Riemannian Ricci curvature lower bounds in metric measure spaces with sigma-finite measure

Published
DOI:10.1090/S0002-9947-2015-06111-X

Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure

Luigi Ambrosio * Nicola Gigli † Andrea Mondino ‡ Tapio Rajala §

July 24, 2012

Contents

1 Introduction **2**

2 Preliminaries **4**
2.1 Metric structure . . . 5
2.2 Optimal transport . . . 6

3 Weak gradients and weighted Cheeger energies **9**

4 Existence of good geodesics **15**
4.1 Intermediate measures and the existence of minimizers . . . 15
4.2 Localization in transport distance . . . 17
4.3 Density bounds for the minimizers . . . 18
4.4 Construction of the geodesic . . . 21

5 Convergence results **23**

6 Equivalence of the different formulations of $RCD(K,\infty)$ **27**
6.1 Derivative of $W_2^2(\cdot,\sigma)$ along the heat flow . . . 27
6.2 Derivative of the entropy along $\mathrm{Ent}_{\mathfrak{m}}$-convex L^∞-bounded geodesics . . . 29

7 Properties of $RCD(K,\infty)$ spaces **33**
7.1 The heat semigroup and its regularizing properties . . . 33
7.2 Connections with Dirichlet forms and Markov processes . . . 35
7.3 Tensorization . . . 36

*Scuola Normale Superiore, Pisa, l.ambrosio@sns.it
†University of Nice, nicola.gigli@unice.fr
‡Scuola Normale Superiore, Pisa, andrea.mondino@sns.it
§University of Jyväskylä, tapio.m.rajala@jyu.fi

1 Introduction

In a recent paper [4] written jointly with G.Savaré, the first and second author introduced a notion of Riemannian Ricci lower bound for metric measure spaces $(X, \mathsf{d}, \mathfrak{m})$, relying on the calculus tools developed in [3]. This definition, in the spirit of the $CD(K, N)$ theory proposed by Lott-Villani [22] and Sturm [29, 30] relies on optimal transportation tools and suitable convexity properties of the relative entropy functional $\mathrm{Ent}_{\mathfrak{m}}$. In the framework of [4], these conditions are enforced adding the assumption that the so-called Cheeger energy (playing here the role of the classical Dirichlet energy) is quadratic.
More precisely, the class of $RCD(K, \infty)$ spaces of [4] can be defined in 3 equivalent ways thanks to this equivalence result (see below for the precise meaning of the various symbols):

Theorem 1.1. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a metric measure space with (X, d) complete and separable, $\mathfrak{m}(X) \in (0, \infty)$ and $\operatorname{supp} \mathfrak{m} = X$. Then the following are equivalent.*

(i) $(X, \mathsf{d}, \mathfrak{m})$ *is a strict $CD(K, \infty)$ space and the W_2-gradient flow $\mathscr{H}_t$ of $\mathrm{Ent}_{\mathfrak{m}}$ on $\mathscr{P}_2(X)$ is additive.*

(ii) $(X, \mathsf{d}, \mathfrak{m})$ *is a strict $CD(K, \infty)$ space and Ch is a quadratic form on $L^2(X, \mathfrak{m})$.*

(iii) $(X, \mathsf{d}, \mathfrak{m})$ *is a length space and any $\mu \in \mathscr{P}_2(X)$ is the starting point of an EVI_K gradient flow of $\mathrm{Ent}_{\mathfrak{m}}$.*

This equivalence is crucial for the study of the spaces $RCD(K, \infty)$: for instance the fine properties of the heat flow and the Bakry-Emery condition obtained in [4] need (ii), while stability of $RCD(K, \infty)$ spaces under Sturm's convergence [30] of metric measure spaces (a variant of measured Gromov-Hausdorff convergence) depends on a crucial way on (iii) and on the stability properties of EVI_K flows.

Aim of this paper is the extension of the theory of $RCD(K, \infty)$ spaces to a class of σ-finite metric measure spaces. This extension includes fundamental examples such as the Lebesgue measure in $\mathbb{R}^n$ and noncompact Riemannian manifolds with bounded geometry. In our class of spaces we obtain the perfect analogue of Theorem 1.1 (see Theorem 6.1). Actually, even in the finite case we improve Theorem 1.1, replacing strict $CD(K, \infty)$ with $CD(K, \infty)$ in (i) and (ii): this is possible mainly thanks to the fine results of Section 4. The measures $\mathfrak{m}$ we will be dealing with satisfy the quantitative σ-finiteness condition

$$\int_X \mathrm{e}^{-\mathsf{c}\mathsf{d}^2(x_0,x)}\,\mathrm{d}\mathfrak{m}(x) < \infty \tag{1.1}$$

for some $x_0 \in X$ and $\mathsf{c} \in (0, \infty)$. As illustrated in [3, Remark 4.21] this condition is already needed and close to being sharp for stochastic completeness (i.e. mass conservation for the heat flow) and it is also a consequence of the $CD(K, \infty)$ condition as formulated in [29], so it is very natural within this theory.

Let us now briefly and informally explain the terminology implicit in Theorem 1.1 and the technical difficulties arising when one considers σ-finite reference measures $\mathfrak{m}$. Cheeger's energy Ch can be defined in $L^2(X, \mathfrak{m})$ by a relaxation procedure

$$\mathsf{Ch}(f) := \frac{1}{2}\inf\left\{\liminf_{h\to\infty}\int_X |Df_h|^2\,\mathrm{d}\mathfrak{m} :\ f_h \text{ locally Lipschitz, } f_h \to f \text{ in } L^2(X, \mathfrak{m})\right\},$$

where $|Df|$ is the slope, see (2.4). Associated to this procedure is a notion of weak upper gradient $|Df|_w$, that provides integral representation to Ch, namely $\mathsf{Ch}(f) = \frac{1}{2}\int |Df|_w^2 \,\mathrm{d}\mathfrak{m}$ whenever $\mathsf{Ch}(f) < \infty$. Since Ch is convex and lower semicontinuous on $L^2(X, \mathfrak{m})$, its gradient flow $\mathsf{h}_t f$ is well defined starting from any initial condition, and one of the main results of [3] is the coincidence of h_t with the quadratic optimal transport distance semigroup $\mathscr{H}_t$ (the W_2 gradient flow of $\mathrm{Ent}_\mathfrak{m}$) under the $CD(K, \infty)$ assumption: more precisely, if $f \in L^2(X, \mathfrak{m})$ and $\int f(x)\mathsf{d}^2(x, x_0)\,\mathrm{d}\mathfrak{m}(x)$ is finite, then $\mathscr{H}_t(f\mathfrak{m}) = (\mathsf{h}_t f)\mathfrak{m}$. This explains the connection between (i) and (ii), where finiteness of $\mathfrak{m}$ does not play any role. Passing to the EVI_K condition, deeply studied in [2, 15], it amounts to a family of differential inequalities indexed by $\sigma \in \mathscr{P}_2(X)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W_2^2(\mu_t, \sigma) \leq \mathrm{Ent}_\mathfrak{m}(\sigma) - \mathrm{Ent}_\mathfrak{m}(\mu_t) - \frac{K}{2}W_2^2(\mu_t, \sigma) \quad \text{for a.e. } t \in (0, \infty). \tag{1.2}$$

Set $\mu_t = (\mathsf{h}_t f)\mathfrak{m}$ and let φ_t be Kantorovich potentials from μ_t to σ. The analysis in [4] shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W_2^2(\mu_t, \sigma) \leq \lim_{\varepsilon\downarrow 0}\frac{\mathsf{Ch}(f_t - \varepsilon\varphi_t) - \mathsf{Ch}(f_t)}{\varepsilon} \tag{1.3}$$

on the one hand, and that the $CD(K, \infty)$ condition gives

$$\lim_{\varepsilon\downarrow 0}\frac{\mathsf{Ch}(\varphi_t - \varepsilon f_t) - \mathsf{Ch}(\varphi_t)}{\varepsilon} \leq \mathrm{Ent}_\mathfrak{m}(\sigma) - \mathrm{Ent}(\mu_t) - \frac{K}{2}W_2^2(\mu_t, \sigma) \tag{1.4}$$

on the other hand. If Ch is quadratic, then we can formally write that both the right hand side in (1.3) and the left hand side in (1.4) coincide with $-\int_X Df_t \cdot D\varphi_t \,\mathrm{d}\mathfrak{m}$, thus providing the connection from (ii) to (iii). However, in the derivation of (1.4) a key role is played by the Sobolev regularity of $\log f_t$, that can be easily achieved if $f_t \geq c > 0$. But, this assumption is not compatible with the σ-finite case, since f_t is a probability density, and even local space-time lower bounds on f_t can hardly be obtained in our framework, where no finite dimensionality assumption on $(X, \mathsf{d}, \mathfrak{m})$ is made. It turns out that this derivation is still possible, but working so to speak in a time-dependent weighted Sobolev space: formally we write

$$\int_X Df_t \cdot D\varphi_t \,\mathrm{d}\mathfrak{m} = \int_X D\log f_t \cdot D\varphi_t \,\mathrm{d}(f_t\mathfrak{m})$$

and, thanks to the energy dissipation estimate

$$\mathrm{Ent}_\mathfrak{m}(f_T\mathfrak{m}) + \int_0^T\int_X \frac{|Df_t|_w^2}{f_t}\,\mathrm{d}\mathfrak{m}\,\mathrm{d}t \leq \mathrm{Ent}_\mathfrak{m}(f\mathfrak{m}),$$

we know that $\log f_t$ belongs for a.e. t to the Sobolev space with weight f_t. Then we prove that for a.e. $t > 0$ the first inequality (1.3) holds, when written in terms of weighted Sobolev spaces, for any choice of the Kantorovich potential φ_t, while the second inequality (1.4) holds for at least one. This suffices for the derivation of (1.2).

Besides the application to σ-finite $RCD(K, \infty)$ spaces, several results of this paper have an independent interest and do not rely on curvature assumptions: for instance the compactness properties of Kantorovich potentials and the analysis of weighted Cheeger energies performed in Section 3. Also, it is worthwhile to mention that existence of geodesics with L^∞ bounds of Section 4 applies to σ-finite $CD(K, \infty)$ spaces, i.e. no quadratic assumption on Ch is needed for the results of the section. Also, since finiteness of $\mathfrak{m}$ was used in [4] essentially only for the equivalence of Theorem 1.1, we describe in the last section the properties of $RCD(K,\infty)$ spaces proved in [4], whose proof extends with no additional effort to the σ-finite case: among them we just mention the Bakry-Emery condition

$$|D(\mathsf{h}_t f)|_w^2 \le \mathrm{e}^{-2Kt}|Df|_w^2 \qquad \mathfrak{m}\text{-a.e. in } X.$$

In connection with the Bakry-Emery condition, we also mention the forthcoming paper [6]. In connection with stability, instead, the extension to the σ-finite case is far from being trivial. We devoted to this a separate paper [7].

The paper is organized as follows. In Section 2 we gather a few facts on relative entropy and optimal transportation, mostly stated without proofs (standard references are [1], [2], [31]); the only original contribution is a compactness result for Kantorovich potentials via De Giorgi's Γ-convergence stated in Lemma 2.3.

In Section 3 we recall the main results of the theory of weak gradients as developed in [3], emphasizing also the connections with the points of view developed in [13, 20, 27]. The main result of the section is that, for probability densities $\rho = g\mathfrak{m}$ with $g \in L^\infty(X,\mathfrak{m})$ and $\mathsf{Ch}(\sqrt{g}) < \infty$, roughly speaking weak gradients w.r.t to $\mathfrak{m}$ and weak gradients with respect to ρ are the same, even though no (local) lower bound on g is assumed. Furthermore, Cheeger's energy Ch_ρ induced by ρ is quadratic if Ch is quadratic. Section 4 is crucial for the development of (short time) L^∞ estimates for displacement interpolation which are new in the situation when (X,d) is unbounded and $\mathfrak{m}$ is not finite. These estimates, which hold when the density of the first measure decays at least as $c_1\mathrm{e}^{-c_2\mathsf{d}^2(x,x_0)}$ for some c_1, $c_2 > 0$ and the second measure has bounded density and support, are obtained combining carefully entropy minimization (an approach proposed by Sturm and then developed in [25, 24]) and splitting of optimal geodesic plans. Section 5 is devoted to the proof of some auxiliary convergence results dealing with entropy, difference quotients of probability densities and Kantorovich potentials, bilinear form Ch_ρ associated to a measure $\rho \in \mathscr{P}_2(X)$ as in Section 3. Section 6 contains the proof of the equivalence result analogous to Theorem 1.1 in the present σ-finite setting.

Acknowledgement. The authors warmly thank G.Savaré for his detailed and helpful comments on a preliminary version of this paper. The authors acknowledge the support of the ERC ADG GeMeThNES. T.R. acknowledges the support of the Academy of Finland, project no. 137528.

2 Preliminaries

We assume throughout the paper that $(X,\mathsf{d},\mathfrak{m})$ is a metric measure space with (X,d) complete and separable. We assume that $\mathfrak{m}$ is a nonnegative Borel measure satisfying $\operatorname{supp}\mathfrak{m} = X$ and (1.1) for some $\mathsf{c} > 0$ and $x_0 \in X$. This assumption includes finite measures and large classes of measures finite on bounded sets as the Lebesgue measure or the Riemannian volume measure of manifolds with bounded geometry.

We denote by $\mathscr{P}(X)$ the space of Borel probability measures on (X,d) and set

$$\mathscr{P}_2(X) := \Big\{\mu \in \mathscr{P}(X) \ : \ \int_X \mathsf{d}^2(x_0,x)\,\mathrm{d}\mu(x) < \infty \ \text{ for some (and hence all) } x_0 \in X\Big\}.$$

According to (1.1) we denote $z = \int_X \mathrm{e}^{-\mathsf{c}\mathsf{d}^2(x,x_0)}\,\mathrm{d}\mathfrak{m}$ and

$$\tilde{\mathfrak{m}} = \frac{1}{z}\mathrm{e}^{-\mathsf{c}\mathsf{d}^2(x,x_0)}\mathfrak{m} \in \mathscr{P}(X), \qquad V(x) = \mathsf{d}(x,x_0). \tag{2.1}$$

Given a nonnegative Borel measure $\mathfrak{n}$, the *relative entropy functional* $\mathrm{Ent}_{\mathfrak{n}} : \mathscr{P}_2(X) \to (-\infty, \infty]$ with respect to $\mathfrak{n}$ is defined by

$$\mathrm{Ent}_{\mathfrak{n}}(\rho) := \begin{cases} \int_X f \log f \,\mathrm{d}\mathfrak{n} & \text{if } \rho = f\mathfrak{n}; \\ \infty & \text{otherwise.} \end{cases} \tag{2.2}$$

By Jensen's inequality, this functional is nonnegative when $\mathfrak{n} \in \mathscr{P}(X)$. More generally, we recall (see [3, Lemma 7.2] for the simple proof) that when $\mathfrak{n} = \mathfrak{m}$ the formula above makes sense on measures $\rho = f\mathfrak{m} \in \mathscr{P}_2(X)$, thanks to the fact that the negative part of $f \log f$ is $\mathfrak{m}$-integrable. More precisely, defining $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ as in (2.1), the following formula for the change of reference measure will be useful:

$$\mathrm{Ent}_{\mathfrak{m}}(\rho) = \mathrm{Ent}_{\tilde{\mathfrak{m}}}(\rho) - c \int_X V^2 \,\mathrm{d}\rho - \log z. \tag{2.3}$$

In the sequel we shall denote by $D(\mathrm{Ent}_{\mathfrak{m}}) \subset \mathscr{P}_2(X)$ the finiteness domain of $\mathrm{Ent}_{\mathfrak{m}}$.

2.1 Metric structure

We shall denote by $\mathrm{Lip}(X)$ the space of Lipschitz functions $f : X \to \mathbb{R}$ and by $\mathrm{Lip}_b(X)$ the subspace of bounded Lipschitz functions.

Given $f : X \to \mathbb{R}$ we define its slope $|Df|$ by

$$|Df| := \limsup_{y \to x} \frac{|f(y) - f(x)|}{\mathsf{d}(y, x)}. \tag{2.4}$$

We shall also use, in connection with Kantorovich potentials, the one-sided counterparts of the slope, namely the ascending slope and descending slopes:

$$|D^+ f(x)| := \limsup_{y \to x} \frac{[f(y) - f(x)]^+}{\mathsf{d}(y, x)}, \qquad |D^- f(x)| := \limsup_{y \to x} \frac{[f(y) - f(x)]^-}{\mathsf{d}(y, x)}. \tag{2.5}$$

Given an open interval $J \subset \mathbb{R}$, an exponent $p \in [1, \infty]$ and $\gamma : J \to X$, we say that γ belongs to $AC^p(J; X)$ if

$$\mathsf{d}(\gamma_s, \gamma_t) \le \int_s^t g(r) \,\mathrm{d}r \qquad \forall s,\, t \in J,\ s < t$$

for some $g \in L^p(J)$. The case $p = 1$ corresponds to *absolutely continuous* curves. It turns out that, if γ belongs to $AC^p(J; X)$, there is a minimal function g with this property, called *metric derivative* and given for a.e. $t \in J$ by

$$|\dot{\gamma}_t| := \lim_{s \to t} \frac{\mathsf{d}(\gamma_s, \gamma_t)}{|s - t|}.$$

See [2, Theorem 1.1.2] for the simple proof. We say that an absolutely continuous curve γ_t has *constant speed* if $|\dot{\gamma}_t|$ is (equivalent to) a constant.

We call (X, d) *a geodesic space* if for any $x_0,\, x_1 \in X$ there exists $\gamma : [0, 1] \to X$ satisfying $\gamma_0 = x_0$, $\gamma_1 = x_1$ and

$$\mathsf{d}(\gamma_s, \gamma_t) = |t - s| \mathsf{d}(\gamma_0, \gamma_1) \qquad \forall s,\, t \in [0, 1]. \tag{2.6}$$

We will denote by $\mathrm{Geo}(X)$ the space of all constant speed geodesics $\gamma : [0,1] \to X$, namely $\gamma \in \mathrm{Geo}(X)$ if (2.6) holds. Recall also that the weaker notion of *length* space: for all $x_0, x_1 \in X$ and $\varepsilon > 0$ there exists $\gamma \in AC([0,1];X)$ such that $\int_0^1 |\dot\gamma_t|\,\mathrm{d}t < \mathsf{d}(x_0,x_1) + \varepsilon$.

From the measure-theoretic point of view, when considering measures on $AC^p(J;X)$ (resp. $\mathrm{Geo}(X)$), we shall consider them as measures on the Polish space $C(J;X)$ endowed with the sup norm, concentrated on the Borel set $AC^p(J;X)$ (resp. closed set $\mathrm{Geo}(X)$). We shall also use the notation $\mathrm{e}_t : C(J;X) \to X$, $t \in J$, for the evaluation map at time t, namely $\mathrm{e}_t(\gamma) := \gamma_t$.

2.2 Optimal transport

Given μ, $\nu \in \mathscr{P}_2(X)$, we define the quadratic optimal transport distance W_2 between them as

$$W_2^2(\mu,\nu) := \inf \int_{X\times X} \mathsf{d}^2(x,y)\,\mathrm{d}\boldsymbol{\gamma}(x,y), \tag{2.7}$$

where the infimum is taken among all Kantorovich transport plans, namely probability measures $\boldsymbol{\gamma}$ on $X \times X$ such that

$$\pi^1_\sharp \boldsymbol{\gamma} = \mu, \qquad \pi^2_\sharp \boldsymbol{\gamma} = \nu.$$

Here, for $\mu \in \mathscr{P}(X)$, a topological space Y and a μ-measurable map $T : X \to Y$, the push-forward measure $T_\sharp\mu \in \mathscr{P}(Y)$ is defined by $T_\sharp\mu(B) := \mu(T^{-1}(B))$ for every Borel set $B \subset Y$.

Since (X,d) is complete and separable, the space $(\mathscr{P}_2(X), W_2)$ is complete and separable. Since the cost d^2 is lower semicontinuous, the infimum in the definition (2.7) of W_2^2 is attained and we call optimal the plans $\boldsymbol{\gamma}$ realizing the minimum; in addition, Kantorovich's duality formula holds:

$$\frac{1}{2}W_2^2(\mu,\nu) = \sup\left\{ \int_X \varphi\,\mathrm{d}\mu + \int_X \psi\,\mathrm{d}\nu :\ \varphi(x) + \psi(y) \le \frac{1}{2}\mathsf{d}^2(x,y) \right\}, \tag{2.8}$$

where the functions φ and ψ in the supremum are respectively in $L^1(X,\mu)$ and in $L^1(X,\nu)$.

Recall that the *c-transform* φ^c of $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is defined by

$$\varphi^c(y) := \inf\left\{ \frac{\mathsf{d}^2(x,y)}{2} - \varphi(x) :\ x \in X \right\}$$

and that ψ is said to be *c-concave* if $\psi = \varphi^c$ for some φ.

Definition 2.1 (Kantorovich potential)**.** *We say that a map $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ is a Kantorovich potential relative to (μ,ν) if:*

(i) there exists a Borel map $\psi : X \to \mathbb{R} \cup \{-\infty\}$ such that $\psi \in L^1(X,\nu)$ and $\varphi = \psi^c$;

(ii) $\varphi \in L^1(X,\mu)$ and the pair (φ,ψ) maximizes (2.8)*.*

Notice that the inequality $\varphi(x) + \psi(y) \le \frac{1}{2}\mathsf{d}^2(x,y)$, when integrated against an optimal plan $\boldsymbol{\gamma}$, forces the integrability of the positive part of φ. For this reason, in (ii) we may equivalently require integrability of the negative part of φ only.

Proposition 2.2 (Existence and choice of gauge of Kantorovich potentials). *If μ, $\nu \in \mathscr{P}_2(X)$, then Kantorovich potentials φ relative to (μ,ν) exist and satisfy*

$$\varphi(x) + \psi(y) = \tfrac{1}{2}\mathsf{d}^2(x,y) \qquad \boldsymbol{\gamma}\text{-a.e. in } X \times X \tag{2.9}$$

for any optimal Kantorovich plan $\boldsymbol{\gamma}$ and

$$|D^+\varphi|(x) \le \mathsf{d}(x,y) \qquad \text{for } \boldsymbol{\gamma}\text{-a.e. } (x,y). \tag{2.10}$$

In addition, if $\operatorname{supp}\nu \subset \overline{B}_R(y_0)$ for some $R \ge 1$, then a locally Lipschitz Kantorovich potential $\varphi = \psi^c$ exists with $\psi \equiv -\infty$ on $X \setminus \operatorname{supp}\nu$, $\psi \le R^2/2$ on $\operatorname{supp}\nu$ and

$$|D\varphi|(x) \le R + \mathsf{d}(x,y_0), \qquad |\varphi(x)| \le 2R^2(1 + \mathsf{d}^2(x,y_0)). \tag{2.11}$$

Proof. Since any complete and separable metric space can be isometrically embedded in a complete, separable and geodesic metric space we can assume with no loss of generality that the space (X,d) is geodesic. The existence part is well known, so let us discuss briefly (2.10), the choice of gauge and the regularity properties of φ when ν has bounded support. From (2.9) and the inequality $\varphi + \varphi^c \le \mathsf{d}^2/2$ we get

$$\varphi(z) - \varphi(x) \le \frac{1}{2}\big(\mathsf{d}^2(z,y) - \mathsf{d}^2(x,y)\big) \qquad \text{for all } z$$

for $\boldsymbol{\gamma}$-a.e. (x,y), so that $|D^+\varphi|(x) \le \mathsf{d}(x,y)$ for γ-a.e. (x,y).

Now, let us set

$$\tilde\psi(x) := \begin{cases} \psi(x) & \text{if } x \in \operatorname{supp}\nu; \\ -\infty & \text{otherwise,} \end{cases} \qquad \tilde\varphi := (\tilde\psi)^c.$$

Since $\tilde\varphi \ge \varphi$, it is obvious that its negative part is μ-integrable and that $(\tilde\varphi,\tilde\psi)$ is a maximizing pair, so that $\tilde\varphi$ is a Kantorovich potential. From

$$\tilde\varphi(x) = \inf_{y \in \operatorname{supp}\nu} \frac{1}{2}\mathsf{d}^2(x,y) - \tilde\psi(y)$$

and the inclusion $\operatorname{supp}\nu \subset B_R(y_0)$ it is immediate to obtain the linear growth of $|D\tilde\varphi|$, in the form stated in (2.11). Finally, possibly adding and subtracting the same constant to the potentials in the maximizing pair, we can assume that $\tilde\varphi(y_0) = 0$. Then, the inequality $\tilde\psi \le \frac{1}{2}\mathsf{d}^2(y_0,\cdot)$ gives $\tilde\psi \le R^2/2$ on $\operatorname{supp}\nu$ and the linear growth of $|D\tilde\varphi|$ gives the quadratic growth of $|\varphi|$, since (X,d) is geodesic. □

Lemma 2.3 (Compactness of Kantorovich potentials). *Consider probability densities σ, $\eta = f\mathfrak{m}$, $\eta_n = f_n\mathfrak{m} \in \mathscr{P}_2(X)$ satisfying the following conditions:*

(a) σ has compact support;

(b) $f_n \to f$ $\mathfrak{m}$-a.e. in X and $\sup_n f_n(x)(1 + \mathsf{d}^2(x,x_0)) \in L^1(X,\mathfrak{m})$ for some $x_0 \in X$.

Let $\varphi_n = \psi_n^c$ be Kantorovich potentials relative to (η_n,σ), in the sense of Definition 2.1, satisfying

$$|\varphi_n(x)| \le C(1 + \mathsf{d}^2(x,x_0)) \tag{2.12}$$

and

$$\psi_n \equiv -\infty \quad \text{on } X \setminus \operatorname{supp}\sigma, \quad \psi_n \le C \tag{2.13}$$

for some constant C independent of n. Then there exist a subsequence $n(k)$ and a Kantorovich potential $\varphi = \psi^c$ of the transportation problem relative to (η, σ) such that $\varphi_{n(k)} \to \varphi$ pointwise. In addition (2.12) *is fulfilled by φ and $\psi \le C$.*

Proof. In this proof we shall use De Giorgi's Γ-convergence (strictly speaking, the Γ^--convergence, the one designed for convergence of minimum problems, see [14]). Since X is separable, by compactness of Γ-convergence (see for instance Proposition 1.42 in [11]) we can assume with no loss of generality that $-\psi_n$ Γ-converges as $n \to \infty$, and we shall denote by $-\psi$ its Γ-limit. Observe that, since by definition of Γ-convergence for every $x \in X$ there exists a sequence $x_n \to x$ such that $-\psi_n(x_n) \to -\psi(x)$, ψ still satisfies (2.13).

By invariance of Γ-convergence under continuous additive perturbations (see for instance Remark 1.7 in [11]) we get

$$\left(\frac{1}{2}\mathsf{d}^2(x,\cdot) - \psi\right) = \Gamma - \lim_{n\to\infty} \left(\frac{1}{2}\mathsf{d}^2(x,\cdot) - \psi_n\right) \qquad \forall x \in X. \tag{2.14}$$

Because of (2.13) and of the compactness of $\operatorname{supp}\sigma$, the Γ-convergent functionals above are equi-coercive, so that the minimum values of the functionals converge to the minimum of the Γ-limit (see for instance Theorem 1.21 in [11]), yielding

$$\varphi_n(x) = \min_X \left(\frac{1}{2}\mathsf{d}^2(x,\cdot) - \psi_n\right) \to \min_X \left(\frac{1}{2}\mathsf{d}^2(x,\cdot) - \psi\right) = \varphi(x), \tag{2.15}$$

where the last equality has to be understood as the definition of $\varphi(x)$. Obviously (2.12) is fulfilled by φ, so that $\varphi \in L^1(X, f\mathfrak{m})$. In connection with ψ, obviously its positive part is σ-integrable.

Now we claim that $\varphi = \psi^c$ is a Kantorovich potential for the limit transportation problem $(f\mathfrak{m}, \sigma)$; we have to prove that

$$\int_X \varphi \, \mathrm{d}(f\mathfrak{m}) + \int_X \psi \, \mathrm{d}\sigma \ge \frac{1}{2} W_2^2(f\mathfrak{m}, \sigma), \tag{2.16}$$

since this inequality provides at the same time also integrability of the negative part of ψ. Since by assumption $\varphi_n = \psi_n^c$ is a Kantorovich potential for $(f_n\mathfrak{m}, \sigma)$, we already know that

$$\int_X \varphi_n \, \mathrm{d}(f_n\mathfrak{m}) + \int_X \psi_n \, \mathrm{d}\sigma = \frac{1}{2} W_2^2(f_n\mathfrak{m}, \sigma). \tag{2.17}$$

Using (b) it is immediate to check the weak convergence of $f_n\mathfrak{m}$ to $f\mathfrak{m}$, so that (see for instance Proposition 2.5 in [1])

$$W_2^2(f\mathfrak{m}, \sigma) \le \liminf_n W_2^2(f_n\mathfrak{m}, \sigma). \tag{2.18}$$

Moreover, using (b) and (2.12), the dominated convergence theorem gives

$$\int_X \varphi_n \, \mathrm{d}(f_n\mathfrak{m}) \to \int_X \varphi \, \mathrm{d}(f\mathfrak{m}). \tag{2.19}$$

Finally, by the very definition of Γ-limit we have

$$-\psi(x) = \inf\left\{\liminf_{n\to\infty} -\psi_n(x_n) \,|\, x_n \to x\right\} \le \liminf_{n\to\infty} -\psi_n(x).$$

Moreover, by assumption (2.13), $-\psi_n \ge -C$. Hence Fatou's lemma gives

$$\limsup_{n\to\infty} \int_X \psi_n \, d\sigma \le \int_X \psi \, d\sigma. \tag{2.20}$$

Putting together (2.17), (2.18), (2.19) and (2.20) we get (2.16) as desired. □

Let us close this section by discussing the geodesic structure of $(\mathscr{P}_2(X), W_2)$, see [1, Theorem 2.10] or [21]. If μ_0, $\mu_1 \in \mathscr{P}_2(X)$ are connected by a constant speed geodesic μ_t in $(\mathscr{P}_2(X), W_2)$, then there exists $\boldsymbol{\pi} \in \mathscr{P}(\mathrm{Geo}(X))$ with $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \mu_t$ for all $t \in [0,1]$ and

$$W_2^2(\mu_s, \mu_t) = \int_{\mathrm{Geo}(X)} \mathsf{d}^2(\gamma_s, \gamma_t) \, d\boldsymbol{\pi}(\gamma) = (s-t)^2 \int_{\mathrm{Geo}(X)} \ell^2(\gamma) \, d\boldsymbol{\pi}(\gamma) \qquad \forall s,\, t \in [0,1],$$

where $\ell(\gamma) = \mathsf{d}(\gamma_0, \gamma_1)$ is the length of the geodesic γ. The collection of all the measures $\boldsymbol{\pi}$ with the above properties is denoted by $\mathrm{OptGeo}(\mu, \nu)$. The measure $\boldsymbol{\pi}$ is not uniquely determined by μ_t, unless (X, d) is non-branching, while the relation between optimal geodesic plans and optimal Kantorovich plans is given by the fact that $\gamma := (\mathrm{e}_0, \mathrm{e}_1)_\sharp \boldsymbol{\pi}$ is optimal whenever $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu, \nu)$.

3 Weak gradients and weighted Cheeger energies

In the next two definitions we consider test plans and "Sobolev" functions with respect to a reference nonnegative Borel measure $\mathfrak{n}$ in X, finite on bounded sets. In the sequel we shall denote by $\mathcal{M}$ this class of measures, including both probability measures and our reference measure $\mathfrak{m}$.

Definition 3.1 (Test plan). *We say that $\boldsymbol{\pi} \in \mathscr{P}(C([0,1]; X))$ is a 2-test plan relative to $\mathfrak{n} \in \mathcal{M}$ if:*

(i) $\boldsymbol{\pi}$ is concentrated on $AC^2([0,1]; X)$ and the 2-action of $\boldsymbol{\pi}$ is finite:

$$\mathcal{A}_2(\boldsymbol{\pi}) := \int \int_0^1 |\dot\gamma_t|^2 \, dt \, d\boldsymbol{\pi}(\gamma) < \infty.$$

(ii) There exists $C \ge 0$ such that $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} \le C\mathfrak{n}$ for all $t \in [0,1]$.

The following definition is inspired by the classical concept [18] of upper gradient, that we now illustrate. A Borel function $G : X \to [0, \infty]$ is an upper gradient of a Borel function $f : X \to \mathbb{R}$ if

$$|f(\gamma_b) - f(\gamma_a)| \le \int_a^b G(\gamma_s) |\dot\gamma_s| \, ds$$

for any absolutely continuous curve $\gamma : [a,b] \to X$. Being the inequality invariant under reparameterization one can also reduce to curves defined in $[0,1]$.

Let $\mathscr{C}(X)$ be the set of continuous parametric curves $C \subset X$ with finite length, where curves equivalent under reparameterization are identified. Recall that any such curve C can be written as $\gamma([0,\ell])$, where ℓ is the length of C and $\gamma : [0,\ell] \to X$ is Lipschitz with $|\dot\gamma| = 1$ a.e. in $[0,\ell]$. We shall denote by $i : AC^2([0,1]; X) \to \mathscr{C}(X)$ the natural surjection. As shown in [27], functions that have an upper gradient in $L^2(X,\mathfrak{n})$ are absolutely continuous along $\mathrm{Mod}_{2,\mathfrak{n}}$-a.e. curve, where

$$\mathrm{Mod}_{2,\mathfrak{n}}(\Gamma) := \inf\left\{\int_X g^2\,\mathrm{d}\mathfrak{n}:\ g: X\to[0,\infty]\text{ Borel},\ \int_\gamma g\geq 1\text{ for all }\gamma\in\Gamma\right\} \tag{3.1}$$

for any $\Gamma\subset\mathscr{C}(X)$. We also recall the following simple consequence of (3.1): for any $\mathrm{Mod}_{2,\mathfrak{n}}$-negligible set Γ there exist Borel functions $r_h: X\to[0,\infty]$ satisfying $\int_X r_h^2\,\mathrm{d}\mathfrak{n}\to 0$ and $\int_\gamma r_h=\infty$ for all $\gamma\in\Gamma$.

Definition 3.2 (The space $\mathcal{S}^2_{\mathfrak{n}}$ and weak upper gradients). *Let $f: X\to\mathbb{R}$, $G: X\to[0,\infty]$ be Borel functions. We say that G is a 2-weak upper gradient relative to $\mathfrak{n}$ of f if*

$$|f(\gamma_1)-f(\gamma_0)|\leq\int_0^1 G(\gamma_s)|\dot\gamma_s|\,\mathrm{d}s \qquad \textit{for }\boldsymbol{\pi}\textit{-a.e. }\gamma$$

for all 2-test plans $\boldsymbol{\pi}$ relative to $\mathfrak{n}$.
We write $f\in\mathcal{S}^2_{\mathfrak{n}}$ if f has a 2-weak upper gradient in $L^2(X,\mathfrak{n})$. The 2-weak upper gradient relative to $\mathfrak{n}$ with minimal $L^2(X,\mathfrak{n})$ norm (the so-called minimal 2-weak upper gradient) will be denoted by $|Df|_{w,\mathfrak{n}}$.

Remark 3.3 (Sobolev regularity along curves). A consequence of $\mathcal{S}^2_{\mathfrak{n}}$ regularity is (see [5, Remark 4.10]) the Sobolev property along curves, namely for any 2-test plan $\boldsymbol{\pi}$ relative to $\mathfrak{n}$ the function $t\mapsto f(\gamma_t)$ belongs to the Sobolev space $W^{1,1}(0,1)$ and

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}f(\gamma_t)\right|\leq|Df|_w(\gamma_t)|\dot\gamma_t| \qquad \text{a.e. in }(0,1)$$

for $\boldsymbol{\pi}$-a.e. γ. Conversely, assume that g is Borel nonnegative, that for any 2-test plan $\boldsymbol{\pi}$ the map $t\mapsto f(\gamma_t)$ is $W^{1,1}(0,1)$ and that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}f(\gamma_t)\right|\leq g(\gamma_t)|\dot\gamma_t| \qquad \text{a.e. in }(0,1)$$

for $\boldsymbol{\pi}$-a.e. γ. Then, the fundamental theorem of calculus in $W^{1,1}(0,1)$ gives that g is a 2-weak upper gradient of f.

Because of the absolute continuity condition $(\mathrm{e}_t)_\sharp\boldsymbol{\pi}\ll\mathfrak{n}$ imposed on test plans, it is immediate to check that the property of being in $\mathcal{S}^2_{\mathfrak{n}}$, as well as $|Df|_{w,\mathfrak{n}}$, are invariant under modifications of f in $\mathfrak{n}$-negligible sets. Furthermore, these concepts are easily seen to be local with respect to $\mathfrak{n}$ in the following sense: if $f\in\mathcal{S}^2_{\mathfrak{n}}$ then $f\in\mathcal{S}^2_{\mathfrak{n}'}$ for all measures $\mathfrak{n}'=\mathfrak{n}\llcorner B$ with $B\subset X$ Borel, and $|Df|_{w,\mathfrak{n}'}\leq|Df|_{w,\mathfrak{n}}$ $\mathfrak{n}'$-a.e. on B: this is due to the fact that test plans relative to $\mathfrak{n}'$ are test plans relative to $\mathfrak{n}$. Conversely,

$$f\in\mathcal{S}^2_{\mathfrak{n}_R}\text{ with }\mathfrak{n}_R:=\mathfrak{n}\llcorner\overline{B}_R(x_0),\ \sup_R\int_X|Df|^2_{w,\mathfrak{n}_R}\,\mathrm{d}\mathfrak{n}_R<\infty \quad\Longrightarrow\quad f\in\mathcal{S}^2_{\mathfrak{n}}. \tag{3.2}$$

This is due to the fact that any curve is bounded, hence any test plan $\boldsymbol{\pi}$ relative to $\mathfrak{n}$ can be monotonically approximated by test plans concentrated on curves contained in a bounded set.

Another property we shall need is the locality with respect to f, see [6] for the simple proof.

Proposition 3.4 (Locality)**.** *Let f_1, $f_2 : X \to \mathbb{R}$ Borel and let $G_1, G_2 \in L^2(X, \mathfrak{n})$ be 2-weak upper gradients of f_1, f_2 relative to $\mathfrak{n}$. Then*

$$\tilde{G}_1 := \begin{cases} G_1 & \text{on } \{f_1 \neq f_2\}; \\ \min\{G_1, G_2\} & \text{on } \{f_1 = f_2\} \end{cases}$$

is a 2-weak upper gradient of f_1. In particular, by minimality we get

$$|Df_1|_{w,\mathfrak{n}} = |Df_2|_{w,\mathfrak{n}} \qquad \mathfrak{n}\text{-a.e. on } \{f_1 = f_2\}. \tag{3.3}$$

Weak gradients share with classical gradients many features, in particular the chain rule [3, Proposition 5.14]

$$|D\phi(f)|_{w,\mathfrak{n}} = \phi'(f)|Df|_{w,\mathfrak{n}} \qquad \mathfrak{n}\text{-a.e. in } X \tag{3.4}$$

for all $\phi : \mathbb{R} \to \mathbb{R}$ Lipschitz and nondecreasing on an interval containing the image of f. By convention, as in the classical chain rule, $\phi'(f)$ is arbitrarily defined at all points x such that ϕ is not differentiable at x, taking into account the fact that $|Df|_{w,\mathfrak{n}} = 0$ $\mathfrak{n}$-a.e. on this set of points.

In the sequel we shall adopt the conventions

$$|Df|_w := |Df|_{w,\mathfrak{m}}, \qquad\qquad \mathcal{S}^2 := \mathcal{S}^2_{\mathfrak{m}},$$

analyzing in detail, in Theorem 3.5 below, the behaviour of $|Df|_{w,\mathfrak{n}}$ and $\mathcal{S}^2_{\mathfrak{n}}$ under modifications of the reference measure $\mathfrak{n}$.

Theorem 3.5. *The following properties hold:*

(a) If $\mathfrak{n} \in \mathcal{M}$ and $\Gamma \subset \mathscr{C}(X)$ is $\mathrm{Mod}_{2,\mathfrak{n}}$-negligible, then any Borel set $\tilde{\Gamma} \subset AC^2([0,1]; X)$ such that $i(\tilde{\Gamma}) \subset \Gamma$ is $\boldsymbol{\pi}$-negligible for any 2-test plan $\boldsymbol{\pi}$ relative to $\mathfrak{n}$. In addition, for any Borel and $\mathfrak{n}$-negligible set $N \subset X$ it holds

$$\mathrm{Mod}_{2,\boldsymbol{n}}\Big(\Big\{\gamma \in \mathscr{C}(X) : \int_{\gamma^{-1}(N)} |\dot{\gamma}| \,\mathrm{d}t > 0\Big\}\Big) = 0.$$

(b) If either $\mathfrak{n} \in \mathscr{P}(X)$ and $f \in \mathcal{S}^2_{\mathfrak{n}}$, or $\mathfrak{n} \in \mathcal{M}$ and $f \in \mathcal{S}^2_{\mathfrak{n}} \cap L^1(X, \mathfrak{n})$, there exist $\phi_n \in \mathrm{Lip}_b(X) \cap L^2(X, \mathfrak{n})$ satisfying $\phi_n \to f$ $\mathfrak{n}$-a.e. in X and $|D\phi_n| \to |Df|_{w,\mathfrak{n}}$ in $L^2(X, \mathfrak{n})$.

(c) If either $\mathfrak{n} \in \mathscr{P}(X)$ and $f \in \mathcal{S}^2_{\mathfrak{n}}$, or $\mathfrak{n} \in \mathcal{M}$ and $f \in \mathcal{S}^2_{\mathfrak{n}} \cap L^1(X, \mathfrak{n})$, then there exists a Borel function $\tilde{f}$ coinciding with f out of an $\mathfrak{n}$-negligible set and having an upper gradient in $L^2(X, \mathfrak{n})$; in addition, there exist upper gradients G_n of $\tilde{f}$ converging to $|Df|_{w,\mathfrak{n}}$ in $L^2(X, \mathfrak{n})$.

Proof. (a) The first statement is a simple consequence of Hölder inequality, see [3, Remark 5.3]. The second one follows just by taking the function g identically equal to ∞ on N and null out of N in (3.1).

(b) Using the chain rule (3.4) we reduce the proof to the case of nonnegative functions f. If f belong to $L^2(X, \mathfrak{n})$ the existence of ϕ_n is one of the main results of [3], see Theorem 6.2 therein. In the general case we approximate f by the truncated functions $f_N = \min\{f, N\}$ and use the chain rule again to show $|Df_N|_{w,\mathfrak{n}} \to |Df|_{w,\mathfrak{n}}$ in $L^2(X, \mathfrak{n})$. Then, a diagonal argument provides the result.

(c) This is part of the theory developed in [20, 27]: if $f_n \to f$ $\mathfrak{n}$-a.e. and G_n are upper gradients of f_n weakly convergent to G in $L^2(X,\mathfrak{n})$, then we can find a Borel function $\tilde f$ equal to f $\mathfrak{n}$-a.e. and a Borel function $\tilde G$ equal to G $\mathfrak{n}$-a.e. such that $\tilde G$ satisfies the upper gradient property relative to $\tilde f$ along $\mathrm{Mod}_{2,\mathfrak{n}}$-almost every curve. In our case when $f \in \mathcal{S}^2_{\mathfrak{n}}$ we may apply statement (b) with $G = |Df|_{w,\mathfrak{n}}$ and choose $f_n = \phi_n$ to find $\tilde f$ and $\tilde G$. Then, denoting by Γ the set of curves where the upper gradient property fails and considering

$$G_h := \tilde G + r_h,$$

where $r_h \in L^2(X,\mathfrak{n})$ satisfy $\int_X r_h^2\,\mathrm{d}\mathfrak{n} \to 0$ and $\int_\gamma r_\epsilon = \infty$ for all $\gamma \in \Gamma$, we obtain upper gradients G_h of $\tilde f$ approximating $|Df|_{w,\mathfrak{n}}$ in $L^2(X,\mathfrak{n})$. □

Theorem 3.6 (Change of reference measure)**.** *Assume that* $\rho = g\mathfrak{m} \in \mathscr{P}_2(X)$ *with* $g \in L^\infty(X,\mathfrak{m})$ *and* $\mathsf{Ch}(\sqrt{g}) < \infty$*. Then:*

(a) $f \in \mathcal{S}^2$ *and* $|Df|_w \in L^2(X,\rho)$ *imply* $f \in \mathcal{S}^2_\rho$ *and* $|Df|_{w,\rho} = |Df|_w$ *ρ-a.e. in* X*;*

(b) $\log g \in \mathcal{S}^2_\rho$ *and* $|D\log g|_{w,\rho} = |Dg|_w/g$ *ρ-a.e. in* X*.*

Proof. (a) Thanks to the locality properties with respect to $\mathfrak{m}$ stated after Definition 3.2 (see in particular (3.2)) we can reduce ourselves to the case when $\mathfrak{m}(X) = 1$. Since the statement is invariant under modification of f and g in $\mathfrak{m}$-negligible sets, by Theorem 3.5(b) we can assume that $\sqrt{g}$ and f are absolutely continuous along $\mathrm{Mod}_{2,\mathfrak{m}}$-almost every curve in $\mathscr{C}(X)$; even more, we can assume that f has an upper gradient H with $\int H^2\,\mathrm{d}\mathfrak{m} < \infty$.

Let us prove first the inequality $|Df|_{w,\rho} \le |Df|_w$ ρ-a.e. in X. By a truncation argument we can assume with no loss of generality that f is bounded; under this assumption we can find bounded Lipschitz functions ϕ_n with $|D\phi_n| \to |Df|_w$ in $L^2(X,\mathfrak{m})$. Since g is bounded it follows that $|D\phi_n| \to |Df|_w$ in $L^2(X,\rho)$; we can now use the stability properties of weak upper gradients [3, Theorem 5.12] to obtain that $|Df|_{w,\rho} \le |Df|_w$ ρ-a.e. in X.

In order to prove the converse inequality $|Df|_{w,\rho} \ge |Df|_w$ ρ-a.e. in X, we consider a function $\tilde f$ coinciding with f ρ-a.e. in X and an upper gradient L of $\tilde f$ with $\int L^2\,\mathrm{d}\rho < \infty$. The converse inequality follows by letting $L \to |Df|_{w,\rho}$ in $L^2(X,\rho)$, if we are able to show that

$$L_1(x) := \begin{cases} H(x) & \text{if } g(x) = 0;\\ \min\{H(x), L(x)\} & \text{if } g(x) > 0, \end{cases}$$

is a 2-weak upper gradient of f relative to $\mathfrak{m}$. More precisely, we will prove that the upper gradient inequality with L_1 in the right hand side holds along $\mathrm{Mod}_{2,\mathfrak{m}}$-almost every curve. We notice first that

$$|\tilde f(\gamma_{\ell(\gamma)}) - \tilde f(\gamma_0)| \le \int_\gamma L$$

along $\mathrm{Mod}_{2,\mathfrak{m}}$-a.e. curve γ satisfying $\inf_\gamma g > 0$ (here we are using the invariance under reparameterization, selecting the arclength one, with $\ell(\gamma)$ equal to the length of γ). Indeed, by definition of 2-modulus, the set

$$\left\{\gamma \in \mathscr{C}(X) :\ \inf_\gamma g > 0,\ \int_\gamma L = \infty\right\}$$

is not only $\mathrm{Mod}_{2,\rho}$-negligible, but also $\mathrm{Mod}_{2,\mathfrak{m}}$-negligible. Now, if we write the upper gradient inequality in averaged form

$$\frac{1}{\epsilon\ell(\gamma)}\int_0^{\epsilon\ell(\gamma)}|\tilde f(\gamma_{\ell(\gamma)-r})-\tilde f(\gamma_r)|\,\mathrm{d}r\le\int_\gamma L\quad\text{with}\quad \epsilon<\frac{1}{2}$$

and using Theorem 3.5(a) with the $\mathfrak{m}$-negligible set $N=\{f\neq\tilde f\}\cap\{g>0\}$, we may replace $\tilde f$ with f in the previous inequality. Now we use the absolute continuity of f along $\mathrm{Mod}_{2,\mathfrak{m}}$-a.e. curve and pass to the limit along a sequence $\epsilon_k\downarrow 0$ to get

$$|f(\gamma_b)-f(\gamma_a)|\le\int_\gamma L$$

along $\mathrm{Mod}_{2,\mathfrak{m}}$-a.e. curve $\gamma:[a,b]\to X$ with $\inf_\gamma g>0$.

Now, the set of curves $\gamma\in\mathscr{C}(X)$ containing a subcurve $\gamma':[a,b]\to X$ with $\inf_{\gamma'}g>0$ and $|f(\gamma'_b)-f(\gamma'_a)|>\int_{\gamma'}L$ is $\mathrm{Mod}_{2,\mathfrak{m}}$-negligible as well. If γ does not belong to this set and $f\circ\gamma$ is absolutely continuous, it is immediate to check (recall that g is continuous along $\mathrm{Mod}_{2,\mathfrak{m}}$-almost every curve) that its derivative is bounded a.e. by $L_1\circ\gamma|\dot\gamma|$, whence the upper gradient inequality along γ follows.

(b) We consider the functions $f_\varepsilon=\log(g+\varepsilon)$. Since $|Dg|_w^2/g^2\in L^1(X,\rho)$ it is immediate to check that all functions f_ε satisfy the assumption in (a), hence $f_\varepsilon\in\mathcal{S}^2_\rho$ and $|Df_\varepsilon|_{w,\rho}=|Df_\varepsilon|_w=|Dg|_w/(g+\varepsilon)$ ρ-a.e. in X. We can now pass to the limit as $\varepsilon\downarrow 0$ and use again the stability of weak upper gradients to get $|Df|_{w,\rho}\le|Dg|_w/g$ ρ-a.e. in X. The converse inequality follows by the chain rule (3.4) with $\phi(s):=\log(\mathrm{e}^s+1)$:

$$\frac{|Dg|_w}{g+1}=|Df_1|_{w,\rho}=\phi'(f)|Df|_{w,\rho}=\frac{g}{g+1}|Df|_{w,\rho}.$$

□

Remark 3.7. Notice that for the validity of (a) suffices, as the proof shows, the existence of a nonnegative function $\tilde g$ continuous along $\mathrm{Mod}_{2,\mathfrak{m}}$-a.e. curve and satisfying $\mathfrak{m}(\{g\neq\tilde g\})=0$.

We shall define $\mathsf{Ch}:L^1(X,\mathfrak{m})\to[0,\infty]$, $\mathsf{Ch}_{\mathfrak{n}}:L^1(X,\mathfrak{n})\to[0,\infty]$ by

$$\mathsf{Ch}(f):=\frac{1}{2}\int_X|Df|_w^2\,\mathrm{d}\mathfrak{m}\quad f\in\mathcal{S}^2,\qquad \mathsf{Ch}_{\mathfrak{n}}(f):=\frac{1}{2}\int_X|Df|_{w,\mathfrak{n}}^2\,\mathrm{d}\mathfrak{n}\quad f\in\mathcal{S}^2_{\mathfrak{n}}\tag{3.5}$$

with the conventions $\mathsf{Ch}(f)=\infty$ on $L^1(X,\mathfrak{m})\setminus\mathcal{S}^2$, $\mathsf{Ch}_{\mathfrak{n}}(f)=\infty$ on $L^1(X,\mathfrak{n})\setminus\mathcal{S}^2_{\mathfrak{n}}$. We will choose $\mathfrak{n}$, as explained in the introduction, to be probability measures.

We shall also denote, whenever Ch (resp. $\mathsf{Ch}_{\mathfrak{n}}$) is a quadratic form, by

$$\mathcal{E}(f,g):=\frac{1}{2}\big(\mathsf{Ch}(f+g)-\mathsf{Ch}(f-g)\big)\qquad\Big(\text{resp. }\mathcal{E}_{\mathfrak{n}}(f,g):=\frac{1}{2}\big(\mathsf{Ch}_{\mathfrak{n}}(f+g)-\mathsf{Ch}_{\mathfrak{n}}(f-g)\big)\Big)$$

the associated symmetric bilinear form, defined on $\mathcal{S}^2\cap L^1(X,\mathfrak{m})$ (resp. $\mathcal{S}^2_{\mathfrak{n}}\cap L^1(X,\mathfrak{n})$).

Still under the assumption that Ch is quadratic, as in [4, Definition 4.13] (see also [17] for a more general, non-quadratic framework) we can define

$$G\big(f,g\big):=\lim_{\varepsilon\downarrow 0}\frac{|D(f+\varepsilon g)|_w^2-|Df|_w^2}{2\varepsilon}\qquad f,\,g\in\mathcal{S}^2,\tag{3.6}$$

where the limit takes place in $L^1(X,\mathfrak{m})$. Notice that $G\big(f,f\big)=|Df|_w^2$ $\mathfrak{m}$-a.e. and that $G\big(\cdot,\cdot\big)$ provides integral representation to $\mathcal{E}$, namely

$$\mathcal{E}(f,g)=\int_X G\big(f,g\big)\,\mathrm{d}\mathfrak{m}.$$

The inequality $|D(f+\varepsilon g)|_w^2\le\big(|Df|_w+\varepsilon|Dg|_w\big)^2=|Df|_w^2+2\varepsilon|Df|_w|Dg|_w+\varepsilon^2|Dg|_w^2$ provides the bound

$$\big|G\big(f,g\big)\big|\le|Df|_w|Dg|_w\qquad \mathfrak{m}\text{-a.e. in } X. \tag{3.7}$$

Also, locality of weak gradients gives

$$G\big(f,g\big)=G\big(f,g'\big)\qquad \mathfrak{m}\text{-a.e. on } \{g=g'\}. \tag{3.8}$$

We will need a chain rule with respect to the second argument, see [4, Lemma 4.7] for the simple proof:

$$\int_X G\big(f,\phi(g)\big)=\int_X \phi'(g)G\big(f,g\big)\qquad \mathfrak{m}\text{-a.e. in } X \tag{3.9}$$

for all $\phi:\mathbb{R}\to\mathbb{R}$ nondecreasing and Lipschitz on an interval containing the image of g, with the same convention on the value of $\phi'(g)$ mentioned in (3.4). Finally, we will need the following lemma, whose proof is more delicate: it relies on the chain rule for $G\big(\cdot,\cdot\big)$ also with respect to the first factor and on the Leibniz rule with respect to the second factor (see [4] for finite measures and [17, Proposition 4.20] for the general case).

Lemma 3.8. *If Ch is quadratic, then $G\big(\cdot,\cdot\big)$ is a symmetric bilinear form. In particular $\int|Df|_w^2 g\,\mathrm{d}\mathfrak{m}=\int G\big(f,f\big)g\,\mathrm{d}\mathfrak{m}$ is a quadratic form for any nonnegative $g\in L^\infty(X,\mathfrak{m})$.*

Theorem 3.9 (Weighted Cheeger energy)**.** *Assume that $\rho=g\mathfrak{m}\in\mathscr{P}_2(X)$ with $g\in L^\infty(X,\mathfrak{m})$ and $\mathsf{Ch}(\sqrt{g})<\infty$. If Ch is a quadratic form, then Ch_ρ is a quadratic form and*

$$\mathcal{E}_\rho(\log g,\varphi)=\mathcal{E}(g,\varphi)\qquad \textit{for all } \varphi:X\to\mathbb{R} \textit{ Lipschitz with bounded support.} \tag{3.10}$$

Proof. By Theorem 3.6(a) and Lemma 3.8, Ch_ρ is a quadratic form on bounded Lipschitz functions with bounded support. By approximation Ch_ρ is a quadratic form on bounded Lipschitz functions and eventually, taking Theorem 3.5(b) into account, on $L^2(X,\rho)$.
Let $f_\varepsilon=\log(g+\varepsilon)\in\mathcal{S}^2$. Then, using again the independence of weak gradients upon the reference measure given by Theorem 3.6(a) and (3.9), we get

$$\begin{aligned}\mathcal{E}_\rho(\varphi,f_\varepsilon)&=\lim_{\delta\downarrow 0}\frac{\mathsf{Ch}_\rho(\varphi+\delta f_\varepsilon)-\mathsf{Ch}_\rho(\varphi)}{\delta}=\lim_{\delta\downarrow 0}\int_X\frac{|D(\varphi+\delta f_\varepsilon)|_w^2-|D\varphi|_w^2}{2\delta}\,\mathrm{d}\rho\\&=\int_X G\big(\varphi,f_\varepsilon\big)\,\mathrm{d}\rho=\int_X G\big(\varphi,g\big)\frac{g}{g+\varepsilon}\,\mathrm{d}\mathfrak{m}.\end{aligned}$$

Passing to the limit as $\varepsilon\downarrow 0$ provides the result, since convergence of the right hand sides is obvious, while convergence of the left hand sides can be obtained working in the vector space $H:=L^2(X,\rho')\cap\mathcal{S}^2_\rho$ endowed with the scalar product

$$\langle h,h'\rangle:=\int_X hh'\,\mathrm{d}\rho'+\mathcal{E}_\rho(h,h')\quad\text{with}\quad \rho':=\frac{1}{1+\log^2 g}\rho.$$

This is indeed a Hilbert space because Ch_ρ is easily seen to be lower semicontinuous (since a truncation argument allows the reduction to sequences uniformly bounded in $L^\infty(X,\rho)$) also w.r.t. $L^2(X,\rho')$ convergence; moreover, clearly $f_\varepsilon\to f$ in $L^2(X,\tilde\rho)$ and since their norms are uniformly bounded we have weak convergence in H. Finally $g\mapsto\mathcal{E}_\rho(\varphi,g)$ is continuous in H. □

4 Existence of good geodesics

This section is devoted to the proof of the existence of geodesics in $(\mathscr{P}_2(X), W_2)$ which are (at least for some initial time interval) better than the ones given directly by the usual $CD(K,\infty)$ inequality. Recall that in a $CD(K,\infty)$ space we have for any μ_0, $\mu_1 \in D(\mathrm{Ent}_\mathfrak{m})$ the existence of a geodesic $(\mu_t) \in \mathrm{Geo}(\mathscr{P}_2(X))$ which satisfies the convexity inequality

$$\mathrm{Ent}_\mathfrak{m}(\mu_t) \le (1-t)\mathrm{Ent}_\mathfrak{m}(\mu_0) + t\mathrm{Ent}_\mathfrak{m}(\mu_1) - \frac{K}{2}t(1-t)W_2^2(\mu_0,\mu_1) \tag{4.1}$$

for all $t \in [0,1]$.

The idea of constructing good geodesics in $CD(K,N)$ spaces was recently used in [25]. There the initial motivation was to obtain geodesics good enough so that the approach of [26] for proving local Poincaré inequalities could be adapted to these spaces. Constructions of geodesics by selecting midpoints have been used also earlier, see for example [9].

Here we modify some results from [25] and [24] to the setting of this paper, repeating with some details the arguments because on some occasions the adaptation is not trivial. The version of these results which we will need in the later sections is the following.

Theorem 4.1. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a $CD(K,\infty)$ space and let $\mu_0 = \rho_0\mathfrak{m}$, $\mu_1 = \rho_1\mathfrak{m} \in D(\mathrm{Ent}_\mathfrak{m})$. Assume in addition that μ_1 has bounded support and density and that the density ρ_0 satisfies the growth-bound*

$$\rho_0(x) \le c_1 \mathrm{e}^{-c_2\mathsf{d}^2(x,x_0)} \qquad \forall x \in X \tag{4.2}$$

for some c_1, $c_2 > 0$ and $x_0 \in X$.

Then there exist $t_0 \in (0,1)$ and a geodesic $(\mu_t) \in \mathrm{Geo}(\mathscr{P}_2(X))$ between μ_0, μ_1 satisfying the convexity inequality (4.1) for all $t \in [0,1]$ and the density bound

$$\sup_{t\in[0,t_0]} ||\rho_t||_{L^\infty(X,\mathfrak{m})} < \infty. \tag{4.3}$$

We will construct the geodesic of Theorem 4.1 by connecting measures of minimal entropy. A similar construction was done recently in [24] in $CD^*(K,N)$ spaces.

4.1 Intermediate measures and the existence of minimizers

The measures with minimal entropy will be selected from the set of all intermediate measures. Recall that for any two measures μ_0, $\mu_1 \in \mathscr{P}_2(X)$ the set of all intermediate points (with a parameter $t \in (0,1)$), which is easily seen to be a convex and closed subset of $\mathscr{P}_2(X)$, will be denoted by

$$\mathcal{I}_t(\mu_0,\mu_1) = \{\nu \in \mathscr{P}_2(X) \,:\, W_2(\mu_0,\nu) = tW_2(\mu_0,\mu_1) \text{ and } W_2(\mu_1,\nu) = (1-t)W_2(\mu_0,\mu_1)\}.$$

Even though the selection process is countable, it will define the whole geodesic by completion. To get the convexity inequality (4.1) for all times we will then need the lower semicontinuity of the entropy w.r.t. W_2-convergence (a direct consequence of (2.3) and of the weak lower semicontinuity of $\mathrm{Ent}_\mathfrak{n}$ in $\mathscr{P}(X)$ when $\mathfrak{n} \in \mathscr{P}(X)$) and tightness estimates. Let us now indicate how the first property of the good geodesics follows easily if we define the geodesic by taking any intermediate point where (4.1) is satisfied.

Proposition 4.2. *Let μ_0, $\mu_1 \in \mathscr{P}_2(X)$. Suppose that we have selected inductively at step $(n+1)$ measures $\mu_t \in \mathcal{I}_{\frac{t-s}{r-s}}(\mu_s, \mu_r)$ satisfying*

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \leq \frac{r-t}{r-s}\mathrm{Ent}_{\mathfrak{m}}(\mu_s) + \frac{t-s}{r-s}\mathrm{Ent}_{\mathfrak{m}}(\mu_r) - \frac{K}{2}\frac{t-s}{r-s}\frac{r-t}{r-s}W_2^2(\mu_s, \mu_r),$$

where $s < t < r$ and the times s and r are two consecutive timepoints in the set of times where the measures have already been selected at step n.

Then (4.1) holds for all μ_t chosen at the $(n+1)$-th step. In particular, if the closure of the selected times is the whole interval $[0,1]$, defining μ_t by completion, we have a geodesic between μ_0 and μ_1 along which (4.1) holds.

Proof. Suppose that we have selected a measure $\mu_t \in \mathcal{I}_t(\mu_0, \mu_1)$ satisfying

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \leq (1-t)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}t(1-t)W_2^2(\mu_0, \mu_1)$$

and after it a measure $\mu_{ts} \in \mathcal{I}_s(\mu_0, \mu_t)$ satisfying

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_{ts}) \leq (1-s)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + s\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \frac{K}{2}s(1-s)W_2^2(\mu_0, \mu_t).$$

Then for the measure μ_{ts} we also have $\mu_{ts} \in \mathcal{I}_{ts}(\mu_0, \mu_1)$ and

$$\begin{aligned}
\mathrm{Ent}_{\mathfrak{m}}(\mu_{ts}) &\leq (1-s)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + s\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \frac{K}{2}s(1-s)W_2^2(\mu_0, \mu_t) \\
&\leq (1-s)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + s\left((1-t)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + t\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}t(1-t)W_2^2(\mu_0, \mu_1)\right) \\
&\quad - \frac{K}{2}s(1-s)W_2^2(\mu_0, \mu_t) \\
&= \left((1-s) + s(1-t)\right)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + ts\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}\left(ts(1-t) + t^2s(1-s)\right)W_2^2(\mu_0, \mu_1) \\
&= (1-ts)\mathrm{Ent}_{\mathfrak{m}}(\mu_0) + ts\mathrm{Ent}_{\mathfrak{m}}(\mu_1) - \frac{K}{2}ts(1-ts)W_2^2(\mu_0, \mu_1).
\end{aligned}$$

Therefore the claim holds for all the points t_i. By the lower semicontinuity of the entropy it then holds also for the closure. □

Now that we know from Proposition 4.2 that the first property of the geodesic in Theorem 4.1 is easily satisfied we turn to the more difficult part of obtaining the density bound (4.3). To do this we will not only select intermediate measures that satisfy (4.1), but measures where the entropy is minimal. The obvious first step is then to prove that there indeed exist such minimizers. In general the set $\mathcal{I}_t(\mu_0, \mu_1)$, though closed, is not compact in $(\mathscr{P}_2(X), W_2)$. However, when we consider a subset of $\mathcal{I}_t(\mu_0, \mu_1)$ with the entropy bounded from above, we have compactness. In particular, we therefore have the existence of minimizers.

Lemma 4.3. *Let μ_0, $\mu_1 \in \mathscr{P}_2(X)$. Then for all $t \in [0,1]$ there exists a minimizer of the entropy in $\mathcal{I}_t(\mu_0, \mu_1)$.*

Proof. Without loss of generality we can assume the existence of $\nu \in \mathcal{I}_t(\mu_0, \mu_1)$ with $\mathrm{Ent}_{\mathfrak{m}}(\nu) < \infty$. We know that the entropy is lower semicontinuous and that $\mathcal{I}_t(\mu_0, \mu_1)$ is closed. The claim then follows if we are able to show that the set

$$\mathcal{K} = \{\mu \in \mathcal{I}_t(\mu_0, \mu_1) \,:\, \mathrm{Ent}_{\mathfrak{m}}(\mu) \le \mathrm{Ent}_{\mathfrak{m}}(\nu)\} \subset \mathscr{P}_2(X)$$

is relatively compact in $(\mathscr{P}_2(X), W_2)$. It suffices to prove that the set $\mathcal{K}$ is uniformly 2-integrable and tight, see [2, Proposition 7.15]. Let us first prove the uniform 2-integrability of the set $\mathcal{I}_t(\mu_0, \mu_1)$. This follows from the fact that for any $\mu \in \mathcal{I}_t(\mu_0, \mu_1)$ we have

$$\int_{X \setminus \overline{B}(x_0,k)} \mathsf{d}^2(x_0, x)\,\mathrm{d}\mu \le \int_{X \setminus \overline{B}(x_0,k/2)} 4\mathsf{d}^2(x_0, x)\,\mathrm{d}(\mu_0 + \mu_1) \to 0, \quad \text{as } k \to \infty$$

since μ_0, $\mu_1 \in \mathscr{P}_2(X)$.

Let us next prove that $\mathcal{K}$ is tight. If $\tilde{\mathfrak{m}} \in \mathscr{P}(X)$ is defined as in (2.1), (2.3) shows that $\sup_{\mu \in \mathcal{K}} \mathrm{Ent}_{\tilde{\mathfrak{m}}}(\mu)$ is finite. Then, tightness of $\mathcal{K}$ is a simple consequence of the equi-integrability of the densities w.r.t. $\tilde{\mathfrak{m}}$.

□

As a technical tool we will need the excess mass functional $\mathcal{F}_C \colon \mathscr{P}_2(X) \to [0, 1]$ which is defined for all thresholds $C \ge 0$ as

$$\mathcal{F}_C(\mu) = \|(\rho - C)^+\|_{L^1(X,\mathfrak{m})} + \mu^s(X),$$

where $\mu = \rho\mathfrak{m} + \mu^s$ with $\mu^s \perp \mathfrak{m}$. This functional, lower semicontinuous under weak convergence, was used in [25] to obtain the first good geodesics in $CD(K, N)$ spaces. The motivation for using the excess mass functional is that its variations under perturbation of the minimizer are easier to estimate, since one only cares about the amount of mass exceeding the threshold.

4.2 Localization in transport distance

As we will later see, the task of finding the first good intermediate measure between μ_0 and μ_1 is slightly more difficult than finding the rest of the geodesic. This is due to the fact that after some μ_t with $t \in (0, 1)$ has been fixed we can consider the transport distances to be essentially constant. This useful observation was made in [24]. It follows from two simple statements. First of all when one fixes an intermediate measure, also the lengths of the curves along which the transport is done gets fixed. This is the content of the next proposition which was proved in [24, Proposition 2.5].

Proposition 4.4. *Let μ_0, $\mu_1 \in \mathscr{P}_2(X)$ and $t_0 \in (0, 1)$. Suppose that there exist constants $0 \le C_1 \le C_2 < \infty$ and a measure $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ with*

$$C_1 \le l(\gamma) \le C_2 \qquad \textit{for } \boldsymbol{\pi}\textit{-a.e. } \gamma \in \mathrm{Geo}(X). \tag{4.4}$$

Then the bounds in (4.4) *hold $\tilde{\boldsymbol{\pi}}$-a.e. for any $\tilde{\boldsymbol{\pi}} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ with $(\mathrm{e}_{t_0})_\sharp \tilde{\boldsymbol{\pi}} = (\mathrm{e}_{t_0})_\sharp \boldsymbol{\pi}$.*

In order to use the previous proposition we will need another observation which is a simple consequence of cyclical monotonicity. Namely, when we work on a part of the transport with some bounds on the lengths of the curves, this part will not get mixed with other parts of the measure at any intermediate time. For the proof of this fact see [24, Lemma 2.6].

Lemma 4.5. *Take $0 \le C_1 \le C_2 \le C_3 \le C_4 \le \infty$ and define*

$$A_1 = \{\gamma \in \mathrm{Geo}(X) \,:\, C_1 \le l(\gamma) \le C_2\} \quad \text{and} \quad A_2 = \{\gamma \in \mathrm{Geo}(X) \,:\, C_3 < l(\gamma) \le C_4\}.$$

Then for any $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ and any $t \in (0,1)$ there exists a Borel set $E \subset \mathrm{Geo}(X)$ with $\boldsymbol{\pi}(E) = 0$ such that

$$\{(\gamma, \hat{\gamma}) \in (A_1 \setminus E) \times (A_2 \setminus E) \,:\, \gamma_t = \hat{\gamma}_t\} = \emptyset.$$

4.3 Density bounds for the minimizers

The information from the minimizers of the entropy and of the excess mass functional are obtained with a contradiction argument. First we assume that there exists a minimizer which does not have the desired density bound. After this we isolate the part of the minimizer where the density bound is exceeded and redefine this part of the measure to be something slightly better. If this new measure is again an intermediate point and we have decreased the energy we are minimizing (the entropy or the excess mass) we are done. To prove that we indeed get an intermediate point we use the next lemma, whose proof relies on the joint convexity of $(\mu, \nu) \mapsto W_2^2(\mu, \nu)$, which was again proved in [25, Lemma 3.5].

Lemma 4.6. *Let μ_0, $\mu_1 \in \mathscr{P}_2(X)$. Then for any $\lambda \in (0,1)$, any $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$, any Borel function $f \colon \mathrm{Geo}(X) \to [0,1]$ with $c = (f\boldsymbol{\pi})(\mathrm{Geo}(X)) \in (0,1)$ and any*

$$\nu \in \mathcal{I}_\lambda\left(\frac{1}{c}(\mathrm{e}_0)_\sharp\,(f\boldsymbol{\pi})\,, \frac{1}{c}(\mathrm{e}_1)_\sharp\,(f\boldsymbol{\pi})\right)$$

we have

$$(\mathrm{e}_\lambda)_\sharp\,((1-f)\boldsymbol{\pi}) + c\nu \in \mathcal{I}_\lambda(\mu_0, \mu_1).$$

The first step which uses the minimization of functionals is the same one that was taken in [25, Proposition 3.11]. We repeat some key points of the proof for the convenience of the reader. In [25] the functionals $\mathcal{F}_C$ were minimized only in the bounded case. A reduction to this case can be also made here and so the following proposition which was proved in a slightly different form in [25, Proposition 3.9 and Proposition 3.11] will suffice.

Proposition 4.7. *Assume that (X, d) is a bounded metric space with a finite measure $\mathfrak{m}$. Let ν_0, $\nu_1 \in \mathscr{P}_2(X)$ and $t \in [0,1]$. Suppose that there exists a constant $C > 0$ so that for any $\boldsymbol{\pi} \in \mathrm{OptGeo}(\nu_0, \nu_1)$ and $A \subset X$ Borel with $\boldsymbol{\pi}(\mathrm{e}_t^{-1}(A)) > 0$ we have that for the measures*

$$\hat{\nu}_0 = \frac{1}{\boldsymbol{\pi}(\mathrm{e}_t^{-1}(A))}(\mathrm{e}_0)_\sharp\left(\boldsymbol{\pi} \llcorner \mathrm{e}_t^{-1}(A)\right), \qquad \hat{\nu}_1 = \frac{1}{\boldsymbol{\pi}(\mathrm{e}_t^{-1}(A))}(\mathrm{e}_1)_\sharp\left(\boldsymbol{\pi} \llcorner \mathrm{e}_t^{-1}(A)\right) \tag{4.5}$$

there exists a measure $\hat{\nu} \in \mathcal{I}_t(\hat{\nu}_0, \hat{\nu}_1)$ with

$$\mathrm{Ent}_{\mathfrak{m}}(\hat{\nu}) \le \log \frac{C}{\boldsymbol{\pi}(\mathrm{e}_t^{-1}(A))}. \tag{4.6}$$

Then there exists a minimizer of $\mathcal{F}_C$ in $\mathcal{I}_t(\nu_0, \nu_1)$ and the minimum is zero.

Proof. Take a threshold $C' > C$. It suffices to prove that the minimum of $\mathcal{F}_{C'}$ in $\mathcal{I}_t(\nu_0, \nu_1)$ is zero and then let $C' \downarrow C$. Without loss of generality we may assume that all minimizers, whose existence is ensured by tightness of $\mathcal{I}_t(\nu_0, \nu_1)$ in $\mathscr{P}(X)$ and lower semicontinuity, are absolutely continuous with respect to $\mathfrak{m}$. Indeed, suppose that there is a measure $\omega \in \mathcal{I}_t(\nu_0, \nu_1)$ with a singular part. Take as A a $\mathfrak{m}$-negligible Borel set where the singular part of ω is concentrated. By the assumption of the Proposition together with Lemma 4.6 we can then redefine the part of ω which is supported on A to be a measure having finite entropy. In particular it will be absolutely continuous with respect to $\mathfrak{m}$. Since we are redefining only the singular part of ω, the value of the functional $\mathcal{F}_{C'}$ does not increase after the redefinition.

Assume, contrary to the claim, that the infimum of $\mathcal{F}_{C'}$ in $\mathcal{I}_t(\nu_0, \nu_1)$ is positive. Denote by $\mathcal{M}_{\min} \subset \mathcal{I}_t(\nu_0, \nu_1)$ the set of minimizers of $\mathcal{F}_{C'}$ in $\mathcal{I}_t(\nu_0, \nu_1)$. With a similar proof as for [25, Proposition 3.9] we see that the set $\mathcal{M}_{\min}$ is always nonempty. Take $\nu \in \mathcal{M}_{\min}$ for which

$$\mathfrak{m}(\{x \in X \ : \ \rho_\nu(x) > C'\}) \geq \left(\frac{C}{C'}\right)^{\frac{1}{4}} \sup_{\omega \in \mathcal{M}_{\min}} \mathfrak{m}(\{x \in X \ : \ \rho_\omega(x) > C'\}), \tag{4.7}$$

where $\nu = \rho_\nu \mathfrak{m}$ and $\omega = \rho_\omega \mathfrak{m}$. Let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\nu_0, \nu_1)$ be such that $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \nu$.

There exists $\delta > 0$ so that

$$\mathfrak{m}(A) > \left(\frac{C}{C'}\right)^{\frac{1}{2}} \mathfrak{m}(A')$$

with

$$A' = \{x \in X \ : \ \rho_\nu(x) > C'\} \quad \text{and} \quad A = \{x \in A' \ : \ \rho_\nu(x) > C' + \delta\}. \tag{4.8}$$

From the assumption of the proposition we know the existence of a measure $\hat{\nu} = \hat{\rho}\mathfrak{m} \in \mathcal{I}_t(\hat{\nu}_0, \hat{\nu}_1)$ with $\mathrm{Ent}_{\mathfrak{m}}(\hat{\nu}) \leq \log(C/\nu(A))$, where $\hat{\nu}_0$ and $\hat{\nu}_1$ are given by (4.5). By Jensen's inequality we then have

$$\mathfrak{m}(\{\hat{\rho} > 0\}) \geq \frac{\nu(A)}{C} \geq \frac{C'}{C}\mathfrak{m}(A) \geq \left(\frac{C'}{C}\right)^{\frac{1}{2}} \mathfrak{m}(A'). \tag{4.9}$$

We can now consider a new measure $\tilde{\nu} = \tilde{\rho}\mathfrak{m}$ defined as the combination

$$\tilde{\nu} = \nu \llcorner (X \setminus A) + \frac{C'}{C' + \delta} \nu \llcorner A + \frac{\delta}{C' + \delta} \nu(A) \hat{\nu}. \tag{4.10}$$

By Lemma 4.6 and the convexity of $\mathcal{I}_t$ we have $\tilde{\nu} \in \mathcal{I}_t(\nu_0, \nu_1)$. Due to the definition (4.8) we only redistribute some of the mass above the density C' when we replace the measure ν by the measure $\tilde{\nu}$, so that $\tilde{\nu} \in \mathcal{M}_{\min}$. Let us calculate how much the excess mass functional changes in this replacement:

$$\mathcal{F}_{C'}(\nu) - \mathcal{F}_{C'}(\tilde{\nu}) = \int_{\{\rho_\nu < C'\}} \min\left\{C' - \rho_\nu, \frac{\delta}{C' + \delta}\nu(A)\hat{\rho}\right\} \mathrm{d}\mathfrak{m}.$$

Because of the minimality of $\mathcal{F}_{C'}$ at ν this integral must be zero. Therefore $\{\hat{\rho} > 0\} \cap \{\rho_\nu < C'\}$ is $\mathfrak{m}$-negligible. On the other hand, for any $y \in \{\hat{\rho} > 0\} \cap \{\rho_\nu \geq C'\}$ we have $\tilde{\rho}(y) > C'$ (if $y \in X \setminus A$ this is trivial, if $y \in A$ the second term in (4.10) gives a contribution larger than C'). This, together with our choice (4.7) of ν, leads to a contradiction:

$$\mathfrak{m}(\{\tilde{\rho} > C'\}) \geq \mathfrak{m}(\{\hat{\rho} > 0\}) \geq \left(\frac{C'}{C}\right)^{\frac{1}{2}} \mathfrak{m}(A') \geq \left(\frac{C'}{C}\right)^{\frac{1}{4}} \sup_{\omega \in \mathcal{M}_{\min}} \mathfrak{m}(\{\rho_\omega > C'\}).$$

□

Next we make another minimization. This time for the entropy itself. A similar argument was used in [24] to obtain good geodesics in metric spaces satisfying the reduced curvature dimension condition $CD^*(K,N)$.

Proposition 4.8. *Let μ_0, $\mu_1 \in \mathscr{P}_2(X)$ and $t \in [0,1]$. Suppose that there exists a constant $C > 0$ so that for any $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ and $A \subset X$ Borel with $\boldsymbol{\pi}(\mathrm{e}_t^{-1}(A)) > 0$ we have that for the restricted measures $\hat{\nu}_0$, $\hat{\nu}_1$ in (4.5) there exists a measure $\hat{\nu} \in \mathcal{I}_t(\hat{\mu}_0, \hat{\mu}_1)$ satisfying (4.6). Then for any minimizer μ_{min} of the entropy in $\mathcal{I}_t(\mu_0, \mu_1)$ we have $\mu_{min} \le C\mathfrak{m}$.*

Proof. Without loss of generality, we can assume $t \in (0,1)$. Let $\nu = \rho\mathfrak{m}$ be one of the minimizers of the entropy in $\mathcal{I}_t(\mu_0, \mu_1)$, which by Lemma 4.3 we know to exist. By (4.6) with $A = X$ we know that $\mathrm{Ent}_{\mathfrak{m}}(\nu) < \infty$. We will show that $\mathcal{F}_C(\nu) = 0$.

Let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ be such that $(\mathrm{e}_t)_\sharp \boldsymbol{\pi} = \nu$. Suppose now by contradiction that $\mathcal{F}_C(\nu) > 0$, let $\eta > 0$ be such that $\mathfrak{m}(\{\rho > C + 2\eta\}) > 0$ and define

$$C_1 = \frac{1}{\eta}\left[\mathfrak{m}(\{\rho > C + \eta\}) - \mathfrak{m}(\{\rho > C + 2\eta\})\right] \ge 0.$$

Since $\tau \mapsto g(\tau) := \mathfrak{m}(\{\rho \ge C + \tau\})$ is nonincreasing, there exists $\delta \in (\eta, 2\eta)$ such that $-g'(\delta) \le C_1$. In particular, choosing δ in this way and fixing $x_0 \in X$, for $\phi \in (0, \eta/3)$ sufficiently small and $R = R(\phi)$ sufficiently large it holds $\mathfrak{m}(L') < \mathfrak{m}(L) + (1 + C_1)\phi$, where

$$L = \{x \in B(x_0, R) \,:\, \rho(x) > C + \delta\} \qquad \text{and} \qquad L' = \{x \in X \,:\, \rho(x) \ge C + \delta - 3\phi\}.$$

Since L is bounded it follows from cyclical monotonicity (see [31, Theorem 8.22]) that $\boldsymbol{\pi}\llcorner\mathrm{e}_t^{-1}(L)$ is supported in a uniformly bounded set of curves. Therefore we can use Proposition 4.7 with $\nu_i = (\nu(L))^{-1}(\mathrm{e}_i)_\sharp\boldsymbol{\pi}\llcorner\mathrm{e}_t^{-1}(L)$ to find a measure

$$\tilde{\nu} = \tilde{\rho}\mathfrak{m} \in \mathcal{I}_t\left(\frac{(\mathrm{e}_0)_\sharp\boldsymbol{\pi}\llcorner\mathrm{e}_t^{-1}(L)}{\nu(L)}, \frac{(\mathrm{e}_1)_\sharp\boldsymbol{\pi}\llcorner\mathrm{e}_t^{-1}(L)}{\nu(L)}\right)$$

with $\tilde{\rho} \le C/\nu(L)$ $\mathfrak{m}$-a.e. in X.

Now consider a new measure $\hat{\nu} = \hat{\rho}\mathfrak{m}$ defined as the combination

$$\hat{\nu} = \nu\llcorner(X \setminus L) + \frac{C + \delta - \phi}{C + \delta}\nu\llcorner L + \frac{\phi}{C + \delta}\nu(L)\tilde{\nu}.$$

By Lemma 4.6 we have $\hat{\nu} \in \mathcal{I}_t(\mu_0, \mu_1)$.

For $x \in L$ we have the estimates

$$\begin{aligned}\hat{\rho}(x) &\le \frac{C + \delta - \phi}{C + \delta}\rho(x) + \frac{\phi}{C + \delta}\nu(L)\tilde{\rho}(x) \le \frac{(C + \delta - \phi)\rho(x) + C\phi}{C + \delta} \\ &= \rho(x) + \frac{(C - \rho(x))\phi}{C + \delta} < \rho(x) - \frac{\delta\phi}{C + \delta}\end{aligned} \tag{4.11}$$

and

$$\hat{\rho}(x) \ge \frac{C + \delta - \phi}{C + \delta}\rho(x) > C + \delta - \phi. \tag{4.12}$$

For $x \in L' \setminus L$ we have

$$\hat{\rho}(x) \le \rho(x) + \frac{\phi}{C + \delta}\nu(L)\tilde{\rho}(x) \le \rho(x) + \frac{C\phi}{C + \delta} < C + \delta + \phi \tag{4.13}$$

and for $x \in X \setminus L'$ we get

$$\hat{\rho}(x) \leq \rho(x) + \frac{\phi}{C+\delta}\nu(L)\tilde{\rho}(x) \leq C + \delta - 3\phi + \frac{C\phi}{C+\delta} < C + \delta - 2\phi. \tag{4.14}$$

Write $C_2 = \frac{\delta}{C+\delta}\mathfrak{m}(L)$. Let us estimate the change in the entropy when we replace ν by $\hat{\nu}$: using the convexity inequality $x \log x - y \log y \leq (x-y)(\log x + 1)$ we can estimate from above $\mathrm{Ent}_{\mathfrak{m}}(\hat{\nu}) - \mathrm{Ent}_{\mathfrak{m}}(\nu)$ by

$$\int_X (\hat{\rho} - \rho)(\log \hat{\rho} + 1) \, \mathrm{d}\mathfrak{m} = \int_X (\hat{\rho} - \rho) \log \hat{\rho} \, \mathrm{d}\mathfrak{m}.$$

Now, we set $w := \hat{\rho} - \rho$, split X as $L \cup (X \setminus L') \cup (L' \setminus L)$ and use the fact that $w \leq 0$ on L and $w \geq 0$ on $X \setminus L$, the inequalities (4.11), (4.12), (4.13), (4.14) and eventually the concavity of log to get

$$\begin{aligned}
&\int_L w \log (C + \delta - \phi) \, \mathrm{d}\mathfrak{m} + \int_{X \setminus L'} w \log (C + \delta - 2\phi) \, \mathrm{d}\mathfrak{m} + \int_{L' \setminus L} w \log (C + \delta + \phi) \, \mathrm{d}\mathfrak{m} \\
= \; &(\log (C + \delta - \phi) - \log (C + \delta - 2\phi)) \int_L w \, \mathrm{d}\mathfrak{m} + (\log (C + \delta + \phi) - \log (C + \delta - 2\phi)) \int_{L' \setminus L} w \, \mathrm{d}\mathfrak{m} \\
\leq \; &-(\log (C + \delta - \phi) - \log (C + \delta - 2\phi)) \frac{\delta\phi}{C+\delta}\mathfrak{m}(L) \\
&+ (\log (C + \delta + \phi) - \log (C + \delta - 2\phi)) \frac{C\phi}{C+\delta}\mathfrak{m}(L' \setminus L) \\
< \; &-(\log (C + \delta - \phi) - \log (C + \delta - 2\phi)) C_2\phi + (\log (C + \delta + \phi) - \log (C + \delta - 2\phi)) (1 + C_1)\phi^2 \\
\leq &- C_2\phi \frac{\phi}{C + \delta - 2\phi} + (1 + C_1)\phi^2 \frac{3\phi}{C + \delta - 2\phi} < 0
\end{aligned}$$

for small enough $\phi \in (0, \eta/3)$. This contradicts the minimality of the entropy at ν. □

4.4 Construction of the geodesic

Proof of Theorem 4.1. In this proof, to avoid a cumbersome notation, we switch to the exp notation and set $C_1 := \|\rho_1\|_{L^\infty(X,\mathfrak{m})}$. Let $D > 0$ be such that $\mathrm{supp}(\mu_1) \subset B(x_0, D)$. We will prove the claim with

$$t_0 := \min\{\frac{c_2}{2K^-}, \frac{1}{2}\}.$$

The geodesic is constructed as follows. First we fix the measure $\mu_{t_0} = \rho_{t_0}\mathfrak{m} \in \mathcal{I}_{t_0}(\mu_0, \mu_1)$ to be a minimizer of the entropy in $\mathcal{I}_{t_0}(\mu_0, \mu_1)$. After this we define the rest of the geodesic for times $t \in (0, t_0)$ inductively. Suppose that for some $n \in \mathbb{N}$ we have defined $\mu_{k2^{-n}t_0}$ for all $k = 0, 1, \ldots, 2^n$. Then for all odd $k \in \mathbb{N}$ with $0 < k < 2^{n+1}$ we define $\mu_{k2^{-n-1}t_0}$ to be a minimizer of the entropy in $\mathcal{I}_{\frac{1}{2}}(\mu_{(k-1)2^{-n-1}t_0}, \mu_{(k+1)2^{-n-1}t_0})$. We construct the geodesic on the interval $(t_0, 1]$ in a similar way by iteratively selecting the midpoints with minimal entropy. The rest of the geodesic is given by completion. Let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ be such that $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} = \mu_t$ for all $t \in [0, 1]$.

Since we are selecting minimizers of the entropy among all the possible intermediate measures in a $CD(K, \infty)$-space, the selected measures satisfy the convexity inequality (4.1) between the given endpoint measures. Therefore, by Proposition 4.2 the inequality (4.1) holds for all $t \in [0, 1]$.

Let us then concentrate on the entropy estimates assumed in Proposition 4.7 and Proposition 4.8. Let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu_0, \mu_1)$ and $A \subset X$ Borel with $M := \boldsymbol{\pi}(\mathrm{e}_{t_0}^{-1}(A)) > 0$, write

$$\hat{\mu}_0 = \hat{\rho}_0 \mathfrak{m} = \frac{1}{M} (\mathrm{e}_0)_\sharp \left(\boldsymbol{\pi} \llcorner \mathrm{e}_{t_0}^{-1}(A) \right) \quad \text{and} \quad \hat{\mu}_1 = \hat{\rho}_1 \mathfrak{m} = \frac{1}{M} (\mathrm{e}_1)_\sharp \left(\boldsymbol{\pi} \llcorner \mathrm{e}_{t_0}^{-1}(A) \right),$$

and take a measure $\nu \in \mathcal{I}_{t_0}(\hat{\mu}_0, \hat{\mu}_1)$ which satisfies the convexity inequality (4.1) between these measures. Now, using (4.2), we have the estimate (with $V(x) = \mathsf{d}(x, x_0)$)

$$\begin{aligned}
\mathrm{Ent}_{\mathfrak{m}}(\nu) \leq\ & (1-t_0)\mathrm{Ent}_{\mathfrak{m}}(\hat{\mu}_0) + t_0 \mathrm{Ent}_{\mathfrak{m}}(\hat{\mu}_1) + \frac{K^-}{2} t_0 (1-t_0) W_2^2(\hat{\mu}_0, \hat{\mu}_1) \\
\leq\ & t_0 \log\left(\frac{C_1}{M}\right) + (1-t_0) \int_X \hat{\rho}_0(x) \left(\log \hat{\rho}_0(x) + \frac{K^-}{2} t_0 (D + V(x))^2 \right) \mathrm{d}\mathfrak{m}(x) \\
\leq\ & t_0 \log\left(\frac{C_1}{M}\right) + (1-t_0) \int_X \hat{\rho}_0(x) \left(\log\left(\frac{c_1}{M}\right) - c_2 V^2(x) + K^- t_0 (D^2 + V^2(x)) \right) \mathrm{d}\mathfrak{m}(x) \\
\leq\ & \log\left(\frac{\max\{C_1, c_1\}}{M}\right) + K^- D^2 = \log\left(\frac{\max\{C_1, c_1\} \exp[K^- D^2]}{M}\right),
\end{aligned}$$

since $K^- t_0 \leq c_2$ by the choice of t_0. By Proposition 4.8 we then have the estimate

$$\|\rho_{t_0}\|_{L^\infty(X, \mathfrak{m})} \leq \max\{C_1, c_1\} \exp[K^- D^2] \leq \max\{C_1, c_1\} \exp[(2K^- + c_2) D^2] =: C.$$

Next we prove that for all $t \in [0, t_0]$ we have $\mu_t = \rho_t \mathfrak{m}$ with the estimate

$$\rho_t(\gamma_t) \leq C \exp\left[-\frac{1}{2}(1 - \frac{t}{t_0})(c_2 - K^- t t_0) \ell^2(\gamma)\right] \quad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma \in \mathrm{Geo}(X). \tag{4.15}$$

First of all the estimate (4.15) is true for $t = t_0$. For $t = 0$ we have that, thanks to (4.2), $\rho_0(\gamma_0)$ can be estimated from above by

$$c_1 \exp\left[-c_2 \mathsf{d}^2(\gamma_0, x_0)\right] \leq c_1 \exp\left[-c_2([\ell(\gamma) - D]^+)^2\right] \leq c_1 \exp\left[-\frac{c_2}{2} \ell^2(\gamma) + c_2 D^2\right] \leq C \exp(-\frac{c_2}{2} \ell^2(\gamma))$$

and so (4.15) holds also at $t = 0$.

Suppose that for some $n \in \mathbb{N}$ the estimate (4.15) holds for all $t = k 2^{-n} t_0$ with $k = 0, 1, \ldots, 2^n$. Take an odd integer k with $0 < k < 2^{n+1}$. Our aim is to prove (4.15) for $t = k 2^{-n-1} t_0$.

Let $l \in (0, \infty)$ and $\epsilon > 0$ be such that we have $\tilde{M} = \boldsymbol{\pi}(\{\gamma : l \leq l(\gamma) \leq l + \epsilon\}) > 0$. Then by Proposition 4.4 we know that any measure

$$\tilde{\boldsymbol{\pi}} \in \mathrm{OptGeo}\left(\frac{1}{\tilde{M}} (\mathrm{e}_0)_\sharp \boldsymbol{\pi} \llcorner \{\gamma : l \leq \ell(\gamma) \leq l + \epsilon\}, \frac{1}{\tilde{M}} (\mathrm{e}_1)_\sharp \boldsymbol{\pi} \llcorner \{\gamma : l \leq \ell(\gamma) \leq l + \epsilon\} \right)$$

is concentrated on geodesics with lengths in the interval $[l, l + \epsilon]$. On the other hand, by Lemma 4.5 we know that

$$(\mathrm{e}_{k 2^{-n-1} t_0})_\sharp \tilde{\boldsymbol{\pi}} \perp (\mathrm{e}_{k 2^{-n-1} t_0})_\sharp \boldsymbol{\pi} \llcorner \{\gamma : \ell(\gamma) \notin [l, l + \epsilon] \text{ and } \gamma_{k 2^{-n-1} t_0} \in A\}.$$

Therefore, in proving (4.15) we may separately deal with the parts of the measure where all the geodesics have lengths in an interval $[l, l + \epsilon]$. Take now a Borel set $A \subset X$ such that for the measure $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi} \llcorner \{\gamma : l \leq \ell(\gamma) \leq l + \epsilon \text{ and } \gamma_{k 2^{-n-1} t_0} \in A\}$ we have $\hat{M} = \hat{\boldsymbol{\pi}}(\mathrm{Geo}(X)) > 0$.

Suppose that the measure

$$\tilde{\nu} \in \mathcal{I}_{\frac{1}{2}}\left(\frac{1}{\hat{M}}(\mathrm{e}_{(k-1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}, \frac{1}{\hat{M}}(\mathrm{e}_{(k+1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}\right)$$

satisfies the convexity inequality (4.1). Then

$$\begin{aligned}
\mathrm{Ent}_{\mathfrak{m}}(\tilde{\nu}) \le\ & \frac{1}{2}\mathrm{Ent}_{\mathfrak{m}}(\hat{M}^{-1}(\mathrm{e}_{(k-1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}) + \frac{1}{2}\mathrm{Ent}_{\mathfrak{m}}(\hat{M}^{-1}(\mathrm{e}_{(k+1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}) \\
& + \frac{K^-}{8}W_2^2\left(\hat{M}^{-1}(\mathrm{e}_{(k-1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}, \hat{M}^{-1}(\mathrm{e}_{(k+1)2^{-n-1}t_0})_\sharp\hat{\boldsymbol{\pi}}\right) \\
\le\ & \frac{1}{2}\log\frac{C}{\hat{M}} - \frac{1}{4}((1-(k-1)2^{-n-1})(c_2 - K^-(k-1)2^{-n-1}t_0^2)l^2) \\
& + \frac{1}{2}\log\frac{C}{\hat{M}} - \frac{1}{4}((1-(k+1)2^{-n-1})(c_2 - K^-(k+1)2^{-n-1}t_0^2)l^2) \\
& + \frac{K^-}{8}(2^{-n}t_0(l+\epsilon))^2 \\
=\ & \log\frac{C}{\hat{M}} - \frac{1}{2}((1-k2^{-n-1})(c_2 - K^-k2^{-n-1}t_0^2)l^2) + \frac{K^-}{8}2^{-2n}t_0^2(2l+\epsilon)\epsilon.
\end{aligned}$$

Proposition 4.8 then gives

$$\rho_t(\gamma_t) \le C\exp\left[-\frac{1}{2}(1-\frac{t}{t_0})(c_2 - K^- t t_0)l^2 + \frac{K^-}{8}2^{-2n}t_0^2(2l+\epsilon)\epsilon\right]$$

for $\boldsymbol{\pi}$-a.e. $\gamma \in \mathrm{Geo}(X)$ with $\ell(\gamma) \in [l, l+\epsilon]$. By letting $\epsilon \downarrow 0$ we then obtain (4.15) for $t = k2^{-n-1}t_0$.

Notice that the estimate (4.15) gives $\rho_t(\gamma_t) \le C\exp\left[-\frac{1}{2}(1-\frac{t}{t_0})(c_2 - K^- t t_0)\ell^2(\gamma)\right] \le C$ for all $t \in [0, t_0]$ for $\boldsymbol{\pi}$-a.e $\gamma \in \mathrm{Geo}(X)$, which is equivalent to (4.3). □

5 Convergence results

This section is devoted to the proof of some auxiliary convergence results. The first one deals with entropy convergence.

Lemma 5.1. *Let $f_n\mathfrak{m}$, $f\mathfrak{m}$ be positive finite measures in X. If $f_n \uparrow f$ $\mathfrak{m}$-a.e. and $\int fV^2\,\mathrm{d}\mathfrak{m} < \infty$, then*

$$\int_X f_n \log f_n\,\mathrm{d}\mathfrak{m} \to \int_X f\log f\,\mathrm{d}\mathfrak{m}. \tag{5.1}$$

The same conclusion holds if $f_n \downarrow f$ $\mathfrak{m}$-a.e. and $\int f_1 V^2\,\mathrm{d}\mathfrak{m} < \infty$.

Proof. Let us first consider the case $f_n \uparrow f$. In this case we can use formula (2.3) to reduce ourselves to the case when $\mathfrak{m}$ is a finite measure. Observe that the function $t \mapsto t\log t$ is decreasing on $[0, \mathrm{e}^{-1}]$ and increasing on $[\mathrm{e}^{-1}, \infty)$; we write it as the difference $\phi_1 - \phi_2$, with

$$\phi_1(z) := \begin{cases} -\frac{1}{\mathrm{e}} & \text{if } z \in [0, \frac{1}{\mathrm{e}}]; \\ z\log z & \text{if } z \ge \frac{1}{\mathrm{e}}, \end{cases} \qquad \phi_2(z) := \begin{cases} -\frac{1}{\mathrm{e}} - z\log z & \text{if } z \in [0, \frac{1}{\mathrm{e}}]; \\ 0 & \text{if } z \ge \frac{1}{\mathrm{e}}. \end{cases}$$

Notice that ϕ_i are nondecreasing and bounded from below. Therefore we can apply the monotone convergence theorem for $\int \phi_i(f_n)\,\mathrm{d}\mathfrak{m}$ to conclude. In the case $f_n \downarrow f$ the argument is the same, noticing that dominated convergence gives $\int f_n V^2\,\mathrm{d}\mathfrak{m} \to \int fV^2\,\mathrm{d}\mathfrak{m} < \infty$. □

Lemma 5.2. *Let $x_0 \in X$, $f\mathfrak{m}$, $g\mathfrak{m} \in \mathscr{P}_2(X)$ with $f(x) \le c_1 \mathrm{e}^{-c_2 \mathsf{d}^2(x,x_0)}$ for some c_1, $c_2 > 0$, $\inf_{B_R(x_0)} f > 0$ for all $R > 0$ and $g \in L^\infty(X,\mathfrak{m})$ and with bounded support. Let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\mu,\sigma)$ be a good geodesic given by Theorem 4.1. Then:*

(1) For $h \in \mathcal{S}^2$ with $|Dh|_w \in L^2(X, f\mathfrak{m})$ and

$$|Dh|_w^2(x) \le C(1+\mathsf{d}^2(x,x_0)) \quad \text{for any } x \in B^c_{R_*}(x_0) \tag{5.2}$$

for some C, $R_* > 0$, it holds

$$\limsup_{t\downarrow 0} \int \left| \frac{h(\gamma_t)-h(\gamma_0)}{\mathsf{d}(\gamma_t,\gamma_0)} \right|^2 \mathrm{d}\boldsymbol{\pi}(\gamma) \le \int |Dh|_w^2(\gamma_0)\,\mathrm{d}\boldsymbol{\pi}(\gamma). \tag{5.3}$$

(2) For all Kantorovich potentials φ relative to (μ,σ) with $|\nabla\varphi|$ having linear growth it holds

$$\lim_{t\downarrow 0} \frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)} = \lim_{t\downarrow 0} \frac{\mathsf{d}(\gamma_0,\gamma_t)}{t} = |D\varphi|_w(\gamma_0) \quad \text{in } L^2(C([0,1];X),\boldsymbol{\pi}). \tag{5.4}$$

Proof. (1) Call f_t the density of $(\mathrm{e}_t)_\sharp\boldsymbol{\pi}$, i.e. $(\mathrm{e}_t)_\sharp\boldsymbol{\pi} = f_t\mathfrak{m}$; we know that for $t > 0$ sufficiently small, say $t \in (0,t_0)$, f_t exists and there exists a constant C_* such that $f_t \le C_*$ $\mathfrak{m}$-a.e. in X for all $t \in (0,t_0)$. By definition of weak upper gradient, for any $t \in (0,t_0)$ and $\boldsymbol{\pi}$-a.e. γ it holds

$$\left| \frac{h(\gamma_t)-h(\gamma_0)}{\mathsf{d}(\gamma_t,\gamma_0)} \right|^2 \le \frac{\left(\int_0^t |Dh|_w(\gamma_s)|\dot\gamma_s|\mathrm{d}s\right)^2}{\mathsf{d}^2(\gamma_t,\gamma_0)} \le \frac{1}{t}\int_0^t |Dh|_w^2(\gamma_s)\mathrm{d}s,$$

therefore

$$\int \left| \frac{h(\gamma_t)-h(\gamma_0)}{\mathsf{d}(\gamma_t,\gamma_0)} \right|^2 \mathrm{d}\boldsymbol{\pi}(\gamma) \le \int \left(\frac{1}{t}\int_0^t |Dh|_w^2(\gamma_s)\mathrm{d}s\right)\mathrm{d}\boldsymbol{\pi}(\gamma) = \int_X \left(\frac{1}{t}\int_0^t f_s\mathrm{d}s\right)|Dh|_w^2\,\mathrm{d}\mathfrak{m}. \tag{5.5}$$

The conclusion of the lemma follows once the following claim is proved:

$$\lim_{t\downarrow 0}\int_X \left(\frac{1}{t}\int_0^t f_s\mathrm{d}s\right)|Dh|_w^2\,\mathrm{d}\mathfrak{m} = \int_X |Dh|_w^2 f\,\mathrm{d}\mathfrak{m}. \tag{5.6}$$

In order to prove the claim we use both the uniform L^∞ estimates on f_t and the 2-uniform integrability of V^2 w.r.t. $f_t\mathfrak{m}$. Notice first that the local boundedness of f^{-1} implies $|Dh|_w^2 \in L^1(B_R(x_0),\mathfrak{m})$ for all $R > 0$; moreover

$$\bar f_t := \left(\frac{1}{t}\int_0^t f_s\mathrm{d}s\right) \to f \quad \text{in duality with } L^1(B_R(x_0),\mathfrak{m}). \tag{5.7}$$

Indeed the weak convergence $f_t\mathfrak{m} \to f\mathfrak{m}$ implies the weak convergence of $\bar f_t$ to f in the duality with $C_b(B_R(x_0))$; then (5.7) follows by the uniform L^∞ bound on $\bar f_t$. Second, observe that (5.2) gives

$$\begin{aligned}\left|\int_X \bar f_t |Dh|_w^2\,\mathrm{d}\mathfrak{m} - \int_{B_R(x_0)} \bar f_t|Dh|_w^2\,\mathrm{d}\mathfrak{m}\right| &\le \frac{C}{t}\int_0^t\int_{B^c_R(x_0)} (1+\mathsf{d}^2(x,x_0))f_s\,\mathrm{d}\mathfrak{m}\mathrm{d}s \\ &\to 0 \quad \text{as } R\to\infty \text{ uniformly in } t \in (0,t_0);\end{aligned} \tag{5.8}$$

the second line comes from the observation that the geodesic $(f_s\mathfrak{m})_{s\in[0,1]}$ is a compact subset in $(\mathscr{P}_2(X),W_2)$, hence tight and 2-uniformly integrable (see [2, Proposition 7.1.5]). The claim (5.6) follows then combining (5.8) and (5.7).
(2) Observe we are under the assumptions of the Metric Brenier Theorem 10.3 in [3], therefore there exists a Borel function L satisfying $L(\gamma_0) := \mathsf{d}(\gamma_0,\gamma_1)$ for $\boldsymbol{\pi}$-a.e. $\gamma\in\mathrm{Geo}(X)$ and, in addition,

$$|D\varphi|_w(x) = |D^+\varphi|(x) = L(x) \quad \text{for } \mathfrak{m}\text{-a.e. } x\in X. \tag{5.9}$$

It trivially follows that for $\boldsymbol{\pi}$-a.e. $\gamma\in\mathrm{Geo}(X)$

$$|D\varphi|_w(\gamma_0) = \mathsf{d}(\gamma_0,\gamma_1) = \frac{\mathsf{d}(\gamma_0,\gamma_t)}{t} \quad \text{for every } t\in(0,1).$$

The missing part is the L^2 convergence of difference quotients, proved and stated in [3] under a different set of assumptions: we adapt the argument to our case, where $|\nabla\varphi|$ has linear growth. Since by optimality we have for $\boldsymbol{\pi}$-a.e. γ that

$$\varphi(\gamma_0)+\varphi^c(\gamma_1) = \frac{\mathsf{d}^2(\gamma_0,\gamma_1)}{2}, \qquad \varphi(\gamma_t)+\varphi^c(\gamma_1) \le \frac{\mathsf{d}^2(\gamma_t,\gamma_1)}{2},$$

we get with a subtraction that

$$\varphi(\gamma_0)-\varphi(\gamma_t) \ge \frac{1-(1-t)^2}{2}\mathsf{d}^2(\gamma_0,\gamma_1) = \frac{2t-t^2}{2}\mathsf{d}^2(\gamma_0,\gamma_1) \quad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma.$$

Therefore, dividing both sides by $\mathsf{d}(\gamma_t,\gamma_0) = t\mathsf{d}(\gamma_1,\gamma_0)$, for $\boldsymbol{\pi}$-a.e. γ it holds

$$\liminf_{t\downarrow 0}\frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)} \ge \mathsf{d}(\gamma_0,\gamma_1) = |D\varphi|_w(\gamma_0). \tag{5.10}$$

On the other hand, by definition of ascending slope

$$\limsup_{t\downarrow 0}\frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)} \le |D^+\varphi|(\gamma_0). \tag{5.11}$$

So, combining (5.9) and (5.10) with (5.11) we get

$$\lim_{t\downarrow 0}\frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)} = |D\varphi|_w(\gamma_0) \quad \text{for } \boldsymbol{\pi}\text{-a.e. } \gamma. \tag{5.12}$$

Now we claim that

$$\frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)} \rightharpoonup |D\varphi|_w\circ \mathrm{e}_0 \qquad \text{weakly in } L^2(\mathrm{Geo}(X),\boldsymbol{\pi}). \tag{5.13}$$

Since by assumption $|\nabla\varphi|$ has linear growth, by part (1) of the present lemma we have

$$\limsup_{t\downarrow 0}\int\left|\frac{\varphi(\gamma_0)-\varphi(\gamma_t)}{\mathsf{d}(\gamma_0,\gamma_t)}\right|^2 \mathrm{d}\boldsymbol{\pi} \le \int |D\varphi|_w^2(\gamma_0)\,\mathrm{d}\boldsymbol{\pi}. \tag{5.14}$$

If ψ is a weak limit point of the difference quotients as $t\downarrow 0$, by Mazur's lemma a sequence of convex combinations of these difference quotients strongly converges in $L^2(\mathrm{Geo}(X),\boldsymbol{\pi})$ to ψ. Since a further subsequence converges $\boldsymbol{\pi}$-a.e., from (5.12) we obtain that $\psi = |\nabla^+\varphi|$. By weak compactness, the claim follows.

We conclude by observing that the lower semicontinuity of the norm under weak convergence together with (5.14) ensure convergence of the $L^2(\mathrm{Geo}(X),\boldsymbol{\pi})$ norms. Since in Hilbert spaces weak convergence and convergence of the norms give strong convergence, the lemma is proved. □

Our third result deals with weak convergence in the weighted Cheeger space: it will be applied to sequences of Kantorovich potentials. In this and in the next lemma we assume that Ch is quadratic, so that Ch_ρ is quadratic whenever $\rho = g\mathfrak{m} \in \mathscr{P}_2(X)$ with $g \in L^\infty(X, \mathfrak{m})$ and with $\mathsf{Ch}(\sqrt{g}) < \infty$.

Lemma 5.3. *Let $\eta = g\mathfrak{m} \in \mathscr{P}_2(X)$ with $g \in L^\infty(X, \mathfrak{m})$ and $\mathsf{Ch}(\sqrt{g}) < \infty$. Consider a sequence $(f_n) \subset \mathcal{S}^2$ with*

$$\sup_{n\in\mathbb{N}} \int_X |Df_n|_w^2 \,\mathrm{d}\eta < \infty, \qquad \sup_{n\in\mathbb{N}} |f_n|(x) \le C(1 + \mathsf{d}^2(x, x_0)), \tag{5.15}$$

and assume that $f_n \to f$ $\mathfrak{m}$-a.e. in X. Then

$$\lim_{n\to\infty} \mathcal{E}_\eta(f_n, \log g) = \mathcal{E}_\eta(f, \log g). \tag{5.16}$$

Proof. We argue as in Theorem 3.9. Let us consider the weighted measure

$$\tilde\eta := \frac{1}{1 + V^2}\eta$$

and the corresponding weighted Sobolev space $H := L^2(X, \tilde\eta) \cap \mathcal{S}^2_\eta$, endowed with the scalar product

$$\langle f, g\rangle_H := \int_X fg \,\mathrm{d}\tilde\eta + \mathcal{E}_\eta(f, g).$$

Observe that, since $L^2(X, \tilde\eta)$ is a Hilbert space, in order to check the completeness of the norm $\|\cdot\|_H$ induced by this scalar product it is enough to check the lower semicontinuity of $\|\cdot\|_H$ with respect to strong convergence in $L^2(X, \tilde\eta)$; but this is clear since Ch_η is lower semicontinuous with respect to $L^2(X, \eta)$ convergence and, on sequences uniformly bounded in $L^\infty(X, \eta)$, the finiteness of η turns $L^2(X, \tilde\eta)$ convergence into $L^2(X, \eta)$ convergence. By a truncation argument one obtains that Ch_η is $L^2(X, \tilde\eta)$-lower semicontinuous. We conclude that $(H, \langle\cdot,\cdot\rangle_H)$ is a Hilbert space (it is even separable, see [4, Proposition 4.10], but we shall not need this fact in the sequel).

Now since $\eta \in \mathscr{P}(X)$, from the second assumption (5.15) and dominated convergence we have that $f_n \to f$ strongly in $L^2(X, \tilde\eta)$. On the other hand, the first assumption in (5.15) implies that $\|f_n\|_H$ is bounded. By reflexivity if follows that $f_n \to f$ weakly in H. The conclusion follows by noticing that, since $\mathsf{Ch}(\sqrt{g}) < \infty$, the map

$$h \mapsto \mathcal{E}_\eta(h, \log g)$$

is linear and continuous from H to $\mathbb{R}$. □

In this last result we estimate how much $\mathcal{E}_\rho(\log g, \varphi)$ changes under modifications of the density g of ρ.

Lemma 5.4. *Let $\eta = g\mathfrak{m}$, $\eta' = g'\mathfrak{m} \in \mathscr{P}_2(X)$ with g, $g' \in L^\infty(X, \mathfrak{m})$ and $\mathsf{Ch}(\sqrt{g})$, $\mathsf{Ch}(\sqrt{g'})$ finite. Let $\varphi : X \to \mathbb{R}$ be a locally Lipschitz function whose gradient has linear growth. Then, setting $E := \{g \ne g'\}$, it holds*

$$\begin{aligned} &|\mathcal{E}_\eta(\log g, \varphi) - \mathcal{E}_{\eta'}(\log g', \varphi)| \\ \le\;& \left(\int_E |D\sqrt{g}|_w^2 \,\mathrm{d}\mathfrak{m}\right)^{1/2}\left(\int_E |D\varphi|_w^2 \,\mathrm{d}\eta\right)^{1/2} + \left(\int_E |D\sqrt{g'}|_w^2 \,\mathrm{d}\mathfrak{m}\right)^{1/2}\left(\int_E |D\varphi|_w^2 \,\mathrm{d}\eta'\right)^{1/2}. \end{aligned} \tag{5.17}$$

Proof. By Lemma 5.3 we can assume, by a simple approximation argument, that φ has bounded support. Under this assumption the quantity to be estimated reduces, thanks to (3.10) and (3.8), to

$$\left|\int_X G(\varphi,g)-G(\varphi,g')\,\mathrm{d}\mathfrak{m}\right| = \left|\int_E G(\varphi,g)-G(\varphi,g')\,\mathrm{d}\mathfrak{m}\right| \le \int_E \big(|Dg|_w|D\varphi|_w+|Dg'|_w|D\varphi|_w\big)\,\mathrm{d}\mathfrak{m}$$

and, after dividing and multiplying by $\sqrt{g}$ and $\sqrt{g'}$, we can use Hölder's inequality to provide the result. □

6 Equivalence of the different formulations of $RCD(K,\infty)$

In this section we prove this result, extending Theorem 1.1 to a class of σ-finite metric measure spaces.

Theorem 6.1. *Let* $(X,\mathsf{d},\mathfrak{m})$ *be a metric measure space with* (X,d) *complete, separable, satisfying* (1.1) *and* $\operatorname{supp}\mathfrak{m}=X$. *Then the following properties are equivalent.*

(i) $(X,\mathsf{d},\mathfrak{m})$ *is a* $CD(K,\infty)$ *space and the semigroup* $\mathscr{H}_t$ *on* $\mathscr{P}_2(X)$ *is additive.*

(ii) $(X,\mathsf{d},\mathfrak{m})$ *is a* $CD(K,\infty)$ *space and* Ch *is a quadratic form on* $L^2(X,\mathfrak{m})$.

(iii) $(X,\mathsf{d},\mathfrak{m})$ *is a length space and any* $\mu\in\mathscr{P}_2(X)$ *is the starting point of an* EVI_K *gradient flow of* $\mathrm{Ent}_{\mathfrak{m}}$.

Any metric measure space $(X,\mathsf{d},\mathfrak{m})$ *satisfying these assumptions and one of the equivalent properties (i), (ii), (iii) will be called (σ-finite)* $RCD(K,\infty)$ *space.*

Recall that h_t stands for the $L^2(X,\mathfrak{m})$-gradient flow of Ch, while $\mathscr{H}_t$ is the W_2-gradient flow of $\mathrm{Ent}_{\mathfrak{m}}$. As mentioned in the introduction, the key implication from (ii) (or (i)) to (iii) is given by the derivative of quadratic optimal transport distance along the heat flow and of the entropy along a geodesic, estimated in the next two subsections. Consequently we shall always assume in this section that Ch is quadratic.

6.1 Derivative of $W_2^2(\cdot,\sigma)$ along the heat flow

Notice that this result, whose proof is achieved by a duality argument, requires no curvature assumption.

Theorem 6.2. *Let* $\mu=f\mathfrak{m}\in D(\mathrm{Ent}_{\mathfrak{m}})$ *and define* $\mu_t:=(\mathsf{h}_tf)\mathfrak{m}=f_t\mathfrak{m}$. *Let* $\sigma\in\mathscr{P}_2(X)$ *with bounded support. Then, for a.e.* $t>0$ *the following property holds: for any Kantorovich potential* φ_t *relative to* (μ_t,σ) *whose slope has linear growth, it holds*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W_2^2(\mu_t,\sigma) = -\mathcal{E}_{\mu_t}(\varphi_t,\log f_t). \tag{6.1}$$

Proof. By the conservation laws and the energy dissipation estimates proved in [3] (see Theorem 4.20 and Proposition 4.22 therein), we have $\int_0^\infty \mathsf{Ch}(\sqrt{f_t})\,\mathrm{d}t<\infty$ and $f_t\le\|f\|_\infty$ $\mathfrak{m}$-a.e. in X for all $t\ge 0$. Also, by Proposition 2.2 the potential φ_t belongs to $L^1(X,\nu)$ for all $\nu\in\mathscr{P}_2(X)$ and its slope has linear growth. Furthermore, the L^1 estimate is uniform in t and in bounded subsets of $\mathscr{P}_2(X)$ and the estimate on the slope depends on σ only.

Since $t \mapsto \rho_t \mathfrak{m}$ is a locally absolutely continuous curve in $\mathscr{P}_2(X)$, the derivative on the left hand side of (6.1) exists for a.e. $t > 0$. Also, the derivative of $t \mapsto f_t$ exists in $L^2(X, \mathfrak{m})$ for a.e. $t > 0$. Fix $t_0 > 0$ where both derivatives exist which is also a Lebesgue point for $\mathsf{Ch}(\sqrt{f_t})$.

We now claim that

$$\lim_{h \downarrow 0} \int_X \psi \frac{f_{t_0} - f_{t_0 - h}}{h} \, \mathrm{d}\mathfrak{m} = -\mathcal{E}_{\mu_{t_0}}(\psi, \log f_{t_0})$$

for all locally Lipschitz functions ψ whose gradient has linear growth. The proof of the claim is easy if we assume, in addition, that ψ has bounded support. Indeed, $h^{-1}(f_{t_0+h} - f_{t_0}) \to \Delta f_{t_0}$ as $h \to 0$ in $L^2(X, \mathfrak{m})$, so that (3.10) gives

$$\lim_{h \to 0} \int_X \psi \frac{f_{t_0+h} - f_{t_0}}{h} \, \mathrm{d}\mathfrak{m} = \int_X \psi \Delta f_{t_0} \, \mathrm{d}\mathfrak{m} = -\mathcal{E}(\psi, f_{t_0}) = -\mathcal{E}_{\mu_{t_0}}(\psi, \log f_{t_0}).$$

For the general case, let $\chi_N : X \to [0,1]$ be satisfying $\mathrm{Lip}(\chi_N) \leq 1$, $\chi_N \equiv 1$ on $B_N(x_0)$ and $\chi_N \equiv 0$ on $X \setminus B_{2N}(x_0)$ and define $\psi^N := \psi \chi_N$. Applying Lemma 6.3 below with $\varphi_N := \psi - \psi^N$ we get

$$\sup_{|h| < t_0/2} \left| \int \varphi_N \frac{\rho_{t_0+h} - \rho_{t_0}}{h} \, \mathrm{d}\mathfrak{m} \right|^2 \leq \sup_{|h| < t_0/2} \frac{8}{h} \int_{t_0 - |h|}^{t_0 + |h|} \mathsf{Ch}(\sqrt{f_s}) \int_X |D\varphi_N|_w^2 \, \mathrm{d}\mu_s \, \mathrm{d}s$$

and hence (by our choice of t_0 and the 2-uniform integrability of μ_s)

$$\limsup_{N \to \infty} \sup_{|h| < t_0/2} \left| \int_X \varphi_N \frac{f_{t_0+h} - f_{t_0}}{h} \, \mathrm{d}\mathfrak{m} \right| = 0,$$

which, taking into account that $\mathcal{E}_{\mu_{t_0}}(\psi^N, \log f_{t_0}) \to \mathcal{E}_{\mu_{t_0}}(\psi, \log f_{t_0})$ thanks to Lemma 5.3, is sufficient to conclude.

Now, notice that since φ_{t_0} is a Kantorovich potential for (μ_{t_0}, σ) it holds

$$\begin{aligned} \frac{1}{2} W_2^2(\mu_{t_0}, \sigma) &= \int_X \varphi_{t_0} \, \mathrm{d}\mu_{t_0} + \int \varphi_{t_0}^c \, \mathrm{d}\sigma \\ \frac{1}{2} W_2^2(\mu_{t_0 - h}, \sigma) &\geq \int_X \varphi_{t_0} \, \mathrm{d}\mu_{t_0 - h} + \int \varphi_{t_0}^c \, \mathrm{d}\sigma \qquad \text{for all } h \text{ such that } t_0 - h > 0. \end{aligned}$$

Taking the difference between the first identity and the second inequality and using the claim with $\psi = \varphi_{t_0}$ we get

$$\frac{1}{2} W_2^2(\mu_{t_0+h}, \sigma) - \frac{1}{2} W_2^2(\mu_{t_0}, \sigma) \geq -h \mathcal{E}_{\mu_{t_0}}(\log f_{t_0}, \varphi_{t_0}) + o(h).$$

Since $t \mapsto W_2^2(\mu_t, \sigma)$ is differentiable at $t = t_0$ we conclude. □

Lemma 6.3. *Let $\mu_s = f_s \mathfrak{m}$ be as in the previous theorem and let $\varphi : X \to \mathbb{R}$ be locally Lipschitz, with $|D\varphi|$ having linear growth. Then, for $[s,t] \subset (0, \infty)$ it holds*

$$\left| \int \varphi \frac{f_t - f_s}{t - s} \, \mathrm{d}\mathfrak{m} \right|^2 \leq \frac{8}{t - s} \int_s^t \mathsf{Ch}(\sqrt{f_r}) \left(\int |D\varphi|_w^2 \, \mathrm{d}\mu_r \right) \mathrm{d}r. \tag{6.2}$$

Proof. Assume first that $\varphi \in L^2(X,\mathfrak{m})$. Then integrating by parts we get

$$\left|\int \varphi \Delta f_r \, d\mathfrak{m}\right|^2 \le \left(\int |D\varphi|_w \, |Df_r|_w \, d\mathfrak{m}\right)^2 \le \int |D\varphi|_w^2 \, d\mu_r \int \frac{|Df_r|_w^2}{f_r} \, d\mathfrak{m},$$

for all $r > 0$, and the thesis follows by integration in (s,t). For the general case, we approximate φ by $\varphi\chi_N$, with χ_N chosen as in the proof of the previous theorem. □

6.2 Derivative of the entropy along $\mathrm{Ent}_{\mathfrak{m}}$-convex L^∞-bounded geodesics

The goal of this subsection is to prove the following theorem, where the curvature condition plays a role.

Theorem 6.4 (Entropy inequality)**.** *Assume that $(X,\mathsf{d},\mathfrak{m})$ is a $CD(K,\infty)$ space. Let $\eta = f\mathfrak{m}$, $\sigma = g\mathfrak{m} \in \mathscr{P}_2(X)$ with g uniformly bounded and having compact support, f uniformly bounded with $\mathsf{Ch}(\sqrt{f}) < \infty$. Then there exists a Kantorovich potential φ from η to σ such that $|\nabla\varphi|$ has linear growth and*

$$\mathrm{Ent}_{\mathfrak{m}}(\sigma) - \mathrm{Ent}_{\mathfrak{m}}(\eta) - \frac{K}{2} W_2^2(\eta,\sigma) \ge -\mathcal{E}_\eta(\varphi, \log f). \tag{6.3}$$

The proof of Theorem 6.4, carried by approximation, is presented at the end of the subsection; the first crucial step is the following proposition, whose proof relies on Proposition 2.2 and Lemma 5.2.

Proposition 6.5. *Under the assumptions of Theorem 6.4, for $\delta > 0$ call*

$$f_{\delta,n} = c_{\delta,n}[(\chi_n^2)\eta \vee \delta e^{-2cV^2}], \tag{6.4}$$

where c is strictly larger than the constant c in (1.1), $c_{\delta,n}$ is the normalizing constant such that $f_{\delta,n}\mathfrak{m}$ is a probability density, χ_n is a 1-Lipschitz cut-off function equal to 1 on $B_n(x_0)$ and null outside $B_{2n}(x_0)$.
Then there exists a Kantorovich potential $\varphi_{\delta,n}$ from $\eta_{\delta,n} := f_{\delta,n}\mathfrak{m}$ to σ satisfying the growth conditions

$$|\varphi_{\delta,n}(x)| \le C(\sigma)(1 + \mathsf{d}^2(x,x_0)), \qquad |D\varphi_{\delta,n}|(x) \le C(\sigma)(1 + \mathsf{d}(x,x_0)), \tag{6.5}$$

such that

$$\mathrm{Ent}_{\mathfrak{m}}(\sigma) - \mathrm{Ent}_{\mathfrak{m}}(\eta_{\delta,n}) - \frac{K}{2} W_2^2(\eta_{\delta,n},\sigma) \ge -\mathcal{E}_{\eta_{\delta,n}}(\varphi_{\delta,n}, \log f_{\delta,n}). \tag{6.6}$$

Proof. First of all we are under the assumptions of Theorem 4.1, so let $\boldsymbol{\pi} \in \mathrm{OptGeo}(\eta_{\delta,n},\sigma)$ and let $(e_t)_\sharp\boldsymbol{\pi} = \mu_t = f_t\mathfrak{m}$, $t \in [0,1]$, be the associated good geodesic from $\eta_{\delta,n}$ to σ with a uniform L^∞ bound on the density for $t \in (0,t_0)$ and the K-convexity of the entropy. Let also φ be the Kantorovich potential, given by Proposition 2.2, with quadratic growth and whose slope has linear growth.

Let us now check that $f_{\delta,n}$ satisfies the assumptions of Lemma 5.2. Indeed, $|D\log f_{\delta,n}| \le C(1+\mathsf{d}(x,x_0))$ whenever $\mathsf{d}(x,x_0) > 2n$, because in this set f_{δ_n} coincides with $c_{\delta,n}\delta e^{-2cV^2}$; in addition, the locality of weak gradients and the partition $X = \{\chi_n^2\eta > \delta e^{-2cV^2}\} \cup \{\chi_n^2\eta \le \delta e^{-2cV^2}\}$ ensure that $|D\log f_{\delta,n}|_w \in L^2(X,\eta_{\delta,n})$ because the finiteness of $\mathsf{Ch}(\sqrt{f})$ ensures that $|D\log f|_w \in L^2(X,\eta)$.

Observe that the convexity of $z \mapsto z \log z$ gives

$$\frac{\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \mathrm{Ent}_{\mathfrak{m}}(\eta_{\delta,n})}{t} \geq \int_X \log f_{\delta,n} \frac{f_t - f_{\delta,n}}{t} \, \mathrm{d}\mathfrak{m} = \int \frac{\log(f_{\delta,n} \circ \mathrm{e}_t) - \log(f_{\delta,n} \circ \mathrm{e}_0)}{t} \, \mathrm{d}\boldsymbol{\pi}. \tag{6.7}$$

Define the functions F_t, $G_t : AC^2([0,1]; X) \to \mathbb{R}$ as

$$F_t(\gamma) := \frac{\log(f_{\delta,n} \circ \mathrm{e}_0) - \log(f_{\delta,n} \circ \mathrm{e}_t)}{\mathsf{d}(\gamma_0, \gamma_t)}, \qquad G_t(\gamma) := \frac{\varphi \circ \mathrm{e}_0 - \varphi \circ \mathrm{e}_t}{\mathsf{d}(\gamma_0, \gamma_t)}. \tag{6.8}$$

Multiplying and dividing the right hand side of (6.7) by $\mathsf{d}(\gamma_0, \gamma_t)$ we obtain

$$\liminf_{t \downarrow 0} \frac{\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \mathrm{Ent}_{\mathfrak{m}}(\eta_{\delta,n}\mathfrak{m})}{t} \geq -\limsup_{t \downarrow 0} \int F_t(\gamma) \frac{\mathsf{d}(\gamma_0, \gamma_t)}{t} \mathrm{d}\boldsymbol{\pi}(\gamma) = -\limsup_{t \downarrow 0} \int F_t G_t \mathrm{d}\boldsymbol{\pi}, \tag{6.9}$$

where, in the last equality, we used that

$$\lim_{t \downarrow 0} \int \left| G_t(\gamma) - \frac{\mathsf{d}(\gamma_0, \gamma_t)}{t} \right|^2 \mathrm{d}\boldsymbol{\pi} = 0 \quad \text{and} \quad \sup_{t \leq t_0} \int |F_t|^2 \, \mathrm{d}\boldsymbol{\pi} < \infty.$$

The first fact is ensured by (2) of Lemma 5.2, as well as the identity

$$\int |D\varphi|_w^2 \circ \mathrm{e}_0 \, \mathrm{d}\boldsymbol{\pi} = \lim_{t \downarrow 0} \int |G_t|^2 \mathrm{d}\boldsymbol{\pi}. \tag{6.10}$$

The second fact is ensured by (1) of the same lemma applied to $h = \log f_{\delta,n}$.

Now, applying Lemma 5.2 to $h = \varphi + \epsilon \log f_{\delta,n}$ gives that

$$\int |D(\varphi + \epsilon \log f_{\delta,n})|_w^2 \circ \mathrm{e}_0 \mathrm{d}\boldsymbol{\pi} \geq \limsup_{t \downarrow 0} \int |G_t(\gamma) + \epsilon F_t(\gamma)|^2 \mathrm{d}\boldsymbol{\pi}(\gamma). \tag{6.11}$$

Subtracting to (6.11) the equality (6.10) and dividing by ϵ gives

$$\limsup_{t \downarrow 0} \int G_t F_t \, \mathrm{d}\boldsymbol{\pi} \leq \liminf_{\epsilon \downarrow 0} \int_X \frac{|D(\varphi + \epsilon \log f_{\delta,n})|_w^2 - |D\varphi|_w^2}{2\epsilon} f_{\delta,n} \, \mathrm{d}\mathfrak{m} = \mathcal{E}_{\eta_{\delta,n}}(\log f_{\delta,n}, \varphi), \tag{6.12}$$

where we used again the uniform bound on the L^2 norm of F_t. Combining (6.9) and (6.12) we obtain

$$\liminf_{t \downarrow 0} \frac{\mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \mathrm{Ent}_{\mathfrak{m}}(\eta_{\delta,n})}{t} \geq -\mathcal{E}_{\eta_{\delta,n}}(\log f_{\delta,n}, \varphi). \tag{6.13}$$

The conclusion follows by (6.13) recalling that, by construction, the entropy is K-convex along the geodesic $(\mu_t)_{t \in [0,1]}$, see (4.1). □

Proof of Theorem 6.4. For every $\delta \in (0,1)$ define the density

$$\tilde{f}_\delta := f \vee (\delta \mathrm{e}^{-2cV^2}) \quad \text{and } f_\delta := c_\delta \tilde{f}_\delta \text{ with } c_\delta \uparrow 1 \text{ as } \delta \downarrow 0 \tag{6.14}$$

(here $c > 0$ is the constant in (1.1)), so that $\tilde{f}_\delta \geq f$ and c_δ are the normalizing constants. We need a further regularization of f_δ; to this aim, let χ_n be standard cut-off functions, namely

$0 \le \chi_n \le 1$, $\mathrm{Lip}(\chi_n) \le 1$, $\chi_n \equiv 1$ on $B_n(x_0)$ and $\chi_n \equiv 0$ on $B^c_{2n}(x_0)$. Then, for every $n > 1$, $\delta > 0$ we define the densities

$$\tilde f_{\delta,n} := (\chi_n^2 f) \vee (\delta \mathrm{e}^{-2cV^2}) \quad \text{and } f_{\delta,n} := c_{\delta,n} \tilde f_{\delta,n} \text{ with } \ c_{\delta,n} \downarrow c_\delta \text{ as } n \to \infty, \tag{6.15}$$

so that $\tilde f_{\delta,n} \le \tilde f_\delta$ and $c_{\delta,n}$ are the normalizing constants. Of course $f_{\delta,n}$ is uniformly bounded and $\eta_{\delta,n} := f_{\delta,n}\mathfrak{m} \in \mathscr{P}_2(X)$, moreover $\mathsf{Ch}(\sqrt{f_{\delta,n}})$ is finite. Indeed by the chain rule and the locality of the weak gradients we have that

$$\begin{cases} |D\sqrt{f_{\delta,n}}|_w & = \sqrt{c_{\delta,n}}|\nabla(\chi_n\sqrt{f})|_w \\ & \le \sqrt{c_{\delta,n}}\left(\chi_n|D\sqrt{f}|_w + \sqrt{f}|D\chi_n|_w\right) \quad \text{if } \chi_n^2 f \ge \delta \mathrm{e}^{-2cV^2} \\ & \\ |D\sqrt{f_{\delta,n}}|_w & = \sqrt{\delta\, c_{\delta,n}}|D\mathrm{e}^{-2cV^2}|_w \\ & \le 4c\sqrt{\delta\, c_{\delta,n}}\,\mathsf{d}(\cdot,x_0)\, e^{-2cV^2} \qquad\qquad \text{otherwise.} \end{cases}$$

Since by assumption $\mathsf{Ch}(\sqrt{f}) < \infty$, it follows not only that $|D\sqrt{f_{\delta,n}}|_w^2$ are uniformly bounded in $L^1(X,\mathfrak{m})$, but also that they are equi-integrable:

$$\sup_{\delta\in(0,1),\, n\in\mathbb{N}} \mathsf{Ch}(\sqrt{f_{\delta,n}}) < \infty \quad \text{and} \quad E_j \downarrow \emptyset \ \Rightarrow \sup_{\delta\in(0,1),\, n\in\mathbb{N}} \int_{E_j} |D\sqrt{f_{\delta,n}}|_w^2 \,\mathrm{d}\mathfrak{m} \to 0. \tag{6.16}$$

Observe that $(\eta_{\delta,n},\sigma)$ has the structure described in Proposition 6.5, so there exists a Kantorovich potential $\varphi_{\delta,n}$ from $\eta_{\delta,n}$ to σ satisfying the growth conditions (6.5) and such that the entropy inequality holds:

$$\mathrm{Ent}_\mathfrak{m}(\sigma) - \mathrm{Ent}_\mathfrak{m}(\eta_{\delta,n}) - \frac{K}{2}W_2^2(\eta_{\delta,n},\sigma) \ge -\mathcal{E}_{\eta_{\delta,n}}(\varphi_{\delta,n}, \log f_{\delta,n}). \tag{6.17}$$

Passage to the limit as $n \to \infty$. Consider the transportation problem from $\eta_\delta := f_\delta \mathfrak{m}$ to σ. We claim the existence of a Kantorovich potential φ_δ such that

$$\mathrm{Ent}_\mathfrak{m}(\sigma) - \mathrm{Ent}_\mathfrak{m}(\eta_\delta) - \frac{K}{2}W_2^2(\eta_\delta,\sigma) \ge -\mathcal{E}_{\eta_\delta}(\varphi_\delta, \log f_\delta). \tag{6.18}$$

We would like to pass to the limit as $n \to \infty$ in (6.17). Let us start by considering the left hand side: applying Lemma 5.1 to $\tilde\eta_{\delta,n} \uparrow \tilde\eta_\delta$ $\mathfrak{m}$-a.e, and recalling that $c_{\delta,n} \downarrow c_\delta$ as $n \to \infty$, we get

$$\mathrm{Ent}_\mathfrak{m}(\eta_{\delta,n}) \to \mathrm{Ent}_\mathfrak{m}(\eta_\delta) \quad \text{as } n \to \infty. \tag{6.19}$$

It is easy to check that $\eta_{\delta,n}$ weakly converge to η_δ and have uniformly integrable 2-moments, so by [2, Proposition 7.1.5] we have

$$\lim_{n\to\infty} W_2^2(\eta_{\delta,n},\sigma) = W_2^2(\eta_\delta,\sigma). \tag{6.20}$$

Now let us show the convergence of the right hand side of (6.17). To simplify the problem we prove first that

$$\lim_{n\to\infty}\Big|\mathcal{E}_{\eta_{\delta,n}}(\varphi_{\delta,n},\log f_{\delta,n}) - \frac{c_{\delta,n}}{c_\delta}\mathcal{E}_{\eta_\delta}(\varphi_{\delta,n},\log f_\delta)\Big| = 0. \tag{6.21}$$

Notice that, calling $A_\delta := \{x \in X : \ f(x) \ge \delta \mathrm{e}^{-2cV^2(x)}\}$ we have $f_{\delta,n} = \frac{c_{\delta,n}}{c_\delta} f_\delta$ on the complement $(A_\delta \cap B_n(x_0)) \cup A_\delta^c$ of $A_\delta \setminus B_n(x_0)$. Since $A_\delta \setminus B_n(x_0) \downarrow \emptyset$ we can use (5.17) of Lemma 5.4 to obtain (6.21), taking (6.16) into account.

From (6.21), and taking into account that $c_{\delta,n} \to c_\delta$ as $n \to \infty$, in order to prove the convergence of the right hand side of (6.17), it is enough to show the existence of a Kantorovich potential φ_δ for (η_δ, σ) such that

$$\mathcal{E}_{\eta_\delta}(\varphi_{\delta,n}, \log f_\delta) \to \mathcal{E}_{\eta_\delta}(\varphi_\delta, \log f_\delta) \quad \text{as } n \to \infty. \tag{6.22}$$

Now we use in a crucial way Lemma 2.3, which ensures the existence of a Kantorovich potential φ_δ for (η_δ, σ) and of a subsequence $n(k)$ such that $\varphi_{\delta,n(k)} \to \varphi_\delta$ pointwise in X. Recalling that $|\varphi_{\delta,n}| \le C(1+V^2)$ and that $\int |\nabla \varphi_{\delta,n}|_w^2 \, \mathrm{d}\eta_\delta$ is uniformly bounded, we are in position to apply Lemma 5.3 and to conclude that (6.22) holds. Therefore we proved the convergence of all terms in (6.17), so that (6.18) holds.
Passage to the limit as $\delta \downarrow 0$. The inequality (6.18) passes to the limit as $\delta \downarrow 0$: more precisely, we claim the existence of a Kantorovich potential φ from $f\mathfrak{m}$ to σ such that

$$\mathrm{Ent}_{\mathfrak{m}}(\sigma) - \mathrm{Ent}_{\mathfrak{m}}(\eta) - \frac{K}{2} W_2^2(\eta, \sigma) \ge -\mathcal{E}_\eta(\varphi, \log f). \tag{6.23}$$

As in the passage to the limit as $n \to \infty$, Lemma 5.1 easily implies that $\mathrm{Ent}_{\mathfrak{m}}(\eta_\delta) \to \mathrm{Ent}_{\mathfrak{m}}(\eta)$, moreover it is easy to check that η_δ weakly converge to η and have uniformly integrable 2-moments, so [2, Proposition 7.1.5] gives $W_2(\eta_\delta, \sigma) \to W_2(\eta, \sigma)$. In order to show the convergence of the right hand side of (6.23) we first prove that

$$\lim_{\delta \downarrow 0} |\mathcal{E}_{\eta_\delta}(\varphi_\delta, \log f_\delta) - c_\delta \mathcal{E}_\eta(\varphi_\delta, \log f)| = 0. \tag{6.24}$$

First of all notice that, after calling $A_\delta := \{x \in X : f(x) \ge \delta \mathrm{e}^{-2cV^2(x)}\}$, we have $f_\delta = c_\delta f$ on A_δ. Since $X \setminus A_\delta \downarrow \{f = 0\}$ as $\delta \downarrow 0$ and $|Df|_w = 0$ $\mathfrak{m}$-a.e. on $\{f = 0\}$, we can use (5.17) of Lemma 5.4 to show (6.24), taking (6.16) into account.

Now that (6.24) is proved, taking into account that $c_\delta \to 1$ as $\delta \downarrow 0$, it is enough to prove the existence of a Kantorovich potential φ from η to σ such that

$$\lim_{i \to \infty} \mathcal{E}_\eta(\varphi_{\delta_i}, \log f) = \mathcal{E}_\eta(\varphi, \log f). \tag{6.25}$$

for some sequence $\delta_i \downarrow 0$. Recall that φ_δ were constructed using Lemma 2.3, so they still satisfy the growth condition (6.5); applying again Lemma 2.3 we get the existence of a Kantorovich potential φ from η to σ and $\delta_i \downarrow 0$ such that $\varphi_{\delta_i} \to \varphi$ pointwise in X as $i \to \infty$. Moreover, by (2.10) and $f \le c_\delta^{-1} f_\delta \le 2 f_\delta$ for δ small enough, we have

$$\int_X |D\varphi_{\delta_i}|_w^2 \, f \, \mathrm{d}\mathfrak{m} \le 2 \int_X |D\varphi_{\delta_i}|_w^2 f_{\delta_i} \, \mathrm{d}\mathfrak{m} \le 2 W_2^2(\eta_{\delta_i}, \sigma),$$

for i large enough. Hence we can apply Lemma 5.3 and conclude that (6.25) holds. Therefore (6.23) is proved and the proof of Theorem 6.4 is then complete. □

Proof of Theorem 6.1. The implications from (i) to (ii) and from (iii) to (i) can be proved exactly as in [4] (and these proofs need no finiteness assumption on $\mathfrak{m}$), so let us focus on the implication from (ii) to (iii). Let us first remark that a $CD(K, \infty)$ space $(X, \mathsf{d}, \mathfrak{m})$ satisfies the length property, more precisely $\operatorname{supp} \mathfrak{m}$ is a length space if [29, Remark 4.6(iii)] (the proof therein, based on an approximate midpoint construction, does not use the local compactness).

It remains to show that the EVI_K-condition holds assuming the $CD(K, \infty)$ condition and the fact that Ch is quadratic. By the contractivity properties of EVI_K-gradient flows it is sufficient to show that $\mu_t := (\mathsf{h}_t f)\mathfrak{m}$ is an EVI_K gradient flow for $\mathrm{Ent}_{\mathfrak{m}}$ for any initial measure $f\mathfrak{m} \in \mathscr{P}_2(X)$ whose density f is bounded and satisfies $\mathsf{Ch}(\sqrt{f}) < \infty$. By the conservation laws and the energy dissipation estimates proved in [3] (see Theorem 4.20 and Proposition 4.22 therein), these properties are preserved in time, so that $\mathsf{h}_t f \le \|f\|_{L^\infty(X,\mathfrak{m})}$ $\mathfrak{m}$-a.e. in X for all $t \ge 0$, $\{\mu_t : \ t \in [0,T]\}$ is a bounded subset of $\mathscr{P}_2(X)$ for all $T > 0$ and

$$\int_0^\infty \mathsf{Ch}(\sqrt{\mathsf{h}_t f})\,\mathrm{d}t < \infty. \tag{6.26}$$

By a simple density argument on the class of "test" measures σ in (1.2) (see for instance [4, Proposition 2.20]), we can restrict ourselves to measures σ of the form $g\mathfrak{m}$ with $g \in L^\infty(X,\mathfrak{m})$ and $\operatorname{supp}\sigma$ compact.

By (6.1) of Theorem 6.2 we get that for a.e. $t > 0$, for any choice of a Kantorovich potential φ_t from μ_t to σ whose slope has linear growth, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W_2^2(\mu_t,\sigma) = -\mathcal{E}_{\mu_t}(\varphi_t, \log \mathsf{h}_t f). \tag{6.27}$$

Therefore, to conclude that (1.2) holds, it suffices to show for a.e. $t > 0$ the existence of a Kantorovich potential φ_t from μ_t to σ whose slope has linear growth and satisfies

$$-\mathcal{E}_{\mu_t}(\varphi_t, \log \mathsf{h}_t f) \le \mathrm{Ent}_{\mathfrak{m}}(\sigma) - \mathrm{Ent}_{\mathfrak{m}}(\mu_t) - \frac{K}{2}W_2^2(\mu_t,\sigma). \tag{6.28}$$

This is precisely the statement of Theorem 6.4 (with $\eta = \mu_t$) and this concludes the proof.

7 Properties of $RCD(K,\infty)$ spaces

In this section we state without proof some properties of $RCD(K,\infty)$ spaces whose proofs, given in [4], do not rely on the finiteness assumption of $\mathfrak{m}$, referring to [4] for details of proofs and a more complete discussion.

7.1 The heat semigroup and its regularizing properties

In this section we describe more in detail the properties of the L^2-semigroup h_t in a $RCD(K,\infty)$ space and the additional information that one can obtain from the identification with W_2-semigroup $\mathscr{H}_t$. By the definition of $RCD(K,\infty)$ spaces, we know that for any $x \in X$ there exists a unique EVI_K gradient flow $\mathscr{H}_t(\delta_x)$ of $\mathrm{Ent}_{\mathfrak{m}}$ starting from δ_x, related to h_t by

$$(\mathsf{h}_t f)\mathfrak{m} = \int f(x)\,\mathscr{H}_t(\delta_x)\,\mathrm{d}\mathfrak{m}(x) \qquad \forall f \in L^2(X,\mathfrak{m}). \tag{7.1}$$

Since $\mathrm{Ent}_{\mathfrak{m}}(\mathscr{H}_t(\delta_x)) < \infty$ for any $t > 0$, it holds $\mathscr{H}_t(\delta_x) \ll \mathfrak{m}$, so that $\mathscr{H}_t(\delta_x)$ has a density, that we shall denote by $\rho_t[x]$. The functions $\rho_t[x](y)$ are the so-called transition probabilities of the semigroup. By standard measurable selection arguments we can choose versions of these densities in such a way that the map $(x,y) \mapsto \rho_t[x](y)$ is $\mathfrak{m}\times\mathfrak{m}$-measurable for all $t > 0$.

In the next theorem we prove additional properties of the flows. The information on both benefits of the identification theorem: for instance the symmetry property of transition probabilities is not at all obvious when looking at $\mathscr{H}_t$ only from the optimal transport point of view, and heavily relies on (7.1). On the other hand, the regularizing properties of h_t are deduced by duality by those of $\mathscr{H}_t$, using in particular the contractivity estimate

$$W_2(\mathscr{H}_t(\mu), \mathscr{H}_t(\nu)) \le \mathrm{e}^{-Kt} W_2(\mu, \nu) \qquad t \ge 0,\ \mu, \nu \in \mathscr{P}_2(X, \mathfrak{m}) \tag{7.2}$$

and the regularization estimates for the Entropy and its slope

$$\mathrm{I}_K(t)\mathrm{Ent}_{\mathfrak{m}}(\mathscr{H}_t(\mu)) + \frac{(\mathrm{I}_K(t))^2}{2}|D^-\mathrm{Ent}_{\mathfrak{m}}|^2(\mathscr{H}_t(\mu)) \le \frac{1}{2}W_2^2(\mu, \mathfrak{m}) \tag{7.3}$$

which are typical of EVI_K-solutions, with $\mathrm{I}_K(t) := \int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r$. Notice also that (7.2) yields $W_1(\mathscr{H}_t(\delta_x), \mathscr{H}_t(\delta_y)) \le \mathrm{e}^{-Kt}\mathsf{d}(x, y)$ for all $x, y \in X$ and $t \ge 0$. This implies that $RCD(K, \infty)$ spaces have Ricci curvature bounded from below by K according to [23], [19].

Theorem 7.1 (Regularizing properties of the heat flow)**.** *Let $(X, \mathsf{d}, \mathfrak{m})$ be a $RCD(K, \infty)$ space. Then:*

(i) The transition probability densities are symmetric

$$\rho_t[x](y) = \rho_t[y](x) \qquad \mathfrak{m} \times \mathfrak{m}\text{-a.e. in } X \times X, \text{ for all } t > 0, \tag{7.4}$$

and satisfy for all $x \in X$ the Chapman-Kolmogorov formula:

$$\rho_{t+s}[x](y) = \int \rho_t[x](z)\rho_s[z](y)\,\mathrm{d}\mathfrak{m}(z) \qquad \text{for } \mathfrak{m}\text{-a.e. } y \in X, \text{ for all } t, s \ge 0. \tag{7.5}$$

(ii) The formula

$$\tilde{\mathsf{h}}_t f(x) := \int f(y)\,\mathrm{d}\mathscr{H}_t(\delta_x)(y) \qquad x \in X \tag{7.6}$$

provides a version of $\mathsf{h}_t f$ for every $f \in L^2(X, \mathfrak{m})$, an extension of h_t to a continuous contraction semigroup in $L^1(X, \mathfrak{m})$ which is pointwise everywhere defined if $f \in L^\infty(X, \mathfrak{m})$.

(iii) The semigroup $\tilde{\mathsf{h}}_t$ maps contractively $L^\infty(X, \mathfrak{m})$ in $C_b(X)$ and, in addition, $\tilde{\mathsf{h}}_t f(x)$ belongs to $C_b\big((0, \infty) \times X\big)$.

(iv) If $f : X \to \mathbb{R}$ is Lipschitz, then $\tilde{\mathsf{h}}_t f$ is Lipschitz on X as well and $\mathrm{Lip}(\tilde{\mathsf{h}}_t f) \le \mathrm{e}^{-Kt}\,\mathrm{Lip}(f)$.

Theorem 7.2 (Bakry-Émery in $RCD(K, \infty)$ spaces)**.** *For any $f \in L^2(X, \mathfrak{m}) \cap \mathcal{S}^2$ and $t > 0$ we have*

$$|D(\mathsf{h}_t f)|_w^2 \le \mathrm{e}^{-2Kt}\mathsf{h}_t(|Df|_w^2) \qquad \mathfrak{m}\text{-a.e. in } X. \tag{7.7}$$

In addition, if $|Df|_w \in L^\infty(X, \mathfrak{m})$ and $t > 0$, then $\mathrm{e}^{-Kt}\big(\tilde{\mathsf{h}}_t|Df|_w^2\big)^{1/2}$ is an upper gradient of $\tilde{\mathsf{h}}_t f$ on X, so that

$$|D^-\tilde{\mathsf{h}}_t f| \le \mathrm{e}^{-Kt}\big(\tilde{\mathsf{h}}_t|Df|_w^2\big)^{1/2} \quad \text{pointwise in } X, \tag{7.8}$$

and f has a Lipschitz version $\tilde{f} : X \to \mathbb{R}$, with $\mathrm{Lip}(\tilde{f}) \le \||Df|_w\|_\infty$.

The regularization properties (7.3) of EVI_K-flows provide an $L\log L$ regularization of the semigroup $\mathscr{H}_t$ starting from arbitrary measures in $\mathscr{P}_2(X)$. When X is a $RCD(K,\infty)$-space with $K>0$, then combining the slope inequality for K-geodesically convex functionals [2, Lemma 2.4.13]

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) \le \frac{1}{2K}|D^-\mathrm{Ent}_{\mathfrak{m}}|^2(\mu)$$

with the identity $|D^-\mathrm{Ent}_{\mathfrak{m}}|^2(f\mathfrak{m}) = \int |Df|_w^2/f\,\mathrm{d}\mathfrak{m}$ between slope and Fisher information, we get the Logarithmic-Sobolev inequality

$$\int_X f\log f\,\mathrm{d}\mathfrak{m} \le \frac{1}{2K}\int_{f>0}\frac{|Df|_w^2}{f}\,\mathrm{d}\mathfrak{m} \quad \text{if } \sqrt{f}\in W^{1,2}(X,\mathsf{d},\mathfrak{m}),\ f\mathfrak{m}\in\mathscr{P}(X), \tag{7.9}$$

which in particular yields the hypercontractivity of h_t, see e.g. [8]. When h_t is ultracontractive, i.e. there exists $p>1$ such that

$$\|\mathsf{h}_t f\|_p \le C(t)\|f\|_1 \qquad \text{for every } f\in L^2(X,\mathfrak{m}),\ t>0, \tag{7.10}$$

then one can also obtain global Lipschitz regularity for the transition probabilities [4, Proposition 6.4], see also [16, Proposition 4.4]. The stronger regularizing property (7.10) is known to be true, for instance, if doubling and Poincaré hold in $(X,\mathsf{d},\mathfrak{m})$, see [28, Corollary 4.2].

We conclude this section with an example of application of the Bakry-Émery estimate (7.2), which can be proved following the ideas of [10].

Theorem 7.3 (Lipschitz regularization)**.** *If $f\in L^2(X,\mathfrak{m})$ then $\mathsf{h}_t f\in \mathcal{S}^2$ for every $t>0$ and*

$$2\,\mathrm{I}_{2K}(t)|D\mathsf{h}_t f|_w^2 \le \mathsf{h}_t f^2 \quad \mathfrak{m}\text{-a.e. in } X; \tag{7.11}$$

in particular, if $f\in L^\infty(X,\mathfrak{m})$ then $\tilde{\mathsf{h}}_t f\in \mathrm{Lip}(X)$ for every $t>0$ with

$$\sqrt{2\,\mathrm{I}_{2K}(t)}\,\mathrm{Lip}(\tilde{\mathsf{h}}_t f) \le \|f\|_\infty \quad \text{for every } t>0. \tag{7.12}$$

7.2 Connections with Dirichlet forms and Markov processes

Since Ch is quadratic, lower semicontinuous in $L^2(X,\mathfrak{m})$ and since $|Df|_w$ has strong locality properties, it turns out that the bilinear form $\mathcal{E}$ associated to Ch, whose domain is from now on restricted from $L^1(X,\mathfrak{m})\cap\mathcal{S}^2$ to $L^2(X,\mathfrak{m})\cap\mathcal{S}^2$, is a local Dirichlet form. In the theory of Dirichlet forms a canonical object is the induced distance, namely

$$\mathsf{d}_{\mathcal{E}}(x,y) := \sup\{|\tilde g(x)-\tilde g(y)|:\ g\in D(\mathcal{E}),\ [g]\le\mathfrak{m}\} \qquad \forall (x,y)\in X\times X, \tag{7.13}$$

where the function $\tilde g$ is the continuous representative in the Lebesgue class of g, see Theorem 7.2). Another canonical object is the local energy measure, namely the measure $[u]$ defined by

$$[u](\varphi) := \mathcal{E}(u,u\varphi) - \frac{1}{2}\mathcal{E}(u^2,\varphi) \qquad \varphi\in L^2(X,\mathfrak{m})\cap\mathcal{S}^2.$$

A consequence of Lemma 3.8 is that $[u]=|Du|_w^2\mathfrak{m}$ for all $u\in L^2(X,\mathfrak{m})\cap\mathcal{S}^2$. Also the distances can be identified:

Theorem 7.4 (Identification of $\mathsf{d}_{\mathcal{E}}$ and d)**.** *The function $\mathsf{d}_{\mathcal{E}}$ in* (7.13) *coincides with d on $X\times X$.*

Finally, using a tightness property of $\mathcal{E}$, the theory of Dirichlet forms can be applied to obtain the representation of transition probabilities in terms of a continuous Markov process:

Theorem 7.5 (Brownian motion)**.** *Let* $(X, \mathsf{d}, \mathfrak{m})$ *be a* $RCD(K,\infty)$ *space. There exists a unique (in law) Markov process* $\{\mathbf{X}_t\}_{\{t\geq 0\}}$ *in* (X, d) *with continuous sample paths in* $[0,\infty)$ *and transition probabilities* $\mathscr{H}_t(\delta_x)$*, i.e.*

$$\mathbf{P}\big(\mathbf{X}_{s+t} \in A \big| \mathbf{X}_s = x\big) = \mathscr{H}_t(\delta_x)(A) \qquad \forall s,\, t \geq 0,\ A \text{ Borel} \tag{7.14}$$

for $\mathfrak{m}$*-a.e.* $x \in X$*.*

7.3 Tensorization

Recall that a metric space (X, d) is said to be non branching if the map $(\mathrm{e}_0, \mathrm{e}_t) : \mathrm{Geo}(X) \to X^2$ is injective for all $t \in (0,1)$, i.e., geodesics do not split.

Theorem 7.6 (Tensorization)**.** *Let* $(X, \mathsf{d}_X, \mathfrak{m}_X)$*,* $(Y, \mathsf{d}_Y, \mathfrak{m}_Y)$ *be metric measure spaces and define the product space* $(Z, \mathsf{d}, \mathfrak{m})$ *as* $Z := X \times Y$*,* $\mathfrak{m} := \mathfrak{m}_X \times \mathfrak{m}_Y$ *and*

$$\mathsf{d}\big((x,y),(x',y')\big) := \sqrt{\mathsf{d}^2_X(x,x') + \mathsf{d}^2_Y(y,y')}.$$

Assume that both $(X, \mathsf{d}_X, \mathfrak{m}_X)$ *and* $(Y, \mathsf{d}_Y, \mathfrak{m}_Y)$ *are* $RCD(K,\infty)$ *and non branching. Then* $(Z, \mathsf{d}, \mathfrak{m})$ *is* $RCD(K,\infty)$ *and non branching as well.*

In the forthcoming paper [6] we will actually be able to prove that the tensorization property of $RCD(K,\infty)$ persists even when the non branching assumption on the base spaces is removed.

References

[1] L. Ambrosio and N. Gigli, *User's guide to optimal transport theory*, To appear in the CIME Lecture Notes in Mathematics, B.Piccoli and F.Poupaud Eds., (2011).

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[3] ——, *Calculus and heat flows in metric measure spaces with Ricci curvature bounded from below*, Submitted paper, arXiv:1106.2090, (2011).

[4] ——, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Submitted paper, arXiv:1109.0222, (2011).

[5] ——, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, Preprint, (2011). Rev. Mat. Iberoamericana, to appear.

[6] ——, *BE is equivalent to RCD*, Preprint, (2012).

[7] L. Ambrosio, N. Gigli, A. Mondino and G. Savaré, *On the notion of convergence of non-compact metric measure spaces and applications.* Preprint, (2012).

[8] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malreu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, no. 10 in Panoramas et Synthèses, Société Mathématique de France, 2000.

[9] K. Bacher and K.T. Sturm *Localization and Tensorization Properties of the Curvature-Dimension Condition for Metric Measure Spaces*, Journal Funct. Anal. **259** (2010), 28–56.

[10] D. Bakry, *Functional inequalities for Markov semigroups*, in Probability measures on groups: recent directions and trends, Tata Inst. Fund. Res., Mumbai, 2006, 91–147.

[11] A. Braides, *Γ-convergence for Beginners*, Oxford Lecture Series in Mathematics and its Applications, Vol. 22, Oxford University Press, 2002.

[12] H. Brezis, *Analyse Fonctionnelle. Théorie et applications*, Collection Mathématiques Appliquées pour la Maîtrise, Masson, Paris, 1983.

[13] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal., **9** (1999), 428–517.

[14] G. Dal Maso, *An Introduction to* Γ*-convergence*, Birkhäuser, Boston, 1993.

[15] S. Daneri and G. Savaré, *Eulerian calculus for the displacement convexity in the Wasserstein distance*, SIAM J. Math. Anal., **40** (2008), 1104–1122.

[16] N. Gigli, K. Kuwada, and S. Ohta, *Heat flow on Alexandrov spaces*, Submitted paper, (2010).

[17] N. Gigli, *On the differential structure of metric measure spaces and applications*, (2012).

[18] J. Heinonen and P. Koskela, *Quasiconformal maps in metric spaces with controlled geometry*, Acta Math., **181** (1998), 1–61.

[19] A. Joulin, *A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature*, Bernoulli, **15** (2009), 532–549.

[20] P. Koskela and P. MacManus, *Quasiconformal mappings and Sobolev spaces,* Studia Math., **131** (1998), 1–17.

[21] S. Lisini, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations, **28** (2007), 85–120.

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

[23] Y. Ollivier, *Ricci curvature of Markov chains on metric measure spaces*, J. Functional Analysis, **256** (2009), 810–864.

[24] T. Rajala, *Improved geodesics for the reduced curvature-dimension condition in branching metric spaces*, Discrete Contin. Dyn. Syst., to appear.

[25] ——, *Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm*, Journal Funct. Anal., **263** (2012), 896–924.

[26] ——, *Local Poincaré inequalities from stable curvature conditions on metric spaces*, Calc. Var. Partial Differential Equations, **44** (2012), 477–494.

[27] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Rev. Mat. Iberoamericana, **16** (2000), 243–279.

[28] K.-T. Sturm, *Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality*, J. Math. Pures Appl., **75** (1996), 273–297.

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

[30] ——, *On the geometry of metric measure spaces. II*, Acta Math., **196** (2006), 133–177.

[31] C. Villani, *Optimal transport. Old and new*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.