A FLOW TANGENT TO THE RICCI FLOW VIA HEAT KERNELS AND MASS TRANSPORT

NICOLA GIGLI AND CARLO MANTEGAZZA

ABSTRACT. We present a new relation between the short time behavior of the heat flow, the geometry of optimal transport and the Ricci flow. We also show how this relation can be used to define an evolution of metrics on non–smooth metric measure spaces with Ricci curvature bounded from below.

CONTENTS

1. INTRODUCTION

The Ricci flow is possibly the most important and largely studied geometric flow in literature, its relevance is well deserved by the key role it played in solving some long standing open conjectures, in particular, the Poincaré conjecture finally proved by Perelman.

In [11] McCann and Topping noticed an interesting relation between such flow, the heat flow and optimal transport: they proved that a family of metrics g_τ on a smooth and compact differential manifold M is a backward *super* Ricci flow, i.e. it satisfies

$$-\frac{\mathrm{d}g_\tau}{\mathrm{d}\tau} + 2\mathrm{Ric}(g_\tau) \geq 0,$$

if and only if the time dependent *quadratic transportation distance* $W_2^{(M,g_\tau)}$ is not increasing along two solutions of the time dependent heat equation

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mu_\tau = \Delta_{g_\tau}\mu_\tau. \tag{1.1}$$

In particular, a backward Ricci flow can be characterized as the minimal evolution among all the flows for which such *non–expansion* property holds.

Keeping in mind that on a fixed Riemannian manifold (M, g) one always has

$$W_2(\mu_t, \nu_n) \le e^{-Kt} W_2(\mu_0, \nu_0),$$

for each couple of solutions μ_t and ν_t of the heat flow, where K is a lower bound on the Ricci tensor of (M, g), McCann–Topping result can be thought as: the Ricci flow is the evolution that precisely compensate the lack/excess of contraction w.r.t. the distance W_2.

In this paper we propose a different point of view on the same subject. Let (M, g) be a compact Riemannian manifold and $\mathscr{P}(M)$ the space of Borel probability measures on M. We denote by $\mathsf{H}_t : \mathscr{P}(M) \to \mathscr{P}(M)$ the heat semigroup, so that given $\mu \in \mathscr{P}(M)$, the curve $t \mapsto \mathsf{H}_t(\mu)$ is the solution of the heat equation with initial condition μ. Then, for every $t \ge 0$ we have an embedding of M in $\mathscr{P}(M)$ given by

$$M \ni x \qquad \longmapsto \qquad \iota_t(x) := \mathsf{H}_t(\delta_x) \in \mathscr{P}(M). \tag{1.2}$$

Endow $\mathscr{P}(M)$ with the distance W_2 and the image $\iota_t(M)$ with the "intrinsic" distance induced by W_2 (i.e. not the "chord" distance W_2 in $\mathscr{P}(M)$ but the "arc" one, where the distance is defined as the minimal length of the paths lying in $\iota_t(M)$). By the backward uniqueness of the heat flow, we know that the map ι_t is injective and thus the distance on $\iota_t(M)$ can be pulled back to a distance d_t on M, which clearly coincides with the Riemannian distance at time $t = 0$. It is not hard to see that d_t is still a Riemannian distance, namely that there exists a smooth metric tensor g_t on M inducing d_t: shortly said, this comes from the fact that, according to Otto, the space $(\mathscr{P}(M), W_2)$ is an infinite dimensional Riemannian manifold and $\iota_t(M)$ a "smooth" finite dimensional submanifold.

Our main result (Theorem 4.6) is that g_t is an evolution of metrics which is "tangent" at time $t = 0$ to the Ricci flow, the rigorous statement being the following.

Theorem 1.1. *Let* $[0, 1] \ni s \mapsto \gamma_s \in M$ *be a geodesic in* (M, g)*. Then, there holds*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_0^1 g_t(\gamma_s', \gamma_s') \, \mathrm{d}s \Big|_{t=0} = -2 \int_0^1 \mathrm{Ric}_g(\gamma_s' \gamma_s') \, \mathrm{d}s.$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t} g_t(\gamma_s', \gamma_s') \Big|_{t=0} = -2 \mathrm{Ric}_g(\gamma_s' \gamma_s') \qquad a.e. \ s \in [0, 1].$$

Notice that we get an integrated/a.e. version of the result rather than the cleaner formula $\frac{\mathrm{d}}{\mathrm{d}t} g_t \Big|_{t=0} = -2 \mathrm{Ric}_g$ due to some potential lack of smoothness of the evolution that we are not able to fully manage at the moment, see Remark 4.3.

Due to McCann–Topping result, our theorem is in some sense not so surprising, since it states that the infinitesimal behavior of the W_2–distance along the heat flow is driven by the Ricci tensor, which is in the same spirit of their work. Yet, at the technical level there is a difference worth to be underlined: the flow g_t that

we define is *not* the Ricci flow: to see this, notice that since H_t is injective for any $t \geq 0$, the metric tensor g_t is never 0, hence our flow never shrinks distances to 0 in finite time, as opposed to the Ricci flow which shrinks spheres to points. In particular, the evolution we define is not driven by a semigroup, otherwise, due to Theorem 1.1, it should be the Ricci flow. Again, this was expected, as the semigroup H_t that we use to define the distance d_t is the heat flow on the initial manifold (M, g), while if one wants to get the Ricci flow, he should use at each time the corresponding Laplacian, as in formula (1.1). This characteristic, which can be seen as a negative point, actually turns out to be useful if one is interested in defining a flow in a non–smooth setting, as we now explain.

In [11], McCann–Topping noticed that they provided a purely metric characterization of Ricci flow, which therefore can be theoretically used to define what a Ricci flow should be if the initial space is non–smooth: the minimal flow (in the sense that it expands distances no faster than any other flow) among all super Ricci flows, where a super Ricci flow is any flow contracting the time dependent W_2–distance along any two solutions of the heat equation.

Unfortunately, although this approach is very intriguing, it is not clear whether such a flow exists or it is unique for a non–smooth initial datum (it is not even clear if at least *one* super Ricci flow exists). Instead, the embedding in formula (1.2) is well defined as soon as one has the heat kernel at his disposal.

The natural abstract class of spaces where an heat kernel exists and well behaves w.r.t. the distance W_2 is the one of $RCD(K, \infty)$ spaces, introduced in [6]. This is a subclass of the class of $CD(K, \infty)$ spaces introduced by Lott–Sturm–Villani (see [10], [15]) of spaces with Ricci curvature bounded from below: shortly said, $RCD(K, \infty)$ spaces are $CD(K, \infty)$ spaces where the heat flow is linear. This choice rules rules out Finsler–type geometries and ensures, on one hand, the existence of a heat kernel, on the other hand, the W_2–contraction along two heat flows (in [12] Sturm and Ohta proved that on a normed space $(\mathbb{R}^d, \|\cdot\|, \mathcal{L}^d)$ the heat flow never contracts the W_2–distance unless the norm comes from a scalar product, therefore, due to the spirit of the discussion here, it is natural to avoid considering this sort of spaces).

Proceeding as in the smooth case, given an $RCD(K, \infty)$ space $(X, \mathsf{d}, \mathfrak{m})$ we can define an evolution of metrics d_t for any $t \geq 0$, our results being then the following.

(i) The distances d_t are well defined for any $t \geq 0$, so that the flow exists and is unique (Theorem 5.15).
(ii) Under very general assumptions – which cover all the finite dimensional situations – the topology induced by d_t is the same as the one induced by d (Theorem 5.15 and Remark 5.17).
(iii) The flow has some very general weak continuity properties both in time (Theorem 5.18) and w.r.t. measured Gromov–Hausdorff convergence of the initial datum (Theorem 5.19).

Concerning point (ii), notice that although this is a different behavior from the one of Ricci flow, it can turn out to be a useful property in a non–smooth setting. Indeed, given that the Ricci flow can create singularities even with a smooth

initial datum, it is unnatural to expect that a Ricci flow for non–smooth initial data does not create singularities in some short time interval. Thus, a Ricci flow with non–smooth initial data could disrupt the topology even instantaneously, which certainly complicates the analysis. Still, we point out that anyway we do not expect the distances d_t to be bi–Lipschitz equivalent to the original one.

About point (iii), we remark that such a property is strictly related to point (i), as "being well-defined" is very close to "having some weak continuity properties under perturbations". Actually, the problem of defining a true Ricci flow for non–smooth initial data is very much related to the lack of a stability result for the Ricci flow on smooth manifolds under measured Gromov–Hausdorff convergence (as pointed out to us by Sturm).

We conclude observing that the definition of the flow of distances d_t with a non–smooth initial datum opens several non–trivial questions about its behavior, which are not addressed in this paper, in particular:

- Given an $RCD(K, \infty)$ space $(X, \mathsf{d}, \mathfrak{m})$ as initial datum, is it true that $(X, \mathsf{d}_t, \mathfrak{m})$ is an $RCD(K_t, \infty)$ space for some K_t, possibly under some finite dimensionality assumption?
- Is the space $(X, \mathsf{d}_t, \mathfrak{m})$, in any sense, "smoother" that the original one?

2. Setting and Preliminaries

2.1. **Metric spaces and quadratic transportation distance.** We recall here the basic facts about analysis in metric spaces and about the Kantorovich quadratic transportation distance W_2.

Given a metric space (X, d) and a non–trivial interval $I \subset \mathbb{R}$, a curve $I \ni t \mapsto x_t \in X$ is said to be absolutely continuous provided that there exists a function $f \in L^1(I)$ such that

$$\mathsf{d}(x_t, x_s) \leq \int_t^s f(r)\,\mathrm{d}r, \qquad \forall t, s \in I, \ t < s. \tag{2.1}$$

It can be proved that if $t \mapsto x_t$ is absolutely continuous, the limit

$$\lim_{h \to 0} \frac{\mathsf{d}(x_{t+h}, x_t)}{|h|}, \tag{2.2}$$

exists for a.e. $t \in I$. It is called the metric speed of the curve, denoted by $|\dot{x}_t|$, belongs to $L^1(I)$ and is the minimal – in the a.e. sense – L^1 function f that can be put in the right hand side of inequality (2.1) (see Theorem 1.1.2 in [4] for the proof). The length of the absolutely continuous curve $[0, 1] \ni t \mapsto x_t \in X$ is, by definition, $\int_0^1 |\dot{x}_t|\,\mathrm{d}t$ and it is easy to check that it holds

$$\int_0^1 |\dot{x}_t|\,\mathrm{d}t = \sup \sum_{i=0}^{N-1} \mathsf{d}(x_{t_i}, x_{t_{i+1}}), \tag{2.3}$$

the sup being taken among all $N \in \mathbb{N}$ and all partitions $0 = t_0 < t_1 < \ldots < t_N = 1$ of $[0, 1]$. We will often denote a curve $t \mapsto x_t$ with (x_t).

Given a complete and separable metric space (X, d), we denote by $\mathscr{P}(X)$ its set of Borel probability measures and by $\mathscr{P}_2(X) \subset \mathscr{P}(X)$ the subset of measures with finite second moment, i.e. probability measures μ such that

$$\int_X \mathsf{d}^2(\cdot, x_0) \, \mathrm{d}\mu < +\infty, \qquad \text{for some (hence, for every) } x_0 \in X.$$

The space $\mathscr{P}_2(X)$ will be endowed with the *quadratic transportation distance* W_2, defined by

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

the infimum being taken among all *transport plans* $\boldsymbol{\gamma} \in \mathscr{P}(X \times X)$ such that

$$\pi_\sharp^1 \boldsymbol{\gamma} = \mu,$$
$$\pi_\sharp^2 \boldsymbol{\gamma} = \nu,$$

being $\pi^1, \pi^2 : X \times X \to X$ being the projections onto the first and second factor respectively.

We recall that the distance W_2 can be defined also in terms of the dual problem of optimal transport:

$$\frac{1}{2} W_2^2(\mu, \nu) = \sup \int_X \varphi \, \mathrm{d}\mu + \int_X \varphi^c \, \mathrm{d}\nu,$$

where the supremum is taken among all Borel maps $\varphi : X \to \mathbb{R}$ and the c–transform is defined as

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

It turns out that for $\mu, \nu \in \mathscr{P}_2(X)$ the above supremum is always achieved, and that the maximal φ can always be taken to be a c–concave function, i.e. a function φ such that $\varphi^{cc} = \varphi$.

The convergence in $(\mathscr{P}_2(X), W_2)$ is characterized by the following well known result.

Theorem 2.1. *Let $n \mapsto \mu_n \in \mathscr{P}_2(X)$ be a sequence and $\mu \in \mathscr{P}_2(X)$. Then, the following are equivalent.*

(i) $W_2(\mu_n, \mu) \to 0$ *as* $n \to \infty$.

(ii) $\int_X f \, \mathrm{d}\mu_n \to \int_X f \, \mathrm{d}\mu$ *for any* $f \in C_b(X)$ *and* $\int_X \mathsf{d}^2(\cdot, x_0) \, \mathrm{d}\mu_n \to \int_X \mathsf{d}^2(\cdot, x_0) \, \mathrm{d}\mu$ *as* $n \to \infty$ *for some* $x_0 \in X$.

(iii) $\int_X f \, \mathrm{d}\mu_n \to \int_X f \, \mathrm{d}\mu$ *for any continuous function* $f : X \to \mathbb{R}$ *with quadratic growth, i.e. such that for some* $x_0 \in X$ *and* $c > 0$ *there holds*

$$|f(x)| \leq c(1 + \mathsf{d}^2(x, x_0)), \qquad \forall x \in X.$$

2.2. **Optimal transport and heat flow on Riemannian manifolds.**

Throughout all the paper (M, g) will be a given compact, C^∞ Riemannian manifold. The canonical volume measure induced by g will be denoted by Vol. We will sometimes indicate $g(v, w)$ by $v \cdot w$ and $g(v, v)$ by $|v|^2$. The set of Borel probability measures on M is denoted by $\mathscr{P}(M)$. All the differential operators that will appear will be relative to the Levi–Civita covariant derivative ∇ associated to the metric g, that is, in particular $\mathrm{div} = \mathrm{div}_g$ and $\Delta = \Delta_g$.

We will denote by $(0, +\infty) \times M \times M \ni (t, x, y) \mapsto \rho(t, x, y) \in \mathbb{R}^+$ the heat kernel on M and for every $x \in M$, $t \geq 0$ by $\mu_{t,x}$ the probability measure defined by $\mu_{t,x} := \rho(t, x, \cdot)\,\mathrm{Vol}$, for $t > 0$ and $\mu_{0,x} := \delta_x$. For $t \geq 0$ we also denote by $\mathsf{H}_t : \mathscr{P}(M) \to \mathscr{P}(M)$ the heat semigroup acting on probability measures, i.e. for any $\mu \in \mathscr{P}(M)$ and $t \geq 0$ the measure $\mathsf{H}_t(\mu) \in \mathscr{P}(M)$ is given by

$$\int_M f(x)\,\mathrm{d}\mathsf{H}_t(\mu)(x) := \int_M \int_M f(y)\rho(t, x, y)\,\mathrm{dVol}(y)\,\mathrm{d}\mu(x), \qquad \forall t \geq 0.$$

In particular, there holds $\mathsf{H}_t(\delta_x) = \mu_{t,x}$.

Theorem 2.2. *Let $\eta, \rho : M \to \mathbb{R}$ be two C^∞ functions such that $\int_M \eta\,\mathrm{dVol} = 0$ and $\rho > 0$. Then, there exists a unique smooth function $\varphi : M \to \mathbb{R}$ with $\int_M \varphi\,\mathrm{dVol} = 0$ which is a solution of the PDE*

$$\eta = \nabla \cdot (\nabla\varphi\,\rho) = \mathrm{div}(\nabla\varphi\,\rho) = \rho\Delta\varphi + g(\nabla\varphi, \nabla\rho)\,.$$

Moreover, such a function φ smoothly depends on the functions η and ρ.

Proof. By the uniform strict positivity of $\rho \in C^\infty$, as M is compact, the above PDE is equivalent to the linear problem

$$\Delta\varphi = -g(\nabla\varphi, \nabla\log\rho) + \eta/\rho\,,$$

then, the existence/uniqueness of a solution in $W^{1,2}(M)$ follows as in the Euclidean case. Expressing the Laplacian and the metric g in local coordinates, the regularity of the solution is then obtained by a standard bootstrap argument, see for instance the book of Gilbarg and Trudinger [9]. $\qquad\square$

Theorem 2.3 (Backward uniqueness of the heat flow)**.** *Let $(0, +\infty) \times M \ni (t, x) \mapsto f_t(x)$ be a solution of the heat equation*

$$\frac{\mathrm{d}}{\mathrm{dt}} f_t = \Delta f_t,$$

such that for some $t_0 \geq 0$ there holds $f_{t_0} \equiv 0$. Then, $f_t \equiv 0$ for any $t \geq 0$.

Proof. This is a consequence of the fact that the heat semigroup is analytic in $L^2(M, \mathrm{Vol})$, see the details in the proof of Proposition 5.16. $\qquad\square$

Later on, we will find useful the following lemma concerning c–concave functions on M (for a proof, see for instance Lemma 1.34 in [1]).

Lemma 2.4. *Let (M, g) be a smooth, complete Riemannian manifold and $\varphi \in C_c^\infty(M)$. Then, there exists some $\bar{\varepsilon} > 0$ such that for $|\varepsilon| \leq \bar{\varepsilon}$ the following facts are true:*

(i) *The function $\varphi_\varepsilon := \varepsilon\varphi$ is c–concave, and $\varphi_\varepsilon^c \in C_c^\infty(M)$.*

(ii) *The maps $x \mapsto T(x) := \exp_x(-\nabla\varphi_\varepsilon(x))$ and $y \mapsto S(y) := \exp_y(-\nabla\varphi_\varepsilon^c(y))$ are smooth and each one inverse of the other.*

(iii) *For every $x \in M$ the curve $s \mapsto \exp_x(-s\nabla\varphi_\varepsilon(x))$ is the unique minimizing geodesic from x to $T(x)$. Similarly, for any $y \in M$ the curve $s \mapsto \exp_y(-s\nabla\varphi_\varepsilon^c(y))$ is the unique minimizing geodesic from y to $S(y)$.*

iv) *The following two duality formulas hold:*

$$\varphi_\varepsilon^c(T(x)) = \tfrac{1}{2}|\nabla\varphi_\varepsilon|_g^2(x) - \varphi_\varepsilon(x), \qquad \forall x \in M, \tag{2.4}$$

$$\varphi_\varepsilon(S(y)) = \tfrac{1}{2}|\nabla\varphi_\varepsilon^c|_g^2(y) - \varphi_\varepsilon^c(y), \qquad \forall y \in M. \tag{2.5}$$

Such $\overline{\varepsilon} > 0$ depends only on the supremum of $|\varphi|, |\nabla\varphi|_g, |\nabla^2\varphi|_g$, on the modulus of the Riemann tensor Riem of M and on the infimum of the injectivity radius in the compact set $\operatorname{supp}\varphi$.

We remark that although in this paper we will let the metric g vary in time, when speaking about absolute continuity of a curve of measures $t \mapsto \mu_t$ and about its metric speed $|\dot\mu_t|$, we will always refer to the quadratic transportation distance W_2 built on top of the Riemannian distance induced by the initial metric tensor g.

Absolutely continuous curves of measures are related to the continuity equation via the following result.

Theorem 2.5. *Let $s \mapsto \mu_s \in \mathscr{P}(M)$ be a continuous curve w.r.t. weak convergence of measures. Then, the following facts are equivalent:*

(i) *The curve $s \mapsto \mu_s$ is absolutely continuous w.r.t. W_2.*

(ii) *For a.e. s there exists $v_s \in \overline{\{\nabla\varphi \ : \ \varphi \in C^\infty(M)\}}^{L^2(\mu_s)}$ such that the continuity equation*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mu_s + \nabla \cdot (v_s\mu_s) = 0$$

holds in the sense of distributions.

In this case, the vector fields v_s are uniquely defined for a.e. s and there holds $|\dot\mu_s|^2 = \int_M |v_s|^2 \,\mathrm{d}\mu_s$ for a.e. s.

Proof. See Theorem 8.3.1 in [4] for the case $M = \mathbb{R}^d$. The case of general Riemannian manifolds then follows easily from Nash embedding theorem, see e.g. Theorem 2.29 in [1] or Theorem 13.8 in [17]. $\qquad\square$

We also recall that the distance W_2 is "contracting" under a lower Ricci bound (see [16])

Theorem 2.6. *Let $\mu, \nu \in \mathscr{P}(M)$. Then, for every $t \geq 0$ there holds*

$$W_2(\mathsf{H}_t(\mu), \mathsf{H}_t(\nu)) \leq e^{-Kt}W_2(\mu, \nu),$$

where K is a global bound from below on the eigenvalues of the Ricci tensor Ric of M.

In particular, if $s \mapsto \gamma_s \in M$ is a Lipschitz curve, the curve $s \mapsto \mu_{t,\gamma_s} \in \mathscr{P}(M)$ is Lipschitz w.r.t. W_2 and there holds

$$|\dot{\mu}_{t,\gamma_s}| \le e^{-Kt}|\gamma'_s|, \tag{2.6}$$

for a.e. s, where $|\dot{\mu}_{t,\gamma_s}|$ denotes the metric speed of the curve.

Proof. The above K–contraction property of the distance W_2 is a well known consequence of the lower bound on the Ricci tensor. It immediately implies the estimate (2.6) for Lipschitz curves. □

We conclude recalling the definition of the Sasaki metric tensor $\overline{g}$ on the tangent bundle TM of (M, g). Given $(x, v) \in TM$ and $V_1, V_2 \in T_{(x,v)}TM$, we find two smooth curves $t \mapsto (x_{i,t}, v_{i,t}) \in TM$ such that $\frac{d}{dt}(x_{i,t}, v_{i,t})\big|_{t=0} = V_i$, $i = 1, 2$. Then, $\overline{g}(V_1, V_2)$ is defined as

$$\overline{g}(V_1, V_2) := g(x'_{1,0}, x'_{2,0}) + g(\nabla_{x'_{1,0}} v_{1,t}, \nabla_{x'_{2,0}} v_{2,t}),$$

where by $\nabla_{x'_{i,0}} v_{i,t}$ we intend the covariant derivative (w.r.t. g) of the vector field $t \mapsto v_{i,t}$ along the curve $t \mapsto x_{i,t}$ at time $t = 0$, $i = 1, 2$. It is readily checked that this is a good definition and that, denoting by $\overline{d}$ the distance on TM induced by $\overline{g}$, there holds

$$\overline{d}^2\big((y, w), (x, 0)\big) \le d^2(y, x) + g(w, w), \qquad \forall x, y \in M, \ w \in T_y M. \tag{2.7}$$

3. DEFINITION OF THE FLOW

We start collecting some basic consequences of Theorems 2.2, 2.3.

Proposition 3.1. *Let $t > 0$, $x \in M$ and $v \in T_x M$. Then, there exists a unique C^∞ function $\varphi_{t,x,v} : M \to \mathbb{R}$ such that $\int_M \varphi_{t,x,v} \, d\mathrm{Vol} = 0$ and*

$$\nabla_x \rho(t, x, y) \cdot v = -\nabla_y \cdot (\nabla \varphi_{t,x,v}(y) \rho(t, x, y)). \tag{3.1}$$

Such $\varphi_{t,x,v}$ smoothly depends on the data t, x, v.

Moreover, if $v \neq 0$, then $\nabla \varphi_{t,x,v}$ is not identically zero.

Proof. Existence, uniqueness, smoothness and smooth dependence on the data follows directly from Theorem 2.2. For the second part of the statement, assume that $\nabla \varphi_{t,x,v} \equiv 0$, hence, from the uniqueness property of equation (3.1) we get that $\varphi_{t,x,v} \equiv 0$ and $\nabla_x \rho(t, x, \cdot) \cdot v \equiv 0$. Now observe that $(t, y) \mapsto \eta(t, y) := \nabla_x \rho(t, x, y) \cdot v$ is still a solution of the heat equation, hence, by Theorem 2.3 and the fact that $\eta(t, \cdot) \equiv 0$ we get $\eta(\cdot, \cdot) \equiv 0$, which easily implies, taking t small, that $v = 0$. □

For $t > 0$ we define a new metric tensor g_t in the following way.

Definition 3.2. *Let $t > 0$, $x \in M$ and $v, w \in T_x M$. Then, $g_t(v, w)$ is defined as*

$$g_t(v, w) := \int_M \nabla \varphi_{t,x,v}(y) \cdot \nabla \varphi_{t,x,w}(y) \rho(t, x, y) \, d\mathrm{Vol}(y).$$

Remark 3.3. In the above definition as well as in the rest of the paper, by $v \cdot w$ we intend $g(v, w)$, i.e. their scalar product w.r.t. the original metric tensor. Similarly, $|v|^2$ will always denote $g(v, v)$. ∎

Proposition 3.4. g_t *is a* C^∞ *metric tensor for the manifold* M *which varies smoothly in* $t \in (0, +\infty)$.

Proof. Uniqueness in equation (3.1) gives that $\varphi_{t,x,v}$ linearly depends on v, so g_t is a bilinear form, which, by definition, is also symmetric and non–negative. Its smoothness is a direct consequence of the smoothness of the heat kernel and of the smooth dependence of $\varphi_{t,x,v}$ on the data.

Finally, assume that $g_t(v, v) = 0$ and notice that by definition and the fact that $\rho(t, x, y) > 0$, for any $t > 0$ and $x, y \in M$, we deduce $\nabla \varphi_{t,x,v} \equiv 0$. Hence, by the last part of the statement of Proposition 3.1 we conclude that v must be 0 and we are done. □

We try now to give a more concrete description of the distance d_t induced by the metric tensor g_t on M. We have

$$\mathsf{d}_t^2(x, y) := \inf_\gamma \int_0^1 g_t(\gamma_s', \gamma_s') \, \mathrm{d}s,$$

the infimum being taken among all smooth curves $\gamma : [0, 1] \to M$ such that $\gamma_0 = x$, $\gamma_1 = y$.

Proposition 3.5. *Let* $s \mapsto \gamma_s \in M$ *be an absolutely continuous curve. For fixed* $t > 0$, *we define the curve in the space of probability measures* $s \mapsto \mu_s \in \mathscr{P}(M)$ *by* $\mu_s := \mu_{t,\gamma_s}$, *that is, at every* s *we consider the measure whose density (w.r.t. to the fixed measure* Vol*) is the heat kernel centered at* γ_s, *at time* t.
Then, the curve $s \mapsto \mu_s$ *is absolutely continuous w.r.t.* W_2 *and there holds*

$$g_t(\gamma_s', \gamma_s') = |\dot{\mu}_s|^2, \qquad a.e. \ s,$$

where $|\dot{\mu}_s|$ *denotes the metric speed of the curve* $s \mapsto \mu_s$ *computed w.r.t. the distance* W_2.

Proof. As the curve $s \mapsto \gamma_s$ is absolutely continuous, it is easy to see that also the curve of delta measures $s \mapsto \mu_{0,\gamma_s}$ is absolutely continuous in $(\mathscr{P}(M), W_2)$. Then, as $\mu_s = \mu_{t,\gamma_s} = \mathsf{H}_t(\mu_{0,\gamma_s})$, by Theorem 2.6 and the fact that Ricci tensor of M is uniformly bounded from below, we get that $s \mapsto \mu_s$ is absolutely continuous in $(\mathscr{P}(M), W_2)$.

By Theorem 2.5 it follows that for a.e. s there exists $v_s \in \overline{\{\nabla\varphi \ : \ \varphi \in C^\infty(M)\}}^{L^2(\mu_s)}$ such that the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\mu_s + \nabla \cdot (v_s \mu_s) = 0,$$

holds in the sense of distributions and $|\dot{\mu}_s|^2 = \int_M |v_s|^2 \, \mathrm{d}\mu_s$ for a.e. s.

Since we know that $\mu_s = \rho(t, \gamma_s, \cdot)$Vol, the continuity equation reads (distributionally)

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\rho(t, \gamma_s, y) + \nabla_y \cdot (v_s \rho(t, \gamma_s, y)) = \nabla_x \rho(t, x, y)|_{x=\gamma_s} \cdot \gamma_s' + \nabla_y \cdot (v_s \rho(t, \gamma_s, y)),$$

which implies, by the uniqueness part of Theorem 2.5 and Proposition 3.1, that for a.e. s we have $v_s = \nabla \varphi_{t,\gamma_s,\gamma_s'}$. Hence, for a.e. s we conclude

$$|\dot{\mu}_s|^2 = \int_M |\nabla \varphi_{t,\gamma_s,\gamma_s'}|^2 \, \mathrm{d}\mu_s = \int_M |\nabla \varphi_{t,\gamma_s,\gamma_s'}(y)|^2 \, \rho(t,\gamma_s,y) \mathrm{d}\mathrm{Vol}(y) = g_t(\gamma_s',\gamma_s') \, ,$$

recalling the very definition of the metric tensor g_t. $\qquad\square$

A straightforward consequence of this proposition is that

$$\mathsf{d}_t^2(x,y) = \inf_\gamma \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s,$$

with $\mu_s = \rho(t,\gamma_s,\cdot)\mathrm{Vol}$. That is, the distance d_t^2 is the infimum of the metric lengths of the curves of probability measures $\rho(t,\gamma_s,\cdot)\mathrm{Vol}$ in $(\mathscr{P}(M), W_2)$.

Remark 3.6. Fix $t > 0$ and notice that since M is compact and g, g_t are two smooth metric tensors, it certainly holds $cg_t \le g \le Cg_t$ for some $c, C > 0$. Therefore a curve $t \mapsto \gamma_t \in M$ is absolutely continuous w.r.t. the distance induced by g if and only if it is absolutely continuous w.r.t. the distance induced by g_t. Hence in Proposition 3.5 it is not important to mention the distance w.r.t. which we are requiring absolute continuity. $\qquad\blacksquare$

We see now the convergence of g_t to the original metric tensor g as $t \to 0$.

Proposition 3.7. *Let $x \in M$ and $v \in T_x M$. Then, there holds*

$$\lim_{t \downarrow 0} g_t(v,v) = g(v,v). \tag{3.2}$$

Proof. Let $s \mapsto \gamma_s$ be a C^1 curve such that $\gamma_0 = x$ and $\gamma_0' = v$. By Proposition 3.5 and Theorem 2.6 we get that

$$\int_0^S g_t(\gamma_s',\gamma_s') \, \mathrm{d}s = \int_0^S |\dot{\mu}_{t,\gamma_s}|^2 \, \mathrm{d}s \le e^{-2Kt} \int_0^S g(\gamma_s',\gamma_s') \, \mathrm{d}s, \qquad \forall t, S > 0.$$

Dividing by S and letting $S \downarrow 0$ we deduce

$$g_t(v,v) \le e^{-2Kt} g(v,v).$$

Thus, to conclude it is sufficient to show that for any $x \in M$ and $v \in T_x M$ there holds

$$\varliminf_{t \downarrow 0} g_t(v,v) \ge g(v,v).$$

It is easy to see that we have

$$g_t(v,v) = \int_M |\nabla \varphi_{t,x,v}|^2 \, \mathrm{d}\mu_{t,x} = \sup_{\psi \in C_c^\infty(M)} 2 \int_M \nabla \psi \cdot \nabla \varphi_{t,x,v} \, \mathrm{d}\mu_{t,x} - \int_M |\nabla \psi|^2 \, \mathrm{d}\mu_{t,x},$$

and that for any $\psi \in C_c^\infty(M)$ there holds

$$
\begin{aligned}
\int_M \nabla\psi \cdot \nabla\varphi_{t,x,v} \, \mathrm{d}\mu_{t,x} &= \int_M \nabla\psi(y) \cdot \nabla\varphi_{t,x,v}(y)\rho(t,x,y) \, \mathrm{d}\mathrm{Vol}(y) \\
&= -\int_M \psi(y)\nabla_y \cdot \big(\nabla\varphi_{t,x,v}(y)\rho(t,x,y)\big) \, \mathrm{d}\mathrm{Vol}(y) \\
&= \int_M \psi(y)\nabla_x\rho(t,x,y) \cdot v \, \mathrm{d}\mathrm{Vol}(y) \\
&= \nabla_x \left(\int_M \psi(y)\rho(t,x,y) \, \mathrm{d}\mathrm{Vol}(y) \right) \cdot v
\end{aligned}
\tag{3.3}
$$

Thus, we can choose any $\psi \in C_c^\infty(M)$ so that $\nabla\psi(x) = v$ and conclude that

$$
\begin{aligned}
\varliminf_{t\downarrow 0} g_t(v,v) &\geq \lim_{t\to 0} \left(2\int_M \nabla\psi \cdot \nabla\varphi_{t,x,v} \, \mathrm{d}\mu_{t,x} - \int_M |\nabla\psi|^2 \, \mathrm{d}\mu_{t,x} \right) \\
&= \lim_{t\to 0} 2\nabla_x \left(\int_M \psi(y)\rho(t,x,y) \, \mathrm{d}\mathrm{Vol}(y) \right) \cdot v - \lim_{t\to 0} \int_M |\nabla\psi|^2 \, \mathrm{d}\mu_{t,x} \\
&= |v|^2 ,
\end{aligned}
$$

by the standard properties of the heat kernel $\rho(t,x,y)$. $\qquad\square$

For the discussion thereafter we introduce the transport plans $\boldsymbol{\gamma}_{t,x,v} \in \mathscr{P}(TM)$ defined as follows.

Definition 3.8 (The transport plans $\boldsymbol{\gamma}_{t,x,v}$)**.** *Let* $t \geq 0, x \in M$ *and* $v \in T_xM$. *Then,* $\boldsymbol{\gamma}_{0,x,v} := \delta_{(x,v)}$ *and* $\boldsymbol{\gamma}_{t,x,v} := (X_{t,x,v})_\sharp\mu_{t,x}$, *where* $X_{t,x,v}(y) := (y, \nabla\varphi_{t,x,v}(y))$.

The natural projection mapping from TM to M will be denoted by π^M.

Corollary 3.9. *Let* $x \in M$ *and* $v \in T_xM$. *Then, there holds* $W_2(\boldsymbol{\gamma}_{t,x,v}, \boldsymbol{\gamma}_{0,x,v}) \to 0$ *as* $t \to 0$, *where the quadratic Kantorovich distance considered is the one built on* $(TM, \overline{\mathsf{d}})$, $\overline{\mathsf{d}}$ *being the Sasaki metric on* TM *constructed from the metric tensor* g *on* M.

Proof. By Theorem 2.1 we know that the W_2–convergence is characterized by convergence of second moments plus weak convergence.

We compute the second moments w.r.t. the point $(x,0) \in T_xM$ and we start proving that

$$
\varlimsup_{t\downarrow 0} \int_{TM} \overline{\mathsf{d}}^2\big((y,w),(x,0)\big) \, \mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y,w) \leq \int_{TM} \overline{\mathsf{d}}^2\big((y,w),(x,0)\big) \, \mathrm{d}\boldsymbol{\gamma}_{0,x,v}(y,w).
\tag{3.4}
$$

Integrating the bound (2.7) w.r.t. $\boldsymbol{\gamma}_{t,x,v}$, we get

$$
\int_{TM} \overline{\mathsf{d}}^2\big((y,w),(x,0)\big)\,\mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y,w)
$$
$$
\leq \int_M \mathsf{d}^2(y,x)\,\mathrm{d}\pi_\sharp^M \boldsymbol{\gamma}_{t,x,v}(y) + \int_{TM} g(w,w)\,\mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y,w)
$$
$$
= \int_M \mathsf{d}^2(y,x)\,\mathrm{d}\mu_{t,x}(y) + \int_{TM} g(\nabla\varphi_{t,x,v}(y),\nabla\varphi_{t,x,v}(y))\,\mathrm{d}\mu_{t,\overline{x}}(y)
$$
$$
= \int_M \mathsf{d}^2(x,\overline{x})\,\mathrm{d}\mu_{t,\overline{x}} + g_t(v,v).
$$

Thus, noticing that $\lim_{t\downarrow 0}\int_M \mathsf{d}^2(x,\overline{x})\,\mathrm{d}\mu_{t,\overline{x}} = 0$, using the limit (3.2) and the trivial inequality $\overline{\mathsf{d}}^2\big((y,w),(x,0)\big) \geq g(w,w)$ we get formula (3.4).

Taking into account the lower semicontinuity of the second moments, the conclusion will follow if we show that $\boldsymbol{\gamma}_t$ weakly converge to $\boldsymbol{\gamma}_0$ as $t\downarrow 0$. The bound on the second moments gives in particular that the family $\{\boldsymbol{\gamma}_t\}_{t\in(0,1)}$ is tight. Let $t_n\downarrow 0$ be any sequence such that $n\mapsto \boldsymbol{\gamma}_{t_n,x,v}$ weakly converges to some $\widetilde{\boldsymbol{\gamma}}\in \mathscr{P}(TM)$. Clearly, there holds $\pi_\sharp^M\widetilde{\boldsymbol{\gamma}} = \delta_x$, hence, we can write $\widetilde{\boldsymbol{\gamma}} = \delta_x \times \boldsymbol{\sigma}$ for some measure $\boldsymbol{\sigma}\in \mathscr{P}(T_xM)$. To conclude, it is then sufficient to show that $\boldsymbol{\sigma} = \delta_v$.

Let $\psi\in C_c^\infty(M)$ and consider the function $\overline{\psi}: TM\to\mathbb{R}$ given by $\overline{\psi}(y,w) := w\cdot\nabla\psi(y)$. As the function $\overline{\psi}$ is continuous with linear growth, taking into account the uniform bound on the second moments of the $\boldsymbol{\gamma}_t$, it is easy to see that we get $\lim_{n\to\infty}\int_M \overline{\psi}\,\mathrm{d}\boldsymbol{\gamma}_{t_n,x,v} = \int_M \overline{\psi}\,\mathrm{d}\widetilde{\boldsymbol{\gamma}}$, i.e.

$$
\lim_{n\to\infty}\int_M \nabla\varphi_{t_n,x,v}(y)\cdot\nabla\psi(y)\,\mathrm{d}\mu_{t_n,x}(y) = \lim_{n\to\infty}\int_M \overline{\psi}(y,w)\,\mathrm{d}\boldsymbol{\gamma}_{t_n,x,v}(y,w)
$$
$$
= \int_M \overline{\psi}(y,w)\,\mathrm{d}\widetilde{\boldsymbol{\gamma}}(y,w)
$$
$$
= \int_M w\cdot\nabla\psi(x)\,\mathrm{d}\boldsymbol{\sigma}(w).
$$

On the other hand, from equation (3.3), letting $t\downarrow 0$, we deduce

$$
\lim_{n\to\infty}\int_M \nabla\varphi_{t_n,x,v}(y)\cdot\nabla\psi(y)\,\mathrm{d}\mu_{t_n,x}(y) = \nabla\psi(x)\cdot v.
$$

Being these last two identities valid for any $\psi\in C_c^\infty(M)$, we conclude that

$$
\int_M w\,\mathrm{d}\boldsymbol{\sigma}(w) = v.
$$

Finally, from the lower semicontinuity of $\boldsymbol{\gamma} \mapsto \int_M g(w,w)\,\mathrm{d}\boldsymbol{\gamma}(y,w)$ w.r.t. weak convergence of measures and the limit (3.2) we have

$$
\begin{aligned}
g(v,v) &= \lim_{n\to\infty} \int_M g(w,w)\,\mathrm{d}\boldsymbol{\gamma}_{t_n,x,v}(y,w) \\
&\geq \int_M g(w,w)\,\mathrm{d}\widetilde{\boldsymbol{\gamma}}(y,w) \\
&= \int_M g(w,w)\,\mathrm{d}\boldsymbol{\sigma}(w) \\
&\geq g\left(\int_M w\,\mathrm{d}\boldsymbol{\sigma}(w), \int_M w\,\mathrm{d}\boldsymbol{\sigma}(w)\right) \\
&= g(v,v),
\end{aligned}
$$

which forces the inequality

$$
\int_M g(w,w)\,\mathrm{d}\boldsymbol{\sigma}(w) \geq g\left(\int_M w\,\mathrm{d}\boldsymbol{\sigma}(w), \int_M w\,\mathrm{d}\boldsymbol{\sigma}(w)\right)
$$

to be an equality. This can be true only if $\boldsymbol{\sigma} = \delta_v$. $\qquad\square$

4. The Main Result

We start bounding from above the derivative $\frac{\mathrm{d}}{\mathrm{d}t} g_t$. Notice that the computations done in the next lemma are precisely those made by Otto–Westdickenberg in [13], which we report for completeness.

Proposition 4.1. *Let $x \in M$ and $v \in T_x M$. Then, there holds*

$$
\frac{\mathrm{d}}{\mathrm{d}t} \tfrac{1}{2} g_t(v,v) \leq - \int_{TM} \mathrm{Ric}(w,w)\,\mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y,w), \qquad \forall t > 0.
$$

Proof. We know by Proposition 3.4 that $(0,+\infty) \ni t \mapsto g_t(v,v)$ is smooth. Differentiating in time equation (3.1) we obtain

$$
\begin{aligned}
&-\nabla_y \cdot (\nabla \partial_t \varphi_{t,x,v}(y)\rho(t,x,y)) \\
&\quad = \nabla_x(\Delta_y \rho(t,x,y)) \cdot v + \nabla_y \cdot (\nabla\varphi_{t,x,v}(y)\Delta_y \rho(t,x,y)) \\
&\quad = \Delta_y\big(\nabla_x \rho(t,x,y) \cdot v\big) + \nabla_y \cdot (\nabla\varphi_{t,x,v}(y)\Delta_y \rho(t,x,y)) \\
&\quad = -\Delta_y\big(\nabla_y \cdot (\nabla\varphi_{t,x,v}\rho(t,x,y))\big) + \nabla_y \cdot (\nabla\varphi_{t,x,v}(y)\Delta_y \rho(t,x,y)).
\end{aligned}
$$

Therefore, by explicit computation and writing φ in place of $\varphi_{t,x,v}$ and ρ in place of $\rho(t, x, \cdot)$, we get:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} g_t(v, v) &= \frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} \int_M |\nabla\varphi|^2 \rho \, \mathrm{d}\mathrm{Vol} \\
&= \int_M \nabla\varphi \cdot \nabla\partial_t\varphi\rho + \frac{|\nabla\varphi|^2}{2} \Delta\rho \, \mathrm{d}\mathrm{Vol} \\
&= \int_M -\varphi\nabla \cdot (\nabla\partial_t\varphi\rho) + \frac{|\nabla\varphi|^2}{2} \Delta\rho \, \mathrm{d}\mathrm{Vol} \\
&= \int_M -\varphi\Delta(\nabla \cdot (\nabla\varphi\rho)) + \varphi\nabla \cdot (\nabla\varphi\Delta\rho) + \frac{|\nabla\varphi|^2}{2} \Delta\rho \, \mathrm{d}\mathrm{Vol} \\
&= \int_M \left(\nabla\Delta\varphi \cdot \nabla\varphi - \Delta\frac{|\nabla\varphi|^2}{2} \right) \rho \, \mathrm{d}\mathrm{Vol} \\
&= \int_M \left(-|\nabla^2\varphi|^2 - \mathrm{Ric}(\nabla\varphi, \nabla\varphi) \right) \rho \, \mathrm{d}\mathrm{Vol} \\
&\leq -\int_M \mathrm{Ric}(\nabla\varphi, \nabla\varphi)\rho \, \mathrm{d}\mathrm{Vol} \\
&= -\int_{TM} \mathrm{Ric}(w, w) \, \mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y, w),
\end{aligned}
$$

which is the thesis. In the last passage we expressed the result using the transport plans of Definition 3.8. $\qquad\square$

Corollary 4.2. *For any $x \in M$ and $v \in T_x M$ there holds*

$$
\overline{\lim_{t\downarrow 0}} \frac{g_t(v, v) - g(v, v)}{t} \leq -2\,\mathrm{Ric}(v, v).
$$

Proof. From the smoothness of $(0, +\infty) \ni t \mapsto g_t(v, v)$ and its continuity at time 0 we have

$$
\frac{g_t(v, v) - g(v, v)}{t} = \frac{1}{t} \int_0^t \frac{\mathrm{d}}{\mathrm{d}\xi} g_\xi(v, v) \, \mathrm{d}\xi,
$$

therefore, taking Proposition 4.1 into account, to conclude it is sufficient to show that

$$
\lim_{t\to 0} \int_{TM} \mathrm{Ric}(w, w) \, \mathrm{d}\boldsymbol{\gamma}_{t,x,v}(y, w) = \mathrm{Ric}(v, v).
$$

This is a direct consequence of the W_2–convergence of the transport plans $\boldsymbol{\gamma}_{t,x,v}$ to the delta measures $\delta_{x,v}$, given by Corollary 3.9, the fact that the map $TM \ni (x, v) \mapsto \mathrm{Ric}(v, v)$ is continuous with quadratic growth and Theorem 2.1. $\qquad\square$

Remark 4.3. As the computations just done show, in order to conclude that $\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} g_t(v, v)\big|_{t=0} = -\mathrm{Ric}(v, v)$, it would be enough to prove that – in the notation of the proof of Proposition 4.1 – there holds $\lim_{t\downarrow 0} \int_M |\nabla^2\varphi|^2 \rho \, \mathrm{d}\mathrm{Vol} = 0$.
As we are unable to get this convergence directly, we proceed differently. $\qquad\blacksquare$

Lemma 4.4. *Let $x \in M$, $v \in T_x M$ and define $s \mapsto \gamma_s := \exp_x(sv)$. Then, for every function $\varphi \in C^\infty(M)$ such that $\nabla\varphi(x) = -v$ and $\varepsilon \in (0, \overline{\varepsilon})$ there holds*

$$\lim_{t \to 0} \frac{\varepsilon}{2t} \left(\int_0^\varepsilon g_t(\gamma_s', \gamma_s') - g_0(\gamma_s', \gamma_s') \, \mathrm{d}s \right) \geq \Delta(\varepsilon\varphi)(x) + \Delta((\varepsilon\varphi)^c)(\gamma_\varepsilon),$$

where $\overline{\varepsilon} = \overline{\varepsilon}(\varphi)$ is given by Lemma 2.4.

Proof. By the definition of d_t and Proposition 3.5 we know that

$$\varepsilon \int_0^\varepsilon g_t(\gamma_s', \gamma_s') \, \mathrm{d}s \geq \mathrm{d}_t^2(x, \gamma_\varepsilon) \geq W_2^2(\mu_{t,x}, \mu_{t,\gamma_\varepsilon}), \qquad \forall \varepsilon, t > 0,$$

with both equalities when $t = 0$ and every $\varepsilon \in (0, \overline{\varepsilon})$. By the dual formulation of the optimal transport problem we have

$$\frac{1}{2} W_2^2(\mu_{t,x}, \mu_{t,\gamma_s}) \geq \int_M \varepsilon\varphi \, \mathrm{d}\mu_{t,x} + \int_M (\varepsilon\varphi)^c \, \mathrm{d}\mu_{t,\gamma_\varepsilon}, \qquad \forall \varepsilon > 0.$$

For $\varepsilon \in (0, \overline{\varepsilon})$, the identity (2.4) gives

$$\frac{1}{2} W_2^2(\mu_{0,x}, \mu_{0,\gamma_\varepsilon}) = \frac{1}{2} \mathrm{d}_0^2(x, \gamma_\varepsilon) = \frac{\varepsilon^2}{2}|v|^2 = \frac{1}{2}|\nabla(\varepsilon\varphi)|^2(x)$$

$$= \varepsilon\varphi(x) + (\varepsilon\varphi)^c(\gamma_\varepsilon) = \int_M \varepsilon\varphi \, \mathrm{d}\mu_{0,x} + \int_M (\varepsilon\varphi)^c \, \mathrm{d}\mu_{0,\gamma_\varepsilon}.$$

Thus, we get

$$\frac{\varepsilon}{2} \int_0^\varepsilon g_t(\gamma_s', \gamma_s') \, \mathrm{d}s \geq \int_M \varepsilon\varphi \, \mathrm{d}\mu_{t,x} + \int_M (\varepsilon\varphi)^c \, \mathrm{d}\mu_{t,\gamma_\varepsilon}, \qquad \forall \varepsilon \in (0, \overline{\varepsilon}), \ t \geq 0,$$

with equality for $t = 0$ and any $\varepsilon \in (0, \overline{\varepsilon})$. It follows that

$$\frac{\varepsilon}{2t} \left(\int_0^\varepsilon g_t(\gamma_s', \gamma_s') - g_0(\gamma_s', \gamma_s') \, \mathrm{d}s \right) \geq \int_M \varepsilon\varphi \, \mathrm{d}\frac{\mu_{t,x} - \mu_{0,x}}{t} + \int_M (\varepsilon\varphi)^c \, \mathrm{d}\frac{\mu_{t,\gamma_\varepsilon} - \mu_{0,\gamma_\varepsilon}}{t}.$$

Now notice that

$$\int_M \varepsilon\varphi \, \mathrm{d}\frac{\mu_{t,x} - \mu_{0,x}}{t} = \frac{1}{t} \left(\int_M \varepsilon\varphi(y) \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \rho(s, x, y) \, \mathrm{d}s \, \mathrm{d}\mathrm{Vol}(y) \right)$$

$$= \frac{1}{t} \int_0^t \int_M \varepsilon\varphi(y) \Delta_y \rho(s, x, y) \, \mathrm{d}s \, \mathrm{d}\mathrm{Vol}(y)$$

$$= \frac{1}{t} \int_0^t \int_M \Delta(\varepsilon\varphi)(y) \rho(s, x, y) \, \mathrm{d}s \, \mathrm{d}\mathrm{Vol}(y),$$

and this last term converges to $\Delta(\varepsilon\varphi)(x)$ as $t \to 0$.

Similarly, we have $\int_M (\varepsilon\varphi)^c \, \mathrm{d}\frac{\mu_{t,\gamma_\varepsilon} - \mu_{0,\gamma_\varepsilon}}{t} \to \Delta((\varepsilon\varphi)^c)(\gamma_\varepsilon)$ as $t \to 0$ and the thesis follows. $\qquad \square$

Proposition 4.5. *Let $\overline{x} \in M$, $\varphi \in C^\infty(M)$ be such that $\nabla^2\varphi(\overline{x}) = 0$. Put $\varphi_\varepsilon := \varepsilon\varphi$. Then, there holds*

$$(\Delta\varphi_\varepsilon^c)(\exp_{\overline{x}}(-\nabla\varphi(\overline{x}))) = -\varepsilon^2 \mathrm{Ric}(\nabla\varphi(\overline{x}), \nabla\varphi(\overline{x})) + \mathrm{REM}_\varepsilon,$$

where the reminder term REM_ε is bounded by

$$|\mathrm{REM}_\varepsilon| \leq \varepsilon^3 C,$$

the constant C depending only on a bound on the norms of $\nabla\varphi, \nabla^2\varphi$, the Riemann tensor Riem and its first covariant derivative.

Proof. Use Lemma 2.4 to find $\overline{\varepsilon} > 0$ such that points (i), (ii), (iii), (iv) of the statement are true for any $\varepsilon \in (0, \overline{\varepsilon})$. Fix such an ε and use the same notation used there.

Put $\overline{y} := T(\overline{x})$, let y_t be a unit speed geodesic such that $y_0 = \overline{y}$ and define the map $H_\varepsilon : [0, 1]^2 \to M$ by

$$H_\varepsilon(t, s) := \exp_{y_t}(-s\nabla\varphi_\varepsilon^c(y_t)).$$

(notice that for t fixed, the map $s \mapsto H_\varepsilon(t, s)$ is a geodesic, so that H_ε is a geodesic variation). By point (i) of Lemma 2.4 we know that H_ε is C^∞ and from point (iii) of the same lemma there holds

$$H_\varepsilon(t, s) = \exp_{y_t}(-s\nabla\varphi_\varepsilon^c(y_t)) = \exp_{S(y_t)}\big(-(1-s)\nabla\varphi_\varepsilon(S(y_t))\big). \tag{4.1}$$

Differentiating this expression in s we get, as $H_\varepsilon(t, 1) = S(y_t)$, that

$$\partial_s H_\varepsilon\Big|_{s=1} = \nabla\varphi_\varepsilon(H_\varepsilon(t, 1)), \qquad \forall t \in [0, 1]. \tag{4.2}$$

We claim that there holds

$$\left|\partial_t H_\varepsilon\Big|_{t=0}\right|_g \leq C_1, \qquad \left|\partial_s H_\varepsilon\Big|_{t=0}\right|_g \leq \varepsilon C_1, \qquad \left|\nabla_t \partial_s H_\varepsilon\Big|_{t=0}\right|_g \leq \varepsilon C_1, \tag{4.3}$$

for any $s \in [0, 1]$ and some constant C_1 depending only on a bound on $\nabla\varphi, \nabla^2\varphi$ and the Riemann tensor Riem of M. Indeed, the first one is obvious, the second comes from the identity

$$\partial_s H_\varepsilon(t, s) = \mathcal{T}_0^{1-s}\big(\varepsilon\nabla\varphi(S(y_t))\big), \tag{4.4}$$

which follows from relation (4.1), where $\mathcal{T}_0^{1-s}$ is the parallel transport map along the curve $r \mapsto \exp_{S(y_t)}\big(-(1-r)\nabla\varphi_\varepsilon(S(y_t))\big)$ from $r = 0$ to $r = 1 - s$. The last bound in (4.3) follows from formula (4.4) taking into account the smoothness of Jacobi fields.

By points (ii) and (iv) of Lemma 2.4 and the identity $H_\varepsilon(t, 1) = S(y_t)$ we have that

$$\varphi_\varepsilon^c(y_t) = \tfrac{1}{2}|\nabla\varphi_\varepsilon|^2(H_\varepsilon(t, 1)) - \varphi_\varepsilon(H_\varepsilon(t, 1)) \stackrel{(4.2)}{=} \left(\tfrac{1}{2}|\partial_s H_\varepsilon|^2 - \varphi_\varepsilon \circ H_\varepsilon\right)\Big|_{s=1}, \qquad \forall t \in [0, 1].$$

Differentiating once and using identity (4.2) again we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_\varepsilon^c(y_t) = \left(\partial_s H_\varepsilon \cdot \nabla_t \partial_s H_\varepsilon - \nabla\varphi_\varepsilon \circ H_\varepsilon \cdot \partial_t H_\varepsilon\right)\Big|_{s=1}$$
$$= \left(\partial_s H_\varepsilon \cdot \nabla_t \partial_s H_\varepsilon - \partial_s H_\varepsilon \cdot \partial_t H_\varepsilon\right)\Big|_{s=1}.$$

Differentiating a second time we obtain

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi_\varepsilon^c(y_t) = \left(|\nabla_t \partial_s H_\varepsilon|_g^2 + \partial_s H_\varepsilon \cdot \nabla_t \nabla_t \partial_s H_\varepsilon - \nabla_t \partial_s H_\varepsilon \cdot \partial_t H_\varepsilon - \partial_s H_\varepsilon \cdot \nabla_t \partial_t H_\varepsilon\right)\Big|_{s=1}.$$

Evaluating this expression at $t = 0$, recalling that $\nabla^2 \varphi_\varepsilon(\overline{x}) = 0$ and identity (4.2) we get $\nabla_t \partial_s H_\varepsilon \big|_{\substack{t=0 \\ s=1}} = 0$ and thus

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \varphi_\varepsilon^c(y_t) \Big|_{t=0} = \left(\partial_s H_\varepsilon \cdot \nabla_t \nabla_t \partial_s H_\varepsilon - \partial_s H_\varepsilon \cdot \nabla_t \partial_t H_\varepsilon \right) \Big|_{\substack{t=0 \\ s=1}}.$$

To compute this expression let $f, g : [0, 1] \to \mathbb{R}$ be defined as

$$f(s) := \partial_s H_\varepsilon \cdot \nabla_t \nabla_t \partial_s H_\varepsilon \Big|_{t=0}, \qquad g(s) := \partial_s H_\varepsilon \cdot \nabla_t \partial_t H_\varepsilon \Big|_{t=0},$$

so that

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \varphi_\varepsilon^c(y_t) \Big|_{t=0} = f(0) + \int_0^1 f'(\xi) \, \mathrm{d}\xi - g(0) - g'(0) - \int_0^1 \int_0^\xi g''(\eta) \, \mathrm{d}\eta. \qquad (4.5)$$

Since $t \mapsto H_\varepsilon(t, 0) = y_t$ is a geodesic, we have $\nabla_t \partial_t H_\varepsilon(t, 0) = 0$, recalling that also $\nabla_s \partial_s H_\varepsilon(t, s) = 0$ for every $t, s \in [0, 1]$ we get

$$f(0) = \partial_s H_\varepsilon \cdot \nabla_t \nabla_t \partial_s H_\varepsilon \Big|_{\substack{t=0 \\ s=0}}, \qquad\qquad g(0) = 0,$$

$$f'(s) = \partial_s H_\varepsilon \cdot \nabla_s \nabla_t \nabla_t \partial_s H_\varepsilon \Big|_{t=0}, \qquad\qquad g'(0) = \partial_s H_\varepsilon \cdot \nabla_s \nabla_t \partial_t H_\varepsilon \Big|_{\substack{t=0 \\ s=0}},$$

$$g''(s) = \partial_s H_\varepsilon \cdot \nabla_s \nabla_s \nabla_t \partial_t H_\varepsilon \Big|_{t=0}.$$

Hence, using repeatedly the fact that

$$R(\partial_t H_\varepsilon, \partial_s H_\varepsilon) X = \nabla_t \nabla_s (X \circ H_\varepsilon) - \nabla_s \nabla_t (X \circ H_\varepsilon)$$

for any smooth vector field X we get

$$f'(s) = \left(R(\partial_s H_\varepsilon, \partial_t H_\varepsilon)(\nabla_t \partial_s H_\varepsilon) \cdot \partial_s H_\varepsilon + \nabla_t \big(R(\partial_s H_\varepsilon, \partial_t H_\varepsilon) \partial_s H \big) \cdot \partial_s H_\varepsilon \right) \Big|_{t=0},$$

$$g''(s) = \Big(\nabla_s \big(R(\partial_s H_\varepsilon, \partial_t H_\varepsilon) \partial_t H_\varepsilon \big) \cdot \partial_s H_\varepsilon + R(\partial_s H_\varepsilon, \partial_t H_\varepsilon)(\nabla_t \partial_s H_\varepsilon) \cdot \partial_s H_\varepsilon$$

$$+ \nabla_t \big(R(\partial_s H_\varepsilon, \partial_t H_\varepsilon) \partial_s H_\varepsilon \big) \cdot \partial_s H_\varepsilon \Big) \Big|_{t=0}$$

$$= \Big((\nabla_s R)(\partial_s H_\varepsilon, \partial_t H_\varepsilon) \partial_t H_\varepsilon \cdot \partial_s H_\varepsilon + R(\partial_s H_\varepsilon, \nabla_t \partial_s H_\varepsilon) \partial_t H_\varepsilon \cdot \partial_s H_\varepsilon$$

$$+ R(\partial_s H_\varepsilon, \partial_t H_\varepsilon)(\nabla_t \partial_s H_\varepsilon) \cdot \partial_s H_\varepsilon + R(\partial_s H_\varepsilon, \partial_t H_\varepsilon)(\nabla_t \partial_s H_\varepsilon) \cdot \partial_s H_\varepsilon$$

$$+ \nabla_t \big(R(\partial_s H_\varepsilon, \partial_t H_\varepsilon) \partial_s H_\varepsilon \big) \cdot \partial_s H_\varepsilon \Big) \Big|_{t=0}.$$

Therefore the bounds (4.3) imply

$$|f'(s)|, |g''(s)| \leq C_2 \varepsilon^3,$$

for some constant C_2 depending only on a bound on the norms of $\nabla \varphi, \nabla^2 \varphi$, the Riemann tensor Riem and its first covariant derivative.

By equation (4.5) and the fact that

$$f(0) - g(0) - g'(0) = R(\partial_t H_\varepsilon, \partial_s H_\varepsilon) \partial_t H_\varepsilon \cdot \partial_s H_\varepsilon \Big|_{\substack{t=0 \\ s=0}},$$

we obtain

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi_\varepsilon^c(y_t)\Big|_{t=0} = R(\partial_t H_\varepsilon, \partial_s H_\varepsilon)\partial_t H_\varepsilon \cdot \partial_s H_\varepsilon\Big|_{\substack{t=0\\s=0}} + \mathrm{REM}_\varepsilon^1$$

$$= R\big(y_0', \nabla\varphi_\varepsilon^c(y_0)\big)y_0' \cdot \nabla\varphi_\varepsilon^c(y_0) + \mathrm{REM}_\varepsilon^1, \tag{4.6}$$

with $|\mathrm{REM}_\varepsilon^1| \leq C_3\varepsilon^3$, for some constant C_3 depending only on a bound on the norms of $\nabla\varphi, \nabla^2\varphi$, the Riemann tensor Riem and its first covariant derivative.

Now let $t \mapsto y_t^i$, $i = 1, \ldots, \dim(M)$, be a family of unit speed geodesics starting from $\overline{y}$ whose derivatives in 0 form an orthonormal basis of $T_{\overline{y}}M$. Writing equation (4.6) for $y_t := y_t^i$ and summing over the index i, it is easy to see that we get

$$\Delta\varphi_\varepsilon^c(\overline{y}) = -\mathrm{Ric}\big(\nabla\varphi_\varepsilon^c(\overline{y}), \nabla\varphi_\varepsilon^c(\overline{y})\big) + \mathrm{REM}_\varepsilon^2, \tag{4.7}$$

with $|\mathrm{REM}_\varepsilon^2| \leq \dim(M)C_3\varepsilon^3$.

To conclude, let $r \mapsto x_{r,\varepsilon} := \exp_{\overline{x}}(-r\varepsilon\nabla\varphi(\overline{x}))$ and observe that by definition and formula (4.1) there holds $x_{0,\varepsilon}' = -\nabla\varphi_\varepsilon(\overline{x})$ and $x_{1,\varepsilon}' = \nabla\varphi_\varepsilon(\overline{y})$, hence,

$$\mathrm{Ric}\big(\nabla\varphi_\varepsilon^c(\overline{y}), \nabla\varphi_\varepsilon^c(\overline{y})\big) = \mathrm{Ric}\big(\nabla\varphi_\varepsilon(\overline{x}), \nabla\varphi_\varepsilon(\overline{x})\big) + \int_0^1 (\nabla_r\mathrm{Ric})(x_{r,\varepsilon}', x_{r,\varepsilon}')\,\mathrm{d}r$$

$$= \varepsilon^2\mathrm{Ric}\big(\nabla\varphi(\overline{x}), \nabla\varphi(\overline{x})\big) + \int_0^1 (\nabla_r\mathrm{Ric})(x_{r,\varepsilon}', x_{r,\varepsilon}')\,\mathrm{d}r. \tag{4.8}$$

Then, given that $|x_{r,\varepsilon}'|_g = \varepsilon|\nabla\varphi(\overline{x})|_g$, we get $|(\nabla_r\mathrm{Ric})(x_{r,\varepsilon}', x_{r,\varepsilon}')|_g \leq C_4\varepsilon^3$ for some constant C_4 depending only on $|\nabla\varphi(\overline{x})|$ and a bound on the covariant derivative of the Riemann tensor Riem.

The thesis then follows from relations (4.7) and (4.8). $\qquad\square$

We are finally ready to prove our main result.

Theorem 4.6. *Let $s \mapsto \gamma_s$ be a geodesic on M (w.r.t. g_0). Then, there holds*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^1 g_t(\gamma_s', \gamma_s')\,\mathrm{d}s\Big|_{t=0} = -2\int_0^1 \mathrm{Ric}(\gamma_s', \gamma_s')\,\mathrm{d}s, \tag{4.9}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}g_t(\gamma_s', \gamma_s')\Big|_{t=0} = -2\mathrm{Ric}(\gamma_s', \gamma_s'), \qquad a.e.\ s \tag{4.10}$$

Proof. Thanks to Corollary 4.2, equation (4.10) follows directly from formula (4.9), thus, we concentrate on this latter.

Let K be a bound from below on the eigenvalues of the Ricci tensor. Then, for $\overline{x} \in M$, $\overline{v} \in T_{\overline{x}}M$, Proposition 4.1 yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}g_t(\overline{v}, \overline{v}) \leq -K\int_M |w|_g^2\,\mathrm{d}\gamma_{t,\overline{x},\overline{v}} = -Kg_t(\overline{v}, \overline{v}), \qquad \forall t > 0.$$

Thus, Proposition (3.2) and the Gronwall lemma give

$$g_t(\overline{v}, \overline{v}) \leq e^{-2Kt}g(\overline{v}, \overline{v}). \tag{4.11}$$

Therefore, from Corollary 4.2 we deduce

$$\varlimsup_{t\to 0} \int_0^1 \frac{g_t(\gamma_s', \gamma_s') - g(\gamma_s', \gamma_s')}{t}\, \mathrm{d}s \le \int_0^1 \varlimsup_{t\to 0} \frac{g_t(\gamma_s', \gamma_s') - g(\gamma_s', \gamma_s')}{t}\, \mathrm{d}s$$

$$\le -2 \int_0^1 \mathrm{Ric}(\gamma_s', \gamma_s')\, \mathrm{d}s, \tag{4.12}$$

where the use of Fatou lemma in the first inequality is justified by the estimate (4.11).

Using the compactness of the image of γ and a partition of the unity argument, it is not difficult to construct (we omit the details) a family $\{\varphi_t\}_{t\in[0,1]} \subset C_c^\infty(M)$ such that $\nabla\varphi_t(\gamma_t) = \gamma_t'$ and $\nabla^2\varphi_t(\gamma_t) = 0$ for any $t \in [0,1]$, and denoting by $\bar\varepsilon_t$ the value of $\bar\varepsilon$ corresponding to $\varphi := \varphi_t$ in Lemma 2.4 and by C_t the value of C corresponding to $\varphi := \varphi_t$ in Proposition 4.5, there holds

$$\bar\varepsilon := \inf_t \bar\varepsilon_t > 0, \qquad\qquad C := \sup_t C_t < +\infty.$$

Let now $0 = s_0 < s_1 < \cdots < s_N = 1$ be a partition of $[0,1]$ such that $\max_i |s_{i+1} - s_i| < \bar\varepsilon$. For $i = 0, \ldots, N-1$, we apply Lemma 4.4 and Proposition 4.5 to $x := \gamma_{s_i}$, $v := \gamma_{s_i}'$, $\varepsilon := s_{i+1} - s_i$ and $\varphi := \varphi_{s_i}$ to get

$$\varlimsup_{t\to 0}(s_{i+1} - s_i) \int_{s_i}^{s_{i+1}} \frac{g_t(\gamma_s', \gamma_s') - g(\gamma_s', \gamma_s')}{2t}\, \mathrm{d}s$$

$$= \Delta\Big((s_{i+1} - s_i)\varphi_{s_i}\Big)(\gamma_{s_i}) + \Delta\Big(\big((s_{i+1} - s_i)\varphi_{s_i}\big)^c\Big)(\gamma_{s_{i+1}})$$

$$\ge -(s_{i+1} - s_i)^2 \mathrm{Ric}(\gamma_{s_i}', \gamma_{s_i}') - C(s_{i+1} - s_i)^3.$$

Dividing by $(s_{i+1} - s_i)$ and summing over $i = 0, \ldots, N-1$, we get

$$\varlimsup_{t\to 0} \int_0^1 \frac{g_t(\gamma_s', \gamma_s') - g(\gamma_s', \gamma_s')}{2t}\, \mathrm{d}s \ge -\sum_{i=0}^{N-1} (s_{i+1} - s_i)\mathrm{Ric}(\gamma_{s_i}', \gamma_{s_i}') - C(s_{i+1} - s_i)^2.$$

Refining the partition in such a way that $\lim \max_i |s_{i+1} - s_i| \to 0$, we conclude that

$$\varlimsup_{t\to 0} \int_0^1 \frac{g_t(\gamma_s', \gamma_s') - g(\gamma_s', \gamma_s')}{2t}\, \mathrm{d}s \ge -2 \int_0^1 \mathrm{Ric}(\gamma_s', \gamma_s')\, \mathrm{d}s,$$

which, together with inequality (4.12), gives the thesis. $\qquad\square$

5. The Construction in a Non–Smooth Setting

By means of Proposition 3.5 we defined a flow using only the heat kernel and an original distance, independently of the presence of a smooth metric tensor. It is therefore natural to try to apply this construction in a non–smooth setting: the natural one being that of $RCD(K, \infty)$ spaces introduced in [6], i.e. those spaces with a lower Ricci curvature bound, in the sense of Lott–Sturm–Villani (see [10], [15]), and where the heat flow is linear. Indeed, a lower Ricci curvature bound seems necessary due to the fact that the heat flow is well defined and nicely behaves in relation with the W_2–geometry only in presence of the $CD(K, \infty)$ condition (see [5]). On the other hand, one does not only need a heat flow, but

also a heat kernel, and this latter exists only if the heat flow is linear (see [6] and [3]).

5.1. **Setting and preliminaries.**

5.1.1. *The Cheeger energy and the Sobolev space $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$.* Let $(X, \mathsf{d}, \mathfrak{m})$ be a complete and separable metric space endowed with a reference non–negative Radon measure $\mathfrak{m}$. The Cheeger energy functional $\mathrm{Ch} : L^2(X, \mathfrak{m}) \to [0, +\infty]$ is defined as

$$\mathrm{Ch}(f) := \inf \varliminf_{n\to\infty} \frac{1}{2} \int_X |Df_n|^2 \, \mathrm{d}\mathfrak{m},$$

where the infimum is taken among all sequences of Lipschitz functions $f_n \in L^2(X, \mathfrak{m})$, converging to f in $L^2(X, \mathfrak{m})$.

The local Lipschitz constant $|Dg| : X \to [0, +\infty]$ of a function $g : X \to \mathbb{R}$ is defined as

$$|Dg|(x) := \varlimsup_{y\to x} \frac{|g(y) - g(x)|}{\mathsf{d}(x, y)}.$$

It is immediate to check that Ch is convex, lower semicontinuous and with dense domain, therefore, the classical theory of gradient flows in Hilbert spaces ensures that for any $f \in L^2(X, \mathfrak{m})$ there exists a unique gradient flow for Ch starting from f. In general, however, Ch is not a quadratic form (consider for instance the case of finite dimensional Banach spaces), therefore, its gradient flow could be non–linearly dependent on the initial datum.

The Sobolev space $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is then defined as

$$W^{1,2}(X, \mathsf{d}, \mathfrak{m}) := \{f \in L^2(X, \mathfrak{m}) \ : \ \mathrm{Ch}(f) < +\infty\},$$

endowed with the norm

$$\|f\|_{W^{1,2}}^2 := \|f\|_{L^2}^2 + 2\mathrm{Ch}(f).$$

Notice that since in general Ch is not a quadratic form, the space $W^{1,2}$ can fail to be a Hilbert space (while it is always a Banach space).

If Ch is a quadratic form, it is immediate to check that it is actually Dirichlet form. In this case, we denote by Δ its infinitesimal generator, then standard Dirichlet form theory grants that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{2} \|f_t\|_{L^2}^2 &= -\mathrm{Ch}(f_t), && \forall t > 0\,, \\
\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ch}(f_t) &= \|\Delta f_t\|_{L^2}^2, && \forall t > 0\,, \\
t &\mapsto \|\Delta f_t\|_{L^2} && \text{is not increasing,}
\end{aligned}
\tag{5.1}
$$

whenever f_t is a gradient flow for Ch.

5.1.2. $CD(K,\infty)$ *spaces and gradient flow of the relative entropy.* Let $(X, \mathsf{d}, \mathfrak{m})$ be such that for some constant $C > 0$ there holds

$$\int_X e^{-C\mathsf{d}^2(\cdot, x_0)} \, \mathrm{d}\mathfrak{m} < +\infty. \tag{5.2}$$

Then, the relative entropy functional $\mathrm{Ent}_{\mathfrak{m}} : \mathscr{P}_2(X) \to \mathbb{R} \cup \{+\infty\}$ is defined as

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) := \begin{cases} \displaystyle\int_X \rho \log \rho \, \mathrm{d}\mathfrak{m}, & \text{if } \mu = \rho\mathfrak{m}, \\ +\infty, & \text{if } \mu \text{ is not absolutely continuous w.r.t. } \mathfrak{m}. \end{cases}$$

Putting $\widetilde{\mathfrak{m}} := z^{-1} e^{-C\mathsf{d}^2(\cdot, x_0)}\mathfrak{m}$, being $z := \int_X e^{-C\mathsf{d}^2(\cdot, x_0)} \, \mathrm{d}\mathfrak{m}$ the normalization constant, where C is the constant in condition (5.2), we see that there holds

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) = \mathrm{Ent}_{\widetilde{\mathfrak{m}}}(\mu) - C \int_X \mathsf{d}^2(\cdot, x_0) \, \mathrm{d}\mu - \log z,$$

which grants, thanks to the fact that the entropy w.r.t. the probability measure $\widetilde{\mathfrak{m}}$ is non–negative and lower semicontinuous in duality with $C_b(X)$, that $\mathrm{Ent}_{\mathfrak{m}}$ is indeed well defined on $\mathscr{P}_2(X)$ and lower semicontinuous w.r.t. W_2–convergence. The domain $D(\mathrm{Ent}_{\mathfrak{m}})$ of the entropy is the set of $\mu \in \mathscr{P}_2(X)$ such that $\mathrm{Ent}_{\mathfrak{m}}(\mu) < +\infty$.

Definition 5.1 ($CD(K,\infty)$ spaces). *A complete separable metric measure space $(X, \mathsf{d}, \mathfrak{m})$ satisfying condition (5.2) for some $C > 0$ is said $CD(K,\infty)$, for $K \in \mathbb{R}$, provided that the following is true. For any couple of measures $\mu, \nu \in D(\mathrm{Ent}_{\mathfrak{m}})$, there exists a geodesic $\mu_t \subset \mathscr{P}_2(X)$ such that $\mu_0 = \mu$, $\mu_1 = \nu$ and*

$$\mathrm{Ent}_{\mathfrak{m}}(\mu_t) \le (1-t)\mathrm{Ent}_{\mathfrak{m}}(\mu) + t\mathrm{Ent}_{\mathfrak{m}}(\nu) - \frac{K}{2}W_2^2(\mu, \nu).$$

Notice that in a $CD(K,\infty)$ space one always has that $(\mathrm{supp}(\mathfrak{m}), \mathsf{d})$ is a length space, i.e. the distance can be always realized as the infimum of the lengths of the curves.

The following result is proved in [5] (see also [8]).

Theorem 5.2. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a $CD(K,\infty)$ space and $\mu \in D(\mathrm{Ent}_{\mathfrak{m}})$ a measure with finite entropy. Then, there exists a unique locally absolutely continuous curve $[0, +\infty) \ni t \mapsto \mu_t \in \mathscr{P}_2(X)$ such that*

$$\mathrm{Ent}_{\mathfrak{m}}(\mu) = \mathrm{Ent}_{\mathfrak{m}}(\mu_T) + \frac{1}{2}\int_0^T |\dot{\mu}_t|^2 \, \mathrm{d}t + \frac{1}{2}\int_0^T |D^-\mathrm{Ent}_{\mathfrak{m}}|^2(\mu_t) \, \mathrm{d}t, \qquad \forall T > 0,$$

where the slope of the entropy $|D^-\mathrm{Ent}_{\mathfrak{m}}|$ is defined as

$$|D^-\mathrm{Ent}_{\mathfrak{m}}|(\nu) := \overline{\lim_{W_2(\sigma, \nu) \to 0}} \frac{\left(\mathrm{Ent}_{\mathfrak{m}}(\sigma) - \mathrm{Ent}_{\mathfrak{m}}(\nu)\right)^-}{W_2(\sigma, \nu)}.$$

The curves defined by this theorem are called gradient flows of the entropy $\mathrm{Ent}_{\mathfrak{m}}$.

5.1.3. *$RCD(K, \infty)$ spaces.* A crucial result obtained in [5] is the identification of the gradient flow of Ch and the one of $\text{Ent}_{\mathfrak{m}}$ (see also [7] for a survey in the compact case).

Theorem 5.3. *Let $(X, \mathsf{d}, \mathfrak{m})$ be a $CD(K, \infty)$ space and $\mu = f\mathfrak{m} \in D(\text{Ent}_{\mathfrak{m}})$ with $f \in L^2(X, \mathfrak{m})$. Let $[0, \infty) \ni t \mapsto f_t \subset L^2(X, \mathfrak{m})$ be the gradient flow of Ch and $[0, \infty) \ni t \mapsto \mu_t \subset \mathscr{P}_2(X)$ the gradient flow of the entropy, respectively, with $f_0 = f$ and $\mu_0 = \mu$. Then,*

$$\mu_t = f_t \mathfrak{m}, \qquad \forall t \geq 0.$$

Due to this result, the heat flow on a $CD(K, \infty)$ space can be unambiguously defined as the gradient flow of Ch or as the the gradient flow of $\text{Ent}_{\mathfrak{m}}$.

There are $CD(K, \infty)$ spaces such that $W^{1,2}$ is not a Hilbert space (e.g. finite dimensional Banach but non–Hilbert spaces, see the last theorem in [17]), hence, having a nonlinear heat flow. The class of spaces with linear heat flow has been investigated in [6] and [3], the definition being the following.

Definition 5.4 ($RCD(K, \infty)$ spaces). *We say that $(X, \mathsf{d}, \mathfrak{m})$ is an $RCD(K, \infty)$ space provided that it is a $CD(K, \infty)$ space and $W^{1,2}(X, \mathsf{d}, \mathfrak{m})$ is a Hilbert space.*

A non–trivial property of $RCD(K, \infty)$ spaces is that the heat flow contracts the W_2–distance (this is false in non–Hilbert, finite dimensional Banach spaces as shown in [12]).

Proposition 5.5. *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space and $[0, \infty) \ni t \mapsto \mu_t, \nu_t$ two gradient flows of the relative entropy. Then*

$$W_2(\mu_t, \nu_t) \leq e^{Kt} W_2(\mu_0, \nu_0), \qquad \forall t \geq 0.$$

A priori, on a $CD(K, \infty)$ space the gradient flow of the entropy is well defined only when the initial measure has finite entropy (Theorem 5.2), but thanks to this contraction result, there is a natural extension of the flow to initial measures in the W_2–closure of the domain of the entropy. Such closure consists in measures μ in $\mathscr{P}_2(X)$ with $\text{supp}(\mu) \subset \text{supp}(\mathfrak{m})$, we will denote the space of these measures μ by $\mathscr{P}_2(\text{supp}(\mathfrak{m}))$. More precisely, we have the following simple corollary.

Corollary 5.6. *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space. Then, there exist a unique one parameter family of maps $\mathsf{H}_t : \mathscr{P}_2(\text{supp}(\mathfrak{m})) \to \mathscr{P}_2(\text{supp}(\mathfrak{m}))$ such that:*

(i) *for any $\mu, \nu \in \mathscr{P}_2(\text{supp}(\mathfrak{m}))$ there holds*

$$W_2(\mathsf{H}_t(\mu), \mathsf{H}_t(\nu)) \leq e^{-Kt} W_2(\mu, \nu), \qquad \forall t \geq 0, \tag{5.3}$$

(ii) *for any $\mu \in \mathscr{P}_2(\text{supp}(\mathfrak{m}))$ the curve $t \mapsto \mathsf{H}_t(\mu)$ is W_2–continuous,*

(iii) *for any $\mu \in D(\text{Ent}_{\mathfrak{m}})$, the curve $t \mapsto \mathsf{H}_t(\mu)$ is the gradient flow of the entropy starting from μ, according to Theorem 5.2.*

It can be shown that $\mathsf{H}_t(\mu) \ll \mathfrak{m}$ for any $\mu \in \mathscr{P}_2(\text{supp}(\mathfrak{m}))$ and any $t > 0$. Thus, the maps $\mathsf{H}_t : \mathscr{P}_2(\text{supp}(\mathfrak{m})) \to \mathscr{P}_2(\text{supp}(\mathfrak{m}))$ induce maps $\mathsf{h}_t : L^1(X, \mathfrak{m}) \to L^1(X, \mathfrak{m})$ via the formula

$$\mathsf{h}_t(f)\mathfrak{m} := \mathsf{H}_t(f\mathfrak{m}), \qquad \forall f \in L^1(X, \mathfrak{m}) \; : \; f\mathfrak{m} \in \mathscr{P}_2(X).$$

and the requirement that h_t is linear and continuous in L^1.

We recall that $H_t(\mu) \in D(\text{Ent}_{\mathfrak{m}})$ implies the $L^1 \to L \log L$ regularization property

$$h_t(f) \in L \log L(X, \mathfrak{m}), \qquad \forall t > 0, \ f \in L^1(X, \mathfrak{m}).$$

We say that the flow h_t is *ultracontractive* provided that the following stronger regularization holds:

$$\exists p > 1 \quad \text{such that} \quad \|h_t(f)\|_{L^p} \le C(t) \|f\|_{L^1}, \qquad \forall t > 0,$$

or equivalently (by the Young inequality for convolutions) if

$$\|h_t(f)\|_{L^\infty} \le \widetilde{C}(t) \|f\|_{L^1}, \qquad \forall t > 0. \tag{5.4}$$

5.1.4. *Convergence of metric–measure structures.* We recall here some basic concepts regarding convergence of metric–measure structures. The approach that we chose is that of $\mathbb{D}$–*convergence* introduced by Sturm in [15] and of *pointed* $\mathbb{D}$–*convergence* analyzed in [2]. There are strong relations between these notions and those of *measured Gromov–Hausdorff convergence* and *pointed measured Gromov–Hausdorff convergence*, we refer to [2] for a discussion.

We say that a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ is normalized provided that $\mathfrak{m}$ is a probability measure and that it has finite variance if $\mathfrak{m} \in \mathscr{P}_2(X)$. In the following definition and the discussion thereafter we write $\sqcup$ for the disjoint union of two sets.

Definition 5.7 ($\mathbb{D}$–convergence)**.** *Let $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$, $n \in \mathbb{N}$, and $(X, \mathsf{d}, \mathfrak{m})$ be normalized metric measure spaces with finite variance. We say that $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$ converges to $(X, \mathsf{d}, \mathfrak{m})$ in $\mathbb{D}$–sense provided that there exists a metric D on $Y :=$ $\sqcup_n X_n \sqcup X$ which coincides with d_n (resp. d) when restricted to X_n (resp. X) and such that*

$$\lim_{n \to \infty} W_2^{(Y, \mathsf{D})}(\mathfrak{m}_n, \mathfrak{m}) = 0$$

Notice that Sturm in [15] defined a distance $\mathbb{D}$ on the space of normalized metric measure spaces with finite variance, and that convergence w.r.t. this distance means precisely what we just defined: we preferred this point of view because in our discussion the presence of a distance behind a converging sequence is not really important.

While $\mathbb{D}$–convergence is suitable to deal with non–compact spaces (as opposed to measured Gromov–Hausdorff convergence), it requires the measure $\mathfrak{m}$ to be in $\mathscr{P}_2(X)$, which is a quite restrictive assumption in general. To overcome this problem, in [2] a variant of $\mathbb{D}$–convergence has been proposed, called pointed $\mathbb{D}$–convergence.

Definition 5.8 (Pointed $\mathbb{D}$–convergence)**.** *Let $(X_n, \mathsf{d}_n, \mathfrak{m}_n, \overline{x}_n)$, $n \in \mathbb{N}$, and $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$ be pointed metric measure spaces with $\overline{x}_n \in \text{supp}(\mathfrak{m}_n)$, $n \in \mathbb{N}$, $\overline{x} \in \text{supp}(\mathfrak{m})$ and $\mathfrak{m}(X) > 0$. We say that $(X_n, \mathsf{d}_n, \mathfrak{m}_n, \overline{x}_n)$ converges to $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$ in the pointed $\mathbb{D}$–sense provided that there exists a constant $\mathsf{C} \ge 0$ such that the following are true.*

(i)

$$\sup_{n \in \mathbb{N}} \int_X \mathsf{d}^2(\cdot, \overline{x}_n) e^{-\mathsf{Cd}^2(\cdot, \overline{x}_n)} \, \mathrm{d}\mathfrak{m}_n < +\infty.$$

ii)

$$\lim_{n \to \infty} \int_X e^{-\mathsf{Cd}^2(\cdot, \overline{x}_n)} \, \mathrm{d}\mathfrak{m}_n = \int_X e^{-\mathsf{Cd}^2(\cdot, \overline{x})} \, \mathrm{d}\mathfrak{m}.$$

iii) *There exists a metric* D *on* $Y := \sqcup_n X_n \sqcup X$ *which coincides with* d_n *(resp.* d) *when restricted to* X_n *(resp.* X) *and such that*

$$\lim_{n \to \infty} \mathsf{D}(x_n, x) = 0,$$
$$\lim_{n \to \infty} W_2^{(Y, \mathsf{D})}(\widetilde{\mathfrak{m}}_n, \widetilde{\mathfrak{m}}) = 0,$$

where $\widetilde{\mathfrak{m}}_n := z_n^{-1} e^{-\mathsf{Cd}^2(\cdot, \overline{x}_n)} \mathfrak{m}_n$, $n \in \mathbb{N}$, *and* $\widetilde{\mathfrak{m}} := z^{-1} e^{-\mathsf{Cd}^2(\cdot, \overline{x})} \mathfrak{m}$, *being* $z_n := \int_X e^{-\mathsf{Cd}^2(\cdot, \overline{x}_n)} \, \mathrm{d}\mathfrak{m}_n$ *and* $z := \int_X e^{-\mathsf{Cd}^2(\cdot, \overline{x})} \, \mathrm{d}\mathfrak{m}$ *the normalization constants.*

It is not difficult to see that under pointed $\mathbb{D}$–convergence there holds

$$\forall x \in \mathrm{supp}(\mathfrak{m}) \text{ there exists } n \mapsto x_n \in \mathrm{supp}(\mathfrak{m}_n) \text{ such that } \lim_{n \to \infty} \mathsf{D}(x_n, x) = 0,$$

which shows, in particular, that $\mathbb{D}$–convergence is a particular case of pointed $\mathbb{D}$–convergence (just pick $\mathsf{C} = 0$ and use this property to obtain a suitable converging sequence of reference points).

Remark 5.9. The definitions of $\mathbb{D}$–convergence and pointed $\mathbb{D}$–convergence can directly be adapted to pseudo metric spaces, i.e. spaces where the "distance" is not required to be positive at couples of different points. In this case, one just requires D to be a pseudo distance on Y. ∎

Lower Ricci curvature bounds and heat flows are stable w.r.t. $\mathbb{D}$–convergence, as stated in the next propositions (for the proof, see [2]).

Proposition 5.10 (Stability of $RCD(K, \infty)$ spaces). *Let* $(X_n, \mathsf{d}_n, \mathfrak{m}_n, \overline{x}_n)$, $n \in \mathbb{N}$, *be a sequence of pointed metric measure spaces converging to some* $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$ *in the pointed* $\mathbb{D}$–*sense, as in Definition 5.8. Assume that* $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$ *is an* $RCD(K, \infty)$ *space for every* $n \in \mathbb{N}$. *Then,* $(X, \mathsf{d}, \mathfrak{m})$ *is an* $RCD(K, \infty)$ *space as well.*

In the next statement, we will denote with $\mathsf{H}_{n,t}$ the heat flow on X_n and by H_t the one on X.

Proposition 5.11 (Stability of the heat flow). *Let* $(X_n, \mathsf{d}_n, \mathfrak{m}_n, \overline{x}_n)$, $n \in \mathbb{N}$, *be a sequence of pointed metric measure spaces converging to some* $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$ *in the pointed* $\mathbb{D}$–*sense, as in Definition 5.8. Assume that* $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$ *is an* $RCD(K, \infty)$ *space for every* $n \in \mathbb{N}$, *so that also* $(X, \mathsf{d}, \mathfrak{m})$ *is an* $RCD(K, \infty)$ *space, according to Proposition 5.10.*

Let (Y, D) as in the Definition 5.8. Then, for every sequence $n \mapsto x_n \in \text{supp}(\mathfrak{m}_n)$ and point $x \in \text{supp}(\mathfrak{m})$ such that $\mathsf{D}(x_n, x) \to 0$, there holds

$$\lim_{n \to \infty} W_2^{(Y,\mathsf{D})}\big(\mathsf{H}_{n,t}(\delta_{x_n}), \mathsf{H}_t(\delta_x)\big) = 0$$

for every $t \geq 0$

5.2. **Definition of the flow and properties.** We are going to define two families of pseudo-distances $\widetilde{\mathsf{d}}_t$ and d_t: the former corresponds to the 'chord' distance in the embedding (1.2), the latter to the 'arc' one.

Definition 5.12. *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space with $\text{supp}(\mathfrak{m}) = X$ and $t \geq 0$. The function $\widetilde{\mathsf{d}}_t : X \times X \to [0, +\infty]$ is defined as:*

$$\widetilde{\mathsf{d}}_t(x, y) := W_2(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y))$$

It is immediate to check that $\widetilde{\mathsf{d}}_t$ is a pseudo-distance on X (i.e. it shares all the properties of a distance except the fact that it can be 0 at couples of different points), see Theorem 5.15 below for the simple details. For a $\widetilde{\mathsf{d}}_t$–Lipschitz curve $s \mapsto \gamma_s$, we will denote by $|\dot\gamma_s|_t$ its metric speed defined as in (2.2) computed in the pseudo-metric space $(X, \widetilde{\mathsf{d}}_t)$ (it is easily verified that to pass from metric to pseudo-metric spaces creates no problems in the definition).

Observe that Corollary 5.6 ensures that if $t \mapsto \gamma_t \in X$ is a d–Lipschitz curve, then it is also $\widetilde{\mathsf{d}}_t$–Lipschitz. Hence the following definition makes sense:

Definition 5.13. *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space with $\text{supp}(\mathfrak{m}) = X$ and $t \geq 0$. The function $\mathsf{d}_t : X \times X \to [0, +\infty]$ is defined as:*

$$\mathsf{d}_t(x, t) := \inf_\gamma \int_0^1 |\dot\gamma_s|_t \, ds,$$

where the infimum is taken among all d–Lipschitz curves γ on $[0, 1]$ joining x to y.

Remark 5.14. In connection with Remark 3.6 notice that in the non-smooth situation we do not expect the pseudo-distances $\widetilde{\mathsf{d}}_t$ to be bi-Lipschitz w.r.t. the original distance d, therefore in defining the pseudo-distance d_t as infimum of length of curves, the length being measured w.r.t. $\widetilde{\mathsf{d}}_t$, we need to make a choice: either we directly consider $\widetilde{\mathsf{d}}_t$–Lipschitz curves or we consider only those which are also d–Lipschitz.

Both choices seem reasonable, we preferred the second one because it makes simpler to prove the desired weak continuity properties in Theorem 5.18. ∎

Theorem 5.15 (Basic properties of the flow). *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space such that $\text{supp}(\mathfrak{m}) = X$. Then, $\widetilde{\mathsf{d}}_0 = \mathsf{d}_0 = \mathsf{d}$ and for every $t > 0$ the functions $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$ are pseudo distances on X (i.e. they share all the properties of a distance except the fact that they can be 0 at couples of different points).*

Also, if for some $t > 0$ the map $\mathsf{H}_t : \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ is injective, then $\widetilde{\mathsf{d}}_t$ and d_t are distances.

Moreover, if (X, d) is compact, the distances $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$ induce the same topology of d *on* X.

Proof. The fact that $\widetilde{\mathsf{d}}_0 = \mathsf{d}_0 = \mathsf{d}$ is obvious.

By construction, $\widetilde{\mathsf{d}}_t$ and d_t are both symmetric, satisfy the triangular inequality and $\widetilde{\mathsf{d}}_t(x, x) = \mathsf{d}_t(x, x) = 0$ for any $x \in X$ and $t \geq 0$. Also, it clearly holds

$$\widetilde{\mathsf{d}}_t(x, y) \leq \mathsf{d}_t(x, y), \qquad \forall x, y \in X, \ t \geq 0. \tag{5.5}$$

Thus, it remains to prove that $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$ are both real valued. By estimate (5.3) we immediately get

$$\widetilde{\mathsf{d}}_t(x, y) \leq e^{-Kt} \mathsf{d}(x, y), \qquad \forall x, y \in X, \ t \geq 0, \tag{5.6}$$

which directly implies $|\dot{\gamma}_s|_t \leq e^{-Kt} |\dot{\gamma}_s|_0$ for a.e. s for any d–Lipschitz curve $s \mapsto \gamma_s$. Hence from the definition we obtain that

$$\mathsf{d}_t(x, y) \leq e^{-Kt} \mathsf{d}(x, y), \qquad \forall x, y \in X, \ t \geq 0. \tag{5.7}$$

Assume now that $\mathsf{H}_t : \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ is injective for some $t > 0$. Then, since W_2 is a distance on $\mathscr{P}_2(X)$, we have $\mathsf{H}_t(\delta_x) \neq \mathsf{H}_t(\delta_y)$ for any $x \neq y$, $t \geq 0$. Hence $\widetilde{\mathsf{d}}_t(x, y) > 0$ and, by relation (5.5), also $\mathsf{d}_t(x, y) > 0$.

Assume that (X, d) is compact. Given the chain of inequalities

$$\widetilde{\mathsf{d}}_t \leq \mathsf{d}_t \leq e^{-Kt} \mathsf{d},$$

to conclude it is sufficient to prove that $\widetilde{\mathsf{d}}_t$ induces the same topology of d. Let $\iota_t : (X, \mathsf{d}) \to (\mathscr{P}_2(X), W_2) \sim (X, \widetilde{\mathsf{d}}_t)$ be given by $\iota_t(x) := \mathsf{H}_t(\delta_x)$. Our aim is to show that ι_t is a homeomorphism of X with its image $Y_t := \iota_t(X) \subset \mathscr{P}_2(X)$.

Inequality (5.6) grants that ι_t is continuous. It is clearly surjective and, by what we proved, also injective. To conclude, we thus need to prove that $\iota_t^{-1} : Y_t \to X$ is continuous. Let $y_n \subset Y_t$ be a sequence converging to some $y \in Y_t$ and put $x_n := \iota_t^{-1}(y_n)$, $n \in \mathbb{N}$, $x := \iota_t^{-1}(y)$. Since X is compact, up to a subsequence, not relabeled, we can assume that x_n converges to some $x' \in X$. Since ι_t is continuous we have $\iota_t(x') = \lim_n \iota_t(x_n) = \lim_n y_n = y$, which forces $x' = x$. Being this result independent of the converging subsequence chosen, the thesis follows. $\qquad \square$

Now, to prove that $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$ are distances, we need to know that the heat flow is injective on $RCD(K, \infty)$ spaces. Quite surprisingly, this does not seem to be so obvious: we only know a proof in the case of ultracontractive flow, where we can bring the problem to a question in L^2 and then use the analyticity of the flow.

Proposition 5.16 (Injectivity of the heat flow)**.** *Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, \infty)$ space. Assume that the flow h_t is ultracontractive in the sense of inequality (5.4). Then, for $x \neq y$ and $t \geq 0$ there holds $\mathsf{H}_t(\delta_x) \neq \mathsf{H}_t(\delta_y)$.*

Proof. By point (ii) of Corollary 5.6 we have that for $x \neq y$ and $t_0 > 0$ sufficiently close to 0 there holds $\mathsf{H}_{t_0}(\delta_x) \neq \mathsf{H}_t(\delta_y)$. Now we use the ultracontractivity property of the flow to write $\mathsf{H}_{t_0}(\delta_x) = f\mathfrak{m}$ and $\mathsf{H}_{t_0}(\delta_y) = \widetilde{f}\mathfrak{m}$ for some $f, \widetilde{f} \in L^2(X, \mathfrak{m})$, $f \neq \widetilde{f}$. The conclusion then follows from the fact that the flow is analytic in $L^2(X, \mathfrak{m})$, as we now explain in detail.

By Theorem 5.3, the (restriction of the) flow h_t in $L^2(X, \mathfrak{m})$ is linear, strongly continuous and the gradient flow of Ch. Denote by Δ its infinitesimal generator. We claim that

$$t\|\Delta h_t(g)\|_{L^2} \leq \|g\|_{L^2}, \qquad \forall g \in L^2(X, \mathfrak{m}), \ t > 0. \tag{5.8}$$

Indeed, using formula (5.1) we get

$$\frac{t^2}{2}\|\Delta h_t(g)\|_{L^2}^2 \leq \int_0^t s\|\Delta h_s(g)\|_{L^2}^2 \, \mathrm{d}s = -\int_0^t s\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{Ch}\big(h_s(g)\big) \, \mathrm{d}s$$

$$= \int_0^t \mathrm{Ch}(h_s(g)) - \mathrm{Ch}(h_t(g)) \, \mathrm{d}s \leq \int_0^t \mathrm{Ch}(h_s(g)) \, \mathrm{d}s$$

$$= -\int_0^t \frac{\mathrm{d}}{\mathrm{d}s}\frac{1}{2}\|h_s(g)\|_{L^2}^2 \, \mathrm{d}s \leq \frac{1}{2}\|g\|_{L^2}^2,$$

and inequality (5.8) follows. Hence, we also get $\|\Delta\Delta h_t(g)\| = \|\Delta h_{t/2}\Delta h_{t/2}(g)\| \leq \frac{4\|g\|}{t^2}$ and, denoting by $\Delta^{(n)}$ the application of n times the operator Δ, by induction we deduce

$$\|\Delta^{(n)}h_t(g)\|_{L^2} \leq \|g\|_{L^2}\frac{n^n}{t^n}, \qquad \forall g \in L^2(X, \mathfrak{m}), \ t > 0.$$

It is readily checked that this bound implies that for any $t_0 > 0$ the series

$$\sum_{n \geq 0} \frac{(t - t_0)^n}{n!}\Delta^{(n)}h_{t_0}(g),$$

converges for any t in a sufficiently small neighborhood of t_0 and that its sum is precisely $h_t(g)$. Hence, the curve $t \mapsto h_t(g)$ is analytic, as claimed, and the injectivity of the heat flow follows. $\square$

Remark 5.17 (The finite dimensional case). There is a natural way to define $RCD(K, N)$ spaces for finite N: just require that the space is $CD(K, N)$ and that $W^{1,2}$ is Hilbert. The fact that $CD(K, N)$ spaces are *doubling* (in particular, bounded closed sets are compact) and support a *weak local 1-1 Poincaré inequality*, together with the results of [14] yield the following Gaussian estimates for the heat kernel:

$$C'\frac{e^{-\frac{\mathsf{d}^2(x,y)}{4C't}}}{\sqrt{\mathfrak{m}(B_{\sqrt{t}}(x))}} \leq \rho(t, x, y) \leq C\frac{e^{-\frac{\mathsf{d}^2(x,y)}{4t}}}{\sqrt{\mathfrak{m}(B_{\sqrt{t}}(x))\mathfrak{m}(B_{\sqrt{t}}(y))}}\left(1 + \frac{\mathsf{d}^2(x,y)}{t}\right)^{n/2},$$

$$\tag{5.9}$$

where $\rho(t, x, \cdot)$ is the density of $H_t(\delta_x)$, n is the doubling constant and the constants C, C' depend only on the doubling constant and the constant appearing in the Poincaré inequality.

In particular, the upper bound implies that heat flow in $RCD(K, N)$ spaces is always ultracontractive and therefore injective. The lower bound and the fact that $\mathfrak{m}$ is doubling easily yield that if $x_n \subset X$ is such that $H_t(\delta_{x_n})$ is a bounded sequence in $(\mathscr{P}_2(X), W_2)$, then x_n is also bounded. This latter fact then ensures

that the proof of the last part of Theorem 5.15 can be repeated without assuming (X, d) to be compact, thus obtaining the following result.

Let $(X, \mathsf{d}, \mathfrak{m})$ be an $RCD(K, N)$ space, $N < +\infty$, with $\operatorname{supp}(\mathfrak{m}) = X$. Then, all the conclusions of Theorem 5.15 are true. We omit the details. ∎

We now analyze the continuity properties of the flow under $\mathbb{D}$–convergence.

Theorem 5.18 (Continuity in time). *Let $(X, \mathsf{d}, \mathfrak{m})$ be a compact normalized $RCD(K, \infty)$ space with $\operatorname{supp}(\mathfrak{m}) = X$ and such that the heat flow $\mathsf{H}_t : \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ is injective for any $t \geq 0$. Then,*

- *the curve $t \mapsto (X, \widetilde{\mathsf{d}}_t, \mathfrak{m})$ is continuous w.r.t. $\mathbb{D}$–convergence,*
- *the curve $t \mapsto (X, \mathsf{d}_t, \mathfrak{m})$ is right continuous w.r.t. $\mathbb{D}$–convergence.*

Proof. By Theorem 5.15 we know that both $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$ induce the same topology of d, hence the Borel structures are the same. In particular, $\mathfrak{m}$ is a Borel measure in both $(X, \widetilde{\mathsf{d}}_t)$ and (X, d_t) and the statement makes sense.

Fix $t \geq 0$, let $n \mapsto t_n \geq 0$ be any sequence converging to t, let the space X_n be a copy of X endowed with the distance $\widetilde{\mathsf{d}}_{t_n}$, the map $\iota_n : X \to X_n$ the corresponding "identity" map and $\mathfrak{m}_n := (\iota_n)_\sharp \mathfrak{m}$. We define the distance $\widetilde{\mathsf{D}}$ on $Y := \sqcup_n X_n \sqcup X$ by putting, for any $x, y \in X$,

$$
\widetilde{\mathsf{D}}(x, y) := \begin{cases} W_2(\mathsf{H}_{t_n}(\delta_x), \mathsf{H}_{t_m}(\delta_y)), & \text{if } x \in X_n, \ y \in X_m, \\ W_2(\mathsf{H}_{t_n}(\delta_x), \mathsf{H}_t(\delta_y)), & \text{if } x \in X_n, \ y \in X, \\ W_2(\mathsf{H}_t(\delta_x), \mathsf{H}_{t_m}(\delta_y)), & \text{if } x \in X, \ y \in X_m, \\ W_2(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y)), & \text{if } x, y \in X. \end{cases}
$$

Clearly, the embeddings of $(X, \widetilde{\mathsf{d}}_t)$ and $(X_n, \widetilde{\mathsf{d}}_n)$ in $(Y, \widetilde{\mathsf{D}})$ are isometries.

The transport plan $(\mathrm{Id}, \iota_n)_\sharp \mathfrak{m} \in \mathscr{P}_2(X \times X_n) \subset \mathscr{P}_2(Y^2)$ is admissible from $\mathfrak{m}$ to $\mathfrak{m}_n$, being Id the identity map, thus, we have

$$
W_2^{(Y, \widetilde{\mathsf{D}})}(\mathfrak{m}, \mathfrak{m}_n) \leq \sqrt{\int_{Y \times Y} \widetilde{\mathsf{D}}^2(x, y) \, \mathrm{d}(\mathrm{Id}, \mathrm{Id})_\sharp \mathfrak{m}(x, y)} = \sqrt{\int_X W_2^2(\mathsf{H}_t(x), \mathsf{H}_{t_n}(x)) \, \mathrm{d}\mathfrak{m}(x)}.
$$

The compactness of (X, d) ensures that $W_2(\mathsf{H}_t(x), \mathsf{H}_{t_n}(x))$ is uniformly bounded by the diameter of X, while from the continuity of the curve $s \mapsto \mathsf{H}_s(\mu)$, for any $\mu \in \mathscr{P}_2(X)$ (Corollary 5.6), we have that $W_2^2(\mathsf{H}_t(x), \mathsf{H}_{t_n}(x))$ goes to 0 as $n \to \infty$. Hence, the dominate convergence theorem implies

$$
\lim_{n \to \infty} W_2^{(Y, \widetilde{\mathsf{D}})}(\mathfrak{m}, \mathfrak{m}_n) = 0,
$$

which is the first claim.

Concerning the second claim, we start noticing that from the semigroup properties of H_t, for any $t, h \geq 0$ we get

$$
\begin{aligned}
\widetilde{\mathsf{d}}_{t+h}(x, y) &= W_2\big(\mathsf{H}_{t+h}(\delta_x), \mathsf{H}_{t+h}(\delta_y)\big) \\
&= W_2\big(\mathsf{H}_h(\mathsf{H}_t(\delta_x)), \mathsf{H}_h(H_t(\delta_y))\big) \\
&\leq e^{-Kh} W_2\big(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y)\big) \\
&= e^{-Kh} \widetilde{\mathsf{d}}_t(x, y),
\end{aligned}
$$

by means of inequality (5.3). Therefore for any d–Lipschitz curve $s \mapsto \gamma_s$ it holds $|\dot\gamma_s|_{t+h} \leq e^{-Kh}|\dot\gamma_s|_t$ for a.e. s, hence directly from the definition we get

$$\mathsf{d}_{t+h}(x,y) \leq e^{-Kh}\mathsf{d}_t(x,y). \tag{5.10}$$

Now, fix $t \geq 0$, a sequence $t_n \downarrow t$ and $\varepsilon > 0$. We use the definition of d_{t_n} to find $\widetilde{\mathsf{d}}_t$–Lipschitz curves $[0,1] \ni s \mapsto \gamma_{n,s}$, joining x to y, such that

$$\int_0^1 |\dot\gamma_s|_t \, \mathrm{d}s \leq \mathsf{d}_{t_n}(x,y) + \varepsilon. \tag{5.11}$$

From the compactness of (X,d) and the inequality (5.7) we can assume, with a reparametrization argument, that the curves $s \mapsto \gamma_{n,s}$ are L–Lipschitz w.r.t. d for some constant L independent of n. This equi–Lipschitz continuity and the compactness of X imply that there exists a subsequence, not relabeled, and a limit curve γ_s, which is L–Lipschitz w.r.t. d and such that $\lim_n \mathsf{d}(\gamma_{n,s}, \gamma_s) = 0$ for any $s \in [0,1]$.

Use identity (2.3) to find $\overline{N} \in \mathbb{N}$ and a partition $0 = \overline{s}_0 < \ldots < \overline{s}_{\overline{N}} = 1$ of $[0,1]$ such that

$$\int_0^1 |\dot\gamma_s|_t \, \mathrm{d}s \leq \sum_{i=0}^{\overline{N}-1} \widetilde{\mathsf{d}}_t(\gamma_{\overline{s}_i}, \gamma_{\overline{s}_{i+1}}) + \varepsilon,$$

By the definition of d_t we get

$$\mathsf{d}_t(x,y) \leq \int_0^1 |\dot\gamma_s|_t \, \mathrm{d}s \leq \sum_{i=0}^{\overline{N}-1} \widetilde{\mathsf{d}}_t(\gamma_{\overline{s}_i}, \gamma_{\overline{s}_{i+1}}) + \varepsilon. \tag{5.12}$$

Since $\gamma_{n,s} \to \gamma_s$ as $n \to \infty$ for any $s \in [0,1]$, taking into account the continuity of $t \mapsto \mathsf{H}_t(\mu)$, for arbitrary $\mu \in \mathscr{P}_2(X)$, we have that $\lim_n \widetilde{\mathsf{d}}_{t_n}(x,y) = \widetilde{\mathsf{d}}_t(x,y)$ for any $x,y \in X$ and thus

$$\sum_{i=0}^{\overline{N}-1} \widetilde{\mathsf{d}}_t(\gamma_{\overline{s}_i}, \gamma_{\overline{s}_{i+1}}) = \lim_{n \to \infty} \sum_{i=0}^{\overline{N}-1} \widetilde{\mathsf{d}}_{t_n}(\gamma_{\overline{s}_i}, \gamma_{\overline{s}_{i+1}}).$$

Therefore, by inequalities (5.11) and (5.12), we deduce

$$\mathsf{d}_t(x,y) \leq \varliminf_{n \to \infty} \mathsf{d}_{t_n}(x,y) + 2\varepsilon. \tag{5.13}$$

Hence, letting $\varepsilon \downarrow 0$, by inequality (5.10), we conclude

$$\mathsf{d}_t(x,y) = \lim_{n \to \infty} \mathsf{d}_{t_n}(x,y), \qquad \forall x,y \in X.$$

The proof of $\mathbb{D}$–convergence of $(X, \mathsf{d}_{t_n}, \mathfrak{m})$ to $(X, \mathsf{d}_t, \mathfrak{m})$ follows along the same lines, using the dominate convergence theorem and the fact that the spaces $(X, \mathsf{d}_{t_n}, \mathfrak{m})$ are – by estimate (5.7) – uniformly bounded. We omit the details. $\square$

We now discuss the stability properties of this flow of (pseudo) distances w.r.t. pointed $\mathbb{D}$–convergence. Shortly said, $\widetilde{\mathsf{d}}_t$ is continuous and d_t is lower semicontinuous under this convergence. We will denote be $\widetilde{\mathsf{d}}_{n,t}, \mathsf{d}_{n,t}$ such pseudo distances for the space $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$, according to Definitions 5.12, 5.13.

Theorem 5.19 (Stability). *Let $(X_n, \mathsf{d}_n, \mathfrak{m}_n, \overline{x}_n)$, for $n \in \mathbb{N}$, be a sequence of pointed $RCD(K, \infty)$ metric measure spaces, converging to some $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$ in the pointed $\mathbb{D}$–sense, as in Definition 5.8. Assume that $\mathrm{supp}(\mathfrak{m}_n) = X_n$ for every $n \in \mathbb{N}$ and that $\mathrm{supp}(\mathfrak{m}) = X$.*

Let (Y, D) as in Definition 5.8 and let $n \mapsto x_n, y_n \in \mathrm{supp}(\mathfrak{m}_n)$ for $x, y \in \mathrm{supp}(\mathfrak{m})$ such that $\mathsf{D}(x_n, x) \to 0$ and $\mathsf{D}(y_n, y) \to 0$ as $n \to \infty$.

Then, for any $t \geq 0$ there holds

$$\widetilde{\mathsf{d}}_t(x, y) = \lim_{n \to \infty} \widetilde{\mathsf{d}}_{n,t}(x_n, y_n). \tag{5.14}$$

Furthermore, if the bounded closed sets in (Y, D) are compact, then it also holds

$$\mathsf{d}_t(x, y) \leq \varliminf_{n \to \infty} \mathsf{d}_{n,t}(x_n, y_n), \tag{5.15}$$

for any $t \geq 0$, with x_n, y_n, x, y as in formula (5.14).

Finally, if (Y, D) is compact, the measures $\mathfrak{m}_n$ are normalized and for some $t > 0$ the heat flows $\mathsf{H}_{n,t} : \mathscr{P}_2(X_n) \to \mathscr{P}_2(X_n)$, $n \in \mathbb{N}$ and $\mathsf{H}_t : \mathscr{P}_2(X) \to \mathscr{P}_2(X)$ are all injective, then $n \mapsto (X_n, \widetilde{\mathsf{d}}_{n,t}, \mathfrak{m}_n)$ converges to $(X, \widetilde{\mathsf{d}}_t, \mathfrak{m})$ in the $\mathbb{D}$–sense.

Proof. We have

$$\left| \widetilde{\mathsf{d}}_{n,t}(x_n, y_n) - \widetilde{\mathsf{d}}_t(x, y) \right|$$

$$= \left| W_2^{(X_n, \mathsf{d}_n)} \big(\mathsf{H}_{n,t}(\delta_{x_n}), \mathsf{H}_{n,t}(\delta_{y_n}) \big) - W_2^{(X, \mathsf{d})} \big(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y) \big) \right|$$

$$= \left| W_2^{(Y, \mathsf{D})} \big(\mathsf{H}_{n,t}(\delta_{x_n}), \mathsf{H}_{n,t}(\delta_{y_n}) \big) - W_2^{(Y, \mathsf{D})} \big(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y) \big) \right|$$

$$\leq W_2^{(Y, \mathsf{D})} \big(\mathsf{H}_{n,t}(\delta_{x_n}), \mathsf{H}_t(\delta_x) \big) + W_2^{(Y, \mathsf{D})} \big(\mathsf{H}_{n,t}(\delta_{y_n}), \mathsf{H}_t(\delta_y) \big).$$

Thus, limit (5.14) follows from Proposition 5.11.

Inequality (5.15) then follows by the general lower semicontinuity of the length w.r.t. convergence of metric spaces, along the same line of the proof of Theorem 5.18: we just use the stability result (5.14) in place of the continuity of $t \mapsto \mathsf{H}_t(\mu)$ when passing to the limit at the level of $\widetilde{\mathsf{d}}_{n,t}$.

About the last statement, we define the function $\mathsf{D}_t : Y^2 \to [0, +\infty)$ by putting

$$\mathsf{D}_t(x, y) := \begin{cases} W_2^{(Y, \mathsf{D})}(\mathsf{H}_{n,t}(\delta_x), \mathsf{H}_{m,t}(\delta_y)), & \text{if } x \in X_n, \ y \in X_m, \\ W_2^{(Y, \mathsf{D})}(\mathsf{H}_{n,t}(\delta_x), \mathsf{H}_t(\delta_y)), & \text{if } x \in X_n, \ y \in X, \\ W_2^{(Y, \mathsf{D})}(\mathsf{H}_t(\delta_x), \mathsf{H}_{n,t}(\delta_y)), & \text{if } x \in X, \ y \in X_n, \\ W_2^{(Y, \mathsf{D})}(\mathsf{H}_t(\delta_x), \mathsf{H}_t(\delta_y)), & \text{if } x, y \in X. \end{cases}$$

Notice that D_t is a distance because D is a distance and by the injectivity of the flows. Moreover, by Theorem 5.15 and Proposition 5.11 we have that D_t induces the same topology of D.

For every $n \in \mathbb{N}$, let $\gamma_n \in \mathscr{P}(Y^2)$ be a transport plan realizing the minimum in

$$\int_{Y \times Y} \mathsf{d}_n^2(x, y) \, \mathrm{d}\gamma,$$

among all $\boldsymbol{\gamma}$ such that $\pi^1_{\sharp}\boldsymbol{\gamma} = \widetilde{\mathfrak{m}}$, $\pi^2_{\sharp}\boldsymbol{\gamma} = \widetilde{\mathfrak{m}}_n$. Then, for every $n \in \mathbb{N}$ we use a gluing argument to find $\boldsymbol{\alpha}_n \in \mathscr{P}(Y^{n+1})$ such that,

$$(\pi^0, \pi^i)_{\sharp}\boldsymbol{\alpha}_n = \boldsymbol{\gamma}_n, \qquad \forall i = 1, \ldots, n$$

and finally, we use Kolmogorov theorem to find $\boldsymbol{\alpha} \in \mathscr{P}(Y^{\mathbb{N}})$ such that

$$(\pi^0, \ldots, \pi^n)_{\sharp}\boldsymbol{\alpha} = \boldsymbol{\alpha}_n, \qquad \forall n \in \mathbb{N}.$$

Let $f_n : Y^{\mathbb{N}} \to [0, +\infty)$ be given by

$$f_n\big((x_n)\big) := \mathsf{D}(x_0, x_n),$$

by construction there holds

$$\int_{Y^{\mathbb{N}}} f_n^2 \,\mathrm{d}\boldsymbol{\alpha} = \int_{Y \times Y} \mathsf{D}^2(x, y) \,\mathrm{d}\boldsymbol{\gamma}_n(x, y) \to 0,$$

by our assumption of pointed $\mathbb{D}$–convergence. Therefore, up to a subsequence, not relabeled, we have that

$$\boldsymbol{\alpha} - a.e.\ (x_n) \text{ there holds } \lim_{n \to \infty} \mathsf{D}(x_n, x_0) = 0. \qquad (5.16)$$

Define now $g_n : Y^{\mathbb{N}} \to [0, +\infty)$ by

$$g_n\big((x_n)\big) := \mathsf{D}_t(x_0, x_n).$$

and notice that thanks to inequality 5.5 and the fact that (Y, D) is compact, the g_n are uniformly bounded. Using Proposition 5.11 and (5.16) we deduce that for $\boldsymbol{\alpha}$–a.e. x_n there holds $g_n\big((x_n)\big) \to 0$ as $n \to \infty$. Hence the dominate convergence theorem yields

$$W_2^{(Y, \mathsf{D}_t)}(\widetilde{\mathfrak{m}}_n, \widetilde{\mathfrak{m}}) \leq \sqrt{\int_{Y^{\mathbb{N}}} g_n^2 \,\mathrm{d}\boldsymbol{\alpha}} \to 0.$$

Being this result independent of the subsequence chosen, it holds for the full original sequence and the proof is completed. $\qquad\square$

We conclude with some comments about the statement and proof of this theorem. A sufficient condition in order to have that the bounded closed sets in (Y, D) are compact, is that the spaces $(X_n, \mathsf{d}_n, \mathfrak{m}_n)$ are uniformly doubling, in the sense that for some constant $C > 0$ there holds

$$\mathfrak{m}_n(B_{2R}(x)) \leq C\mathfrak{m}_n(B_R(x)), \qquad \forall n \in \mathbb{N},\ x \in X_n,\ R > 0.$$

Indeed, this assumption passes to the limit, on doubling spaces the support of the measure is the whole space and uniformly doubling spaces are uniformly totally bounded.

In the last part of the statement (as well as in Theorem 5.18), there are some hidden non–trivial technical problems. First of all we notice that the only way to get $\mathbb{D}$–convergence is to use the dominate convergence theorem – as we did: this is due to the fact that Proposition 5.11 is not quantitative, hence without a uniform

bound on D_t it seems hard to get the desired W_2–convergence. As soon as (Y, D) is bounded, we can argue as in the proof and obtain

$$\lim_{n \to \infty} \int_{Y^{\mathbb{N}}} \mathsf{D}_t(x_0, x_n) \, \mathrm{d}\boldsymbol{\alpha}\big((x_n)\big) = 0.$$

Yet, this is not enough to conclude that $W_2^{(Y,\mathsf{D}_t)}(\widetilde{\mathfrak{m}}, \widetilde{\mathfrak{m}}_n) \to 0$ because we do not know if $\boldsymbol{\alpha}$ is a Borel transport plan in $Y^{\mathbb{N}}$ when on Y we consider the Borel structure given by D_t. Actually, it is not even clear whether on general pointed $RCD(K, \infty)$ spaces $(X, \mathsf{d}, \mathfrak{m}, \overline{x})$, the measure $\widetilde{\mathfrak{m}}$ defined as in point (iii) of Definition 5.8 is Borel w.r.t. any of the pseudo distances $\widetilde{\mathsf{d}}_t, \mathsf{d}_t$, so the transport problem does not really make sense, at least in classical terms (of course this measure as well as the cost function are Borel w.r.t. the original distance d, but these are not the terms under which the W_2–distance is defined). It is for this reason that we added some assumptions granting that the topology – and a fortiori the Borel structures – of (Y, D) and (Y, D_t) coincide.

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, A. Mondino, and G. Savaré, *On the notion of convergence of non-compact metric measure spaces and applications*, preprint, (2012).

[3] L. Ambrosio, N. Gigli, A. Mondino, R. Tapio, and G. Savaré, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure*, preprint, (2012).

[4] 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, 2005.

[5] L. Ambrosio, N. Gigli, and G. Savaré, *Calculus and heat flows in metric measure spaces with Ricci curvature bounded from below*, Submitted paper, arXiv:1106.2090, (2011).

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

[7] ———, *Heat flow and calculus on metric measure spaces with Ricci curvature bounded below - the compact case*, Accepted for publication at Bull. UMI, (2012).

[8] N. Gigli, *On the heat flow on metric measure spaces: existence, uniqueness and stability*, Calc. Var. PDE, 39 (2010), pp. 101–120.

[9] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations*, Springer, Berlin, 2nd ed., 1983.

[10] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), 169 (2009), pp. 903–991.

[11] R. McCann and P. Topping, *Ricci flow, entropy and optimal transportation*, Amer. J. of Math., 132 (2010), pp. 711–730.

[12] S.-i. Ohta and K.-T. Sturm, *Non-contraction of heat flow on Minkowski spaces*, to appear in Arch. Ration. Mech. Anal., arXiv:1009.2312, (2010).

[13] F. Otto and M. Westdickenberg, *Eulerian calculus for the contraction in the Wasserstein distance*, SIAM J. Math. Anal., 37 (2005), pp. 1227–1255 (electronic).

[14] K. T. Sturm, *Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality*, J. Math. Pures Appl. (9), 75 (1996), pp. 273–297.

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

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

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

(Nicola Gigli) Université de Nice, Park Valrose, 06108 Nice, France
E-mail address, N. Gigli: `gigli@unice.fr`

(Carlo Mantegazza) Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, 56126 Pisa, Italy
E-mail address, C. Mantegazza: `c.mantegazza@sns.it`