Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVIII (2018), 845-880

Equivalence between dimensional contractions in Wasserstein distance and the curvature-dimension condition

FRANÇOIS BOLLEY, IVAN GENTIL, ARNAUD GUILLIN
AND KAZUMASA KUWADA

Abstract. The curvature-dimension condition is a generalization of the Bochner inequality to weighted Riemannian manifolds and general metric measure spaces. It is now known to be equivalent to evolution variational inequalities for the heat semigroup, and quadratic Wasserstein distance contraction properties at different times. On the other hand, in a compact Riemannian manifold, it implies a same-time Wasserstein contraction property for this semigroup. In this work we generalize the latter result to metric measure spaces and more importantly prove the converse: contraction inequalities are equivalent to curvature-dimension conditions. Links with functional inequalities are also investigated.

Mathematics Subject Classification (2010): 58J65 (primary); 58J35, 53B21 (secondary).

1. Introduction

The von Renesse-Sturm theorem (see [27]) ensures that a Wasserstein distance contraction property between solutions to the heat equation on a Riemannian manifold is equivalent to a lower curvature condition. This result is one of the first equivalence results relating the Wasserstein distance and a curvature condition. Recent works have been devoted to a more precise curvature-dimension condition instead of a sole curvature condition. In this work, and in a fairly general framework, we derive new *dimensional* contraction properties under a curvature-dimension condition and we show that they are all equivalent to it.

Let Δ be the Laplace-Beltrami operator on a smooth Riemannian manifold $(\mathbf{M}, \mathcal{G})$ and let $(P_t h)_{t\geqslant 0}$ be the solution to the heat equation $\partial_t u = \Delta u$ with h as the initial condition. Many of the coming notions and results have been considered in a more general setting, but for simplicity in the introduction we focus on this case.

This research was supported by the French ANR-12-BS01-0019 STAB project and JSPS Grant-in-Aid for Young Scientist (A) 26707004.

Received November 16, 2016; accepted in revised form February 15, 2017.
Published online May 2018.

The Bochner identity states that

$$\frac{1}{2}\Delta|\nabla f|^2 - \nabla f\cdot\nabla\Delta f = |\nabla\nabla f|^2 + \mathrm{Ric}(\nabla f,\nabla f),$$

where Ric is the Ricci curvature of $(\mathbf{M},\mathcal{G})$. The manifold associated with its Laplacian is said to satisfy the $CD(R,m)$ curvature-dimension condition if its Ricci curvature is uniformly bounded from below by $R\in\mathbb{R}$ and its dimension is smaller than $m\in(0,+\infty]$. In this case

$$\frac{1}{2}\Delta|\nabla f|^2 - \nabla f\cdot\nabla\Delta f \geqslant \frac{1}{m}(\Delta f)^2 + R|\nabla f|^2 \tag{1.1}$$

by the Cauchy-Schwarz inequality. The $CD(R,m)$ condition and (1.1) are the starting point of many comparison theorems, functional and geometrical inequalities, bounds on the heat kernel, etc. (see, *e.g.*, [8,13,26,28]).

In this work we focus on the link between the curvature-dimension condition and Wasserstein distance contraction properties of the heat semigroup. The von Renesse-Sturm theorem [27] states that: the $CD(R,\infty)$ condition holds if and only if

$$W_2^2(P_t g dx, P_t h dx) \le e^{-2Rt} W_2^2(g dx, h dx) \tag{1.2}$$

for all $t\geqslant 0$ and probability densities g,h with respect to the Riemannian measure dx. Here W_2 is the Wasserstein distance with quadratic cost.

There are many proofs of this result as well as extensions to more general evolutions and spaces, see for instance [2,8,9,15,17,23,28,29]. Following the seminal papers [21,25], attention has been drawn to taking the *dimension* of the manifold into account.

A first way of including the dimension is to use *two different times s and t* in the inequality (1.2). It is proved in [9,18] that the $CD(0,m)$ condition implies

$$W_2^2(P_s g dx, P_t h dx) \le W_2^2(g dx, h dx) + 2m\big(\sqrt{t}-\sqrt{s}\big)^2 \tag{1.3}$$

for all $s,t\geqslant 0$ and all probability densities g,h. A non zero lower bound on the curvature and the equivalence have been further considered in [13,18]:

- In [18], the fourth author proved that the $CD(R,m)$ condition holds if and only if

$$W_2^2(P_t g dx, P_s h dx) \le A(s,t,R,m) W_2^2(g dx, h dx) + B(s,t,m,R) \tag{1.4}$$

  for all $s,t\geqslant 0$ and all probability densities g,h, and for appropriate positive functions A,B;
- In [13], the authors proved that the $CD(R,m)$ condition holds if and only if

$$\begin{aligned} s_{\frac{R}{m}}\left(\frac{1}{2}W_2(P_t g dx, P_s h dx)\right)^2 \le e^{-R(t+s)}\, s_{\frac{R}{m}}\left(\frac{1}{2}W_2(g dx, h dx)\right)^2 \\ + \frac{m}{R}\left(1-e^{-R(s+t)}\right)\frac{\big(\sqrt{t}-\sqrt{s}\big)^2}{2(t+s)} \end{aligned} \tag{1.5}$$

for all $s, t \geqslant 0$ and all probability densities g, h. Here $s_r(x) = \sin(\sqrt{r}x)/\sqrt{r}$ if $r > 0$, $s_r(x) = \sinh(\sqrt{|r|}x)/\sqrt{|r|}$ if $r < 0$ and $s_0(x) = x$, hence recovering (1.3) when $R = 0$. Both inequalities (1.4) and (1.5) are extensions of (1.2) and (1.3), taking the dimension into account.

Contraction properties with the *same time* have been derived in [11] for the Euclidean heat equation in $\mathbb{R}^m$, and then extended by the third author in [14] to a compact Riemannian manifold. Let $\mathrm{Ent}_{dx} h = \int h \log h \, dx$ be the entropy of a probability density h. Then the $CD(R, m)$ condition implies that

$$W_2^2(P_t g dx, P_t h dx) \leq e^{-2Rt} W_2^2(g dx, h dx) - \frac{2}{m} \int_0^t e^{-2R(t-u)} (\mathrm{Ent}_{dx} P_u g - \mathrm{Ent}_{dx} P_u h)^2 du$$

for all $t \geqslant 0$ and all probability densities g, h. This bound has also been proved in [11] for the Markov transportation distance instead of the W_2 distance. This distance differs from W_2 and has actually been tailored to Markov semigroups and the Bakry-Émery Γ_2 calculus. Dimensional contraction properties for a Wasserstein distance defined with an adapted cost have also been derived in [29].

In this paper we derive diverse *same time* contraction inequalities under a general $CD(R, m)$ curvature-dimension condition, and in fact prove that they are all *equivalent* to this condition. The results and the proofs will be given in the two settings of a smooth Riemannian manifold and of a more general Riemannian energy measure space, which is introduced in [6] and closely related to the so-called RCD metric measure spaces (see [5] and also [1,13]).

The paper is organized as follows. In Section 2, we state and explain the context of our main result, Theorem 2.1. In Section 3, we present the strategy of our proof, motivated by the elementary gradient flow approach in Euclidean space. The main issue, from the weakest contraction to the curvature-dimension condition, is proved on a Riemannian manifold in Section 4, and on a Riemannian energy measure space in Section 5. The general strategy is the same in both settings, and it could seem redundant to give both proofs. However the proof in the Riemannian setting is rather simpler, presents the most important steps of the argument and thus gives a way to get it in a more general space. We believe that it is an opportunity to emphasize, in our example, the main issues arising in transferring a proof in the Riemannian setting to the abstract measure space setting. Indeed, there, regularity is no more available "for free", and our proof will crucially use a whole panel of powerful tools developed by L. Ambrosio, N. Gigli, G. Savaré, K.-T. Sturm and coauthors to overcome this difficulty, in particular localization and mollification by semigroup.

The easier implications in Theorem 2.1 are directly proved on a Riemannian energy measure space in Section 5. The last section gives a new and simple derivation of a classical entropy-energy inequality, as well as dimensional HWI inequalities: for this we start from our contraction inequalities instead of the curvature-dimension condition, as in earlier works.

ACKNOWLEDGEMENTS. The authors are grateful to the referees for a careful reading of the manuscript and helpful comments and questions which improved the paper.

2. Main result

Our main theorem states that, in a quite general framework, a curvature-dimension condition is equivalent to same time Wasserstein distance contraction inequalities.

Let $(\mathbf{X}, d)$ be a Polish metric space, $\mathcal{P}(\mathbf{X})$ be the set of Borel probability measures on $\mathbf{X}$ and $\mathcal{P}_2(\mathbf{X})$ be the set of all $\mu \in \mathcal{P}(\mathbf{X})$ such that $\int d(x_0, x)^2 \, d\mu(x) < \infty$ for some $x_0 \in \mathbf{X}$. The (quadratic) Wasserstein distance between ν_1 and ν_2 in $\mathcal{P}_2(\mathbf{X})$ is defined by

$$W_2(\nu_1, \nu_2) = \inf_{\pi} \sqrt{\iint d(x, y)^2 \, d\pi(x, y)},$$

where the infimum runs over all probability measures π on $\mathbf{X} \times \mathbf{X}$ with marginals ν_1 and ν_2.

A fundamental tool is the Kantorovich dual representation: for $\nu_1, \nu_2 \in \mathcal{P}_2(\mathbf{X})$,

$$\frac{W_2^2(\nu_1, \nu_2)}{2} = \sup_{\psi} \left\{ \int Q\psi \, d\nu_1 - \int \psi \, d\nu_2 \right\}. \tag{2.1}$$

Here the supremum runs over all bounded Lipschitz functions ψ (in this case [26, Theorem 5.10] can be extended to Lipschitz instead of continuous functions, see [17, Remark 3.6]) and $Q\psi$ is the inf-convolution of ψ, defined on $\mathbf{X}$ by

$$Q\psi(x) = \inf_{y \in \mathbf{X}} \left\{ \psi(y) + \frac{d(x, y)^2}{2} \right\}.$$

The Wasserstein space $(\mathcal{P}_2(\mathbf{X}), W_2)$ is described in the reference books [2] and [26]. We shall define the entropy $\mathrm{Ent}_\mu f$ of a probability density f with respect to a (finite or not) measure μ by $\mathrm{Ent}_\mu f = \int f \log f \, d\mu$ if $f(\log f)_+ \in \mathbb{L}^1(\mu)$ and ∞ otherwise.

Our result will be stated in the two settings of a Riemannian Markov triple $(\mathbf{M}, \mu, \Gamma)$ (RMT in short), and a Riemannian energy measure space $(\mathbf{X}, \tau, \mu, \mathcal{E})$ (REM in short). These settings will be described in detail in Sections 4 and 5 respectively. A REM space is a particular metric measure space, developed in [6]. A RMT is a smooth Riemannian manifold equipped with a weighted Laplacian (see [8]) and is a particular example of REM space.

Even if a RMT is a REM space we prefer to state and prove our result in both settings since the argument is a little simpler in the Riemannian case. We also believe that it emphasizes the main difficulties when generalizing a result from a smooth setting to an abstract metric measure space. In both spaces, $(P_t)_{t \geqslant 0}$ denotes the associated Markov semigroup. It is defined through the weighted Laplacian in the RMT case, and through the Dirichlet form in the REM case.

The $CD(R,m)$ curvature-dimension condition is defined using the Bochner inequality (1.1) in a Riemannian manifold and in a weak form in a metric measure space (see Definitions 4.1 and 5.1).

Recall finally that for $r \in \mathbb{R}$ the map s_r is defined on $\mathbb{R}$ by

$$s_r(x) = \begin{cases} \sin(\sqrt{r}\,x)/\sqrt{r} & \text{if } r > 0 \\ \sinh(\sqrt{|r|}\,x)/\sqrt{|r|} & \text{if } r < 0 \\ x & \text{if } r = 0. \end{cases}$$

Theorem 2.1 (Equivalence between contractions and $CD(R,m)$ condition). *Consider a RMT or REM space as in Sections 4 and 5, with (finite or not) reference measure μ and associated semigroup $(P_t)_{t\geqslant 0}$. Let $R \in \mathbb{R}$ and $m > 0$. Then the following properties are equivalent:*

(i) *The $CD(R,m)$ (or weak $CD(R,m)$ in a REM space) curvature-dimension condition holds;*

(ii) *For any $t \geqslant 0$ and any probability densities g, h with respect to μ, there holds*

$$\begin{aligned} s_{\frac{R}{m}}\left(\frac{1}{2}W_2(P_t g\mu, P_t h\mu)\right)^2 &\le e^{-2Rt}\, s_{\frac{R}{m}}\left(\frac{1}{2}W_2(g\mu, h\mu)\right)^2 \\ &\quad - 2m\int_0^t e^{-2R(t-u)}\sinh^2\left(\frac{\mathrm{Ent}_\mu P_u g - \mathrm{Ent}_\mu P_u h}{2m}\right)du; \end{aligned} \tag{2.2}$$

(iii) *For any $t \geqslant 0$ and any probability densities g, h with respect to μ,*

$$\begin{aligned} W_2^2(P_t g\mu, P_t h\mu) &\le e^{-2Rt}\, W_2^2(g\mu, h\mu) \\ &\quad - \frac{2}{m}\int_0^t e^{-2R(t-u)}\left(\mathrm{Ent}_\mu P_u g - \mathrm{Ent}_\mu P_u h\right)^2 du. \end{aligned} \tag{2.3}$$

See Theorems 4.3 and 5.3 for a more precise framework of Theorem 2.1.

A bound with the same additional term as in (ii) has also been derived in [10] for some specific instances of symmetric Fokker-Planck equations in $\mathbb{R}^m$, for which the generator only satisfies a $CD(R,\infty)$ condition. Combined with a deficit in the Talagrand inequality, it has led to refined convergence estimates on the solutions.

The more difficult (iii) $\Rightarrow$ (i) is proved in both RMT and REM spaces, in Sections 4 and 5 respectively. The easier (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) are directly proved on a REM space in Section 5.

3. Strategy of the proofs

3.1. Example of a gradient flow in $\mathbb{R}^d$

Let us first present the easiest case of a smooth gradient flow in $\mathbb{R}^d$. There we shall see that the equivalence between the contraction inequality (2.3) and the $CD(R,m)$

curvature-dimension condition is natural. It gives a way to understand the general case.

Let $F : \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{C}^2$ smooth function, and let $(X_t)_{t \geqslant 0}$ be a gradient flow for the function F, that is, a solution to the differential equation

$$\frac{dX_t}{dt} = -\nabla F(X_t). \tag{3.1}$$

Following [13], the function F satisfies a $CD(R, m)$ curvature-dimension condition for $R \in \mathbb{R}$ and $m > 0$ if for any $x, h \in \mathbb{R}^d$, the map $[0, 1] \ni s \mapsto \varphi(s) = F(x+sh)$ satisfies the convexity inequality

$$\varphi''(s) \geqslant R||h||^2 + \frac{1}{m}\big(\varphi'(s)\big)^2. \tag{3.2}$$

Here $|| \cdot ||$ is the Euclidean norm in $\mathbb{R}^d$. Since the path $(x + sh)_{s\in[0,1]}$ is a geodesic between x and $x+h$, this means that F satisfies a (R, m)-convexity condition along geodesics.

Let now $(X_t)_{t\geqslant 0}$ and $(Y_t)_{t\geqslant 0}$ be two solutions to (3.1) with initial conditions X_0 and Y_0 respectively. Let also $\varphi_t(s) = F(X_t + s(Y_t - X_t))$, so that $\varphi_t'(s) = \nabla F(X_t + s(Y_t - X_t)) \cdot (Y_t - X_t)$. Then the function $\Lambda(t) = ||X_t - Y_t||^2$ satisfies

$$\Lambda'(u) = -2(X_u - Y_u) \cdot (\nabla F(X_u) - \nabla F(Y_u)) = -2 \int_0^1 \varphi_u''(s)ds.$$

If now the function F satisfies the above $CD(R, m)$ condition (3.2), then

$$\Lambda'(u) \leq -2R||X_u - Y_u||^2 - \frac{2}{m}\int_0^1 \big(\varphi_u'(s)\big)^2 du \leq -2R\Lambda(u) - \frac{2}{m}(\varphi_u(1) - \varphi_u(0))^2$$

by the Cauchy-Schwarz inequality. Integrating over the interval $[0, t]$, we get

$$||X_t - Y_t||^2 \leq e^{-2Rt}||X_0 - Y_0||^2 - \frac{2}{m}\int_0^t e^{-2R(t-u)}(F(X_u) - F(Y_u))^2 du. \tag{3.3}$$

Conversely, let us assume that the gradient flow driven by F satisfies the property (3.3) for any $t \geqslant 0$ and any initial conditions X_0 and Y_0. Then F satisfies the $CD(R, m)$ condition (3.2). For, taking the time derivative of (3.3) at $t = 0$ implies

$$-(X_0 - Y_0) \cdot (\nabla F(X_0) - \nabla F(Y_0)) \leq -R||X_0 - Y_0||^2 - \frac{1}{m}(F(X_0) - F(Y_0))^2.$$

Let then x, h in $\mathbb{R}^d$ and $s \in [0, 1]$ be fixed. A Taylor expansion for $Y_0 = x+(s+\varepsilon)h$ tending to $X_0 = x + sh$ (along a geodesic), so for $\varepsilon \to 0$, gives

$$-h \cdot \nabla^2 F(x + sh)h \leq -R||h||^2 - \frac{1}{m}(\nabla F(x + sh) \cdot h)^2.$$

This is exactly the $CD(R, m)$ condition (3.2).

Let us observe that inequality (3.3) is exactly (2.3) when replacing $\mathbb{R}^d$ with the space of probability densities, the Euclidean norm with the Wasserstein distance, F with the entropy, $(X_t)_{t\geqslant 0}$ with the semigroup $(P_t)_{t\geqslant 0}$ and the $CD(R,m)$ condition (3.2) with the corresponding Bakry-Émery condition, which is equivalent to the (R,m)-convexity of the entropy (see [13]). Of course, this computation is natural since the considered evolution is the gradient flow of the entropy with respect to the Wasserstein distance, see [2,16].

We now want to mimic the above proof for a smooth gradient flow on $\mathbb{R}^d$ to the setting of a general semigroup on $(\mathcal{P}_2(\mathbf{X}), W_2)$. As here in the smooth case, we shall see in the coming section that geodesics play a fundamental role.

3.2. How to adapt the gradient flow proof to the general case?

The most natural method to prove that a contraction inequality in Wasserstein distance, as in (1.2), implies a curvature condition is to use close Dirac measures as initial data (see, *e.g.*, [9]). In our case, this can not be performed since the entropy of a Dirac measure is infinite. There seems to be hope since we consider the entropy of the heat kernel in positive time, when it becomes finite. However, it does not work again if we are on a homogeneous space. For instance, on $\mathbb{R}^d$, the entropy of the heat kernel $p_t(x,\cdot)$ does not depend on x and the dimensional corrective terms in Theorem 2.1 vanish if we consider two Dirac measures as initial data.

To solve this issue we shall consider as initial data a probability density g (with respect to μ) and a perturbation of it, both in sufficiently wide classes of functions. The perturbation will be built by means of a geodesic in the Wasserstein space $(\mathcal{P}_2(\mathbf{X}), W_2)$. Of course the best way would be to consider directly a geodesic in the Wasserstein space as it was first used in [16]. In our general setting of a RMT or a REM space, it is difficult to deal with such a geodesic due to the lack of regularity. That is why we use a "smooth" modification of a geodesic path. More precisely, given such a g, we are looking for a path $(g_s)_{s\geqslant 0}$ of probability densities whose Taylor expansion for small s is a geodesic in $\mathcal{P}_2(\mathbf{X})$ with a direction given by a function f. We explain the idea on a RMT.

For that, consider the generator $L^g = L + \Gamma(\log g,\cdot)$ (see (4.1) for the definition of Γ) with associated semigroup $(P_t^g)_{t\geqslant 0}$. Given a direction function f, there are two ways of defining the path $(g_s)_{s\geqslant 0}$, both admitting the same Taylor expansion for small s:

- One can first consider the path $g_s = g(1 - sL^g f)$ for small s and a smooth and compactly supported function f. The function g_s is a smooth, bounded and compactly supported perturbation of g. This path will be used on a RMT since such functions are adapted to the Riemannian setting;
- One can also consider the path $\tilde{g}_s = g(1 + f - P_s^g f)$, again for s small and "nice" $f \in \mathbb{L}^\infty(\mu)$. The path $(\tilde{g}_s)$ has the same Taylor expansion as (g_s) since $f - P_s^g f = -sL^g f + o(s)$. This path will be used on REM spaces. Indeed, regularity of functions (such as g_s above) is clearly a difficult issue in the setting of metric measure spaces, and $\mathbb{L}^\infty(\mu)$ functions are much more adapted to them.

By using the semigroup $(P_s^g)_{s\geq 0}$ instead of the generator L^g, we can apply the maximum principle which preserves (essential) boundedness of functions.

Remark 3.1. Let us see, formally and in the Euclidean space $\mathbb{R}^d$, why the probability measure $g_s dx$ has the same first-order Taylor expansion as the geodesic in the Wasserstein space. Let ν_0 be a probability measure in $\mathbb{R}^d$ being absolutely continuous with respect to the Lebesgue measure, $\psi : \mathbb{R}^d \to \mathbb{R}$ be a convex map, and

$$\nu_s = ((1-s)\mathrm{Id} + s\nabla\psi)_\# \nu_0$$

for $s \in [0,1]$. The path $(\nu_s)_{s\in[0,1]}$ is a geodesic path between ν_0 and ν_1 in the Wasserstein space, that is for any $s, t \in [0,1]$,

$$W_2(\nu_s, \nu_t) = |t-s|\, W_2(\nu_0, \nu_1).$$

Moreover, for any test function $H : \mathbb{R}^d \to \mathbb{R}$, and by a formal Taylor expansion when s goes to 0,

$$\begin{aligned}\int H d\nu_s &= \int H((1-s)x + s\nabla\psi(x)) d\nu_0(x)\\ &= \int [H(x) + s\nabla H(x)\cdot(\nabla\psi(x) - x) + o(s)] d\nu_0(x).\end{aligned}$$

Assume now that $d\nu_0 = g dx$ for a function g. Then, by integration by parts as in (4.4) below,

$$\int H d\nu_s = \int H d\nu_0 - s\int H\, L^g(f)\, d\nu_0 + o(s) = \int H\, g_s dx + o(s),$$

where $f(x) = \psi(x) - |x|^2/2$.

In conclusion, the path $(g_s)_{s\geqslant 0}$ appears as a (smooth) first-order Taylor expansion of the W_2-geodesic path $(\nu_s)_{s\geqslant 0}$.

Observe that in a general setting we cannot expect a sufficient level of smoothness of the Kantorovich potential ψ, even on a Riemannian manifold.

4. The Riemannian Markov triple context

In this section we prove the implication (iii) ⇒ (i) of Theorem 2.1 in the context of a Riemannian manifold, in the form of Theorem 4.3 below.

4.1. Framework and results

Let $(\mathbf{M}, \mathcal{G})$ be a connected complete $\mathcal{C}^\infty$-Riemannian manifold. Let V be a $\mathcal{C}^\infty$ function on $\mathbf{M}$ and consider the Markov semigroup $(P_t)_{t\geqslant 0}$ with generator $L = \Delta - \nabla V \cdot \nabla$, where Δ is the Laplace-Beltrami operator. Let also $d\mu = e^{-V} dx$

where dx is the Riemannian measure and Γ be the *carré du champ* operator, defined by

$$\Gamma(f,g)=\frac{1}{2}(L(fg)-fLg-gLf) \tag{4.1}$$

for any smooth f, g. We let $\Gamma(f)=\Gamma(f,f)=|\nabla f|^2$ where $|\nabla f|$ stands for the length of ∇f with respect to the Riemannian metric $\mathcal{G}$.

We assume that $(\mathbf{M},\mu,\Gamma)$ is a full Markov triple in a Riemannian manifold, as in [8, Chapter 3], and in this work we call it a *Riemannian Markov triple* (RMT). It has to be mentioned that we need an additional hypothesis to obtain a full Markov triple: the hypothesis proposed in [8, Chapter 3] is a uniform lower bound on the Ricci curvature of M plus the Hessian of the function V: there exists a constant $\rho\in\mathbb{R}$ such that $\mathrm{Ric}_{\mathcal{G}}+\nabla^2 V\geqslant\rho\mathrm{Id}$. A more general statement will be given in Section 5.

The measure μ is reversible with respect to the semigroup, that is, for any $t\geqslant 0$, P_t is a self-adjoint operator in $\mathbb{L}^2(\mu)$. Moreover the integration by parts formula

$$\int fLg\,d\mu=-\int\Gamma(f,g)d\mu$$

holds for all f, g in the set $\mathcal{C}_c^\infty(\mathbf{M})$ of infinitely differentiable and compactly supported functions on $\mathbf{M}$. The generator L satisfies the diffusion property, that is, for any smooth functions φ, f, g,

$$L(\varphi(f))=\varphi'(f)Lf+\varphi''(f)\Gamma(f),$$

or equivalently

$$\Gamma(\varphi(f),g)=\varphi'(f)\Gamma(f,g). \tag{4.2}$$

In other words, the carré du champ operator is a derivation operator for each component.

The map $(x,t)\mapsto P_th(x)$ is simply the solution to the parabolic equation $\partial_t u=Lu$ with h as the initial condition.

Definition 4.1 (**$CD(R,m)$ condition).** Let $R\in\mathbb{R}$ and $m\in(0,\infty]$. We say that the RMT $(\mathbf{M},\mu,\Gamma)$ satisfies a $CD(R,m)$ curvature-dimension condition if

$$\Gamma_2(f)\geqslant R\Gamma(f)+\frac{1}{m}(Lf)^2$$

for any smooth function f, say in $\mathcal{C}_c^\infty(\mathbf{M})$, where

$$\Gamma_2(f)=\frac{1}{2}(L\Gamma(f)-2\Gamma(f,Lf)). \tag{4.3}$$

Let us notice that m can be different from the dimension of the manifold $\mathbf{M}$. The $CD(R,m)$ curvature-dimension condition is called the Bakry-Émery or Γ_2 condition and has been introduced in [7] (see also the recent [8]).

Example 4.2. On a d-dimensional Riemannian manifold $(\mathbf{M}, \mathcal{G})$

- The operator $L = \Delta$ satisfies a $CD(R, m)$ condition if $m \geqslant d$ and the Ricci curvature of the manifold is bounded from below by R;
- More generally, the operator $L = \Delta - \nabla V \cdot \nabla$ satisfies a $CD(R, m)$ condition if $m \geqslant d$ and

$$\mathrm{Ric} + \mathrm{Hess}(V) \geqslant R\,\mathcal{G} + \frac{1}{m-d} \nabla V \otimes \nabla V,$$

  where Ric is the Ricci tensor of $(\mathbf{M}, \mathcal{G})$, see for instance [8, Section C6] (when $m = d$ then we need $V = 0$).

In a RMT, the following result gives the implication (iii) $\Rightarrow$ (i) in Theorem 2.1:

Theorem 4.3. *Let* $(\mathbf{M}, \mu, \Gamma)$ *be a Riemannian Markov triple and* $(P_t)_{t \geqslant 0}$ *its associated Markov semigroup. Let* $R \in \mathbb{R}$ *and* $m > 0$. *If the inequality* (2.3) *holds for any* $t \geqslant 0$ *and any smooth functions* g, h *on* $\mathbf{M}$ *with* $g\mu, h\mu$ *in* $\mathcal{P}_2(\mathbf{M})$, *then the* $CD(R, m)$ *condition of Definition* 4.1 *holds.*

4.2. Proof of Theorem 4.3

It is based on the approximation of geodesics introduced in Section 3.2 (see Remark 3.1), properties of the Hopf-Lax solution of the Hamilton-Jacobi equation, and an adapted class of test functions.

Let f be in $\mathcal{C}_c^\infty(\mathbf{M})$. Let also g be a smooth and positive function on $\mathbf{M}$ such that $g\mu \in \mathcal{P}_2(\mathbf{M})$,

$$\int g \,|\log g|\, d\mu < \infty \quad \text{and} \quad \int \frac{\Gamma(g)}{g} d\mu < \infty.$$

Let us define the generator L^g by

$$L^g h = Lh + \Gamma(\log g, h)$$

on smooth functions h. Since $g > 0$, then L^g is well defined on the set $\mathcal{C}_c^\infty(\mathbf{M})$ and $L^g h \in \mathcal{C}_c^\infty(\mathbf{M})$ for any $h \in \mathcal{C}_c^\infty(\mathbf{M})$. Moreover, the generator L^g satisfies an integration by parts formula with respect to the probability measure $g\mu$: for $h, k \in \mathcal{C}_c^\infty(\mathbf{M})$ (one of them can be with non compact support)

$$\int h\, L^g k \, g d\mu = -\int \Gamma(h, k)\, g d\mu. \tag{4.4}$$

For any $s \geqslant 0$, let us define $g_s = g(1 - sL^g f)$. The function $L^g f$ is in $\mathcal{C}_c^\infty(\mathbf{M})$, so bounded, and we can let $N = ||L^g f||_\infty$. We shall frequently use the bounds $(1 - sN)g \leq g_s \leq (1+sN)g$. In particular $g_s > 0$ for $s < 1/N$. Moreover $\int g_s d\mu = 1$. Hence, for s small enough, which we now assume, $g_s\mu$ is in $\mathcal{P}_2(\mathbf{M})$ with a smooth and positive density. The proof of Theorem 4.3 consists in applying (2.3) with g_s instead of f, dividing by $2s^2$ and letting s go to 0. For this we shall estimate the three terms in the inequality.

A key tool is the Hopf-Lax semigroup defined on bounded Lipschitz functions ψ by

$$Q_s\psi(x) := \inf_{y\in\mathbf{M}}\left\{\psi(y) + \frac{d(x,y)^2}{2s}\right\}, \quad s>0,\ x\in\mathbf{M}. \tag{4.5}$$

The map $x\mapsto Q_s\psi(x)$ is Lipschitz for every $s\geqslant 0$, and the map $(s,x)\mapsto Q_s\psi(x)$ satisfies the Hamilton-Jacobi equation

$$\partial_s Q_s\psi + \frac{1}{2}|\nabla Q_s\psi|^2 = 0, \quad \lim_{s\to 0} Q_s\psi = \psi$$

in a sense given in [26, Theorem 22.46 and 30.30] for instance. We observe that $sQ_s(\psi) = Q_1(s\psi) = Q(s\psi)$, so for $s>0$ the Kantorovich duality (2.1) can be written as

$$\frac{W_2^2(\nu_1,\nu_2)}{2s^2} = \frac{1}{s}\sup_{\psi}\left[\int Q_s\psi\, d\nu_1 - \int \psi\, d\nu_2\right]. \tag{4.6}$$

Estimate on the term on the left-hand side of (2.3). Letting $\psi = f$ in (4.6), we obtain

$$\frac{W_2^2(P_tg_s\mu, P_tg\mu)}{2s^2} \geqslant \int \frac{Q_sf\,P_tg_s - f\,P_tg}{s}d\mu. \tag{4.7}$$

Since f is Lipschitz, almost everywhere in $\mathbf{M}$ we have

$$\lim_{s\to 0}\frac{Q_sf\,P_tg_s - f\,P_tg}{s} = -\frac{1}{2}\Gamma(f)P_tg - fP_t(gL^gf)$$

by (vii') in [26, Theorem 30.30]. But, by the definition of Q_sf and since f is bounded,

$$Q_sf(x) = \inf_{y\in B(x,\sqrt{4s\|f\|_\infty})}\left\{f(y) + \frac{d(x,y)^2}{2s}\right\}.$$

Thus, for the Lipschitz seminorm $\|\cdot\|_{\mathrm{Lip}}$,

$$\begin{aligned}
0 &\geqslant \frac{Q_sf(x)-f(x)}{s}\\
&\geqslant \inf_{y\in B(x,\sqrt{4s\|f\|_\infty})\setminus\{x\}}\left\{\frac{f(y)-f(x)}{d(x,y)}\frac{d(x,y)}{s} + \frac{d(x,y)^2}{2s^2}\right\}\\
&\geqslant -\frac{1}{2}\sup_{y\in B(x,\sqrt{4s\|f\|_\infty})\setminus\{x\}}\left(\frac{f(y)-f(x)}{d(x,y)}\right)^2 \geqslant -\frac{1}{2}\|f\|_{\mathrm{Lip}}^2
\end{aligned} \tag{4.8}$$

(see also [26, page 585]). Moreover $||Q_s f||_\infty \le ||f||_\infty$, so, adding and subtracting $Q_s f P_t g$,

$$\left|\frac{Q_s f P_t g_s - f P_t g}{s}\right| \le ||Q_s f||_\infty \, |P_t(gL^g f)| + P_t g \, \frac{f - Q_s f}{s}$$
$$\le \left(||f||_\infty \, ||L^g f||_\infty + \frac{||f||^2_{\mathrm{Lip}}}{2}\right) P_t g.$$

The right-hand side is in $\mathbb{L}^1(\mu)$, so by the Lebesgue dominated convergence theorem

$$\liminf_{s\to 0} \frac{W_2^2(P_t g_s \mu, P_t g\mu)}{2s^2} \geqslant \int \left(-\frac{1}{2}\Gamma(f) P_t g - f P_t(gL^g f)\right) d\mu.$$

Now, by reversibility of the measure μ and the integration by parts formula (4.4),

$$\int f P_t(gL^g f) d\mu = \int P_t f L^g(f) \, g d\mu = -\int \Gamma(f, P_t f) \, g d\mu.$$

Thus we obtain our first estimate:

$$\liminf_{s\to 0} \frac{W_2^2(P_t g_s \mu, P_t g\mu)}{2s^2} \geqslant -\frac{1}{2}\int P_t(\Gamma(f)) g d\mu + \int \Gamma(f, P_t f) g d\mu. \quad (4.9)$$

Estimate on the first term on the right-hand side. According to (4.6) we need an upper bound on the quantities $\int Q_s(\psi) g_s d\mu - \int \psi g d\mu$, independent of the bounded Lipschitz function ψ.

First of all, for $0 < t < s$,

$$\frac{d}{dt}\int Q_t\psi \, g_t \, d\mu = \int \left[-\frac{1}{2}\Gamma(Q_t\psi)(1 - tL^g f) - Q_t\psi L^g f\right] g d\mu. \quad (4.10)$$

This is justified by item (vii) in [26, Theorem 22.46 and 30.30] and the properties that $g\mu \in \mathcal{P}(\mathbf{M})$, $L^g f$ is bounded, $||Q_t\psi||_\infty \le ||\psi||_\infty$ and $||Q_t\psi||_{\mathrm{Lip}} \le ||\psi||_{\mathrm{Lip}}$ for any t.

Now the integration by parts formula (4.4) gives $-\int Q_t\psi \, L^g f \, g d\mu = \int \Gamma(Q_t\psi, f) g d\mu$. Recall that $L^g f$ is bounded and that we have let $N = ||L^g f||_\infty$. For $t < s < 1/N$ we obtain

$$\frac{d}{dt}\int Q_t\psi \, g_t \, d\mu \le \int \left[-\frac{1}{2}\Gamma(Q_t\psi)(1 - sN) + \Gamma(Q_t\psi, f)\right] g d\mu$$
$$= \int \left[-\frac{1 - sN}{2}\Gamma\left(Q_t\psi - \frac{1}{1 - sN} f\right) + \frac{1}{2(1 - sN)}\Gamma(f)\right] g d\mu$$
$$\le \frac{1}{2(1 - sN)} \int \Gamma(f) g d\mu.$$

Integrating over the set $t \in [0, s]$:

$$\int Q_s\psi\, g_s d\mu - \int \psi g d\mu \leq \frac{s}{2(1-sN)} \int \Gamma(f) g d\mu.$$

Finally the Kantorovich duality (4.6) gives our second estimate:

$$\limsup_{s\to 0} \frac{W_2^2(g_s\mu, g\mu)}{2s^2} \leq \frac{1}{2}\int \Gamma(f) g d\mu. \tag{4.11}$$

Estimate on the second term on the right-hand side. Let $u > 0$ and let us compute the limit of $\frac{1}{s}\big(\mathrm{Ent}_\mu P_u g_s - \mathrm{Ent}_\mu P_u g\big)$ when s goes to 0. First, for any $s > 0$,

$$\frac{d}{ds} P_u(g_s)\log P_u(g_s) = -(1+\log P_u g_s)\; P_u\big(gL^g f\big).$$

Then, for $0 < s < 1/N$,

$$\big|(1+\log P_u g_s) P_u(gL^g f)\big| \leq N P_u g\; (1+\log(1+N) + |\log P_u(g)|).$$

Forgetting the dimensional corrective term in (2.3), by the von Renesse-Sturm theorem [27] the RMT satisfies a $CD(R,\infty)$ condition. In particular, and since $\int \Gamma(g)/g\, d\mu < \infty$, one can use a local logarithmic Sobolev inequality [8, Theorem 5.5.2] to deduce $\int P_u g\, |\log P_u g|\, d\mu < \infty$. In particular the right-hand side in the last inequality is in $L^1(\mu)$. Then, by the Lebesgue convergence theorem and (4.4),

$$\begin{aligned}\lim_{s\to 0} \frac{\mathrm{Ent}_\mu P_u g_s - \mathrm{Ent}_\mu P_u g}{s} &= -\int (1+\log P_u g) P_u\big(gL^g f\big) d\mu \\ &= -\int P_u(\log P_u g) L^g f g d\mu = \int \Gamma(P_u(\log P_u g), f) g d\mu.\end{aligned}$$

By the Fatou lemma we obtain the third estimate:

$$\begin{aligned}&\limsup_{s\to 0} -\frac{1}{m}\int_0^t e^{-2R(t-u)} \left[\frac{\mathrm{Ent}_\mu P_u g_s - \mathrm{Ent}_\mu P_u g}{s}\right]^2 du \\ &\leq -\frac{1}{m}\int_0^t e^{-2R(t-u)} \left(\int \Gamma(P_u(\log P_u g), f) g d\mu\right)^2 du.\end{aligned} \tag{4.12}$$

Conclusion. Dividing the inequality (2.3) by $2s^2$, letting s go to 0 and using the three estimates (4.9), (4.11) and (4.12), we get

$$\begin{aligned}&-\frac{1}{2}\int P_t\Gamma(f)\, g d\mu + \int \Gamma(f, P_t f) g d\mu \\ &\leq \frac{e^{-2Rt}}{2}\int \Gamma(f) g d\mu - \frac{1}{m}\int_0^t e^{-2R(t-u)} \left(\int \Gamma(P_u(\log P_u g), f) g d\mu\right)^2 du.\end{aligned}$$

This inequality is an equality when $t = 0$, and since $f \in \mathcal{C}_c^\infty(\mathbf{M})$, its derivative at $t = 0$ implies

$$-\frac{1}{2}\int L\Gamma(f)\, g d\mu + \int \Gamma(f, Lf) g d\mu \leq -R\int \Gamma(f) g d\mu - \frac{1}{m}\left(\int \Gamma(\log g, f) g d\mu\right)^2.$$

Since $\int \Gamma(\log g, f) g d\mu = \int \Gamma(g, f) d\mu = -\int g L f d\mu$ and by definition of the Γ_2 operator we get

$$\int \Gamma_2(f) g d\mu \geqslant R \int \Gamma(f)\, g d\mu + \frac{1}{m}\left(\int Lf\, g d\mu\right)^2 \tag{4.13}$$

for any $f \in \mathcal{C}_c^\infty(\mathbf{M})$ and any positive smooth probability density g with finite $\int g|\log g| d\mu$ and $\int \frac{\Gamma(g)}{g} d\mu$.

Inequality (4.13) appears as a weak form of the $CD(R, m)$ condition. Again from the $CD(R, \infty)$ condition, it is a consequence of Wang's Harnack inequality (see [8, Theorem 5.6.1] and [28]) that there exist $\alpha_0 > 0$ and $o \in \mathbf{M}$ such that

$$\int \exp(-\alpha_0 d(o, x)^2)\, d\mu(x) < \infty. \tag{4.14}$$

Then, for given $x \in \mathbf{M}$, for any $p > \alpha_0$ the function g_p defined by $g_p(y) = Z_p e^{-pd(x,y)^2}$ for a normalisation constant Z_p is such that $g_p \mu \in \mathcal{P}_2(\mathbf{X})$ and $\int g_p |\log g_p| d\mu$, $\int \Gamma(g_p)/g_p d\mu < \infty$. Moreover $(g_p)_p$ converges to the Dirac measure δ_x at x, so replacing g by g_p in (4.13) and letting $p \to +\infty$ we get

$$\Gamma_2(f) \geqslant R\Gamma(f) + \frac{1}{m}(Lf)^2$$

at any $x \in \mathbf{M}$ and for any function $f \in \mathcal{C}_c^\infty(\mathbf{M})$. This is the $CD(R, m)$ condition as in Definition 4.1, and this finishes the proof of Theorem 4.3.

5. The Riemannian energy measure space context

In this section we prove Theorem 2.1 in the context of a Riemannian energy measure (REM) space. The proof goes along the same overall strategy as in the manifold case of Section 4.2. However, to overcome the lack of differentiability, it will require several tools and results from optimal transport and heat distributions on metric measure spaces.

The framework is stated in Section 5.1. As an intermezzo, in Sections 5.2 and 5.3 we give the proofs of (i) ⇒ (ii) ⇒ (iii) in Theorem 2.1. The main implication (iii) ⇒ (i) is stated and proved in Section 5.4, in the form of Theorem 5.3. The path $(\tilde{g}_s)_{s\geqslant 0}$ is constructed in Section 5.4.1, the three key estimates are given in Section 5.4.2, finally the main proof is given in Section 5.4.3.

5.1. Framework

As a natural framework, we state our result on a Riemannian energy measure space, as introduced in [6]. Let $(\mathbf{X}, \tau)$ be a Polish topological space and μ a locally finite Borel measure on $\mathbf{X}$ with a full support. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a strongly local symmetric Dirichlet form on $\mathbb{L}^2(\mu)$. Let finally $(P_t)_{t\geqslant 0}$ be its associated semigroup and L its generator, with domain $\mathcal{D}(L) \subset \mathbb{L}^2(\mu)$. As for a Markov triple, see [8], and since P_t is symmetric and sub-Markovian, we can extend P_t to a semigroup of contractions on $\mathbb{L}^p(\mu)$ for $p \in [1, \infty]$. We also let $\mathcal{E}(f) := \mathcal{E}(f, f)$ and

$$\|f\|_{\mathcal{E}}^2 := \|f\|_{\mathbb{L}^2(\mu)}^2 + \mathcal{E}(f)$$

for $f \in \mathcal{D}(\mathcal{E})$. In this work we assume that $(\mathbf{X}, \tau, \mu, \mathcal{E})$ is a Riemannian energy measure space in the sense of [6, Definition 3.16]. A basic example of a REM space is a Riemannian Markov triple as in Section 4. In this case, $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is canonically defined by completion of $(f, f) \mapsto \int |\nabla f|^2 \, d\mu$. RCD spaces introduced in [1,5] are another important class of REM spaces. In this case, $\mathcal{E}/2$ is given by the $\mathbb{L}^2$-Cheeger energy functional. As we will see below, our REM space becomes an $\mathsf{RCD}(R, \infty)$ space in an appropriate sense under one of the conditions in Theorem 2.1 (see the argument in section 5.1 below): hence our argument falls into the framework of a RCD space and it would make no difference to state or to prove our result in the framework of a RCD space instead of a REM space. However our conditions in Theorem 2.1 are described in terms of the Markov semigroup $(P_t)_{t\geqslant 0}$ and its infinitesimal generator L, so we thought that the framework of a REM space was natural and adapted, and preferred it rather than a RCD space as a starting point.

To make this presentation concise, we prefer to state the crucial properties of a REM space instead of its precise definition. Indeed the definition consists in several notions, which will be used only indirectly through these properties:

- The intrinsic distance $d_{\mathcal{E}}$ associated with $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$, in the sense of [6, Section 3.3], becomes a distance function, further denoted d. It is compatible with the topology τ and the space $(\mathbf{X}, d)$ is complete [6, Definition 3.6] and length metric [6, Theorem 3.10].

We let $\mathrm{Lip}_b(\mathbf{X})$ denote the set of bounded Lipschitz functions on $\mathbf{X}$ (with respect to d). Let $|\nabla f| : \mathbf{X} \to \mathbb{R}$ be the local Lipschitz constant of a Lipschitz function f on $\mathbf{X}$:

$$|\nabla f|(x) := \limsup_{y\to x} \frac{|f(y) - f(x)|}{d(x, y)}.$$

- $\mathcal{E}/2$ coincides with the $\mathbb{L}^2$-Cheeger energy associated with d, defined for $f \in \mathbb{L}^2(\mu)$ by

$$\mathrm{Ch}(f) := \inf \left\{ \liminf_{n\to\infty} \frac{1}{2} \int |\nabla f_n|^2 d\mu \, ; \; f_n \in \mathrm{Lip}_b(\mathbf{X}), \; f_n \to f \text{ in } \mathbb{L}^2(\mu) \right\}.$$

As a result, $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ admits a carré du champ, *i.e.* there is a symmetric bilinear map $\Gamma : \mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E}) \to \mathbb{L}^1(\mu)$ such that

$$\mathcal{E}(f, g) = \int \Gamma(f, g)\, d\mu.$$

As on smooth spaces, L and Γ satisfy the diffusion property (4.2). The coincidence of $\mathcal{E}/2$ and the Cheeger energy makes many connections between d and Γ. For instance, $\mathcal{D}(\mathcal{E}) \cap \mathrm{Lip}_b(\mathbf{X})$ is dense in $\mathcal{D}(\mathcal{E})$ with respect to $\| \cdot \|_{\mathcal{E}}$. In addition,

$$\Gamma(f) \leq |\nabla f|^2 \quad \mu\text{-a.e.} \tag{5.1}$$

for any Lipschitz $f \in \mathcal{D}(\mathcal{E})$. See [6, Theorem 3.12] and [6, Theorem 3.14] for all these facts.

Note that $\mathcal{D}(\mathcal{E}) \cap \mathbb{L}^\infty(\mu)$ is an algebra and Γ satisfies the Leibniz rule:

$$\Gamma(fg, h) = f\Gamma(g, h) + g\Gamma(f, h) \quad \text{for } f, g \in \mathcal{D}(\mathcal{E}) \cap \mathbb{L}^\infty(\mu) \text{ and } h \in \mathcal{D}(\mathcal{E}).$$

We state further assumptions for our main theorem. Fix a reference point $o \in \mathbf{X}$.

Regularity assumption.

(Reg1) There is $\alpha_0 > 0$ such that (4.14) holds.
(Reg2) $(\mathbf{X}, \tau)$ is locally compact.

Assumption (Reg1): is equivalent to the condition (MD.exp) in [6] (see, *e.g.*, the comments after Equation (3.13) in [6]). This integrability condition yields the conservativity of P_t, *i.e.*

$$\int P_t f\, d\mu = \int f\, d\mu$$

for $f \in \mathbb{L}^1(\mu)$ (see [6, Theorem 3.14]). This is equivalent to $P_t 1 = 1$ μ-a.e, that is, the semigroup is Markovian (instead of sub-Markovian). In fact (4.14) is a nearly optimal condition to ensure that the semigroup is conservative (see [3, Remark 4.21]). Thus it is not restrictive.

Assumption (Reg2): implies that any closed bounded set in $\mathbf{X}$ is compact (see, *e.g.*, [12, Proposition 2.5.22]). Moreover, $(\mathbf{X}, d)$ is a geodesic space (see, *e.g.*, [12, Theorem 2.5.23]). As a result, $(\mathcal{P}_2(\mathbf{X}), W_2)$ is also a geodesic space (see, *e.g.*, [20, Corollary 1 and Proposition 1]).

In this framework, we should be careful when defining the operator Γ_2 in (4.3) since $\Gamma(f)$ may not belong to $\mathcal{D}(L)$ even for a sufficiently nice f. To avoid such a technical difficulty, and following [6, Definition 2.4], we employ a weak form of the $CD(R, m)$ condition

Definition 5.1 (Weak $CD(R,m)$ condition). Let $R \in \mathbb{R}$ and $m > 0$. We say that the REM space $(\mathbf{X}, \tau, \mu, \mathcal{E})$ satisfies a weak $CD(R,m)$ condition if, for all $f \in \mathcal{D}(L)$ with $Lf \in \mathcal{D}(\mathcal{E})$ and all $g \in \mathcal{D}(L) \cap \mathbb{L}^\infty(\mu)$ with $g \geqslant 0$ and $Lg \in \mathbb{L}^\infty(\mu)$,

$$\frac{1}{2}\int \Gamma(f) Lg \, d\mu - \int \Gamma(f, Lf) g \, d\mu \geqslant R \int \Gamma(f) g \, d\mu + \frac{1}{m}\int (Lf)^2 g \, d\mu. \quad (5.2)$$

The proof of (iii) $\Rightarrow$ (i) (and also of (ii) $\Rightarrow$ (iii)) of Theorem 2.1 will need further regularity properties on the space and semigroup, which will in fact be consequences of (iii) (or (ii)).

Note indeed that (2.3) in (iii) yields a W_2-contraction

$$W_2^2(P_t g d\mu, P_t h d\mu) \le e^{-2Rt} W_2^2(g d\mu, h d\mu) \quad (5.3)$$

by neglecting the term involving m. Then, by [6, Corollary 3.18], (5.3) implies a $CD(R,\infty)$ condition in the sense of (5.2). This fact is very helpful for further discussion in the sequel since it ensures regularity of the space in many respects. As a regularization property of P_t, we have

$$P_t h \in \mathrm{Lip}_b(\mathbf{X}) \text{ for } h \in \mathbb{L}^2(\mu) \cap \mathbb{L}^\infty(\mu), \ t > 0 \quad (5.4)$$

(see [6, Theorem 3.17]); more precisely, $P_t h$ has a version which belongs to $\mathrm{Lip}_b(\mathbf{X})$). In addition, $(\mathbf{X}, d, \mu)$ becomes an $\mathsf{RCD}(R,\infty)$ space (see [6, Theorem 4.17]). Then, for a probability density h with respect to μ, $((P_t h)\mu)_{t \geqslant 0}$ is a gradient flow of Ent_μ in the sense of the R-evolution variational inequality [1, Theorem 6.1]. As a consequence, we obtain the following properties:

- We can extend the action of P_t to $\nu \in \mathcal{P}_2(\mathbf{X})$ in the sense that $P_t \nu$ is a solution to the R-evolution variational inequality and that $P_t \nu = (P_t h)\mu$ if $\nu = h\mu$. In particular, $(P_t \nu)_{t \geqslant 0}$ becomes a continuous curve in $(\mathcal{P}_2(\mathbf{X}), W_2)$, see [1, Theorem 6.1]. In addition, $\nu \mapsto P_t \nu$ is a continuous map from $(\mathcal{P}_2(\mathbf{X}), W_2)$ to itself, see [1, Equation (7.2)];
- $P_t \nu \ll \mu$ for $\nu \in \mathcal{P}_2(\mathbf{X})$ and $t > 0$, and its density ρ_t satisfies $\mathrm{Ent}_\mu \rho_t \in \mathbb{R}$. This property is included in the definition of the R-evolution variational inequality, see, *e.g.*, [1, Definition 2.5]. Recall that, under (4.14), $\mathrm{Ent}_\mu \rho$ is well-defined and $\mathrm{Ent}_\mu \rho \in (-\infty, \infty]$ for $\rho : \mathbf{X} \to [0,\infty]$ with $\rho\mu \in \mathcal{P}_2(\mathbf{X})$, see, *e.g.*, [3, Section 7];
- There is a positive symmetric measurable function $p_t(x,y)$ such that P_t coincides with the integral operator associated with p_t, see [1, Theorem 7.1];
- For any bounded measurable h and $\nu \in \mathcal{P}_2(\mathbf{X})$, we have

$$\int h \, dP_t \nu = \int P_t h \, d\nu, \quad (5.5)$$

  see [6, Proposition 3.2]. By the monotone convergence theorem, we can extend this identity to those h which are bounded only from below (or above);
- For any $f \in \mathcal{D}(L)$ and $h \in \mathcal{D}(\mathcal{E})$ we have the integration by parts formula

$$\int \Gamma(h, f) \, d\mu = -\int h \, Lf \, d\mu. \quad (5.6)$$

5.2. Proof of (i) ⇒ (ii) in Theorem 2.1

In [13] M. Erbar, K.-T. Sturm and the fourth author of this paper have proved an *Evolutional variational inequality* (EVI in short) in the REM spaces. Let g, h be probability densities with respect to μ and let $U_m = \exp(-\mathrm{Ent}_\mu \cdot / m)$. Then, under the weak $CD(R, m)$ condition as in (i),

$$\frac{d}{dt} s_{\frac{R}{m}} \left(\frac{1}{2} W_2(P_t h\mu, g\mu)\right)^2 + R\, s_{\frac{R}{m}} \left(\frac{1}{2} W_2(P_t h\mu, g\mu)\right)^2 \le \frac{m}{2}\left(1 - \frac{U_m(g)}{U_m(P_t h)}\right). \quad (5.7)$$

But it is classical, see, *e.g.*, [2], how to deduce a contraction property in W_2 distance between solutions $(P_t h)_{t \geqslant 0}$ and $(P_t g)_{t \geqslant 0}$ from an EVI: one applies the EVI to the curve $(P_t h)_{t \geqslant 0}$ and $P_s g$ for a given s, and then (with the time variable s) to the curve $(P_s g)_{s \geqslant 0}$ and $P_t h$ for a given t; then one adds both inequalities, takes $t = s$ and integrate in time. Then one obtains (ii).

To sum up, it turns out that the EVI (5.7) not only leads to the property (1.5), as observed in [13], but also to the *same-time* contraction property (ii).

5.3. Proof of (ii) ⇒ (iii) in Theorem 2.1

We first observe that $\sinh^2(x) \geqslant x^2$ for any x, so (ii) in Theorem 2.1 implies the same bound with $\sinh^2(x)$ replaced by x^2 in the integral. Then the implication (ii) ⇒ (iii) is a consequence of the following result, which we prove in the general context of a geodesic space.

Proposition 5.2. *Let (Y, d_Y) be a geodesic metric space, $U : Y \to (-\infty, \infty]$ and $\varphi_t : Y \to Y$ ($t \ge 0$) a one-parameter family of maps. Suppose that $t \mapsto \varphi_t(y)$ is continuous for all $y \in Y$ and $U(\varphi_t(y)) \in \mathbb{R}$ for all $t > 0$ and $y \in Y$. Suppose also that for $y_0, y_1 \in Y$ and $t > 0$,*

$$\begin{aligned} s_{\frac{R}{m}} \left(\frac{1}{2} d_Y(\varphi_t(y_0), \varphi_t(y_1))\right)^2 &\le e^{-2Rt}\, s_{\frac{R}{m}} \left(\frac{1}{2} d_Y(y_0, y_1)\right)^2 \\ &\quad - \frac{1}{2m} \int_0^t e^{-2R(t-u)} (U(\varphi_u(y_0)) - U(\varphi_u(y_1)))^2 du. \end{aligned} \quad (5.8)$$

Then

$$\begin{aligned} d_Y(\varphi_t(y_0), \varphi_t(y_1))^2 &\le e^{-2Rt}\, d_Y(y_0, y_1)^2 \\ &\quad - \frac{2}{m} \int_0^t e^{-2R(t-u)} (U(\varphi_u(y_0)) - U(\varphi_u(y_1)))^2 du. \end{aligned}$$

Proof. We adapt the argument of [13, Proposition 2.22]. Let $(y_s)_{s \in [0,1]}$ be a geodesic from y_0 to y_1 in Y, and let $t > 0$ be fixed. For any n and $1 \le i \le n$, let $x_i^n = d_Y(\varphi_t(y_{(i-1)/n}), \varphi_t(y_{i/n}))$. Then

$$d_Y(\varphi_t(y_0), \varphi_t(y_1))^2 \le \left(\sum_{i=1}^n x_i^n\right)^2 \le n \sum_{i=1}^n (x_i^n)^2$$

for any n. In particular

$$d_Y(\varphi_t(y_0), \varphi_t(y_1))^2 \le \limsup_{n\to\infty} n \sum_{i=1}^{n} \left(x_i^n\right)^2.$$

Now, by neglecting the second term in the right-hand side of (5.8) and by geodesic property,

$$s_{\frac{R}{m}}\left(\frac{x_i^n}{2}\right) \le e^{-Rt} s_{\frac{R}{m}}\left(\frac{1}{2} d_Y(y_{(i-1)/n}, y_{i/n})\right) = e^{-Rt} s_{\frac{R}{m}}\left(\frac{1}{2n} d_Y(y_0, y_1)\right).$$

It follows, as in [13, (2.32)], that there exists a constant c such that $x_i^n \le c/n$ for large n and any $1 \le i \le n$. Moreover $s_{\frac{R}{m}}(x)^2 = x^2 - Rx^4/(3m) + O(x^6)$ as x tends to 0, so that

$$\limsup_{n\to\infty} n \sum_{i=1}^{n} \left(x_i^n\right)^2 = 4 \limsup_{n\to\infty} n \sum_{i=1}^{n} s_{\frac{R}{m}}\left(x_i^n/2\right)^2. \tag{5.9}$$

As a consequence

$$\begin{aligned} &d_Y(\varphi_t(y_0), \varphi_t(y_1))^2 \\ &\le 4 \limsup_{n\to\infty} n \sum_{i=1}^{n} s_{\frac{R}{m}}\left(\frac{1}{2} d_Y(\varphi_t(y_{(i-1)/n}), \varphi_t(y_{i/n}))\right)^2 \\ &\le 4 \limsup_{n\to\infty} \Bigg(n \sum_{i=1}^{n} e^{-2Rt} s_{\frac{R}{m}}\left(\frac{1}{2} d_Y(y_{(i-1)/n}, y_{i/n})\right)^2 \\ &\qquad - \frac{1}{2m} \int_0^t e^{-2R(t-u)} n \sum_{i=1}^{n} \left(U(\varphi_u(y_{(i-1)/n})) - U(\varphi_u(y_{i/n}))\right)^2 du \Bigg) \end{aligned}$$

by assumption (5.8).

Then the conclusion follows from this estimate by using (5.9) with $d_Y(y_{(i-1)/n}, y_{i/n})$ in place of x_i^n in the first term, and the Cauchy-Schwarz inequality in the second term. □

Let us return to the proof of (ii) ⇒ (iii) in Theorem 2.1. We first check that (2.2) yields (5.3). As we derived (5.3) from (2.3), the estimate (2.2) yields

$$s_{\frac{R}{m}}\left(\frac{1}{2} W_2(P_t h\mu, P_t g\mu)\right)^2 \le e^{-2Rt} s_{\frac{R}{m}}\left(\frac{1}{2} W_2(h\mu, g\mu)\right)^2 \tag{5.10}$$

by neglecting the term involving m. From this inequality, we can extend P_t to a map from $\mathcal{P}_2(\mathbf{X})$ to itself, in a canonical way. Moreover, in (5.10) we can replace $h\mu$ and $g\mu$ with any $\nu_0, \nu_1 \in \mathcal{P}_2(\mathbf{X})$ respectively. Then we obtain (5.3) by a similar argument as in Proposition 5.2. Thus, as discussed in Section 5.1, $(\mathbf{X}, d, \mu)$ is an

$\mathsf{RCD}(R,\infty)$ space and all properties at the end of Section 5.1 become available. We remark that the extension of P_t given on the basis of (5.10) coincides with the one given by the $\mathsf{RCD}(R,\infty)$ property.

Now we only need to show that P_t fulfills all the assumptions for φ_t in Proposition 5.2 with $(Y, d_Y) = (\mathcal{P}_2(\mathbf{X}), W_2)$ and $U = \mathrm{Ent}_\mu$. Here we are extending the definition of Ent_μ so that, for $\nu \in \mathcal{P}_2(\mathbf{X})$, $\mathrm{Ent}_\mu \nu = \mathrm{Ent}_\mu d\nu/d\mu$ if $\nu \ll \mu$ and $\mathrm{Ent}_\mu \nu = \infty$ otherwise. By taking observations at the beginning of this section into account, it suffices to prove that (2.2) implies

$$s_{\frac{R}{m}}\left(\frac{1}{2}W_2(P_t\nu_0, P_t\nu_1)\right)^2 \le e^{-2Rt} s_{\frac{R}{m}}\left(\frac{1}{2}W_2(\nu_0,\nu_1)\right)^2 - \frac{1}{2m}\int_0^t e^{-2R(t-u)}\big(\mathrm{Ent}_\mu P_u\nu_0 - \mathrm{Ent}_\mu P_u\nu_1\big)^2 du$$

for $\nu_0, \nu_1 \in \mathcal{P}_2(\mathbf{X})$ and $t > 0$. But this is true since $P_\delta\nu_0, P_\delta\nu_1 \ll \mu$ for any $\delta \in (0, t)$, so that

$$s_{\frac{R}{m}}\left(\frac{1}{2}W_2(P_t\nu_0, P_t\nu_1)\right)^2 \le e^{-2R(t-\delta)} s_{\frac{R}{m}}\left(\frac{1}{2}W_2(P_\delta\nu_0,P_\delta\nu_1)\right)^2 - \frac{1}{2m}\int_\delta^t e^{-2R(t-u)}\big(\mathrm{Ent}_\mu P_u\nu_0 - \mathrm{Ent}_\mu P_u\nu_1\big)^2 du$$

by (2.2) and the bound $\sinh^2(x) \geqslant x^2$; moreover $P_\delta\nu_i \to \nu_i$ in W_2 as $\delta \downarrow 0$ for $i = 0, 1$: this gives the assertion. Hence the proof of (ii) $\Rightarrow$ (iii) in Theorem 2.1 is completed. □

5.4. Proof of (iii) ⇒ (i) in Theorem 2.1

In this section we prove the main implication (iii) $\Rightarrow$ (i) in Theorem 2.1, in the following form

Theorem 5.3. *Let* $(\mathbf{X}, \tau, \mu, \mathcal{E})$ *be a Riemannian energy measure space satisfying the above regularity assumptions (Reg1) and (Reg2). Let* $R \in \mathbb{R}$ *and* $m > 0$.

If inequality (2.3) *holds for any* $t \geqslant 0$ *and probability densities* $g, h \in \mathbb{L}^1(\mu)$ *with* $g\mu, h\mu \in \mathcal{P}_2(\mathbf{X})$*, then the weak* $CD(R,m)$ *condition of Definition* 5.1 *holds.*

In particular, the conditions (ii) and (iii) in Theorem 2.1 are equivalent to the weak $CD(R, m)$ condition.

5.4.1. *Construction of the path* $(\tilde{g}_s)_{s\geqslant 0}$

In this section, we build the path $\tilde{g}_s$ mentioned in Section 3.2, under (2.3). Recall that $(\mathbf{X}, d, \mu)$ is now an $\mathsf{RCD}(R,\infty)$ space as remarked at the end of Section 5.1. For $x \in \mathbf{X}$ and $r > 0$, we denote the open ball of radius r centered at x by $B_r(x)$.

For this we first define $g(= \tilde{g}_0)$. We take g in a more tractable (but large enough) class than the full class of Definition 5.1. Fix $\alpha > \alpha_0$ with α_0 as in (4.14), $\lambda \in (0, 1)$ and $g_0 : \mathbf{X} \to \mathbb{R}$ Lipschitz with compact support. Let us define g as follows:

$$g := \frac{1}{Z}\left((1-\lambda)g_0 + \lambda \exp\left(-\alpha d(x, o)^2\right)\right) \tag{5.11}$$

where $Z > 0$ is a normalizing constant such that $g\mu \in \mathcal{P}(\mathbf{X})$. Note that (4.14) yields $g\mu \in \mathcal{P}_2(\mathbf{X})$. We fix g until the end of the proof of Proposition 5.10 below. We can define the $\mathbb{L}^2$-Cheeger energy functional $\mathcal{E}_g/2$ associated with d and the probability measure $g\mu$. Let $\mathcal{D}(\mathcal{E}_g)$ be the set of $f \in \mathbb{L}^2(g\mu)$ with $\mathcal{E}_g(f) < \infty$. Recall that $\mathcal{D}(\mathcal{E}_g)$ is complete with respect to $\|\cdot\|_{\mathcal{E}_g}$.

To define the path $(\tilde{g}_s)_{s\geqslant 0}$ we need the corresponding generator L^g, and for this we show the following auxiliary lemma.

Lemma 5.4. *In the above notation, $\mathcal{D}(\mathcal{E}) \subset \mathcal{D}(\mathcal{E}_g)$ and*

$$\mathcal{E}_g(f) = \int \Gamma(f) g \, d\mu \tag{5.12}$$

for $f \in \mathcal{D}(\mathcal{E})$. In addition, $(\mathcal{E}_g, \mathcal{D}(\mathcal{E}_g))$ is bilinear.

We do not know whether (5.12) is valid for any $f \in \mathcal{D}(\mathcal{E}_g)$. Thus we have to be careful when we apply the integration by parts formula (4.4) for L^g.

Proof. The former assertion follows from [3, Lemma 4.11]. For the latter assertion, take $f, \tilde{f} \in \mathcal{D}(\mathcal{E}_g)$. For each $n \in \mathbb{N}$, take also $\chi_n \in \mathrm{Lip}_b(\mathbf{X})$ with $0 \le \chi_n \le 1$, $\chi|_{B_n(o)} \equiv 1$ and $\chi|_{B_{n+1}(o)^c} \equiv 0$.

Since, for each $n \in \mathbb{N}$, g is bounded away from 0 on $B_n(o)$, we have $f_n := f\chi_n \in \mathcal{D}(\mathcal{E})$ by the locality of the Cheeger energy, see [3, Proposition 4.8 (b)] and [3, Lemma 4.11]. Moreover, $(f_n)_{n\in\mathbb{N}}$ forms a Cauchy sequence with respect to $\|\cdot\|_{\mathcal{E}_g}$ and hence $\|f_n - f\|_{\mathcal{E}_g} \to 0$. By the same argument, we have $\|\tilde{f}_n - \tilde{f}\|_{\mathcal{E}_g} \to 0$ for $\tilde{f}_n := \tilde{f}\chi_n$. By (5.12), and recalling that Γ is symmetric bilinear, we have

$$\mathcal{E}_g(f_n + \tilde{f}_n) + \mathcal{E}_g(f_n - \tilde{f}_n) = 2\left(\mathcal{E}_g(f_n) + \mathcal{E}_g(\tilde{f}_n)\right).$$

Therefore the conclusion holds by letting $n \to \infty$. □

By Lemma 5.4, $(\mathcal{E}_g, \mathcal{D}(\mathcal{E}_g))$ is a closed bilinear form on $\mathbb{L}^2(g\mu)$. Hence there are an associated $\mathbb{L}^2$-semigroup P_t^g of symmetric linear contraction and its generator L^g. By [3, Proposition 4.8 (b)], $\mathcal{E}_g$ is sub-Markovian. Thus P_t^g satisfies the maximum principle, *i.e.* $P_t^g f \le c$ if $f \le c$ for $f \in \mathbb{L}^2(g\mu)$ and $c \in \mathbb{R}$. In addition, $\mathrm{Lip}_b(\mathbf{X}) \cap \mathcal{D}(\mathcal{E}_g)$ is dense in $\mathcal{D}(\mathcal{E}_g)$ with respect to $\|\cdot\|_{\mathcal{E}_g}$. Note that we can define P_t^g and L^g without bilinearity of $\mathcal{E}_g$ (see [3, Section 4] and references therein). However, then they can be nonlinear and the integration by parts formula (4.4) may not hold.

Lemma 5.5. *In the above notation:*

(i) $g \in \mathcal{D}(\mathcal{E}) \cap \mathbb{L}^\infty(\mu)$ *and* $\log g \in \mathcal{D}(\mathcal{E}_g)$*;*
(ii) $\mathcal{D}(L) \subset \mathcal{D}(L^g)$.

Proof. (i) The first claim follows from (5.1) and (4.14). For the second one, note that

$$\mathcal{E}_g(\log g) \leq \int |\nabla \log g|^2 g \, d\mu.$$

It is the integrated form of (5.1) for $\mathcal{E}_g$ instead of $\mathcal{E}$. Then the claim follows from (4.14).

(ii) Let $f \in \mathcal{D}(L)$ and $h \in \mathcal{D}(\mathcal{E}_g)$. Take $h_n \in \mathrm{Lip}_b(\mathbf{X}) \cap \mathcal{D}(\mathcal{E}_g)$ for $n \in \mathbb{N}$ such that $\|h_n - h\|_{\mathcal{E}_g} \to 0$. By a truncation argument used in the proof of Lemma 5.4, we may assume that each h_n is supported on a bounded set, without loss of generality. Then $h_n \in \mathcal{D}(\mathcal{E}) \cap \mathbb{L}^\infty(\mu)$ and hence $h_n g \in \mathcal{D}(\mathcal{E})$. Thus the Leibniz rule, the assertion (i), (5.6) and (5.1) imply

$$\begin{aligned}
\left| \int \Gamma(h_n, f) g \, d\mu \right| &= \left| \int \Gamma(h_n g, f) \, d\mu - \int h_n \Gamma(g, f) \, d\mu \right| \\
&\leq \left| \int h_n (Lf) g \, d\mu \right| + \left| \int h_n \Gamma(\log g, f) g \, d\mu \right| \\
&\leq \|h_n\|_{L^2(g\mu)} \left(\|g\|_\infty \|Lf\|_{\mathbb{L}^2(\mu)} + \left\| \frac{|\nabla g|^2}{g} \right\|_\infty \mathcal{E}(f)^{1/2} \right).
\end{aligned}$$

The definition of g yields $\||\nabla g|^2/g\|_\infty < \infty$. Thus there is $C > 0$ independent of h and n such that

$$\left| \mathcal{E}_g(h_n, f) \right| \leq C \|h_n\|_{\mathbb{L}^2(g\mu)}.$$

Here we used Lemma 5.4. By letting $n \to \infty$, we can replace h_n with h in this inequality. Hence $f \in \mathcal{D}(L^g)$ since h is arbitrary in $\mathcal{D}(\mathcal{E}_g)$. □

We can now define the path $(\tilde{g}_s)_{s \geqslant 0}$. Let $f \in \mathcal{D}(L) \cap \mathrm{Lip}_b(\mathbf{X})$ with $\|f\|_\infty \leq 1/4$. We fix f until the end of the following section, and observe that $f \in \mathbb{L}^2(g\mu)$. Then we let

$$\tilde{g}_s := g\big(1 + f - P_s^g f\big). \tag{5.13}$$

By the $\mathbb{L}^\infty$-bound on f and the maximum principle for P_s^g, we have

$$\frac{1}{2} g \leq \tilde{g}_s \leq 2g. \tag{5.14}$$

In what follows, we may assume without loss of generality that $L^g f$ is not identically 0. For, by (5.6) and Lemma 5.5,

$$\int Lf \; g \, d\mu = -\int \Gamma(f, g) \, d\mu = -\int \Gamma(f, \log g) g \, d\mu = \int L^g f \; \log g \; g \, d\mu. \tag{5.15}$$

Thus, if $L^g f$ is identically 0, then $\int Lf\, g\, d\mu = 0$; hence (5.19) below holds in this specific case (without the next section) since the $CD(R,\infty)$ condition holds on our $\mathsf{RCD}(R,\infty)$ space.

5.4.2. *Three key estimates*

The proof of Theorem 5.3 is based on (5.19) in Proposition 5.10 below. In turn, this bound is based on the three key estimates in Lemmas 5.6, 5.7 and 5.9, which in the manifold case of Section 4.2 correspond to (4.9), (4.11) and (4.12). The proofs are a bit different since we use $\tilde{g}_s$ instead of g_s.

The Hopf-Lax semigroup $(Q_s)_{s\geqslant 0}$ given by (4.5) will again play a crucial role. Required properties for Q_s in this framework are given in [3, Section 3] or [4, Section 3] for instance.

We begin with the *first estimate*, corresponding to (4.9):

Lemma 5.6 (First estimate).

$$\liminf_{s\to 0} \frac{W_2^2(P_t\tilde{g}_s\mu, P_t g\mu)}{2s^2} \geqslant -\frac{1}{2}\int P_t(|\nabla f|^2) g\, d\mu + \int \Gamma(f, P_t f) g\, d\mu.$$

Proof. It suffices to prove an lower bound on the right-hand side of (4.7). By a rearrangement,

$$\begin{aligned}\int \frac{Q_s f\, P_t\tilde{g}_s - f\, P_t g}{s} d\mu = &\int \frac{Q_s f - f}{s} P_t(\tilde{g}_s - g)\, d\mu \\ &+ \int \frac{Q_s f - f}{s} P_t g\, d\mu + \int f\, \frac{P_t(\tilde{g}_s - g)}{s}\, d\mu.\end{aligned} \tag{5.16}$$

Since $g\mu \in \mathcal{P}(\mathbf{X})$, the Cauchy-Schwarz inequality yields $s^{-1}(\tilde{g}_s - g) \to -g\, L^g f$ in $\mathbb{L}^1(\mu)$. Thus the last term in (5.16) converges to $-\int f\, P_t(g L^g f)\, d\mu$. By Lemma 5.4, and as in Section 4.2, this quantity is equal to the second term on the right-hand side of the assertion. Moreover, by the general bound (4.8), the first term on the right-hand side of (5.16) goes to 0. Finally, by (4.8) and the Lebesgue dominated convergence theorem we conclude on the second term as in the Riemannian case of Section 4.2. More precisely, we have

$$\begin{aligned}&\liminf_{s\to 0} \int \frac{Q_s f(x) - f(x)}{s} P_t g(x)\, \mu(dx) \\ &\quad \geqslant -\frac{1}{2} \limsup_{s\to 0} \int \sup_{y\in B(x,\sqrt{4s\|f\|_\infty})\setminus\{x\}} \left(\frac{f(y) - f(x)}{d(x,y)}\right)^2 P_t g(x)\mu(dx) \\ &\quad = -\frac{1}{2}\int |\nabla f|^2 P_t g\, d\mu.\end{aligned}$$

Thus the assertion holds. □

Next lemma deals with the *second estimate* and corresponds to (4.11).

Lemma 5.7 (Second estimate).

$$\limsup_{s\to 0} \frac{W_2^2(\tilde{g}_s\mu, g\mu)}{2s^2} \le \frac{1}{2(1-2\|f\|_\infty)} \int \Gamma(f) g \, d\mu.$$

Proof. Again, by the dual form (4.6), we need to bound $\int Q_s\psi \, \tilde{g}_s d\mu - \int \psi g d\mu$ uniformly from above on the bounded Lipschitz functions ψ. We can assume that ψ is moreover supported on a bounded set. Then the function $(s_1, s_2) \mapsto \int Q_{s_1}(\psi)\tilde{g}_{s_2} \, d\mu$ satisfies the assumption of [2, Lemma 4.3.4] since we have (5.14) and $\|Q_{s_1}\psi\|_\infty \le \|\psi\|_\infty$. Thus, instead of (4.10), we obtain

$$\begin{aligned} \frac{d}{ds}\int Q_s(\psi)\tilde{g}_s \, d\mu &\le \frac{d}{ds}\int Q_s(\psi)\tilde{g}_{s_0} \, d\mu\bigg|_{s_0=s} + \frac{d}{ds}\int Q_{s_0}(\psi)\tilde{g}_s \, d\mu\bigg|_{s_0=s} \\ &= \int \left[-\frac{1}{2}|\nabla Q_s\psi|^2\big(1 + f - P_s^g f\big) - Q_s\psi \; L^g P_s^g f \right] g \, d\mu \end{aligned}$$

for a.e. $s > 0$. Here the equality follows from [4, Theorem 3.6], the properties $\|Q_s\psi\|_{\mathrm{Lip}} < \infty$, $\|Q_s\psi\|_\infty < \infty$ and the Lebesgue dominated convergence theorem. Note that $Q_s\psi \in \mathcal{D}(\mathcal{E}_g)$ since $Q_s\psi$ is Lipschitz with a bounded support. Thus, by virtue of Lemma 5.4 and (5.1),

$$\begin{aligned} -\int Q_s\psi\big(L^g P_s^g f\big) g \, d\mu = \mathcal{E}_g\big(Q_s\psi, P_s^g f\big) &\le \sqrt{\mathcal{E}_g(Q_s\psi)\mathcal{E}_g\big(P_s^g f\big)} \\ &\le \sqrt{\int |\nabla Q_s\psi|^2 g \, d\mu \; \mathcal{E}_g\big(P_s^g f\big)}. \end{aligned}$$

By combining this estimate with the last one, we obtain

$$\begin{aligned} \frac{d}{ds}\int Q_s(\psi)\tilde{g}_s \, d\mu &\le \frac{1}{2(1-2\|f\|_\infty)}\mathcal{E}_g\big(P_s^g f\big) \le \frac{1}{2(1-2\|f\|_\infty)}\mathcal{E}_g(f) \\ &= \frac{1}{2(1-2\|f\|_\infty)} \int \Gamma(f) g \, d\mu. \end{aligned}$$

Here the second inequality follows from the spectral decomposition for quadratic forms and the equality follows from Lemma 5.4 again since $f \in \mathcal{D}(L) \subset \mathcal{D}(\mathcal{E})$. Thus the conclusion follows by integrating this estimate, as in the proof of (4.11). □

For the *third estimate*, we still require some preparation. We call $\mathcal{C}_2(\mathbf{X})$ the set of continuous functions ψ on $\mathbf{X}$ for which there exists $C > 0$ such that $|\psi(x)| \le C(1 + d(o,x)^2)$. For $\psi \in \mathcal{C}_2(\mathbf{X})$ and $\nu \in \mathcal{P}_2(\mathbf{X})$, we have $\psi \in \mathbb{L}^1(\nu)$. By assumption on g, $\psi \in \mathbb{L}^p(g\mu)$ for any $\psi \in \mathcal{C}_2(\mathbf{X})$ and $p \in [1, \infty)$. The following lemma ensures integrability properties required in the proof of Lemma 5.9 below.

Lemma 5.8. *In the above notation:*

(i) $\psi g\mu \in \mathcal{P}_2(\mathbf{X})$ *for any* $\psi \in \mathbb{L}^2(g\mu)$ *with* $\psi g\mu \in \mathcal{P}(\mathbf{X})$*;*
(ii) $\log P_u g \in \mathcal{C}_2(\mathbf{X})$ *for* $u \geqslant 0$.

Proof. (i) Using Assumption (Reg1) and (5.11), this follows from

$$\int d(o,x)^2\psi(x)g(x)\,\mu(dx) \le \left(\int d(o,x)^4 g(x)\,\mu(dx)\right)^{1/2}\left(\int \psi^2 g\,d\mu\right)^{1/2} < \infty.$$

(ii) By (5.11) this is obvious for $u = 0$ and hence we consider the case $u > 0$. First of all, $\log P_u g$ is continuous on $\mathbf{X}$ since $P_u g > 0$. Moreover, since $(\mathbf{X}, d, \mu)$ is an $\mathsf{RCD}(R,\infty)$ space, we have the log-Harnack inequality

$$P_u(\log g)(o) - \frac{Rd(x,o)^2}{2(e^{2Ru}-1)} \le \log P_u g(x) \le \log \|g\|_\infty$$

(see [6, Lemma 4.6] or [19, Proposition 4.1]). Moreover $\log g \in \mathcal{C}_2(\mathbf{X})$ and $P_u\delta_o \in \mathcal{P}_2(\mathbf{X})$ by the properties after (5.4), so we have $\int \log g\,dP_u\delta_o = P_u(\log g)(o) \in \mathbb{R}$. Thus $\log P_u g \in \mathcal{C}_2(\mathbf{X})$. □

We recall characterizations of convergence in W_2 for later use. Let $\nu_n \in \mathcal{P}_2(\mathbf{X})$, $n \in \mathbb{N}$ and $\nu \in \mathcal{P}_2(\mathbf{X})$. Then $W_2(\nu_n, \nu) \to 0$ is equivalent to either of the following (see, *e.g.*, [26, Theorem 6.9]):

- $\nu_n \to \nu$ weakly and $\int d(o,x)^2\nu_n(dx) \to \int d(o,x)^2\nu(dx)$;
- $\int \psi\,d\nu_n \to \int \psi\,d\nu$ for any $\psi \in \mathcal{C}_2(\mathbf{X})$.

We now turn to the *third estimate*.

Lemma 5.9 (Third estimate).

$$\liminf_{s\to 0} \frac{1}{s^2}\int_0^t e^{-2R(t-u)}\big[\mathrm{Ent}_\mu P_u\tilde{g}_s - \mathrm{Ent}_\mu P_u g\big]^2 du \geqslant \int_0^t e^{-2R(t-u)}\left[\int P_u\big(gL^g f\big)\log P_u g\,d\mu\right]^2 du.$$

Proof. By the Fatou lemma, it suffices to show that

$$\liminf_{s\to 0}\left[\frac{\mathrm{Ent}_\mu P_u\tilde{g}_s - \mathrm{Ent}_\mu P_u g}{s}\right]^2 \geqslant \left[\int P_u\big(gL^g f\big)\log P_u g\,d\mu\right]^2$$

for each $u > 0$. By (5.14) and since $\mathrm{Ent}_\mu P_u g \in \mathbb{R}$, we have $P_u\tilde{g}_s \log P_u g$, $P_u g\log P_u g \in \mathbb{L}^1(\mu)$. Moreover $a^2 \geqslant (a+b)^2/(1+\delta) - b^2/\delta$ for $\delta > 0$ and

$$0 \le x\log x - x + 1 \le (x-1)^2$$

for $x \geqslant 0$, so

$$\big(\mathrm{Ent}_\mu P_u\tilde{g}_s - \mathrm{Ent}_\mu P_u g\big)^2 \geqslant \frac{1}{1+\delta}\left(\int (P_u\tilde{g}_s - P_u g)\log P_u g\, d\mu\right)^2 - \frac{1}{\delta}\left(\int \frac{(P_u\tilde{g}_s - P_u g)^2}{P_u g} d\mu\right)^2.$$

By the Cauchy-Schwarz inequality for P_u,

$$\limsup_{s\to 0} \frac{1}{s}\int \frac{(P_u\tilde{g}_s - P_u g)^2}{P_u g} d\mu \le \limsup_{s\to 0}\frac{1}{s}\int P_u\left(\frac{(\tilde{g}_s - g)^2}{g}\right) d\mu = \limsup_{s\to 0}\ s\int \left|\frac{P_s^g f - f}{s}\right|^2 g\, d\mu = 0.$$

Since $\delta > 0$ is arbitrary, it suffices to show

$$\lim_{s\to 0}\frac{1}{s}\int P_u\big(g(P_s^g f - f)\big)\log P_u g\, d\mu = \int P_u\big(gL^g f\big)\log P_u g\, d\mu \tag{5.17}$$

in order to complete the proof. Here the well-definedness of the right-hand side is included in the assertion. Since $r \mapsto r_+$ is 1-Lipschitz, $s^{-1}(P_s^g f - f)_+ = (s^{-1}(P_s^g f - f))_+$ converges to $(L^g f)_+$ in $\mathbb{L}^2(g\mu)$ and hence in $\mathbb{L}^1(g\mu)$. By [3, Theorem 4.16 (d)], $\int L^g f\, g\, d\mu = 0$. Hence $\|(L^g f)_+\|_{\mathbb{L}^1(g\mu)} > 0$ since $L^g f$ is not identically 0 (as assumed at the end of Section 5.4.1). Thus $\|(P_s^g f - f)_+\|_{\mathbb{L}^1(g\mu)} > 0$ for sufficiently small $s > 0$. Let us now define $\nu_s^f, \nu_0^f \in \mathcal{P}(\mathbf{X})$ as follows:

$$\nu_s^f := \frac{(P_s^g f - f)_+}{\|(P_s^g f - f)_+\|_{\mathbb{L}^1(g\mu)}} g\mu, \qquad \nu_0^f := \frac{(L^g f)_+}{\|(L^g f)_+\|_{\mathbb{L}^1(g\mu)}} g\mu.$$

Then $\nu_s^f \to \nu_0^f$ weakly in $\mathcal{P}(\mathbf{X})$ as $s \to 0$. Moreover, by (i) in Lemma 5.8, $\nu_s^f \in \mathcal{P}_2(\mathbf{X})$ for $s \geqslant 0$ since $f, P_s^g f, L^g f \in \mathbb{L}^2(g\mu)$. Furthermore $W_2(\nu_s^f, \nu_0^f) \to 0$ as $s \to 0$ by the remark after Lemma 5.8: for

$$\left|\int d(o,\cdot)^2 d\nu_s^f - \int d(o,\cdot)^2 d\nu_0^f\right| \le \left(\int d(o,\cdot)^4 g\, d\mu\right)^{1/2} \left\| \frac{s}{\big\|\big(P_s^g f - f\big)_+\big\|_{\mathbb{L}^1(g\mu)}}\frac{(P_s^g f - f)_+}{s} - \frac{1}{\|(L^g f)_+\|_{\mathbb{L}^1(g\mu)}}(L^g f)_+\right\|_{L^2(g\mu)} \to 0$$

as again $s^{-1}(P_s^g f - f)_+ \to (L^g f)_+$ in $\mathbb{L}^2(g\mu)$ (and hence in $\mathbb{L}^1(g\mu)$).

Then, likewise, $P_u \nu_s^f \in \mathcal{P}_2(\mathbf{X})$ for $u, s \geqslant 0$ and

$$\lim_{s\to 0} W_2\left(P_u \nu_s^f, P_u \nu_0^f\right) = 0 \tag{5.18}$$

by (5.3). By Lemma 5.8 again, $\log P_u g \in \mathcal{C}_2(\mathbf{X})$ and in particular $\log P_u g \in \mathbb{L}^1(P_u \nu_0^f)$. Hence, by (5.18) and the remark after Lemma 5.8, we obtain

$$\begin{aligned}
&\lim_{s\to 0} \frac{1}{s} \int P_u \left(g\left(P_s^g f - f\right)_+\right) \log P_u g \, d\mu \\
&= \lim_{s\to 0} \frac{\left\|\left(P_s^g f - f\right)_+\right\|_{\mathbb{L}^1(g\mu)}}{s} \int \log P_u g \, d P_u \nu_s^f \\
&= \|(L^g f)_+\|_{\mathbb{L}^1(g\mu)} \int \log P_u g \, d P_u \nu_0^f = \int P_u \left(g\left(L^g f\right)_+\right) \log P_u g \, d\mu \in \mathbb{R}.
\end{aligned}$$

We can apply the same argument to $(P_s^g f - f)_-$ instead of $(P_s^g f - f)_+$ to show the corresponding assertion. In particular, the integral in the right-hand side of (5.17) is well-defined and these two claims yield (5.17). □

5.4.3. *Conclusion of the proof of Theorem* 5.3

Let g be as in the last section, that is, given by (5.11). To proceed, we recall the notion of semigroup mollification introduced in [6, Section 2.1]. Let $\kappa \in \mathcal{C}_c^\infty((0,\infty))$ with $\kappa \geqslant 0$ and $\int_0^\infty \kappa(r)\,dr = 1$. For $\varepsilon > 0$ and $f \in \mathbb{L}^p(\mu)$ with $p \in [1,\infty]$, we define $\mathfrak{h}_\varepsilon f$ by

$$\mathfrak{h}_\varepsilon f := \frac{1}{\varepsilon} \int_0^\infty P_r f \, \kappa\left(\frac{r}{\varepsilon}\right) dr.$$

It is immediate that $\|\mathfrak{h}_\varepsilon f - f\|_{\mathcal{E}} \to 0$ as $\varepsilon \to 0$ for $f \in \mathcal{D}(\mathcal{E})$. Moreover, for $f \in \mathbb{L}^2(\mu) \cap \mathbb{L}^\infty(\mu)$, $\mathfrak{h}_\varepsilon f, L(\mathfrak{h}_\varepsilon f) \in \mathcal{D}(L) \cap \mathrm{Lip}_b(\mathbf{X})$. Here the latter one comes from the following representation:

$$L \mathfrak{h}_\varepsilon f = -\frac{1}{\varepsilon^2} \int_0^\infty P_r f \, \kappa'\left(\frac{r}{\varepsilon}\right) dr.$$

Proposition 5.10. *Following the same assumptions as in Theorem* 5.3*, let* $f = \mathfrak{h}_\varepsilon f_0$ *for some* $\varepsilon > 0$ *and* $f_0 \in \mathbb{L}^2(\mu) \cap \mathbb{L}^\infty(\mu)$*. Then* $\Gamma(f) \in \mathcal{D}(\mathcal{E})$*, and for* g *as above*

$$\frac{1}{2}\int \Gamma(\Gamma(f), g)\, d\mu + \int \Gamma(f, Lf) g \, d\mu \le -R \int \Gamma(f) g \, d\mu - \frac{1}{m}\left(\int Lf \, g \, d\mu\right)^2. \tag{5.19}$$

Proof. By assumption, $f \in \mathcal{D}(L) \cap \mathrm{Lip}_b(\mathbf{X})$. Moreover, $\Gamma(f) = |\nabla f|^2$ μ-a.e. by [6, Theorem 3.17]. Let $\eta > 0$ be so small that $\eta \|f\|_\infty \le 1/4$. By applying

Lemma 5.6, Lemma 5.7 and Lemma 5.9 to ηf instead of f in (2.3),

$$-\frac{\eta^2}{2}\int P_t\Gamma(f)\, g\, d\mu + \eta^2\int \Gamma(f, P_t f)g\, d\mu \le \frac{e^{-2Rt}\eta^2}{2(1-2\eta\|f\|_\infty)}\int \Gamma(f)g\, d\mu - \frac{\eta^2}{m}\int_0^t e^{-2R(t-u)}\Big(\int P_u((L^g f)g)\log P_u g\, d\mu\Big)^2 du.$$

By dividing this inequality by η^2 and letting $\eta \to 0$,

$$\begin{aligned}&-\frac{1}{2}\int P_t\Gamma(f)\, g\, d\mu + \int\Gamma(f, P_t f)g\, d\mu \\ &\le \frac{e^{-2Rt}}{2}\int\Gamma(f)g\, d\mu - \frac{1}{m}\int_0^t e^{-2R(t-u)}\Big(\int P_u\big((L^g f)g\big)\log P_u g d\mu\Big)^2 du.\end{aligned} \tag{5.20}$$

By virtue of mollification by $\mathfrak{h}_\varepsilon$, we have $Lf \in \mathcal{D}(\mathcal{E})$ and

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0}\int \Gamma(f, P_t f)g\, d\mu &= -\frac{1}{\varepsilon^2}\int_0^\infty \kappa'\Big(\frac{r}{\varepsilon}\Big)\int \Gamma(f, P_r f_0)g\, d\mu dr \\ &= \int \Gamma(f, Lf)g\, d\mu.\end{aligned}$$

Note that $\Gamma(f) \in \mathcal{D}(\mathcal{E})$ (hence the left-hand side of (5.19) is well-defined). This fact follows from [24, Lemma 3.2] with the aid of mollification by $\mathfrak{h}_\varepsilon$. Then, by Lemma 5.11 below, we can differentiate (5.20) at $t = 0$ to obtain

$$\begin{aligned}&\frac{1}{2}\int \Gamma(\Gamma(f), g)\, d\mu + \int \Gamma(f, Lf)g\, d\mu \\ &\le -R\int \Gamma(f)g\, d\mu - \frac{1}{m}\Big(\int (L^g f)g\log g\, d\mu\Big)^2 \\ &= -R\int \Gamma(f)g\, d\mu - \frac{1}{m}\Big(\int (Lf)g\, d\mu\Big)^2.\end{aligned}$$

Here we have used (5.15) also in the last equality. This is nothing but the desired inequality. □

Lemma 5.11. *For* $\psi \in L^2(g\mu)$,

$$\lim_{u\to 0}\int P_u(\psi g)\log P_u g\, d\mu = \int \psi g\log g\, d\mu.$$

Proof. We may assume $\psi \ge 0$ and $\psi g\mu \in \mathcal{P}(\mathbf{X})$ without loss of generality. Then in particular $\psi g\mu \in \mathcal{P}_2(\mathbf{X})$ by Lemma 5.8 (i). First of all,

$$\int P_u(\psi g)|\log P_u g|\, d\mu < \infty$$

by a similar argument as in Lemma 5.8. Thus

$$\int P_u(\psi g) \log P_u g \, d\mu = \int \psi g P_u(\log P_u g) \, d\mu \leq \int \psi g \log P_{2u} g \, d\mu$$

by the Fubini theorem and the Jensen inequality for P_u as integral operator. Now, for each x, $\lim_{u \to 0} W_2(P_u \delta_x, \delta_x) = 0$ by the remarks in the end of Section 5.1, and g is bounded and continuous, so $P_u g(x) = \int g d P_u \delta_x \to g(x)$. Moreover $\log P_{2u} g \leq \log ||g||_\infty$ and $\psi g \mu$ is a probability measure, so by the Fatou lemma

$$\limsup_{u \to 0} \int P_u(\psi g) \log P_u g \, d\mu \leq \int \psi g \log g \, d\mu. \tag{5.21}$$

For the opposite bound, again by the Jensen inequality for P_u,

$$\int P_u(\psi g) \log P_u g \, d\mu \geqslant \int P_u(\psi g) P_u(\log g) \, d\mu = \int \log g P_{2u}(\psi g) \, d\mu.$$

Moreover $\log g$ is in $\mathcal{C}_2(\mathbf{X})$ and $W_2(P_{2u}(\psi g)\mu, \psi g \mu) \to 0$ as $u \to 0$, again by the remarks in the end of Section 5.1. Hence, by the remark after Lemma 5.8, we obtain

$$\liminf_{u \to 0} \int P_u(\psi g) \log P_u g \, d\mu \geqslant \lim_{u \to 0} \int P_{2u}(\psi g) \log g \, d\mu = \int \psi g \log g \, d\mu. \tag{5.22}$$

Hence the conclusion follows from the combination of (5.21) and (5.22). □

Now we are in turn to complete the proof of Theorem 5.3.

Proof of Theorem 5.3. The last crucial step consists in transforming $(\int (Lf) g \, d\mu)^2$ into $\int (Lf)^2 g \, d\mu$ which will be done by a localization procedure. Let f be as in Proposition 5.10.

Remark first that, by letting $\lambda \to 0$ in the definition (5.11), we obtain (5.19) for g_0 instead of the function g of (5.11). To put the square inside the integral in (5.19), we need to *localize* this inequality, and thus we employ a partition of unity. Let $\eta > 0$. Since $Lf \in \mathrm{Lip}_b(\mathbf{X})$, we can take $\delta > 0$ sufficiently small so that $|Lf(x) - Lf(y)| < \eta$ for any $x, y \in \mathbf{X}$ with $d(x, y) < 4\delta$. Since $\operatorname{supp} g_0$ is compact, there is $\{x_i\}_{i=1}^n \subset \mathbf{X}$ such that $\operatorname{supp} g_0 \subset \bigcup_{i=1}^n B_\delta(x_i)$ (note that we require the regularity assumption (Reg2) only at this point). Let us define $\tilde{\psi}_i$ $(i = 1, \ldots, n)$ by $\tilde{\psi}_i(x) := 0 \vee (2\delta - d(x_i, x))$ and

$$\psi_i(x) := \begin{cases} \dfrac{\tilde{\psi}_i(x)}{\sum_{j=1}^n \tilde{\psi}_j(x)} & \text{if} \quad \tilde{\psi}_i(x) \neq 0, \\ 0 & \text{if} \quad \tilde{\psi}_i(x) = 0. \end{cases}$$

Then $\psi_i \in \mathrm{Lip}(\mathrm{supp}\, g_0), 0 \le \psi_i \le 1, \mathrm{supp}\, \psi_i \subset B_{2\delta}(x_i)$ and $\sum_{i=1}^n \psi_i(x) = 1$ for $x \in \mathrm{supp}\, g_0$. By applying (5.19) for $\psi_i g_0 / \|\psi_i g_0\|_{\mathbb{L}^1(\mu)}$ instead of g_0, we have

$$\begin{aligned}
&\frac{1}{2}\int \Gamma(\Gamma(f), g_0)\, d\mu + \int \Gamma(f, Lf) g_0\, d\mu \\
&= \sum_{i=1}^n \left(\frac{1}{2}\int \Gamma(\Gamma(f), \psi_i g_0)\, d\mu + \int \Gamma(f, Lf)\psi_i g_0\, d\mu \right) \\
&\le -R\int \Gamma(f) g_0\, d\mu - \frac{1}{m}\sum_{i=1}^n \frac{1}{\|\psi_i g_0\|_{\mathbb{L}^1(\mu)}} \left(\int (Lf)\psi_i g_0\, d\mu \right)^2 .
\end{aligned}$$

By the choice of δ and $\{\psi_i\}_{i=1}^n$, with $\eta < 1$,

$$\begin{aligned}
&\sum_{i=1}^n \frac{1}{\|\psi_i g_0\|_{\mathbb{L}^1(\mu)}} \left(\int (Lf)\psi_i g_0\, d\mu \right)^2 \geqslant (1-\eta) \sum_{i=1}^n \|\psi_i g_0\|_{\mathbb{L}^1(\mu)} Lf(x_i)^2 - \eta \\
&\geqslant (1-\eta)\int (Lf)^2 g_0\, d\mu - \eta - 2\eta(1-\eta)\|Lf\|_\infty .
\end{aligned}$$

By letting $\eta \to 0$,

$$-\frac{1}{2}\int \Gamma(\Gamma(f), g_0)\, d\mu - \int \Gamma(f, Lf) g_0\, d\mu \geqslant R\int \Gamma(f) g_0\, d\mu + \frac{1}{m}\int (Lf)^2 g_0\, d\mu .$$

Let now $g \in \mathcal{D}(L) \cap \mathbb{L}^\infty(\mu)$ with $g \geqslant 0$ and $Lg \in \mathbb{L}^\infty(\mu)$, as in Theorem 5.3. By virtue of mollification by $\mathfrak{h}_\varepsilon$, (5.1) and (5.4), we have $\Gamma(f), \Gamma(f, Lf), (Lf)^2 \in \mathbb{L}^1(\mu) \cap \mathbb{L}^\infty(\mu)$. Thus we can replace g_0 in the last inequality with $g_1 \in \mathrm{Lip}_b(\mathbf{X}) \cap \mathcal{D}(\mathcal{E})$, by a standard truncation argument. Then we can replace g_1 with g since $\mathcal{D}(\mathcal{E}) \cap \mathrm{Lip}_b(\mathbf{X})$ is dense in $\mathcal{D}(\mathcal{E})$ with respect to $\|\cdot\|_{\mathcal{E}}$.

Finally, we remove the mollification $\mathfrak{h}_\varepsilon$. Let $f \in \mathcal{D}(L)$ with $Lf \in \mathcal{D}(\mathcal{E})$ and $f_n := (-n) \vee f \wedge n$. Then we have, from the integration by parts formula (5.6),

$$\begin{aligned}
&\frac{1}{2}\int \Gamma(\mathfrak{h}_\varepsilon f_n) Lg\, d\mu - \int \Gamma(\mathfrak{h}_\varepsilon f_n, L\mathfrak{h}_\varepsilon f_n) g\, d\mu \\
&\geqslant R\int \Gamma(\mathfrak{h}_\varepsilon f_n) g\, d\mu + \frac{1}{m}\int (L\mathfrak{h}_\varepsilon f_n)^2 g\, d\mu .
\end{aligned}$$

By virtue of mollification by $\mathfrak{h}_\varepsilon$, $\|\mathfrak{h}_\varepsilon f_n - \mathfrak{h}_\varepsilon f\|_{\mathcal{E}} \to 0$ and $\|L\mathfrak{h}_\varepsilon f_n - L\mathfrak{h}_\varepsilon f\|_{\mathcal{E}} \to 0$ as $n \to \infty$. Thus we obtain (5.2) by letting $n \to \infty$ and $\varepsilon \to 0$ after it, with taking $L\mathfrak{h}_\varepsilon f = \mathfrak{h}_\varepsilon Lf$ into account. □

6. Links with functional inequalities

A new proof of the entropy-energy inequality We now consider the case where $R > 0$ and μ is a probability measure. It is classical that the $CD(R, m)$ condition implies the entropy-energy inequality

$$\mathrm{Ent}_\mu h \le \frac{m}{2} \log\left(1 + \frac{1}{mR} I(h)\right) \tag{6.1}$$

for any function h such that $\int h d\mu = 1$. Here $I(h) = \int \Gamma(h)/h \, d\mu$ is the Fisher information of h. This inequality is given in [8, Theorem 6.8.1] for instance, and also in [13, Corollary 3.28] via the (R, m)-convexity of Ent_μ.

Inequality (6.1) improves upon the standard non dimensional logarithmic Sobolev inequality $\mathrm{Ent}_\mu h \le I(h)/2R$, a consequence of the $CD(R, \infty)$ condition. It leads for example to a sharp bound on the instantaneous creation of the entropy of the heat semigroup in $\mathcal{P}_2(\mathbf{X})$, namely

$$\mathrm{Ent}_\mu P_t h \le \frac{m}{2} \log \frac{1}{1 - e^{-2Rt}}$$

for all h and $t > 0$. For similar bounds, see also [13, Proposition 2.17] for a gradient flow argument starting from the (R, m)-convexity of Ent_μ, and [10, Proposition 3.1] for Fokker-Planck equations on $\mathbb{R}^m$ with R-convex potentials.

The two approaches of [8] and [13] are rather involved, and we now give a formal (and below rigorous) and direct way of recovering (6.1) from the contraction inequality (2.3) in Theorem 2.1 (which is equivalent to the $CD(R, m)$ condition). The key point is the (formal) identity

$$\limsup_{\delta \downarrow 0} \frac{W_2^2(P_{\delta+t} h\mu, P_t h\mu)}{\delta^2} = I(P_t h) \tag{6.2}$$

(see, *e.g.*, [22, Equation (26)]) and the classical identity $\frac{d}{du}\mathrm{Ent}_\mu P_u h = -I(P_u h)$. Indeed, from inequality (2.3) and the Fatou lemma, for any $0 \le s < t$,

$$\begin{aligned}
I(P_t h) = \limsup_{\delta \downarrow 0} \frac{W_2^2(P_{t+\delta} h\mu, P_t h\mu)}{\delta^2} &\le e^{-2R(t-s)} \limsup_{\delta \downarrow 0} \frac{W_2^2(P_{s+\delta} h\mu, P_s h\mu)}{\delta^2} \\
&\quad - \frac{2}{m} \int_s^t e^{-2R(t-u)} \liminf_{\delta \downarrow 0} \left(\frac{\mathrm{Ent}_\mu P_{u+\delta} h - \mathrm{Ent}_\mu P_u h}{\delta}\right)^2 du \\
&= e^{-2R(t-s)} I(P_s h) - \frac{2}{m} \int_s^t e^{-2R(t-u)} I(P_u h)^2 du.
\end{aligned}$$

This yields the differential inequality

$$\frac{d}{dt} I(P_t h) \le -2R \, I(P_t h) - \frac{2}{m} I(P_t h)^2$$

and then

$$I(P_t h) \le \frac{mRI(h)}{e^{2Rt}(I(h) + mR) - I(h)} \tag{6.3}$$

by integration on $[0, t]$. The entropy-energy inequality (6.1) follows by further integrating (6.3) on $[0, +\infty)$ and using $\lim_{t\to\infty} \mathrm{Ent}_\mu P_t h = 0$.

Before making this argument rigorous we give a formal argument to (6.2) at $t = 0$, alternative to [22]. For simplicity, assume that $\mu = dx$ is the Riemannian measure and $(P_t)_{t\geqslant 0}$ is the heat semigroup associated with the Laplace-Beltrami operator $L = \Delta$. Let h be a probability density with respect to dx. First

$$\partial_s P_{s\delta} h + \nabla \cdot (w_s P_{s\delta} h) = 0,$$

where $w_s = -\delta \nabla \log P_{s\delta} h$. Then one can check that at the first order in δ, the couple $(P_{s\delta} h, w_s)_{s\in[0,1]}$ is optimal between $P_\delta h\mu$ and $h\mu$ in the Benamou-Brenier formulation (see [26, Chapter 7]). Hence

$$\frac{W_2^2(P_\delta h\mu, h\mu)}{\delta^2} = \int_0^1 \int |\nabla \log P_{s\delta} h|^2 P_{s\delta} h \, d\mu ds + o(1) \to I(h), \quad \delta \to 0.$$

Theorem 6.1. *In a REM space as in Section 5, the contraction inequality* (2.3) *implies the entropy-energy inequality* (6.1).

Proof. Let h be a probability density with $h\mu \in \mathcal{P}_2(\mathbf{X})$ and $I(h) < \infty$, as we can assume. Recall that $(\mathbf{X}, d, \mu)$ is a $\mathsf{RCD}(R, \infty)$ space under our assumption (2.3). Thus, by [3, Theorem 9.3 (i) and Theorem 8.5 (i)],

$$-\frac{d}{du} \mathrm{Ent}_\mu P_u h = I(P_u h) = \limsup_{\delta\downarrow 0} \frac{W_2^2(P_{u+\delta} h\mu, P_u h\mu)}{\delta^2} \tag{6.4}$$

for a.e. $u \in (0, +\infty)$. In particular, (6.2) holds almost everywhere and, proceeding as above,

$$I(P_t h) \le e^{-2R(t-s)} I(P_s h) - \frac{2}{m} \int_s^t e^{-2R(t-u)} I(P_u h)^2 du \tag{6.5}$$

for any $t > s > 0$ where (6.4) is valid.

We now prove that (6.5) holds for all $t > s \geqslant 0$. For this, set $\psi(t) := e^{2Rt} I(P_t h)$. Then ψ is non-increasing on $[0, \infty)$ by a standard argument: indeed, by $CD(R, \infty)$ with the self-improvement argument in [24], we have $\sqrt{\Gamma(P_t h)} \le e^{-Rt} P_t(\sqrt{\Gamma(h)})$ for all $t \geqslant 0$. It yields

$$\frac{\Gamma(P_t h)}{P_t h} \le e^{-2R(t-s)} \frac{\big(P_{t-s}(\sqrt{\Gamma(P_s h)})\big)^2}{P_{t-s}(P_s h)} \le e^{-2R(t-s)} P_{t-s}\left(\frac{\Gamma(P_s h)}{P_s h}\right).$$

Thus the claim follows by integrating this inequality by μ. Moreover $t \mapsto I(P_t h)$ is lower semi-continuous (see, *e.g.*, [3, Lemma 4.10]). Thus ψ is lower semi-continuous and non-increasing on $[0, \infty)$, so also right-continuous. This implies that (6.5) holds for $t > s \geqslant 0$.

Let now $\delta > 0$. By dividing (6.5) by $e^{-2Rt}(\psi(t)+\delta)(\psi(s)+\delta)$, for $t > s > 0$,

$$\frac{2}{m(\psi(s)+\delta)(\psi(t)+\delta)}\int_s^t e^{-2Ru}\psi(u)^2\,du \leq \frac{1}{\psi(t)+\delta} - \frac{1}{\psi(s)+\delta}. \tag{6.6}$$

We claim

$$\frac{2(1-\delta)}{m}\int_0^t e^{-2Ru}\left(\frac{\psi(u)}{\psi(u)+\delta}\right)^2 du \leq \frac{1}{\psi(t)+\delta} - \frac{1}{\psi(0)+\delta} \tag{6.7}$$

for any $t \in [0,\infty)$. For the proof of the claim, we let J be the subset of $t \in [0,\infty)$ satisfying (6.7) and prove $J = [0,\infty)$. First, $0 \in J$ obviously holds and hence $J \neq \emptyset$. Second, if $t \in J$ and $t' \in (t,\infty)$ with $t'-t$ sufficiently small, then $t' \in J$. Indeed, by the right continuity of ψ, we have $\psi(u)+\delta \geqslant (1-\delta)(\psi(t)+\delta)$ for any $u > t$ being sufficiently close to t. We take $t' > t$ so that this holds for all $u \in (t,t')$. Thus (6.7) for this t, (6.6) and ψ being non-increasing yield

$$\begin{aligned}
&\frac{2(1-\delta)}{m}\int_0^{t'} e^{-2Ru}\left(\frac{\psi(u)}{\psi(u)+\delta}\right)^2 du\\
\leq&\frac{1}{\psi(t)+\delta} - \frac{1}{\psi(0)+\delta} + \frac{2}{m(\psi(t)+\delta)(\psi(t')+\delta)}\int_t^{t'} e^{-2Ru}\psi(u)^2\,du\\
\leq&\frac{1}{\psi(t')+\delta} - \frac{1}{\psi(0)+\delta}
\end{aligned}$$

and hence $t' \in J$. Third, J is closed under increasing sequences. That is, for any bounded increasing sequence $(t_n)_{n\in\mathbb{N}}$ in J, then $\lim_{n\to\infty} t_n \in J$. This property follows from the fact that ψ is lower semi-continuous. Now these three properties imply $J = [0,\infty)$ and hence the claim holds.

Finally we obtain (6.3) for all $t \geqslant 0$ by taking $\delta \downarrow 0$ and rearranging terms in (6.7). But

$$\mathrm{Ent}_\mu h - \mathrm{Ent}_\mu P_t h = \int_0^t I(P_s h)ds \tag{6.8}$$

for all t by [3, Theorem 9.3 (i) and 8.5 (i)] again. Hence integrating (6.3) in t concludes the proof. □

A dimensional HWI type inequality For R being 0 or negative, no logarithmic Sobolev inequality for μ holds in general, and following [22] it can be replaced by a HWI interpolation inequality with an additional W_2 term: this is inequality giving an upper bound on the entropy H in terms of the distance W_2 and the Fisher information I. As above, let us see how to derive a dimensional form of this inequality from the contraction property (2.2) in Theorem 2.1.

In a REM space as in Section 5, with a reference measure μ in $\mathcal{P}_2(\mathbf{X})$, assume the contraction property (2.2) with $R = 0$. Let g, h such that $g\mu, h\mu \in \mathcal{P}_2(\mathbf{X})$, $I(h) < \infty$ and $g\mu$ has bounded support. Recall first that $(\mathbf{X}, d, \mu)$ is a $\mathsf{RCD}(0,\infty)$

space under our assumption (2.2). In particular $I(P_t h) \leq I(h)$ for all $t \geqslant 0$. Then [1, Theorem 6.3] and the Cauchy-Schwarz inequality yield

$$\frac{1}{2}\frac{d}{dt}W_2^2(P_t h\mu, g\mu) \geqslant -W_2(P_t h\mu, g\mu)\sqrt{I(P_t h)}$$

for almost every $t > 0$. In particular

$$\begin{aligned}\frac{1}{2}W_2^2(P_t h\mu, g\mu) - \frac{1}{2}W_2^2(h\mu, g\mu) &\geqslant -\int_0^t W_2(P_s h\mu, g\mu)\sqrt{I(P_s h)}\, ds \\ &\geqslant -\int_0^t W_2(P_s h\mu, g\mu)\sqrt{I(h)}\, ds\end{aligned}$$

for all $t \geqslant 0$.

If now g converges to 1 in such a way that $g\mu$ converges to μ in the W_2 distance, then using the triangular inequality

$$\left|W_2(P_s h\mu, g\mu) - W_2(P_s h\mu, \mu)\right| \leq W_2(g\mu, \mu)$$

for any $0 \leq s \leq t$ one can pass to the limit above, leading to

$$\frac{1}{2}W_2^2(P_t h\mu, \mu) - \frac{1}{2}W_2^2(h\mu, \mu) \geqslant -\int_0^t W_2(P_s h\mu, \mu)\sqrt{I(h)}\, ds.$$

Now by (2.2) the left-hand side is bounded from above by

$$-4m\int_0^t \sinh^2\left(\frac{\mathrm{Ent}_\mu P_s h}{2m}\right) ds.$$

Finally $s \mapsto W_2(P_s h\mu, \mu)$ and $s \mapsto \mathrm{Ent}_\mu P_s h$ are continuous on $[0, t]$, so one can let t go to 0 and obtain

$$\sinh^2\left(\frac{\mathrm{Ent}_\mu h}{2m}\right) \leq \frac{1}{4m}W_2(h\mu, \mu)\sqrt{I(h)}. \tag{6.9}$$

In [13], they proved the HWI inequality

$$\exp\left(\frac{\mathrm{Ent}_\mu g_1 - \mathrm{Ent}_\mu g_0}{m}\right) \leq 1 + \frac{1}{m}W_2(g_1\mu, g_0\mu)\sqrt{I(g_1)},$$

and we can obtain (6.9) also from this inequality by considering the cases $(g_0, g_1) = (h, 1)$ and $(g_0, g_1) = (1, h)$ in this inequality and summing them up. The inequality obtained in the case $(g_0, g_1) = (1, h)$ is in fact better than (6.9) (and has the same behaviour for large $\mathrm{Ent}_\mu h$).

Here is a possible application of (6.9): in the above notation and assumptions (with $R = 0$), there exists a positive numerical constant C such that

$$\mathrm{Ent}_\mu P_t h \leq \frac{m}{2} \max \left\{ C, \log \frac{W_2^2(h\mu, \mu)}{mt} \right\}, \quad t > 0$$

for all h with $h\mu \in \mathcal{P}_2$. This bound is a consequence of (6.8), (6.9) with $P_t h$ instead of h, the bounds $W_2(P_t h\mu, \mu) \leq W_2(h\mu, \mu)$ and $\sinh^4(x) \geqslant e^{4x}/32$ for x large enough.

For short time, this gives a regularization bound of the entropy as $m/2 \log(1/t)$, which is exactly the behaviour observed above for $R > 0$, and also for the heat kernel on $\mathbb{R}^m$; it also improves on the corresponding bound $m \log(1/t)$ in [13, Proposition 2.17, (ii)].

References

[1] L. AMBROSIO, N. GIGLI, A. MONDINO and T. RAJALA, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure*, Trans. Amer. Math. Soc. **367** (2015), 4661–4701.

[2] L. AMBROSIO, N. GIGLI and G. SAVARÉ, "Gradient Flows in Metric Spaces and in the Space of Probability Measures", Lectures in Math. ETH Zürich. Birkhäuser, Basel, 2008.

[3] L. AMBROSIO, N. GIGLI and G. SAVARÉ, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math. **195** (2014), 289–391.

[4] L. AMBROSIO, N. GIGLI and G. SAVARÉ, *Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces*, Rev. Mat. Iberoam. **29** (2013), 969–996.

[5] L. AMBROSIO, N. GIGLI and G. SAVARÉ, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J. **163** (2014), 1405–1490.

[6] L. AMBROSIO, N. GIGLI and G. SAVARÉ, *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*, Ann. Probab. **43** (2015), 339–404.

[7] D. BAKRY and M. ÉMERY, *Diffusions hypercontractives*, In: "Séminaire de Probabilités, XIX, 1983/84", Lecture Notes in Math. 1123, Springer, Berlin, 1985, 177–206.

[8] D. BAKRY, I. GENTIL and M. LEDOUX, "Analysis and Geometry of Markov Diffusion Operators", Vol. 348 of Grund. Math. Wiss., Springer, Cham, 2014.

[9] D. BAKRY, I. GENTIL and M. LEDOUX, *On Harnack inequalities and optimal transportation*, Ann. Sc. Norm. Super. Pisa, Cl. Sci. **14** (2015), 705–727.

[10] F. BOLLEY, I. GENTIL and A. GUILLIN, *Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities*, Ann. Probab. **46** (2018), 261–301.

[11] F. BOLLEY, I. GENTIL and A. GUILLIN, *Dimensional contraction via Markov transportation distance*, J. Lond. Math. Soc. (2) **90** (2014), 309–332.

[12] D. BURAGO, YU. BURAGO and S. IVANOV, "A Course in Metric Geometry", Grad. studies math., Vol. 33, Amer. Math. Soc., Providence, 2001.

[13] M. ERBAR, K. KUWADA and K.-T. STURM, *On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces*, Invent. Math. **201** (2015), 993–1071.

[14] I. GENTIL, *Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold*, Potential Anal. **42** (2015), 861–873.

[15] N. GIGLI, K. KUWADA and S.-I. OHTA, *Heat flow on Alexandrov spaces*, Comm. Pure Appl. Math. **66** (2013), 307–331.

[16] R. JORDAN, D. KINDERLEHRER and F. OTTO, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), 1–17.

[17] K. KUWADA, *Duality on gradient estimates and Wasserstein controls*, J. Funct. Anal. **258** (2010), 3758–3774.

[18] K. KUWADA, *Space-time Wasserstein controls and Bakry-Ledoux type gradient estimates*, Calc. Var. Partial Differential Equations **54** (2015), 127–161.

[19] H.-Q. LI, *Dimension-free Harnack inequalities on RCD* (K,∞) *spaces*, J. Theoret. Probab. **29** (2016), 1280–1297.

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

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

[22] F. OTTO and C. VILLANI, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), 361–400.

[23] F. OTTO and M. WESTDICKENBERG, *Eulerian calculus for the contraction in the Wasserstein distance*, SIAM J. Math. Anal. **37** (2005), 1227–1255.

[24] G. SAVARÉ, *Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in RCD*(K,∞) *metric measure spaces*, Discrete Contin. Dyn. Syst. A **34** (2014), 1641–1661.

[25] K.-T. STURM, *On the geometry of metric measure spaces. I and II*, Acta Math. **196** (2006), 65–177.

[26] C. VILLANI, "Optimal Transport, Old and New", Grund. Math. Wiss., Vol. 338, Springer, Berlin, 2009.

[27] M.-K. VON RENESSE and K.-T. STURM, *Transport inequalities, gradient estimates, entropy and Ricci curvature*, Comm. Pure Appl. Math. **68** (2005), 923–940.

[28] F.-Y. WANG, "Functional Inequalities, Markov Processes, and Spectral Theory", Science Press, Beijing, 2004.

[29] F.-Y. WANG, *Equivalent semigroup properties for the curvature-dimension condition*, Bull. Sci. Math. **135** (2011), 803–815.

Laboratoire de Probabilités
et Modèles Aléatoires
Umr Cnrs 7599
Univ. Pierre et Marie Curie
Paris 6, France
francois.bolley@upmc.fr

Institut Camille Jordan
Umr Cnrs 5208
Univ. Claude Bernard
Lyon 1, France
gentil@math.univ-lyon1.fr

Laboratoire de Mathématiques
Umr Cnrs 6620
Univ. Clermont Auvergne, France
guillin@math.univ-bpclermont.fr

Tokyo Institute of Technology
Japan
kuwada@math.titech.ac.jp