theorem 4.10 it follows that one can replace (2) (and analogously (3)) with the following (a priori stronger, but in fact equivalent) statement. For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r\downarrow 0} \epsilon(r) = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.

   For every $r \in (0, R)$ and every $\mu_{1/2} \in \mathcal{P}(M)$ with $\mu_{1/2} \ll \text{vol}_g$ and $\text{supp}\,\mu_{1/2} \subset B_{r^4}^{g,E}(x)$, setting $y = \exp_x^g(rv)$, there exists a regular c_p-optimal dynamical plan $\Pi = \Pi(\mu_{1/2}, v, r)$ with associated curve

$$(\mu_t := (e_t)_\sharp \Pi)_{t \in [0,1]} \subset \mathcal{P}(M)$$

of probability measures such that

$$\text{supp}\,\mu_1 \subset \{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$$

and satisfying (4.33).

Remark 4.12 (An equivalent statement via an auxiliary Riemannian metric h). Given an auxiliary Riemannian metric h on M, with arguments as in the proof below, condition (3) in Theorem 4.10 can be replaced by

(3)′ For every $p \in (0, 1)$ and every $x \in M$ there exist $R = R(x) > 0$ and a function $\epsilon = \epsilon_x : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$ such that for all $v \in \text{Int}(\mathcal{C}_x)$ with $h(v, v) \leq R^2$ the following assertion holds.

For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v satisfying (4.33).

Moreover, both in (3) and (3)′ one can replace $B_{r^4}^{g,E}(x)$ by $B_{r^4}^{h}(x)$ and

$$\{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}$$

by $B_{r^2}^{h}(y)$.

Remark 4.13 (The tensor $\tilde{T}$). As mentioned above, we will assume that the cosmological constant Λ and the energy momentum tensor T are given, say from physics and/or mathematical general relativity. Given g, Λ and T, for convenience of notation we let $\tilde{T}$ be defined in (1.6).

Let us stress that not any symmetric bilinear form $\tilde{T}$ would correspond to a physically meaningful situation; in order to be physically relevant, it is crucial that $\tilde{T}$ is given by (1.6) where T is a physical energy-momentum tensor (in particular T has to satisfy $\nabla^a T_{ab} = 0$, i.e. be "freely gravitating", it has to satisfy some suitable energy condition like the "dominant energy condition", etc.).

Proof of Theorem 4.10. (1)$\Rightarrow$(2) is trivial.

(2)$\Rightarrow$(1) follows from the identity theorem for polynomials and the fact that $\mathcal{C}$ has non-empty open interior.

(2)$\Rightarrow$(3). From the implication (1)$\Rightarrow$(2) in Theorem 4.7, we get a regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ as in (3) such that the upper bound in (4.33) holds. Moreover, from (4.25) it follows that

$$d_h^{TM}(\dot{\gamma}_t, rv) \leq r\epsilon(r) \quad \text{for } \Pi\text{-a.e. } \gamma \text{ and all } t \in [0, 1]. \tag{4.34}$$

Recalling that the implication (1)$\Rightarrow$(2) in Theorem 4.3 gives the convexity property (4.10) of the entropy along *every* regular c_p-optimal dynamical plan, and using (4.34), we conclude that also the lower bound in (4.33) holds.

(3)$\Rightarrow$(4) is trivial.

(4)$\Rightarrow$(2). The fact that

$$\text{Ric}(v, v) \leq \tilde{T}(v, v), \quad \forall v \in \mathcal{C},$$

follows directly from Theorem 4.7. The fact that

$$\text{Ric}(v, v) \geq \tilde{T}(v, v), \quad \forall v \in \mathcal{C},$$

can be showed following arguments already used in this paper; let us briefly discuss it. Fix $p \in (0, 1)$ given by (4) and assume for contradiction that there are $x \in M$, $\delta > 0$ and $v \in T_x M \cap \mathcal{C}$ with $-g(v, v) > 0$ such that

$$\mathrm{Ric}(v, v) \leq (\tilde{T} + 3\delta g)(v, v).$$

Thanks to Lemma 3.2, up to replacing v with sv for some $s \in (0, 1)$ small enough, we can construct a c_p-convex function $\phi : M \to \mathbb{R}$, smooth in a neighbourhood of x and satisfying

$$\nabla_g^q \phi(x) = v, \quad \mathrm{Hess}_\phi(x) = 0.$$

Then, by continuity,

- we can find a relatively compact open neighbourhood $E \subset\subset \mathrm{Int}(\mathcal{C})$ of v in TM such that

$$\mathrm{Ric}(w, w) \leq (\tilde{T} + 2\delta g)(w, w), \quad \forall w \in E; \tag{4.35}$$

- ϕ is smooth on $B_{100r}^h(x)$ with

$$\nabla_g^q \phi \in E \text{ on } B_{10r}^h(x), \quad \mathrm{d}_h^{TM}(\nabla_g^q \phi, v) \leq \epsilon(r) \text{ on } B_{10r}^h(x),$$
$$|\mathrm{Hess}_\phi|_h \leq \epsilon(r) \quad \text{on } B_{10r}^h(x), \quad \lim_{r \to 0} \epsilon(r) = 0.$$

For $t \in [0, 1]$, consider the map

$$\Psi_{1/2}^t : z \mapsto \exp_z(r(t - 1/2)\nabla_g^q \phi(z)).$$

Notice that

$$\Psi_{1/2}^t(B_{r^4}^{g,E}(x)) \subset B_{10r}^h(x), \quad \forall t \in [0, 1].$$

Set

$$\mu_{1/2} = \mathrm{vol}_g(B_{r^4}^{g,E}(x))^{-1} \mathrm{vol}_g \llcorner B_{r^4}^{g,E}(x), \quad \mu_t := (\Psi_{1/2}^t)_\#(\mu_{1/2}) \text{ for } t \in [0, 1].$$

By the properties of ϕ, the plan Π representing the curve $(\mu_t)_{t \in [0,1]}$ of probability measures is a regular c_p-optimal dynamical plan and

$$\mathrm{supp}\,\mu_1 \subset \{\exp_y^g(r^2 w) \mid w \in T_y M \cap \mathcal{C}, g(w, w) = -1\}.$$

Moreover

$$\frac{1}{r}\frac{d}{dt}\Psi_{1/2}^t(z) \in E, \quad \forall z \in B_r^h(x), \forall t \in [0, 1]. \tag{4.36}$$

We can now follow verbatim the arguments in (1)$\Rightarrow$(2) of Theorem 4.7 by using (4.35) and (4.36), obtaining a function $\epsilon(r) \to 0$ as $r \to 0$ such that

$$\mathrm{Ent}(\mu_1|\mathrm{vol}_g) - 2\,\mathrm{Ent}(\mu_{1/2}|\mathrm{vol}_g) + \mathrm{Ent}(\mu_0|\mathrm{vol}_g) \leq \frac{(\tilde{T} + \delta g)(v, v) + \epsilon(r)}{4}r^2.$$

The last inequality clearly contradicts the lower bound in (4.33). ∎

In the vacuum case $T \equiv 0$ with cosmological constant $\Lambda \in \mathbb{R}$, the Einstein equations read

$$\text{Ric} \equiv \frac{\Lambda}{n/2 - 1} g, \tag{4.37}$$

for an n-dimensional space-time $(M, g, \mathcal{C})$. Specializing Theorem 4.10 with the choice

$$\tilde{T} = \frac{\Lambda}{n/2 - 1} g$$

and using Corollary 4.4 to sharpen the lower bound in (4.38) for the constant case, we obtain the following optimal transport formulation of the Einstein vacuum equations.

Theorem 4.14. *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 3$ and let $\Lambda \in \mathbb{R}$. Then the following assertions are equivalent:*

(1) *The space-time $(M, g, \mathcal{C})$ satisfies the Einstein vacuum equations of general relativity (4.37) corresponding to the cosmological constant equal to Λ.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r) = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.*

For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) satisfying

$$-\frac{\Lambda}{n/2 - 1} \leq \frac{4}{r^2} [\text{Ent}(\mu_1 | \text{vol}_g) - 2 \text{Ent}(\mu_{1/2} | \text{vol}_g) + \text{Ent}(\mu_0 | \text{vol}_g)]$$

$$\leq -\frac{\Lambda}{n/2 - 1} + \epsilon(r). \tag{4.38}$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in (2) holds true.*

It is worth isolating the case of zero cosmological constant.

Corollary 4.15. *Let $(M, g, \mathcal{C})$ be a space-time of dimension $n \geq 2$. Then the following assertions are equivalent:*

(1) *The space-time $(M, g, \mathcal{C})$ satisfies the Einstein vacuum equations of general relativity with zero cosmological constant, i.e. $\text{Ric} \equiv 0$.*

(2) *For every $p \in (0, 1)$ and every relatively compact open subset $E \subset\subset \text{Int}(\mathcal{C})$ there exist $R = R(E) \in (0, 1)$ and a function $\epsilon = \epsilon_E : (0, \infty) \to (0, \infty)$ with $\lim_{r \downarrow 0} \epsilon(r)/r^2 = 0$ such that for all $x \in p_{TM \to M}(E)$ and $v \in T_x M \cap E$ with $g(v, v) = -R^2$ the following assertion holds.*

For every $r \in (0, R)$, there exists an r-concentrated regular c_p-optimal dynamical plan $\Pi = \Pi(x, v, r)$ in the direction of v (with respect to E) satisfying

$$0 \leq \text{Ent}(\mu_1 | \text{vol}_g) - 2 \text{Ent}(\mu_{1/2} | \text{vol}_g) + \text{Ent}(\mu_0 | \text{vol}_g) \leq \epsilon(r). \tag{4.39}$$

(3) *There exists $p \in (0, 1)$ such that the assertion as in (2) holds true.*

Appendix A. A q-Bochner identity in Lorentzian setting

In this section we prove a Bochner type identity in Lorentzian setting for the linearization of the q-Box operator, the Lorentzian analog of the q-Laplacian; let us mention that related results have been obtained in the Riemannian [59,77] and Finsler settings [83,84] but to the best of our knowledge the Lorentzian $\mathscr{L}_p$ framework has not been considered before.

Throughout the section, $(M, g, \mathscr{C})$ is a space-time, $U \subset M$ is an open subset and $q \in (-\infty, 0)$ is fixed. Let $\phi \in C^3(U)$ satisfy

$$-\nabla_g \phi \in \mathrm{Int}(\mathscr{C}) \quad \text{on } U.$$

Denote by

$$\nabla_g^q \phi := -|g(\nabla_g \phi, \nabla_g \phi)|^{\frac{q-2}{2}} \nabla_g \phi$$

the q-gradient, by

$$\Box_g^q \phi := \mathrm{div}(-\nabla_g^q \phi)$$

the q-Box operator of ϕ and by L_ϕ^q the linearization of the q-Box operator at ϕ defined by the following relation:

$$\frac{d}{dt}\bigg|_{t=0} \Box_g^q(\phi + tu) = L_\phi^q u, \quad \forall u \in C_c^\infty(M). \tag{A.1}$$

The ultimate goal of the section is to prove the following result.

Proposition A.1. *Under the above notation, the following q-Bochner identity holds:*

$$-L_\phi^q \left(\frac{(-g(\nabla_g \phi, \nabla_g \phi))^{q/2}}{q} \right) = g(\nabla_g \Box_g^q \phi, -\nabla_g^q \phi)$$

$$+ (q-2)^2 |g(\nabla_g \phi, \nabla_g \phi)|^{q-2} \left(\frac{\mathrm{Hess}_\phi(\nabla_g \phi, \nabla_g \phi)}{|g(\nabla_g \phi, \nabla_g \phi)|} \right)^2$$

$$+ |g(\nabla_g \phi, \nabla_g \phi)|^{q-2} \left(\mathrm{Ric}(\nabla_g \phi, \nabla_g \phi) + g(\mathrm{Hess}_\phi, \mathrm{Hess}_\phi) \right)$$

$$- 2(q-2) |g(\nabla_g \phi, \nabla_g \phi)|^{q-3} \mathrm{Hess}_\phi(\nabla_g \phi, \mathrm{Hess}_\phi(\nabla_g \phi)). \tag{A.2}$$

The proof of Proposition A.1 requires some preliminary lemmas. First of all we derive an explicit expression for the operator L_ϕ^q.

Lemma A.2. *Under the above notation,*

$$L_\phi^q u = (-g(\nabla_g \phi, \nabla_g \phi))^{\frac{q-2}{2}} \left(\Box_g u - (q-2) \frac{\mathrm{Hess}_u(\nabla_g \phi, \nabla_g \phi)}{-g(\nabla_g \phi, \nabla_g \phi)} \right)$$

$$+ (2-q)(-g(\nabla_g \phi, \nabla_g \phi))^{-1} g(\nabla_g \phi, \nabla_g u) \Box_g^q \phi$$

$$+ 2(2-q) (-g(\nabla_g \phi, \nabla_g \phi))^{\frac{q-4}{2}} \mathrm{Hess}_\phi \left(\nabla_g \phi, \nabla_g u + \frac{g(\nabla_g \phi, \nabla_g u)}{-g(\nabla_g \phi, \nabla_g \phi)} \nabla_g \phi \right). \tag{A.3}$$

Proof. By the very definitions of $L_\phi^q u$ and $\square_g^q \phi$, we have

$$
\begin{aligned}
L_\phi^q u &= \operatorname{div}\left(\frac{d}{dt}\bigg|_{t=0}(-g(\nabla_g(\phi+tu),\nabla_g(\phi+tu)))^{\frac{q-2}{2}}\nabla_g(\phi+tu)\right) \\
&= \operatorname{div}\big((2-q)(-g(\nabla_g\phi,\nabla_g\phi))^{\frac{q-4}{2}}g(\nabla_g\phi,\nabla_g u)\nabla_g\phi \\
&\qquad\qquad\qquad\qquad\qquad + (-g(\nabla_g\phi,\nabla_g\phi))^{\frac{q-2}{2}}\nabla_g u\big).
\end{aligned} \tag{A.4}
$$

In order to make the last formula explicit, we compute

$$
\nabla_g(-g(\nabla_g\phi,\nabla_g\phi))^\alpha = -2\alpha\,(-g(\nabla_g\phi,\nabla_g\phi))^{\alpha-1}\operatorname{Hess}_\phi(\nabla_g\phi), \tag{A.5}
$$

$$
\nabla_g(g(\nabla_g\phi,\nabla_g u)) = \operatorname{Hess}_u(\nabla_g\phi) + \operatorname{Hess}_\phi(\nabla_g u). \tag{A.6}
$$

Plugging (A.5) and (A.6) into (A.4) gives (A.3). ∎

We next show a q-Bochner identity for the operator A_ϕ^q defined as

$$
A_\phi^q(u) := (-g(\nabla_g\phi,\nabla_g\phi))^{\frac{q-2}{2}}\left(\square_g u - (q-2)\frac{\operatorname{Hess}_u(\nabla_g\phi,\nabla_g\phi)}{-g(\nabla_g\phi,\nabla_g\phi)}\right). \tag{A.7}
$$

Lemma A.3. *Under the above notation, the following identity holds:*

$$
\begin{aligned}
-L_\phi^q&\left(\frac{(-g(\nabla_g\phi,\nabla_g\phi))^{q/2}}{q}\right) + g(\nabla_g\square_g^q\phi,\nabla_g^q\phi) \\
&= (q-2)|g(\nabla_g\phi,\nabla_g\phi)|^{\frac{q-4}{2}}\square_g^q\phi\,\operatorname{Hess}_\phi(\nabla_g\phi,\nabla_g\phi) \\
&\quad + |g(\nabla_g\phi,\nabla_g\phi)|^{q-2}\bigg(q(q-2)\left(\frac{\operatorname{Hess}_\phi(\nabla_g\phi,\nabla_g\phi)}{|g(\nabla_g\phi,\nabla_g\phi)|}\right)^2 \\
&\qquad\qquad\qquad\qquad + \operatorname{Ric}(\nabla_g\phi,\nabla_g\phi) + g(\operatorname{Hess}_\phi,\operatorname{Hess}_\phi)\bigg).
\end{aligned} \tag{A.8}
$$

Proof. We perform the computation at an arbitrary point $x_0 \in U$. In order to simplify the computations, we consider normal coordinates (x^i) in a neighbourhood of x_0 with $\frac{\partial}{\partial x^1} \in \mathcal{C}$. We have

$$
\begin{aligned}
-\frac{\square_g(-g(\nabla_g\phi,\nabla_g\phi))^{q/2}}{q} &= g^{ij}\partial_i\big((-g(\nabla_g\phi,\nabla_g\phi))^{\frac{q-2}{2}}g^{kl}\partial_{jk}\phi\,\partial_l\phi\big) \\
&= (-g(\nabla_g\phi,\nabla_g\phi))^{\frac{q-2}{2}}\bigg(-\frac{q-2}{-g(\nabla_g\phi,\nabla_g\phi)}g^{ij}g^{mn}\partial_{im}\phi\partial_n\phi g^{kl}\partial_{jk}\phi\,\partial_l\phi \\
&\qquad\qquad + g^{ij}g^{kl}\partial_{ijk}\phi\,\partial_l\phi + g^{ij}g^{kl}\partial_{jk}\phi\,\partial_{li}\phi\bigg).
\end{aligned} \tag{A.9}
$$

Now, from the symmetry of second order derivatives and the very definition of the Riemann tensor (2.7), we have

$$
\partial_{ijk}\phi = \partial_{ikj}\phi = g(R(\partial_{x^i},\partial_{x^k})\nabla_g\phi,\partial_{x^j}) + \partial_{kij}\phi. \tag{A.10}
$$

Thus

$$g^{ij} g^{kl} \partial_{ijk}\phi \, \partial_l\phi = g(\nabla_g \Box_g \phi, \nabla_g \phi) + \mathrm{Ric}(\nabla_g \phi, \nabla_g \phi),$$

and we can rewrite (A.9) as

$$
-\frac{\Box_g(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q}
$$
$$
= (-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} \bigg((2-q)\frac{g(\mathrm{Hess}_\phi(\nabla_g\phi), \mathrm{Hess}_\phi(\nabla_g\phi))}{-g(\nabla_g\phi, \nabla_g\phi)}
$$
$$
+ g(\nabla_g \Box_g\phi, \nabla_g\phi) + \mathrm{Ric}(\nabla_g\phi, \nabla_g\phi) + g(\mathrm{Hess}_\phi, \mathrm{Hess}_\phi) \bigg). \qquad (A.11)
$$

We now compute the second part of $-L_\phi^q(\frac{(-g(\nabla_g\phi,\nabla_g\phi))^{q/2}}{q})$. To this end observe that

$$
\nabla_g\left(\frac{(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q} \right) = -(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} \, \mathrm{Hess}_\phi(\nabla_g\phi), \qquad (A.12)
$$
$$
\mathrm{Hess}_{\frac{(-g(\nabla_g\phi,\nabla_g\phi))^{q/2}}{q}}(\nabla_g\phi, \nabla_g\phi)
$$
$$
= (q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-4}{2}}[\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)]^2
$$
$$
- (-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} g(\nabla_{\nabla_g\phi}\nabla_{\nabla_g\phi}\nabla_g\phi, \nabla_g\phi). \qquad (A.13)
$$

It is useful to express $\Box_g^q$ in terms of $\Box_g$:

$$
\Box_g^q \phi := \mathrm{div}\big((-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} \nabla_g\phi \big)
$$
$$
= (-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} \left(\Box_g\phi - (q-2)\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{-g(\nabla_g\phi, \nabla_g\phi)} \right). \qquad (A.14)
$$

Using (A.14), we can write

$$
g(\nabla_g \Box_g^q\phi, \nabla_g\phi)
$$
$$
= g\left(\nabla_g\left((-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} \left(\Box_g\phi - (q-2)\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{-g(\nabla_g\phi, \nabla_g\phi)} \right) \right), \nabla_g\phi \right)
$$
$$
= (-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}} g(\nabla_g \Box_g\phi, \nabla_g\phi) - (q-2)\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{-g(\nabla_g\phi, \nabla_g\phi)} \Box_g^q\phi
$$
$$
- (q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-4}{2}} \big(g(\nabla_{\nabla_g\phi}\nabla_{\nabla_g\phi}\nabla_g\phi, \nabla_g\phi)
$$
$$
+ g(\mathrm{Hess}_\phi(\nabla_g\phi), \mathrm{Hess}_\phi(\nabla_g\phi)) \big)
$$
$$
- 2(q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-6}{2}}[\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)]^2. \qquad (A.15)
$$

Plugging (A.11) and (A.13) into (A.7), and simplifying using (A.15), gives the desired (A.8). ∎

Proof of Proposition A.1. Combining the expression of $L^q_\phi u$ as in (A.3) with the definition of $A^q_\phi u$ as in (A.7) we can write

$$
\begin{aligned}
L^q_\phi u = A^q_\phi u &+ (2-q)(-g(\nabla_g\phi, \nabla_g\phi))^{-1} g(\nabla_g\phi, \nabla_g u)\Box^q_g\phi \\
&+ 2(2-q)\,(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-4}{2}} \mathrm{Hess}_\phi\left(\nabla_g\phi, \nabla_g u + \frac{g(\nabla_g\phi, \nabla_g u)}{-g(\nabla_g\phi, \nabla_g\phi)}\nabla_g\phi\right).
\end{aligned}
\tag{A.16}
$$

Specializing (A.16) with $u = \frac{(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q}$ and using (A.12) gives

$$
\begin{aligned}
-L^q_\phi\left(\frac{(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q}\right) = &-A^q_\phi\left(\frac{(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q}\right) \\
&- (q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{\frac{q-2}{2}}\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{-g(\nabla_g\phi, \nabla_g\phi)}\Box^q_g\phi \\
&- 2(q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{q-3}\,\mathrm{Hess}_\phi\big(\nabla_g\phi, \mathrm{Hess}_\phi(\nabla_g\phi)\big) \\
&- 2(q-2)(-g(\nabla_g\phi, \nabla_g\phi))^{q-2}\left(\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{-g(\nabla_g\phi, \nabla_g\phi)}\right)^2.
\end{aligned}
\tag{A.17}
$$

Now, the combination of (A.17) and (A.8) yields

$$
\begin{aligned}
- L^q_\phi\left(\frac{(-g(\nabla_g\phi, \nabla_g\phi))^{q/2}}{q}\right) = &-g(\nabla_g\Box^q_g\phi, \nabla^q_g\phi) \\
&+ (q-2)^2|g(\nabla_g\phi, \nabla_g\phi)|^{q-2}\left(\frac{\mathrm{Hess}_\phi(\nabla_g\phi, \nabla_g\phi)}{|g(\nabla_g\phi, \nabla_g\phi)|}\right)^2 \\
&+ |g(\nabla_g\phi, \nabla_g\phi)|^{q-2}(\mathrm{Ric}(\nabla_g\phi, \nabla_g\phi) + g(\mathrm{Hess}_\phi, \mathrm{Hess}_\phi)) \\
&- 2(q-2)\,|g(\nabla_g\phi, \nabla_g\phi)|^{q-3}\mathrm{Hess}_\phi(\nabla_g\phi, \mathrm{Hess}_\phi(\nabla_g\phi)).
\end{aligned}
\tag{A.18}
$$

∎

Appendix B. A synthetic formulation of Einstein's vacuum equations in a non-smooth setting

B.1. The Lorentzian synthetic framework

The goal of this appendix in to give a synthetic formulation of Einstein's vacuum equations (i.e. zero stress-energy tensor $T \equiv 0$ but possibly non-zero cosmological constant Λ) and show its stability under a natural adaptation for Lorentzian synthetic spaces of the measured Gromov–Hausdorff convergence (classically designed as a notion of convergence for metric measure spaces).

This appendix was written in early 2021, more than two years after the rest of the paper was posted on arXiv (in October 2018). One of the reasons is that we will build on

the synthetic time-like Ricci curvature lower bounds in a non-smooth setting developed by the first author in collaboration with Cavalletti in [24] (posted on arXiv in 2020).

The general framework is given by Lorentzian pre-length/geodesic spaces. Let us start by recalling some basics of the theory, following the approach of Kunzinger–Sämann [54].

A *causal space* $(X, \ll, \leq)$ is a set X endowed with a preorder $\leq$ and a transitive relation $\ll$ contained in $\leq$.

We write $x < y$ if $x \leq y$ and $x \neq y$. When $x \ll y$ (resp. $x \leq y$), we say that x and y are *time-like* (resp. *causally*) related.

We define the *chronological* (resp. *causal*) *future* of a subset $E \subset X$ as

$$I^+(E) := \{y \in X \mid \exists x \in E, \, x \ll y\}, \quad J^+(E) := \{y \in X \mid \exists x \in E, \, x \leq y\}$$

respectively. Analogously, we define $I^-(E)$ (resp. $J^-(E)$) the *chronological* (resp. *causal*) *past* of E. In case $E = \{x\}$ is a singleton, with a slight abuse of notation, we will write $I^\pm(x)$ (resp. $J^\pm(x)$) instead of $I^\pm(\{x\})$ (resp. $J^\pm(\{x\})$).

Recall that a metric space (X, d) is said to be *proper* if closed and bounded subsets are compact.

Definition B.1 (Lorentzian pre-length space $(X, \mathsf{d}, \ll, \leq, \tau)$). A *Lorentzian pre-length space* $(X, \mathsf{d}, \ll, \leq, \tau)$ is a casual space $(X, \ll, \leq)$ additionally equipped with a proper metric d and a lower semicontinuous function $\tau : X \times X \to [0, \infty]$, called a *time-separation function*, satisfying

$$\tau(x, y) + \tau(y, z) \leq \tau(x, z), \quad \forall x \leq y \leq z \quad \text{(reverse triangle inequality)}$$
$$\tau(x, y) = 0 \text{ if } x \nleq y, \quad \tau(x, y) > 0 \Leftrightarrow x \ll y. \tag{B.1}$$

We will consider X endowed with the metric topology induced by d.

We say that X is (resp. *locally*) *causally closed* if $\{x \leq y\} \subset X \times X$ is a closed subset (resp. every point $x \in X$ has a neighbourhood U such that $\{x \leq y\} \cap \bar{U} \times \bar{U}$ is closed in $\bar{U} \times \bar{U}$).

Throughout this appendix, $I \subset \mathbb{R}$ will denote an arbitrary interval. A non-constant curve $\gamma : I \to X$ is called (future-directed) *time-like* (resp. *causal*) if γ is locally Lipschitz continuous (with respect to d) and $\gamma_{t_1} \ll \gamma_{t_2}$ (resp. $\gamma_{t_1} \leq \gamma_{t_2}$) for all $t_1, t_2 \in I$ with $t_1 < t_2$.

The length of a causal curve is defined via the time-separation function, in analogy to the theory of length metric spaces: for a future-directed causal curve $\gamma : [a, b] \to X$ we set

$$\mathrm{L}_\tau(\gamma) := \inf \left\{ \sum_{i=0}^{N-1} \tau(\gamma_{t_i}, \gamma_{t_{i+1}}) \, \Big| \, a = t_0 < t_1 < \cdots < t_N = b, \, N \in \mathbb{N} \right\}.$$

A future-directed causal curve $\gamma : [a, b] \to X$ is *maximal* (or *geodesic*) if it realizes the time separation, i.e. $\mathrm{L}_\tau(\gamma) = \tau(\gamma_a, \gamma_b)$.

A Lorentzian pre-length space $(X, \mathsf{d}, \ll, \leq, \tau)$ is called

- *non-totally imprisoning* if for every compact set $K \Subset X$ there is constant $C > 0$ such that the d-arclength of all causal curves contained in K is bounded by C;

- *globally hyperbolic* if it is non-totally imprisoning and for any $x, y \in X$ the set $J^+(x) \cap J^-(y)$ is compact in X;

- $\mathcal{K}$-*globally hyperbolic* if it is non-totally imprisoning and for any compact subsets $K_1, K_2 \Subset X$, the set $J^+(K_1) \cap J^-(K_2)$ is compact in X;

- *geodesic* if for all $x, y \in X$ with $x \leq y$ there is a future-directed causal curve γ from x to y with $\tau(x, y) = L_\tau(\gamma)$, i.e. a (maximizing) geodesic from x to y.

For a globally hyperbolic Lorentzian geodesic (actually length would suffice) space $(X, \mathsf{d}, \ll, \leq, \tau)$, the time-separation function τ is finite and continuous [54, Theorem 3.28]. Moreover, any globally hyperbolic Lorentzian length space (for the definition of Lorenzian length space see [54, Definition 3.22], we omit it for brevity since we will not use it) is geodesic [54, Theorem 3.30].

A *measured Lorentzian pre-length space* $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ is a Lorentzian pre-length space endowed with a Radon non-negative measure $\mathfrak{m}$ with supp $\mathfrak{m} = X$.

Examples entering the class of Lorentzian synthetic spaces

In this short section we briefly recall some notable examples entering the aforementioned framework of Lorentzian synthetic spaces.

Space-times with a continuous Lorentzian metric. Let M be a smooth manifold endowed with a continuous Lorentzian metric g. Assume that (M, g) is time-oriented, i.e. there is a continuous time-like vector field. Observe that, for C^0 metrics, the natural class of differentiability of the underlying manifolds is C^1; now, C^1 manifolds always admit a C^∞ subatlas, and one can pick such a subatlas whenever convenient.

A causal (respectively time-like) curve in M is by definition a locally Lipschitz curve whose tangent vector is causal (resp. time-like) almost everywhere. One could also start from absolutely continuous (AC for short) curves, but since causal AC curves always admit a Lipschitz reparametrization [62, Section 2.1, Remark 2.3], we do not loose any generality with the above convention. Set $L_g(\gamma)$ to be the g-length of a causal curve $\gamma : I \to M$, i.e.

$$L_g(\gamma) := \int_I \sqrt{-g(\dot{\gamma}, \dot{\gamma})} \, dt.$$

The time-separation function $\tau : M \times M \to [0, \infty]$ is then defined in the classical way, i.e.

$$\tau(x, y) := \sup \{L_g(\gamma) : \gamma \text{ is future-directed causal from } x \text{ to } y\} \quad \text{if } x \leq y,$$

and $\tau(x, y) = 0$ otherwise. The reverse triangle inequality (B.1) follows directly from the definition. Moreover, every L_g-maximal curve γ is also L_τ-maximal, and $L_g(\gamma) = L_\tau(\gamma)$ (see for instance [54, Remark 5.1]). In order to have an underlying metric structure, we also fix a complete Riemannian metric h on M and denote by d^h the associated distance function.

For a space-time with a Lorentzian C^0 metric:

- Global hyperbolicity implies causal closedness and $\mathcal{K}$-global hyperbolicity [71, Proposition 3.3 and Corollary 3.4].

- Recall that a *Cauchy hypersurface* is a subset which is met exactly once by every inextendible causal curve. Every Cauchy hypersurface is a closed casual topological hypersurface [71, Proposition 5.2]. Global hyperbolicity is equivalent to the existence of a Cauchy hypersurface [71, Theorems 5.7 and 5.9].

- By [54, Proposition 5.8], if g is a causally plain (or, more strongly, locally Lipschitz) Lorentzian C^0 metric on M then the associated synthetic structure is a pre-length Lorentzian space. More strongly, from [54, Theorems 3.30 and 5.12], if g is a globally hyperbolic and causally plain Lorentzian C^0 metric on M then the associated synthetic structure is a causally closed, $\mathcal{K}$-globally hyperbolic Lorentzian geodesic space.

Closed cone structures. Closed cone structures provide a rich source of examples of Lorentzian pre-length and length spaces, which can be seen as the synthetic-Lorentzian analogue of Finsler manifolds (see [54, Section 5.2] for more details). See Minguzzi's review paper [62] for a comprehensive analysis of causality theory in the framework of closed cone structures, including embedding and singularity theorems.

Some examples towards quantum gravity. The framework of Lorentzian pre-length spaces allows one to handle situations where one may not have the structure of a manifold or a Lorentz(–Finsler) metric. A remarkable example is given by certain approaches to quantum gravity (see for instance [58]) where it is shown that it is possible to reconstruct a globally hyperbolic space-time and the causality relation from a countable dense set of events, in a purely order-theoretic manner.

Two approaches to quantum gravity, particularly linked to Lorentzian pre-length spaces, are the theory of *causal Fermion systems* [34,35] and the *theory of causal sets* [13]. The basic idea in both cases is that the structure of space-time needs to be adjusted on a microscopic scale to include quantum effects. This leads to non-smoothness of the underlying geometry, and the classical structure of Lorentzian manifold emerges only in the macroscopic regime. For the connection to the theory of Lorentzian (pre-)length spaces we refer to [54, Section 5.3], [35, Section 5.1].

B.2. The p-Lorentz–Wasserstein distance on a Lorentzian pre-length space

We denote with $\mathcal{P}(X)$ (resp. $\mathcal{P}_c(X)$, or $\mathcal{P}_{ac}(X)$) the collection of all Borel probability measures (resp. with compact support, or absolutely continuous with respect to $\mathfrak{m}$).

Given a probability measure $\mu \in \mathcal{P}(X)$ we define its relative entropy by

$$\mathrm{Ent}(\mu|\mathfrak{m}) = \int_X \varrho \log(\varrho) \, d\mathfrak{m}$$

if $\mu = \varrho \, \mathfrak{m} \in \mathcal{P}_{ac}(X)$ and $(\varrho \log(\varrho))_+$ is $\mathfrak{m}$-integrable. Otherwise we set $\mathrm{Ent}(\mu|\mathfrak{m}) = +\infty$. We let $\mathrm{Dom}(\mathrm{Ent}(\cdot|\mathfrak{m}))$ be the finiteness domain of $\mathrm{Ent}(\cdot|\mathfrak{m})$.

We next recall some basics of optimal transport theory in Lorentzian pre-length spaces, following the approach and the notation in [24].

Given $\mu, \nu \in \mathcal{P}(X)$, denote

$$\Pi(\mu, \nu) := \{\pi \in \mathcal{P}(X \times X) \mid (p_1)_\sharp \pi = \mu, \ (p_2)_\sharp \pi = \nu\},$$
$$\Pi_\leq(\mu, \nu) := \{\pi \in \Pi(\mu, \nu) \mid \pi(X_\leq^2) = 1\},$$
$$\Pi_\ll(\mu, \nu) := \{\pi \in \Pi(\mu, \nu) \mid \pi(X_\ll^2) = 1\},$$

where $X_\leq^2 := \{(x, y) \in X^2 \mid x \leq y\}$ and $X_\ll^2 := \{(x, y) \in X^2 \mid x \ll y\}$.

Definition B.2 (The p-Lorentz–Wasserstein distance). Let $(X, \mathsf{d}, \ll, \leq, \tau)$ be a Lorentzian pre-length space and let $p \in (0, 1]$. Given $\mu, \nu \in \mathcal{P}(X)$, the p-*Lorentz–Wasserstein distance* is defined by

$$\ell_p(\mu, \nu) := \sup_{\pi \in \Pi_\leq(\mu, \nu)} \left(\int_{X \times X} \tau(x, y)^p \, d\pi(x, y) \right)^{1/p}. \tag{B.2}$$

When $\Pi_\leq(\mu, \nu) = \emptyset$ we set $\ell_p(\mu, \nu) := -\infty$.

Note that Definition B.2 extends to Lorentzian pre-length spaces the corresponding notion given in the smooth Lorentzian setting in [28] (see also [61], and [76] for $p = 1$); when $\Pi_\leq(\mu, \nu) = \emptyset$ we adopt the convention of McCann [61] (note that [28] sets $\ell_p(\mu, \nu) = 0$ in this case).

A key property of ℓ_p is the reverse triangle inequality. This was proved in the smooth Lorentzian setting by Eckstein–Miller [28, Theorem 13] and in the present synthetic setting in [24, Proposition 2.5]. Such a property is the natural Lorentzian analogue of the fact that the Kantorovich–Rubinstein–Wasserstein distances W_p, $p \geq 1$, in the metric space setting satisfy the usual triangle inequality (see for instance [80, Section 6]).

Proposition B.3 (ℓ_p satisfies the reverse triangle inequality). *Let $(X, \mathsf{d}, \ll, \leq, \tau)$ be a Lorentzian pre-length space and let $p \in (0, 1]$. Then ℓ_p satisfies the reverse triangle inequality*

$$\ell_p(\mu_0, \mu_1) + \ell_p(\mu_1, \mu_2) \leq \ell_p(\mu_0, \mu_2), \quad \forall \mu_0, \mu_1, \mu_2 \in \mathcal{P}(X), \tag{B.3}$$

where we adopt the convention that $\infty - \infty = -\infty$ on the left hand side.

A coupling $\pi \in \Pi_\leq(\mu, \nu)$ maximizing (B.2) is called ℓ_p-*optimal*. The set of ℓ_p-optimal couplings from μ to ν is denoted by $\Pi_\leq^{p\text{-opt}}(\mu, \nu)$. We also set $\Pi_\ll^{p\text{-opt}}(\mu, \nu) := \Pi_\ll(\mu, \nu) \cap \Pi_\leq^{p\text{-opt}}(\mu, \nu)$.

We say that $(\mu_s)_{s \in [0,1]} \subset \mathcal{P}(X)$ is an ℓ_p-*geodesic* if

$$\ell_p(\mu_s, \mu_t) = (t - s)\ell_p(\mu_0, \mu_1) \in (0, \infty) \quad \text{for all } 0 \leq s < t \leq 1. \tag{B.4}$$

Note that, with the convention above, ℓ_p-geodesics are implicitly future-directed and time-like.

The next lemma follows by the classical theory of optimal transport, since the assumptions allow one to localize the argument on $\operatorname{supp} \mu_0 \times \operatorname{supp} \mu_1 \subset X_\leq^2$ where the optimal transport problem becomes standard.

Lemma B.4. *Assume* $\operatorname{supp} \mu \times \operatorname{supp} \nu \subset X_{\leq}^2$.

(1) *Every ℓ_p-optimal coupling $\pi \in \Pi_{\leq}^{p\text{-opt}}(\mu, \nu)$ is also τ^p-cyclically monotone.*

(2) *If $\pi \in \Pi_{\leq}(\mu, \nu)$ is τ^p-cyclically monotone then $\pi \in \Pi_{\leq}^{p\text{-opt}}(\mu, \nu)$ is ℓ_p-optimal.*

B.3. Synthetic time-like Ricci upper bounds

Inspired by Sturm's approach to Riemannian/metric Ricci curvature upper bounds [75] and by Theorem 4.7 (see also Remark 4.8), we propose the following synthetic notion of time-like Ricci curvature bounded above. We denote by $B_r^{\mathsf{d}}(x)$ the metric ball in (X, d) with centre x and radius r.

Definition B.5 (Synthetic time-like Ricci curvature bounded above). Fix $p \in (0, 1)$ and $K \in \mathbb{R}$. We say that a measured Lorentzian pre-length space $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ *has time-like Ricci curvature bounded above by K in the synthetic sense* if there exist $r_0 > 0$ and a function $\omega : [0, r_0) \to [0, \infty)$ with $\lim_{r \downarrow 0} \omega(r) = 0$ such that for every $r \in [0, r_0)$ the following holds.

For any $x, y \in X$ with $\mathsf{d}(x, y) = r > 0$ and such that $B_{r4}^{\mathsf{d}}(x) \times B_{r2}^{\mathsf{d}}(y) \subset X_{\ll}^2$, and for every $\mu_0 \in \operatorname{Dom}(\operatorname{Ent}(\cdot|\mathfrak{m})$ with $\operatorname{supp} \mu_0 \subset B_{r4}^{\mathsf{d}}(x)$, there exists an ℓ_p-geodesic $(\mu_t)_{t \in [-1,1]}$ satisfying

- $\operatorname{supp} \mu_1 \subset B_{r2}^{\mathsf{d}}(y)$,
- $\operatorname{supp} \mu_{-1} \times \operatorname{supp} \mu_1 \subset X_{\leq}^2$,
- $\bigcup_{t \in [-1,1]} \operatorname{supp} \mu_t \subset B_{10 r_0}^{\mathsf{d}}(x)$,
- $\operatorname{Ent}(\mu_{-1}|\mathfrak{m}) - 2 \operatorname{Ent}(\mu_0|\mathfrak{m}) + \operatorname{Ent}(\mu_1|\mathfrak{m}) \leq (K + \omega(r)) r^2$.

A key property of the above notion of time-like Ricci bounded above is the stability under weak convergence of measured Lorentzian pre-length spaces. The stability of Riemannian/metric Ricci upper bounds via optimal transport was proved by Sturm [75]. The notion of convergence we use below is a slight reinforcement (we ask that the immersion maps are isometries with respect to the metric structures instead of merely topological embedding maps) of the weak convergence used in [24] to show stability of synthetic time-like Ricci *lower* bounds; in any case it is a natural adaptation to the Lorentzian setting of the mGH convergence used for metric measure spaces (see for instance [38, Section 3] for an overview of equivalent formulations of convergence for metric measure spaces).

Theorem B.6 (Stability of time-like Ricci curvature upper bounds). *Let $\{(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)\}_{j \in \mathbb{N} \cup \{\infty\}}$ be a sequence of measured Lorentzian geodesic spaces satisfying the following properties:*

(1) *There exists a locally causally closed, globally hyperbolic Lorentzian geodesic space $(X, \mathsf{d}, \ll, \leq, \tau)$ such that each $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$, $j \in \mathbb{N} \cup \{\infty\}$, is isomorphically embedded in it, i.e. there exist inclusion maps $\iota_j : X_j \hookrightarrow X$ such that for all $x_j^1, x_j^2 \in X_j$ and $j \in \mathbb{N} \cup \{\infty\}$, the following holds:*

- $\mathsf{d}(\iota_j(x_j^1), \iota_j(x_j^2)) = \mathsf{d}_j(x_j^1, x_j^2)$;

- $x_j^1 \leq_j x_j^2$ if and only if $\iota_j(x_j^1) \leq \iota_j(x_j^2)$;
- $\tau(\iota_j(x_j^1), \iota_j(x_j^2)) = \tau_j(x_j^1, x_j^2)$.

(2) *The measures $(\iota_j)_\sharp \mathfrak{m}_j$ converge to $(\iota_\infty)_\sharp \mathfrak{m}_\infty$ weakly in duality with $C_c(X)$ in X, i.e.*

$$\int \varphi \, d(\iota_j)_\sharp \mathfrak{m}_j \to \int \varphi \, d(\iota_\infty)_\sharp \mathfrak{m}_\infty, \quad \forall \varphi \in C_c(X), \tag{B.5}$$

where $C_c(X)$ denotes the set of continuous functions with compact support.

(3) *Volume non-collapsing: there exists a function $v : (0, \infty) \to (0, \infty)$ such that*

$$\mathfrak{m}_j(B_r^{\mathsf{d}_j}(x_j)) \geq v(r) > 0 \quad \text{for every } x_j \in X_j.$$

(4) *There exists a function $\omega : [0, \infty) \to [0, \infty)$ with $\lim_{r\downarrow 0} \omega(r) = 0$ and there exist $p \in (0, 1)$, $K \in \mathbb{R}$ such that $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$ has time-like Ricci curvature bounded above by K with respect to $p \in (0, 1)$, $r_0 > 0$ and with remainder function ω in the synthetic sense of Definition B.5.*

Then also the limit space $(X_\infty, \mathsf{d}_\infty, \mathfrak{m}_\infty, \ll_\infty, \leq_\infty, \tau_\infty)$ has time-like Ricci curvature bounded above by K with respect to $p \in (0, 1)$, $r_0 + 1$ and with remainder function ω in the synthetic sense.

Proof. For simplicity of notation, we will identify X_j with its isomorphic image $\iota_j(X_j) \subset X$ and the measure $\mathfrak{m}_j$ with $(\iota_j)_\sharp \mathfrak{m}_j$, for each $j \in \mathbb{N} \cup \{\infty\}$.

We will write $B_r^{\mathsf{d}_j}(x)$ for the ball in (X_j, d_j) centred at $x \in X_j$. Notice that $B_r^{\mathsf{d}_j}(x) = B_r^{\mathsf{d}}(x) \cap X_j$.

Fix any $x_\infty, y_\infty \in X_\infty$ with $\mathsf{d}_\infty(x_\infty, y_\infty) = r > 0$ and $B_{r^4}^{\mathsf{d}_\infty}(x_\infty) \times B_{r^2}^{\mathsf{d}_\infty}(y_\infty) \subset (X_\infty^2)_\ll$. Since $(X_\infty^2)_\ll \subset X_\infty \times X_\infty$ is an open subset, there exists $\varepsilon > 0$ such that

$$B_{(r+\varepsilon)^4}^{\mathsf{d}_\infty}(x_\infty) \times B_{(r+\varepsilon)^2}^{\mathsf{d}_\infty}(y_\infty) \subset (X_\infty^2)_\ll. \tag{B.6}$$

Fix $\mu_0^\infty \in \text{Dom}(\text{Ent}(\cdot|\mathfrak{m}_\infty))$ with $\text{supp}\,\mu_0^\infty \subset B_{r^4}^{\mathsf{d}_\infty}(x_\infty)$.

Step 1. We claim that, for $j \in \mathbb{N}$ sufficiently large, there exist $x_j, y_j \in X_j$ such that

(1) $x_j \to x_\infty$ and $y_j \to y_\infty$ in (X, d);

(2) $\mathsf{d}_j(x_j, y_j) =: r_j \to r$;

(3) $B_{r_j^4}^{\mathsf{d}_j}(x_j) \times B_{r_j^2}^{\mathsf{d}_j}(y_j) \subset (X_j^2)_\ll$.

Since by assumption $\mathfrak{m}_j \rightharpoonup \mathfrak{m}_\infty$ weakly as measures in (X, d), by the lower semicontinuity over an open subset we find that for every $r > 0$,

$$0 < \mathfrak{m}_\infty(B_r^{\mathsf{d}_\infty}(x_\infty)) = \mathfrak{m}_\infty(B_r^{\mathsf{d}}(x_\infty)) \leq \liminf_{j \to \infty} \mathfrak{m}_j(B_r^{\mathsf{d}}(x_\infty)).$$

In particular, for every $r > 0$ there exists $J(r) > 0$ such that $B_r^{\mathsf{d}}(x_\infty) \cap X_j \neq \emptyset$ for all $j \geq J(r)$. By a diagonal argument, there exist $x_j \in X_j$ with $x_j \to x_\infty$ in (X, d). An

analogous argument gives a sequence $y_j \in X_j$ with $y_j \to y_\infty$. Since the inclusion maps ι_j are isometric, we also get

$$r_j := \mathsf{d}_j(x_j, y_j) = \mathsf{d}(x_j, y_j) \to \mathsf{d}(x_\infty, y_\infty) = r.$$

It remains to show the third claim. For every $j \in \mathbb{N}$, let $x_j' \in B^{\mathsf{d}_j}_{r_j^4}(x_j)$ and $y_j' \in B^{\mathsf{d}_j}_{r_j^2}(y_j)$ be arbitrary. Combining the volume non-collapsing assumption with the weak convergence $\mathfrak{m}_j \rightharpoonup \mathfrak{m}_\infty$ and the properness of the metrics, we find that there exist $x_\infty' \in B^{\mathsf{d}_\infty}_{(r+\varepsilon)^4}(x_\infty)$ and $y_\infty' \in B^{\mathsf{d}_\infty}_{(r+\varepsilon)^2}(y_\infty)$ such that $x_j' \to x_\infty'$ and $y_j' \to y_\infty'$ up to a subsequence. Recalling (B.6) and that ι_∞ is an isomorphic embedding, we infer that $(x_\infty', y_\infty') \in X_\ll^2$. Since $X_\ll^2 \subset X^2$ is an open subset, we conclude that for j sufficiently large,

$$(x_j', y_j') \subset X_\ll^2 \cap X_j^2 = (X_j^2)_\ll$$

as desired.

Step 2. We claim that, for every $j \in \mathbb{N}$, there exists $\mu_0^j \in \text{Dom}(\text{Ent}(\,\cdot\,|\mathfrak{m}_j))$ with $\text{supp}\,\mu_0^j \subset B^{\mathsf{d}_j}_{r_j^4}(x_j)$ such that $\mu_0^j \to \mu_0^\infty$ narrowly and

$$\text{Ent}(\mu_0^\infty|\mathfrak{m}_\infty) \geq \limsup_{j \to \infty} \text{Ent}(\mu_0^j|\mathfrak{m}_j). \tag{B.7}$$

Since $\mathfrak{m}_j \rightharpoonup \mathfrak{m}_\infty$ weakly, there exists $R \in (2r^4, 3r^4)$ such that

$$z_j^{-1} := \mathfrak{m}_j(B^{\mathsf{d}}_R(x_\infty)) \to \mathfrak{m}_\infty(B^{\mathsf{d}}_R(x_\infty)) =: z_\infty^{-1}. \tag{B.8}$$

For $j \in \mathbb{N} \cup \{\infty\}$, denote $\tilde{\mathfrak{m}}_j := z_j \, \mathfrak{m}_j \llcorner B^{\mathsf{d}}_R(x_\infty)$ and observe that $\tilde{\mathfrak{m}}_j \rightharpoonup \tilde{\mathfrak{m}}_\infty$ weakly and thus (since now the supports are uniformly bounded in X) in $W_q^{(X,\mathsf{d})}$ for some (or equivalently every) $q \in [1, \infty)$. In particular $\tilde{\mathfrak{m}}_j \to \tilde{\mathfrak{m}}_\infty$ in $W_2^{(X,\mathsf{d})}$. Suppose

$$\eta_j \in \Pi(\tilde{\mathfrak{m}}_\infty, \tilde{\mathfrak{m}}_j) \text{ is a } W_2^{(X,\mathsf{d})}\text{-optimal coupling.} \tag{B.9}$$

Let us first consider the case $\mu_0^\infty = \rho \, \tilde{\mathfrak{m}}_\infty$ with $\rho \in L^\infty(\tilde{\mathfrak{m}}_\infty)$. Define $\eta_j' \in \mathcal{P}(X)$ as $d\eta_j'(x, y) := \rho(x) \, d\eta_j(x, y)$ and $\tilde{\mu}_0^j := (p_2)_\sharp \eta_j' \in \mathcal{P}(X_j)$. By construction, $\eta_j' \ll \eta_j$, hence $\tilde{\mu}_0^j \ll (p_2)_\sharp \eta_j = \tilde{\mathfrak{m}}_j$. Let $\tilde{\mu}_0^j = \tilde{\rho}_0^j \tilde{\mathfrak{m}}_j$. It is readily checked from the definition that $\tilde{\rho}_0^j(y) = \int \rho(x) \, d(\eta_j)_y(x)$, where $\{(\eta_j)_y\}$ is the disintegration of η_j with respect to the projection on the second marginal. By Jensen's inequality applied to the convex function $u(s) := s \log(s)$, we have

$$
\begin{aligned}
\text{Ent}(\tilde{\mu}_0^j|\tilde{\mathfrak{m}}_j) = \int u(\tilde{\rho}_0^j) \, d\tilde{\mathfrak{m}}_j &= \int u\left(\int \rho(x) \, d(\eta_j)_y(x) \right) d\tilde{\mathfrak{m}}_j(y) \\
&\leq \int \int u(\rho(x)) \, d(\eta_j)_y(x) \, d\tilde{\mathfrak{m}}_j(y) = \int \int u(\rho(x)) \, d\eta_j(x, y) \\
&= \int u(\rho) \, d(p_1)_\sharp \eta_j = \int u(\rho) \, d\tilde{\mathfrak{m}}_\infty = \text{Ent}(\mu_0^\infty|\tilde{\mathfrak{m}}_\infty). \tag{B.10}
\end{aligned}
$$

Since by construction we have $\gamma'_j \in \Pi(\mu_0^\infty, \tilde{\mu}_0^j)$, it follows that

$$W_2(\mu_0^\infty, \tilde{\mu}_0^j)^2 \leq \int \mathsf{d}(x,y)^2 \, d\eta'_j(x,y) = \int \rho(x) \mathsf{d}(x,y)^2 \, d\eta_j(x,y)$$
$$\leq \|\rho\|_{L^\infty(\tilde{\mathfrak{m}}_\infty)} W_2(\tilde{\mathfrak{m}}_\infty, \tilde{\mathfrak{m}}_j)^2,$$

and therefore $W_2(\mu_0^\infty, \tilde{\mu}_0^j) \to 0$. In particular, $\tilde{\mu}_0^j \to \mu_0^\infty$ narrowly in (X, d). Using the fact that $x_j \to x_\infty$, $r_j \to r$ and supp $\mu_0^\infty \subset B_{r^4}^{\mathsf{d}_\infty}(x_\infty)$, it follows that $c_j^{-1} := \tilde{\mu}_0^j(B_{(r'_j)^4}^{\mathsf{d}_j}(x_j)) \to 1$, where we set $r'_j := r_j - 1/j$. Define $\mu_0^j := c_j \, \tilde{\mu}_0^j \llcorner B_{(r'_j)^4}^{\mathsf{d}_j}(x_j)$.

Clearly, $\mu_0^j \in \mathrm{Dom}(\mathrm{Ent}(\,\cdot\,|\mathfrak{m}_j))$, supp $\mu_0^j \subset B_{r_j^4}^{\mathsf{d}_j}(x_j)$, $W_2(\mu_0^\infty, \mu_0^j) \to 0$ and

$$\lim_{j \to \infty} \left| \mathrm{Ent}(\mu_0^j | \tilde{\mathfrak{m}}_j) - \mathrm{Ent}(\tilde{\mu}_0^j | \tilde{\mathfrak{m}}_j) \right| = 0. \tag{B.11}$$

The combination of (B.8), (B.10) and (B.11) gives the desired (B.7).

If ρ is not bounded, for $k \in \mathbb{N}$ define $\rho^k := C_k \min\{\rho, k\}$, C_k being such that $\mu_\infty^k := \rho^k \, \tilde{\mathfrak{m}}_\infty \in \mathscr{P}(X_\infty)$. Clearly,

$$\limsup_{k \to \infty} \mathrm{Ent}(\mu_\infty^k | \mathfrak{m}_\infty) \leq \mathrm{Ent}(\mu_\infty | \mathfrak{m}_\infty), \quad \lim_{k \to \infty} W_2(\mu_\infty^k, \mu_\infty) = 0.$$

Then apply the previous procedure to μ_∞^k and conclude with a diagonal argument.

Step 3. *Conclusion.* By assumption, for every $j \in \mathbb{N}$ the space $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$ has time-like Ricci curvature bounded above by K with respect to $p \in (0, 1)$, $r_0 > 0$ and with remainder function ω. Thus there exists an ℓ_p-geodesic $(\mu_t^j)_{t \in [-1,1]}$ with supp $\mu_1^j \subset B_{r_j^2}^{\mathsf{d}_j}(y_j)$, $\bigcup_{t \in [-1,1]} \mathrm{supp} \, \mu_t^j \subset B_{10r_0}^{\mathsf{d}_j}(x_j)$ such that

$$\mathrm{Ent}(\mu_{-1}|\mathfrak{m}_j) - 2\,\mathrm{Ent}(\mu_0|\mathfrak{m}_j) + \mathrm{Ent}(\mu_1|\mathfrak{m}_j) \leq (K + \omega(r))r^2. \tag{B.12}$$

Since by construction $x_j \to x_\infty$, there exists $J > 0$ such that

$$\bigcup_{j \geq J} \bigcup_{t \in [-1,1]} \mathrm{supp} \, \mu_t^j \subset B_{11r_0}^{\mathsf{d}}(x_\infty). \tag{B.13}$$

Let $\eta_j \in \mathscr{P}(C[-1,1], (X, \mathsf{d}))$ be the dynamical plan associated to the geodesic $(\mu_t^j)_{t \in [-1,1]}$, i.e. $(e_t)_\# \eta_j = \mu_t^j$ for every $t \in [0,1]$, $j \in \mathbb{N}$. From the non-total imprisoning property of X, we infer that there exists $C > 0$ such that

$$\sup \{\mathsf{L}_{\mathsf{d}}(\gamma) \mid \gamma \in \mathrm{supp} \, \eta_j, \; j \in \mathbb{N}\} \leq C < \infty.$$

From [55, Theorem 4], we infer that $(\mu_t^j)_{t \in [-1,1]}$ is an absolutely continuous curve in the W_1-Kantorovich–Wasserstein space $(\mathscr{P}(X), W_1)$ with respect to d, with bounded length:

$$\mathsf{L}_{W_1}((\mu_t^j)_{t \in [-1,1]}) \leq \int \mathsf{L}_{\mathsf{d}}(\gamma) \, d\eta_j(\gamma) \leq C < \infty, \quad \forall j \in \mathbb{N}.$$

By the metric Arzelà–Ascoli Theorem we deduce that there exists a limit continuous curve $(\mu_t^\infty)_{t\in[-1,1]} \subset \mathcal{P}(X_\infty) \cap \mathcal{P}(X, W_1^{(X,d)})$ such that (up to a subsequence) $W_1^{(X,d)}(\mu_t^j, \mu_t^\infty) \to 0$ and thus $\mu_t^j \to \mu_t^\infty$ narrowly in X as $j \to \infty$, for every $t \in [0,1]$. From (B.13) we have

$$\bigcup_{t\in[-1,1]} \operatorname{supp}\mu_t^\infty \subset B_{11r_0}^d(x_\infty) \subset B_{10(r_0+1)}^d(x_\infty).$$

By assumption, we also know that for every $j \in \mathbb{N}$,

$$\operatorname{supp}\mu_{-1}^j \times \operatorname{supp}\mu_1^j \subset (X_j^2)_\leq = X_\leq^2 \cap X_j^2, \quad \forall j \in \mathbb{N}. \tag{B.14}$$

Since by assumption X is locally causally closed, we infer that

$$\operatorname{supp}\mu_{-1}^\infty \times \operatorname{supp}\mu_1^\infty \subset (X_\infty^2)_\leq = X_\leq^2 \cap X_\infty^2. \tag{B.15}$$

For every $j \in \mathbb{N}$, let $\pi_j \in \Pi_\leq^{p\text{-opt}}(\mu_{-1}^j, \mu_1^j)$. Thanks to Lemma B.4 we know that π_j is τ^p-cyclically monotone.

From (B.13), we see that $\operatorname{supp}\pi_j \subset B_{11r_0}^d(x_\infty)^2$ for $j \geq J$ and thus (by the Prokhorov Theorem) there exists $\pi_\infty \in \mathcal{P}(X \times X)$ such that $\pi_j \to \pi_\infty$ narrowly. Using the continuity of the projection maps, it is readily seen that $\pi_\infty \in \Pi(\mu_{-1}^\infty, \mu_1^\infty)$ and thus $\pi_\infty \in \Pi_\leq(\mu_{-1}^\infty, \mu_1^\infty)$, thanks to (B.15).

Since by global hyperbolicity τ^p is continuous, the τ^p-cyclical monotonicity is preserved under narrow convergence. We infer that π_∞ is τ^p-cyclically monotone and thus ℓ_p-optimal thanks to (B.15) and Lemma B.4. Therefore,

$$\ell_p(\mu_{-1}^j, \mu_1^j) = \left(\int \tau(x,y)^p \, d\pi_j(x,y)\right)^{1/p}$$

$$\to \left(\int \tau(x,y)^p \, d\pi_\infty(x,y)\right)^{1/p} = \ell_p(\mu_{-1}^\infty, \mu_1^\infty). \tag{B.16}$$

For intermediate $t \in (-1,1)$, since $(\mu_t^j)_{t\in[-1,1]}$ is an ℓ_p-geodesic, we have

$$\ell_p(\mu_{-1}^\infty, \mu_t^\infty) \geq \lim_{j\to\infty} \ell_p(\mu_{-1}^j, \mu_t^j) = (t+1) \lim_{j\to\infty} \ell_p(\mu_{-1}^j, \mu_1^j)$$

$$= (t+1)\,\ell_p(\mu_{-1}^\infty, \mu_1^\infty). \tag{B.17}$$

By the reverse triangle inequality (B.3), we find that the curve $(\mu_t^\infty)_{t\in[-1,1]}$ is an ℓ_p-geodesic from μ_{-1}^∞ to μ_1^∞.

The joint lower semicontinuity of $\operatorname{Ent}(\cdot|\cdot)$ under narrow convergence of probability measures (a well known fact, for a proof see for instance [3, Lemma 9.4.3]) yields

$$\operatorname{Ent}(\mu_t^\infty|\tilde{\mathfrak{m}}_\infty) \leq \liminf_{j\in\mathbb{N}} \operatorname{Ent}(\mu_t^j|\tilde{\mathfrak{m}}_j), \quad \forall t \in [0,1],$$

and thus, by (B.8),

$$\operatorname{Ent}(\mu_t^\infty|\mathfrak{m}_\infty) \leq \liminf_{j\in\mathbb{N}} \operatorname{Ent}(\mu_t^j|\mathfrak{m}_j), \quad \forall t \in [0,1]. \tag{B.18}$$

The combination of (B.7), (B.12) and (B.18) gives

$$\mathrm{Ent}(\mu_{-1}^{\infty}|\mathfrak{m}_{\infty}) - 2\,\mathrm{Ent}(\mu_0^{\infty}|\mathfrak{m}_{\infty}) + \mathrm{Ent}(\mu_1^{\infty}|\mathfrak{m}_{\infty}) \leq (K + \omega(r))r^2$$

as desired. ∎

B.4. Synthetic time-like Ricci lower bounds

In this subsection we briefly recall some basics of the synthetic theory of time-like Ricci lower bounds developed in [24].

Definition B.7 (Time-like p-dualizable measures). Let $(X, \mathsf{d}, \ll, \leq, \tau)$ be a Lorentzian pre-length space and fix $p \in (0, 1]$. We say that $(\mu, \nu) \in \mathcal{P}(X)^2$ is *time-like p-dualizable* (*by* $\pi \in \Pi_{\ll}(\mu, \nu)$) if

(1) $\ell_p(\mu, \nu) \in (0, \infty)$;

(2) $\pi \in \Pi_{\leq}^{p\text{-opt}}(\mu, \nu)$ and $\pi(X_{\ll}^2) = 1$;

(3) there exist measurable functions $a, b : X \to \mathbb{R}$ with $a \oplus b \in L^1(\mu \otimes \nu)$ such that $\ell^p \leq a \oplus b$ on supp $\mu \times$ supp ν.

The pair $(\mu, \nu) \in (\mathcal{P}(X))^2$ is said to be *strongly time-like p-dualizable* if in addition there exists a measurable ℓ^p-cyclically monotone set $\Gamma \subset X_{\ll}^2 \cap (\text{supp}\,\mu \times \text{supp}\,\nu)$ such that a coupling $\pi \in \Pi_{\leq}(\mu, \nu)$ is ℓ_p-optimal if and only if it is concentrated on Γ, i.e. $\pi(\Gamma) = 1$.

The motivation for considering (strongly) time-like p-dualizable pairs of measures is twofold: firstly, the p-optimal coupling $d\pi(x, y)$ matches events described by $d\mu(x)$ with events described by $d\nu(y)$ so that $x \ll y$; secondly, Kantorovich duality holds (see [76, Proposition 2.7] in smooth Lorentzian setting and in case $p = 1$ and [24, Section 2.4] for the non-smooth setting and general $p \in (0, 1]$).

Motivated by the work of McCann [61] and by Theorem 4.3, in [24] the following synthetic notion of time-like Ricci curvature bounded below by $K \in \mathbb{R}$ and dimension bounded above by $N \in (0, \infty]$ was proposed:

Definition B.8 ($\mathsf{TCD}_p^e(K, N)$ and $\mathsf{wTCD}_p^e(K, N)$ conditions). Fix $p \in (0, 1)$, $K \in \mathbb{R}$, $N \in (0, \infty]$. We say that a measured Lorentzian pre-length space $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ satisfies $\mathsf{TCD}_p^e(K, N)$ (resp. $\mathsf{wTCD}_p^e(K, N)$) if the following holds. For any couple $(\mu_0, \mu_1) \in (\mathrm{Dom}(\mathrm{Ent}(\cdot|\mathfrak{m})))^2$ which is time-like p-dualizable (resp. $(\mu_0, \mu_1) \in [\mathrm{Dom}(\mathrm{Ent}(\cdot|\mathfrak{m})) \cap \mathcal{P}_c(X)]^2$ which is strongly time-like p-dualizable) by some $\pi \in \Pi_{\ll}^{p\text{-opt}}(\mu_0, \mu_1)$, there exists an ℓ_p-geodesic $(\mu_t)_{t \in [0,1]}$ such that the function $[0, 1] \ni t \mapsto e(t) := \mathrm{Ent}(\mu_t|\mathrm{vol}_g)$ is semi-convex (and thus in particular locally Lipschitz in $(0, 1)$) and satisfies

$$e''(t) - \frac{1}{N}e'(t)^2 \geq K \int_{X \times X} \tau(x, y)^2 \, d\pi(x, y), \tag{B.19}$$

in the distributional sense on $[0, 1]$, where we adopt the convention that the second summand on the left hand side is zero if $N = \infty$.

The following stability result for synthetic time-like Ricci lower bounds was proved in [24, Theorem 3.14].

Theorem B.9 (Weak stability of $\mathsf{TCD}_p^e(K, N)$). *Let $\{(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)\}_{j \in \mathbb{N} \cup \{\infty\}}$ be a sequence of measured Lorentzian geodesic spaces satisfying the following properties:*

(1) *There exists a locally causally closed, $\mathcal{K}$-globally hyperbolic Lorentzian geodesic space $(X, \mathsf{d}, \ll, \leq, \tau)$ such that each $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$, $j \in \mathbb{N} \cup \{\infty\}$, is isomorphically embedded in it (as in (1) of Theorem B.6).*

(2) *The measures $(\iota_j)_\# \mathfrak{m}_j$ converge to $(\iota_\infty)_\# \mathfrak{m}_\infty$ weakly in duality with $C_c(X)$ in X, i.e. (B.5) holds.*

(3) *There exist $p \in (0, 1)$, $K \in \mathbb{R}$ and $N \in (0, \infty]$ such that $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$ satisfies $\mathsf{TCD}_p^e(K, N)$, for each $j \in \mathbb{N}$.*

Then the limit space $(X_\infty, \mathsf{d}_\infty, \mathfrak{m}_\infty, \ll_\infty, \leq_\infty, \tau_\infty)$ satisfies $\mathsf{wTCD}_p^e(K, N)$.

B.5. Synthetic vacuum Einstein equations

Combining the synthetic upper and lower bounds on the time-like Ricci curvature, i.e. Definitions B.5 and B.8, it is natural to propose the following synthetic version for the vacuum Einstein equations (with possibly non-zero cosmological constant):

Definition B.10 (Synthetic vacuum Einstein equations). Fix $p \in (0, 1)$, $\Lambda \in \mathbb{R}$ and $N \in (0, \infty]$. We say that the measured Lorentzian pre-length space $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ satisfies the (resp. *weak*) *synthetic formulation of the vacuum Einstein equations* $\mathrm{Ric} \equiv \Lambda$ with cosmological constant $\Lambda \in \mathbb{R}$ and has *synthetic dimension* $\leq N$ if

- $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ satisfies $\mathsf{TCD}_p^e(\Lambda, N)$ (resp. $\mathsf{wTCD}_p^e(\Lambda, N)$);
- there exists $r_0 > 0$ and a function $\omega : [0, r_0) \to [0, \infty)$ with $\lim_{r \downarrow 0} \omega(r) = 0$ such that $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ has time-like Ricci curvature bounded above by Λ with respect to $p \in (0, 1)$, r_0 and ω.

The combination of the stability of time-like Ricci lower and upper bounds (i.e. Theorems B.6 and B.9) gives the stability of the synthetic vacuum Einstein equations under the aforementioned natural Lorentzian variant of measured Gromov–Hausdorff convergence.

Theorem B.11 (Weak stability of synthetic vacuum Einstein equations). *Let $\{(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)\}_{j \in \mathbb{N} \cup \{\infty\}}$ be a sequence of measured Lorentzian geodesic spaces satisfying the following properties:*

(1) *There exists a locally causally closed, $\mathcal{K}$-globally hyperbolic Lorentzian geodesic space $(X, \mathsf{d}, \ll, \leq, \tau)$ such that each $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$, $j \in \mathbb{N} \cup \{\infty\}$, is isomorphically embedded in it (as in (1) of Theorem B.6).*

(2) *The measures $(\iota_j)_\# \mathfrak{m}_j$ converge to $(\iota_\infty)_\# \mathfrak{m}_\infty$ weakly in duality with $C_c(X)$ in X, i.e. (B.5) holds.*

(3) *Volume non-collapsing: there exists a function $v : (0, \infty) \to (0, \infty)$ such that $\mathfrak{m}_j(B_r^{\mathsf{d}_j}(x_j)) \geq v(r) > 0$ for every $x_j \in X_j$.*

(4) *There exist $p \in (0, 1)$, $\Lambda \in \mathbb{R}$, $N \in (0, \infty]$, $r_0 > 0$ and $\omega : [0, r_0) \to [0, \infty)$ with $\lim_{r \downarrow 0} \omega(r) = 0$ such that $(X_j, \mathsf{d}_j, \mathfrak{m}_j, \ll_j, \leq_j, \tau_j)$ satisfies the synthetic formulation*

of the vacuum Einstein equations $\text{Ric} \equiv \Lambda$ *with cosmological constant* $\Lambda \in \mathbb{R}$, *with synthetic dimension* $\leq N$ *with respect to* $p \in (0, 1)$, r_0 *and* ω *as in Definition* B.10.

Then the limit space $(X_\infty, \mathsf{d}_\infty, \mathfrak{m}_\infty, \ll_\infty, \leq_\infty, \tau_\infty)$ *satisfies the weak synthetic formulation of the vacuum Einstein equations* $\text{Ric} \equiv \Lambda$ *with cosmological constant* $\Lambda \in \mathbb{R}$, *with synthetic dimension* $\leq N$ *with respect to* $p \in (0, 1)$, $r_0 + 1$ *and* ω.

Remark B.12. By [24, Theorem 3.1] after [61] (see also Corollary 4.4) and by Theorem 4.7 (see also Remark 4.8), if $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ is (for simplicity say a compact subset in) a smooth Lorentzian manifold, then $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ satisfies the Einstein equations $\text{Ric} \equiv \Lambda$ in the smooth classical sense if and only if $(X, \mathsf{d}, \mathfrak{m}, \ll, \leq, \tau)$ satisfies the Einstein equations in the synthetic sense of Definition B.10. Therefore, Theorem B.11 shows that the corresponding limits of smooth solutions to the Einstein equations $\text{Ric} \equiv \Lambda$ satisfy the weak synthetic Einstein equations $\text{Ric} \equiv \Lambda$ in the sense of Definition B.10. In other terms, the vacuum Einstein equations are stable under the conditions (and with respect to the notion of convergence) of Theorem B.11.

Let us mention that the stability of the Einstein equations under various notions of (weak) convergence is a topic of high interest in general relativity. Classically, the problem is phrased in terms of convergence of a sequence of Lorentzian metrics g_j to a limit Lorentzian metric g_∞, *on a fixed underlying manifold*. It is well known that if g_j are solutions of the vacuum Einstein equations, $g_j \to g_\infty$ in C^0_{loc} and the derivatives of g_j converge in L^2_{loc}, then the limit g_∞ satisfies the vacuum Einstein equations as well.

However, if $g_j \to g_\infty$ in C^0_{loc} and the derivatives of g_j converge *weakly* in L^2_{loc}, explicit examples are known (see [19, 40] for examples in symmetry classes) where the limit g_∞ may satisfy the Einstein equations with a *non-vanishing* stress energy momentum tensor. Burnett [19] conjectured that if there exist $C > 0$ and $\lambda_j \to 0$ such that

$$|g_j - g_\infty| \leq \lambda_j, \quad |\partial g_j| \leq C, \quad |\partial^2 g_j| \leq C\lambda_j^{-1},$$

then g_∞ is isometric to a solution to the Einstein-massless Vlasov system for some appropriate choice of Vlasov field. Such a conjecture remains open, although there has been recent progress [45, 46] when g_j are assumed to be $\mathbb{U}(1)$-symmetric. We also mention the recent work [57] where concentrations (at the level of ∂g_j) are allowed in addition to oscillations.

Theorem B.11 gives a new point of view on the stability of the vacuum Einstein equations. Indeed, while in the aforementioned results the metrics g_j are converging *on a fixed underlying manifold*, in Theorem B.11 also the underlying space X may vary (along the sequence and at the limit), allowing change in topology in the limit, as one may expect in case of formation of singularities. Moreover, the notion of convergence is quite different in spirit: while in the aforementioned results $g_j \to g_\infty$ in a suitable *functional-analytic sense*, in Theorem B.11 the spaces are converging in a *more geometric sense* (inspired by the measured Gromov–Hausdorff convergence).

References

[1] Ambrosio, L., Gigli, N.: A user's guide to optimal transport. In: Modelling and Optimisation of Flows on Networks, Lecture Notes in Math. 2062, Springer, Heidelberg, 1–155 (2013) MR 3050280

[2] Ambrosio, L., Gigli, N., Mondino, A., Rajala, T.: Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure. Trans. Amer. Math. Soc. **367**, 4661–4701 (2015) Zbl 1317.53060 MR 3335397

[3] Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. 2nd ed., Lectures in Math. ETH Zürich, Birkhäuser, Basel (2008) Zbl 1145.35001 MR 2401600

[4] Ambrosio, L., Gigli, N., Savaré, G.: Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. Invent. Math. **195**, 289–391 (2014) Zbl 1312.53056 MR 3152751

[5] Ambrosio, L., Gigli, N., Savaré, G.: Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163**, 1405–1490 (2014) Zbl 1304.35310 MR 3205729

[6] Ambrosio, L., Mondino, A., Savaré, G.: Nonlinear diffusion equations and curvature conditions in metric measure spaces. Mem. Amer. Math. Soc. **262**, no. 1270, v+121 pp. (2019) Zbl 07161621 MR 4044464

[7] Bacher, K., Sturm, K.-T.: Localization and tensorization properties of the curvature-dimension condition for metric measure spaces. J. Funct. Anal. **259**, 28–56 (2010) Zbl 1196.53027 MR 2610378

[8] Bardeen, J. M., Carter, B., Hawking, S. W.: The four laws of black hole mechanics. Comm. Math. Phys. **31**, 161–170 (1973) Zbl 1125.83309 MR 334798

[9] Bekenstein, J. D.: Black holes and entropy. Phys. Rev. D (3) **7**, 2333–2346 (1973) Zbl 1369.83037 MR 366343

[10] Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. Numer. Math. **84**, 375–393 (2000) Zbl 0968.76069 MR 1738163

[11] Bernard, P., Suhr, S.: Lyapounov functions of closed cone fields: from Conley theory to time functions. Comm. Math. Phys. **359**, 467–498 (2018) Zbl 1391.83010 MR 3783554

[12] Bertrand, J., Puel, M.: The optimal mass transport problem for relativistic costs. Calc. Var. Partial Differential Equations **46**, 353–374 (2013) Zbl 1287.49048 MR 3016512

[13] Bombelli, L., Lee, J., Meyer, D., Sorkin, R. D.: Space-time as a causal set. Phys. Rev. Lett. **59**, 521–524 (1987) MR 899046

[14] Bouchitté, G., Buttazzo, G.: Characterization of optimal shapes and masses through Monge–Kantorovich equation. J. Eur. Math. Soc. **3**, 139–168 (2001) Zbl 0982.49025 MR 1831873

[15] Brenier, Y.: Extended Monge–Kantorovich theory. In: Optimal Transportation and Applications (Martina Franca, 2001), Lecture Notes in Math. 1813, Springer, Berlin, 91–121 (2003) Zbl 1064.49036 MR 2006306

[16] Brenier, Y.: A homogenized model for vortex sheets. Arch. Ration. Mech. Anal. **138**, 319–353, (1977) Zbl 0962.35140 MR 1467558

[17] Brenier, Y., Frisch, U., Hénon, M., Loeper, G., Matarrese, S., Mohayaee, R., Sobolevskiǐ, A.: Reconstruction of the early universe as a convex optimization problem. Monthly Notices Roy. Astronom. Soc. **346**, 501–524 (2003)

[18] Brué, E., Semola, D.: Constancy of the dimension for RCD(K, N) spaces via regularity of Lagrangian flows. Comm. Pure Appl. Math. **73**, 1141–1204 (2020) Zbl 1442.35054 MR 4156601

[19] Burnett, G. A.: The high-frequency limit in general relativity. J. Math. Phys. **30**, 90–96 (1989) Zbl 0664.53041 MR 974460

[20] Cannarsa, P., Sinestrari, C.: Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control. Progr. Nonlinear Differential Equations Appl. 58, Birkhäuser Boston, Boston, MA (2004) Zbl 1095.49003 MR 2041617

[21] Carfora, M.: The Wasserstein geometry of nonlinear σ models and the Hamilton–Perelman Ricci flow. Rev. Math. Phys. **29**, art. 1750001, 71 pp. (2017) Zbl 1360.53067 MR 3595479

[22] Cavalletti, F., Milman, E.: The globalization theorem for the Curvature–Dimension condition. Invent. Math. **226**, 1–137 (2021) MR 4309491

[23] Cavalletti, F., Mondino, A.: Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. Invent. Math. **208**, 803–849 (2017) Zbl 1375.53053 MR 3648975

[24] Cavalletti, F., Mondino, A.: Optimal transport in Lorentzian synthetic spaces, synthetic time-like Ricci curvature lower bounds and applications. arXiv:2004.08934 (2020)

[25] Cordero-Erausquin, D., McCann, R. J., Schmuckenschläger, M.: A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. Invent. Math. **146**, 219–257 (2001) Zbl 1026.58018 MR 1865396

[26] Cordero-Erausquin, D., McCann, R. J., Schmuckenschläger, M.: Prékopa–Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport. Ann. Fac. Sci. Toulouse Math. (6) **15**, 613–635 (2006) Zbl 1125.58007 MR 2295207

[27] Cordero-Erausquin, D., Nazaret, B., Villani, C.: A mass-transportation approach to sharp Sobolev and Gagliardo–Nirenberg inequalities. Adv. Math. **182**, 307–332 (2004) Zbl 1048.26010 MR 2032031

[28] Eckstein, M., Miller, T.: Causality for nonlocal phenomena. Ann. Henri Poincaré **18**, 3049–3096 (2017) Zbl 1412.60008 MR 3685983

[29] Ehrlich, P. E., Jung, Y.-T., Kim, S.-B.: Volume comparison theorems for Lorentzian manifolds. Geom. Dedicata **73**, 39–56 (1998) Zbl 0939.53036 MR 1651891

[30] Erbar, M., Kuwada, K., Sturm, K.-T.: On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201**, 993–1071 (2015) Zbl 1329.53059 MR 3385639

[31] Erbar, M., Sturm, K.-T.: Rigidity of cones with bounded Ricci curvature. J. Eur. Math. Soc. **23**, 219–235 (2021) Zbl 07328111 MR 4186467

[32] Figalli, A., Guionnet, A.: Universality in several-matrix models via approximate transport maps. Acta Math. **217**, 81–176 (2016) Zbl 1402.60009 MR 3646880

[33] Figalli, A., Maggi, F., Pratelli, A.: A mass transportation approach to quantitative isoperimetric inequalities. Invent. Math. **182**, 167–211 (2010) Zbl 1196.49033 MR 2672283

[34] Finster, F.: The Continuum Limit of Causal Fermion Systems. Fundamental Theories Phys. 186, Springer (2016) MR 3525361

[35] Finster, F.: Causal Fermion systems: a primer for Lorentzian geometers. J. Phys. Conf. Ser. **968**, art. 012004, 14 pp. (2018) MR 3919947

[36] Frisch, U., Matarrese, S., Mohayaee, R., Sobolevskii, A.: A reconstruction of the initial conditions of the Universe by optimal mass transportation. Nature **417**, 260–262 (2002)

[37] Gigli, N.: On the differential structure of metric measure spaces and applications. Mem. Amer. Math. Soc. **236**, no. 1113, vi+91 pp. (2015) Zbl 1325.53054 MR 3381131

[38] Gigli, N., Mondino, A., Savaré, G.: Convergence of pointed non-compact metric measure spaces and stability of Ricci curvature bounds and heat flows. Proc. London Math. Soc. (3) **111**, 1071–1129 (2015) Zbl 1398.53044 MR 3477230

[39] Graf, M.: Singularity theorems for C^1-Lorentzian metrics. Comm. Math. Phys. **378**, 1417–1450 (2020) Zbl 1445.53052 MR 4134950

[40] Green, S. R., Wald, R. M.: Examples of backreaction of small-scale inhomogeneities in cosmology. Phys. Rev. D **87**, art. 124037 (2013)

[41] Haslhofer, R., Naber, A.: Ricci curvature and Bochner formulas for martingales. Comm. Pure Appl. Math. **71**, 1074–1108 (2018) Zbl 1393.60042 MR 3794527

[42] Hawking, S. W.: The occurrence of singularities in cosmology. I. Proc. Roy. Soc. London Ser. A **294**, 511–521 (1966) Zbl 0139.45803 MR 208978

[43] Hawking, S. W., Ellis, G. F. R.: The Large Scale Structure of Space-Time. Cambridge Monogr. Math. 1, Cambridge Univ. Press, London (1973) Zbl 0265.53054 MR 0424186

[44] Hawking, S. W., Penrose, R.: The singularities of gravitational collapse and cosmology. Proc. Roy. Soc. London Ser. A **314**, 529–548 (1970) Zbl 0954.83012 MR 264959

[45] Huneau, C., Luk, J.: High-frequency backreaction for the Einstein equations under polarized $\mathbb{U}(1)$-symmetry. Duke Math. J. **167**, 3315–3402 (2018) Zbl 1412.35327 MR 3881199

[46] Huneau, C., Luk, J.: Trilinear compensated compactness and Burnett's conjecture in general relativity. arXiv:1907.10743 (2019)

[47] Jacobson, T.: Thermodynamics of spacetime: the Einstein equation of state. Phys. Rev. Lett. **75**, 1260–1263 (1995) Zbl 1020.83609 MR 1343440

[48] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**, 1–17 (1998) Zbl 0915.35120 MR 1617171

[49] Kell, M.: On interpolation and curvature via Wasserstein geodesics. Adv. Calc. Var. **10**, 125–167 (2017) Zbl 1454.58017 MR 3630091

[50] Kell, M., Suhr, S.: On the existence of dual solutions for Lorentzian cost functions. Ann. Inst. H. Poincaré Anal. Non Linéaire **37**, 343–372 (2020) Zbl 1439.49060 MR 4072806

[51] Ketterer, C., Mondino, A.: Sectional and intermediate Ricci curvature lower bounds via optimal transport. Adv. Math. **329**, 781–818 (2018) Zbl 1387.53027 MR 3783428

[52] Klainerman, S., Nicolò, F.: The Evolution Problem in General Relativity. Progr. Math. Phys. 25, Birkhäuser Boston, Boston, MA (2003) Zbl 1010.83004 MR 1946854

[53] Klartag, B.: Needle decompositions in Riemannian geometry. Mem. Amer. Math. Soc. **249**, no. 1180, v+77 pp. (2017) Zbl 1457.53028 MR 3709716

[54] Kunzinger, M., Sämann, C.: Lorentzian length spaces. Ann. Global Anal. Geom. **54**, 399–447 (2018) Zbl 06970105 MR 3867652

[55] Lisini, S.: Characterization of absolutely continuous curves in Wasserstein spaces. Calc. Var. Partial Differential Equations **28**, 85–120 (2007) Zbl 1132.60004 MR 2267755

[56] Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. of Math. (2) **169**, 903–991 (2009) Zbl 1178.53038 MR 2480619

[57] Luk, J., Rodnianski, I.: High-frequency limits and null dust shell solutions in general relativity. arXiv:2009.08968 (2020)

[58] Martin, K., Panangaden, P.: A domain of spacetime intervals in general relativity. Comm. Math. Phys. **267**, 563–586 (2006) Zbl 1188.83071 MR 2249782

[59] Matei, A.-M.: First eigenvalue for the p-Laplace operator. Nonlinear Anal. **39**, 1051–1068 (2000) Zbl 0948.35090 MR 1735181

[60] McCann, R. J.: A convexity principle for interacting gases. Adv. Math. **128**, 153–179 (1997) Zbl 0901.49012 MR 1451422

[61] McCann, R. J.: Displacement convexity of Boltzmann's entropy characterizes the strong energy condition from general relativity. Cambridge J. Math. **8**, 609–681 (2020) Zbl 1454.53058 MR 4192570

[62] Minguzzi, E.: Causality theory for closed cone structures with applications. Rev. Math. Phys. **31**, art. 1930001, 139 pp. (2019) Zbl 1432.53102 MR 3955368

[63] Minguzzi, E., Sánchez, M.: The causal hierarchy of spacetimes. In: Recent Developments in Pseudo-Riemannian Geometry, ESI Lect. Math. Phys., Eur. Math. Soc., Zürich, 299–358 (2008) MR 2436235

[64] Mondino, A., Naber, A.: Structure theory of metric measure spaces with lower Ricci curvature bounds. J. Eur. Math. Soc. **21**, 1809–1854 (2019) Zbl 1468.53039 MR 3945743

[65] Naber, A.: Characterizations of bounded Ricci curvature on smooth and non-smooth spaces, arXiv:1306.6512 (2013)

[66] O'Neill, B.: Semi-Riemannian Geometry. Pure Appl. Math. 103, Academic Press, New York (1983) MR 719023

[67] Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. Comm. Partial Differential Equations **26**, 101–174 (2001) Zbl 0984.35089 MR 1842429

[68] Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**, 361–400 (2000) Zbl 0985.58019 MR 1760620

[69] Padmanabhan, T.: Emergent gravity paradigm: recent progress. Modern Phys. Lett. A **30**, art. 1540007, 21 pp. (2015) Zbl 1311.83047 MR 3305080

[70] Penrose, R.: Gravitational collapse and space-time singularities. Phys. Rev. Lett. **14**, 57–59 (1965) Zbl 0125.21206 MR 172678

[71] Sämann, C.: Global hyperbolicity for spacetimes with continuous metrics. Ann. Henri Poincaré **17**, 1429–1455 (2016) Zbl 1342.83014 MR 3500220

[72] Santambrogio, F.: Optimal Transport for Applied Mathematicians. Progr. Nonlinear Differential Equations Appl. 87, Birkhäuser/Springer, Cham (2015) Zbl 1401.49002 MR 3409718

[73] Sturm, K.-T.: On the geometry of metric measure spaces. I. Acta Math. **196**, 65–131 (2006) Zbl 1105.53035 MR 2237206

[74] Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. **196**, 133–177 (2006) Zbl 1106.53032 MR 2237207

[75] Sturm, K.-T.: Remarks about synthetic upper Ricci bounds for metric measure spaces. arXiv:1711.01707 (2017)

[76] Suhr, S.: Theory of optimal transport for Lorentzian cost functions. Münster J. Math. **11**, 13–47 (2018) Zbl 1410.49054 MR 3873093

[77] Valtorta, D.: Sharp estimate on the first eigenvalue of the p-Laplacian. Nonlinear Anal. **75**, 4974–4994 (2012) Zbl 1247.35081 MR 2927560

[78] Verlinde, E.: On the origin of gravity and the laws of Newton. J. High Energy Phys. **2011**, no. 4, art. 029, 27 pp. Zbl 1260.81284 MR 2833282

[79] Villani, C.: Topics in Optimal Transportation. Grad. Stud. Math. 58, Amer. Math. Soc., Providence, RI (2003) Zbl 1106.90001 MR 1964483

[80] Villani, C.: Optimal Transport. Grundlehren Math. Wiss. 338, Springer, Berlin (2009) Zbl 1156.53003 MR 2459454

[81] von Renesse, M.-K., Sturm, K.-T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature. Comm. Pure Appl. Math. **58**, 923–940 (2005) Zbl 1078.53028 MR 2142879

[82] Wald, R. M.: General Relativity. Univ. of Chicago Press, Chicago, IL (1984) Zbl 0549.53001 MR 757180

[83] Xia, Q.: Sharp spectral gap for the Finsler p-Laplacian. Sci. China Math. **62**, 1615–1644 (2019) Zbl 1423.53094 MR 3984392

[84] Xia, Q.: Geometric and functional inequalities on Finsler manifolds. Preprint