ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Advances in Mathematics 231 (2012) 2755–2801

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

Geometry and analysis of Dirichlet forms

Pekka Koskela[a], Yuan Zhou[a,b,*]

[a] *Department of Mathematics and Statistics, P.O. Box 35 (MaD), FI-40014, University of Jyväskylä, Finland*
[b] *Department of Mathematics, Beijing University of Aeronautics and Astronautics, Beijing 100191, PR China*

Received 3 November 2011; accepted 6 August 2012
Available online 7 September 2012

Communicated by D. Stroock

Abstract

Let $\mathscr{E}$ be a regular, strongly local Dirichlet form on $L^2(X, m)$ and d the associated intrinsic distance. Assume that the topology induced by d coincides with the original topology on X, and that X is compact, satisfies a doubling property and supports a weak $(1, 2)$-Poincaré inequality. We first discuss the (non-) coincidence of the intrinsic length structure and the gradient structure. Under the further assumption that the Ricci curvature of X is bounded from below in the sense of Lott–Sturm–Villani, the following are shown to be equivalent:

(i) the heat flow of $\mathscr{E}$ gives the unique gradient flow of $\mathscr{U}_\infty$,
(ii) $\mathscr{E}$ satisfies the Newtonian property,
(iii) the intrinsic length structure coincides with the gradient structure.

Moreover, for the standard (resistance) Dirichlet form on the Sierpinski gasket equipped with the Kusuoka measure, we identify the intrinsic length structure with the measurable Riemannian and the gradient structures. We also apply the above results to the (coarse) Ricci curvatures and asymptotics of the gradient of the heat kernel.

Keywords: Dirichlet form; Intrinsic distance; Length structure; Differential structure; Sierpinski gasket; Gradient flow; Ricci curvature; Poincaré inequality; Metric measure space

* Corresponding author at: Department of Mathematics, Beijing University of Aeronautics and Astronautics, Beijing 100191, PR China.
E-mail addresses: pkoskela@maths.jyu.fi (P. Koskela), yuanzhou@buaa.edu.cn (Y. Zhou).

0001-8708/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.aim.2012.08.004

Contents

1. Introduction

It is well known that on $\mathbb{R}^n$, associated to the Dirichlet energy

$$\int_{\mathbb{R}^n} |\nabla f(x)|^2 dx,$$

there is a naturally defined heat semigroup (flow). Jordan et al. [17] and Otto [35] understood this heat flow as a gradient flow of the Boltzmann–Shannon entropy with respect to the L^2-Wasserstein metric on the space of probability measures on $\mathbb{R}^n$. Since then this has been extended to Riemannian manifolds, Finsler manifolds, Heisenberg groups, Alexandrov spaces and metric measure spaces; see, for example, [35,1,51,9,18,33,13,2]. The gradient flow has also attracted considerable attention in various settings; see, for example, [1,13,51,12] and the reference therein. In particular, the works [1,12,13] in abstract setting motivate one to extend the above phenomenon of [17] to settings such as metric measure spaces with Ricci curvatures of Lott et al. [49,50,29] bounded from below.

Moreover, a heat semigroup (flow) is naturally associated to any given Dirichlet form. Via this, a notion of Ricci curvature bounded from below was introduced by Bakry and Emery [4]. Observe that the Ricci curvature of Bakry–Emery essentially depends on the differential (gradient) structure. On the other hand, under some additional assumptions on the underlying metric measure space, a notion of Ricci curvature bounded from below was introduced by Lott et al. [29,49,50], purely in terms of the length structure. It is then natural to analyze the connections between these different approaches; see [13,2] for seminal studies in this direction. In this paper, we consider the intrinsic length structures and gradient structures of Dirichlet forms.

Let X be a locally compact, connected and separable Hausdorff space and m a nonnegative Radon measure with support X. Let $\mathscr{E}$ be a regular, strongly local Dirichlet form on $L^2(X)$, Γ the squared gradient and d the intrinsic distance induced by $\mathscr{E}$. We always assume that the topology induced by d coincides with the original topology on X.

In Section 2, we establish the coincidence of the intrinsic length structure and the gradient structure of Dirichlet forms under a doubling property, a weak Poincaré inequality and the Newtonian property. Indeed, we prove that if (X, d, m) satisfies the doubling property, then for every $u \in \mathrm{Lip}(X)$, the energy measure $\Gamma(u,u)$ is absolutely continuous with respect to m and $\frac{d}{dm}\Gamma(u,u) \le (\mathrm{Lip}\, u)^2$ almost everywhere; see Theorem 2.1. If we further assume that (X, d, m) supports a weak $(1,p)$-Poincaré inequality for some $p \in [1,\infty)$ and that $(X, \mathscr{E}, m)$ satisfies the Newtonian property introduced in this paper, then $\frac{d}{dm}\Gamma(u,u) = (\mathrm{Lip}\, u)^2$ almost everywhere; see Theorem 2.2.

In Section 3, by perturbing the classical Dirichlet energy form of $\mathbb{R}^2$ on a large Cantor set, we construct a simple example that satisfies a doubling property and a weak Poincaré inequality, but so that the intrinsic length structure does not coincide with the gradient structure; see Proposition 3.1. This shows that a doubling property and a weak Poincaré inequality are not sufficient to guarantee that $\frac{d}{dm}\Gamma(u,u) = (\mathrm{Lip}\, u)^2$ almost everywhere. A more general construction can be found in [47]. Moreover, the gradient (differential) structure of our perturbed Dirichlet form does not coincide with the distinguished gradient (differential) structure of Cheeger's; see Proposition 3.2. Recall that if (X, d, m) satisfies a doubling property and a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$, then Cheeger [7] constructed a differential structure equipped with a distinguished inner product norm, which coincides with the gradient structure of Γ if $(X, \mathscr{E}, m)$ further satisfies the Newtonian property; see Corollary 3.1.

In Section 4, with the aid of the above results, for the standard (resistance) Dirichlet form on the standard Sierpinski gasket equipped with the Kusuoka measure, we identify the intrinsic length structure with the measurable Riemannian and the gradient structures. In particular, some refined Rademacher theorems are established. See Theorems 4.1–4.3.

In Section 5, we assume that (X, d, m) is compact and satisfies a doubling property. If the entropy $\mathscr{U}_\infty$ is weakly λ-displacement convex for some $\lambda \in \mathbb{R}$, then we obtain the equivalence of the following:

(i) for all Lipschitz functions u, $\frac{d}{dm}\Gamma(u,u) = (\mathrm{Lip}\, u)^2$ almost everywhere,
(ii) $(X, \mathscr{E}, m)$ satisfies the Newtonian property,
(iii) the heat flow of $\mathscr{E}$ gives the unique gradient flow of $\mathscr{U}_\infty$;

see Theorems 5.1 and 5.2. Recall that the existence and uniqueness of the gradient flow of $\mathscr{U}_\infty$ was already established in [1,12].

In Section 6, applying the results of Section 2, we first obtain a dual formula related to Kuwada's dual theorem and the boundedness from below of the coarse Ricci curvature of Ollivier [34]; this does not require the Newtonian property. Moreover, with some additional assumptions, relying on [39], we obtain that if the Ricci curvature of (X, d) is bounded from below in the sense of Lott et al. [49,50,29], then the Ricci curvature of $(X, \mathscr{E})$ is bounded from below in the sense of Bakry–Emery [3,4].

In Section 7, assuming that $(X, \mathscr{E}, m)$ is compact and has a spectral gap, we show that the identity $\Gamma(d_x, d_x) = m$ for all $x \in X$ actually reflects some short time asymptotics of the gradient of the heat kernel.

Finally, we state some *conventions*. Throughout the paper, we denote by C a *positive constant* which is independent of the main parameters, but which may vary from line to line. Constants with subscripts, such as C_0, do not change in different occurrences. The *notation* $A \lesssim B$ or $B \gtrsim A$ means that $A \le CB$. If $A \lesssim B$ and $B \lesssim A$, we then write $A \sim B$. Denote by $\mathbb{N}$ the *set of positive integers*. For any locally integrable function f, we denote by $\fint_E f\, d\mu$ the *average of f on E*, namely, $\fint_E f\, d\mu \equiv \frac{1}{\mu(E)} \int_E f\, d\mu$.

2. Dirichlet forms: $\frac{d}{dm}\Gamma(u,u) = (\mathrm{Lip}\, u)^2$

The main aim of this section is to establish the coincidence of the intrinsic length structure and the gradient structure of Dirichlet forms under a doubling property, a weak Poincaré inequality and the Newtonian property; see Theorems 2.1 and 2.2.

Let X be a locally compact, connected and separable Hausdorff space and m be a nonnegative Radon measure with support X. In this paper, $L^p(X)$ with $p \in (1, \infty]$ is the space of integrable functions of order p on X; $\mathscr{C}(X)$ (resp. $\mathscr{C}_0(X)$) the collection of all continuous functions (with compact supports) on X, and $\mathscr{M}(X)$ the collection of all signed Radon measures on X.

Recall that a *Dirichlet form* $\mathscr{E}$ on $L^2(X)$ is a closed, nonnegative definite and symmetric bilinear form defined on a dense linear subspace $\mathbb{D}$ of $L^2(X)$, that satisfies the *Markov property*: for any $u \in \mathbb{D}$, $v = \min\{1, \max\{0, u\}\}$, we have $\mathscr{E}(v, v) \le \mathscr{E}(u, u)$. Then $\mathscr{E}$ is said to be *strongly local* if $\mathscr{E}(u, v) = 0$ whenever $u, v \in \mathbb{D}$ with u a constant on a neighborhood of the support of v; to be *regular* if there exists a subset of $\mathbb{D} \cap \mathscr{C}_0(X)$ which is both dense in $\mathscr{C}_0(X)$ with uniform norm and in $\mathbb{D}$ with the norm $\|\cdot\|_{\mathbb{D}}$ defined by $\|u\|_{\mathbb{D}} = [\|u\|^2_{L^2(X)} + \mathscr{E}(u, u)]^{1/2}$ for each $u \in \mathbb{D}$. Beurling and Deny [6] showed that a regular, strongly local Dirichlet form $\mathscr{E}$ can be written as

$$\mathscr{E}(u, v) = \int_X d\Gamma(u, v)$$

for all $u, v \in \mathbb{D}$, where Γ is an $\mathscr{M}(X)$-valued nonnegative definite and symmetric bilinear form defined by the formula

$$\int_X \phi \, d\Gamma(u, v) \equiv \frac{1}{2}\left[\mathscr{E}(u, \phi v) + \mathscr{E}(v, \phi u) - \mathscr{E}(uv, \phi)\right] \tag{2.1}$$

for all $u, v \in \mathbb{D} \cap L^\infty(X)$ and $\phi \in \mathbb{D} \cap \mathscr{C}_0(X)$. We call $\Gamma(u, v)$ the *Dirichlet energy measure (squared gradient)* and $\sqrt{\frac{d}{dm}\Gamma(u, u)}$ the *length of the gradient*.

Observe that, since $\mathscr{E}$ is strongly local, Γ is local and satisfies the Leibniz rule and the chain rule; see for example [11]. Then both $\mathscr{E}(u, v)$ and $\Gamma(u, v)$ can be defined for $u, v \in \mathbb{D}_{\mathrm{loc}}$, the *collection of all* $u \in L^2_{\mathrm{loc}}(X)$ *satisfying* that for each relatively compact set $K \subset X$, there exists a function $w \in \mathbb{D}$ such that $u = w$ almost everywhere on K. With this, the *intrinsic distance on* X *associated to* $\mathscr{E}$ is defined by

$$d(x, y) \equiv \sup\{u(x) - u(y) : u \in \mathbb{D}_{\mathrm{loc}} \cap \mathscr{C}(X),\ \Gamma(u, u) \le m\}. \tag{2.2}$$

Here $\Gamma(u, u) \le m$ means that $\Gamma(u, u)$ is absolutely continuous with respect to m and $\frac{d}{dm}\Gamma(u, u) \le 1$ almost everywhere.

In this paper, we always assume that $\mathscr{E}$ is a regular, strongly local Dirichlet form on $L^2(X)$, and that the topology induced by d is equivalent to the original topology on X. Notice that, under this assumption, d is a distance, $d(x, y) < \infty$ for all $x, y \in X$, and (X, d) is a length space; see [44,48,42].

For such a space, the very first question is the coincidence of the gradient structure of Γ and the length structure of d. It is well known that for all $x \in X$, $\Gamma(d_x, d_x) \le m$ as proved in [44]. Very recently, it was observed in [10] (see also [42]) that, for $u \in \mathrm{Lip}(X)$ with Lipschitz constant 1, we have $\Gamma(u, u) \le m$. Moreover, under a doubling assumption, we are able to establish a pointwise relation between $\frac{d}{dm}\Gamma(u, u)$ and $\mathrm{Lip}\, u$ as follows. Here and in what follows, for a measurable function u, its *pointwise Lipschitz constant* is defined as

$$\mathrm{Lip}\, u(x) \equiv \limsup_{y \to x} \frac{|u(x) - u(y)|}{d(x, y)},$$

and $\mathrm{Lip}(X)$ stands for the *collection of all measurable functions* u with

$$\|u\|_{\mathrm{Lip}(X)} \equiv \sup_{x, y \in X, x \ne y} \frac{|u(x) - u(y)|}{d(x, y)} < \infty.$$

When it is necessary, we also write Lip as Lip_d to specify the distance d. We say that (X, d, m) satisfies a *doubling property* if there exists a constant $C_0 > 1$ such that for all $x \in X$ and $r > 0$,

$$m(B(x, 2r)) \le C_0 m(B(x, r)) < \infty. \tag{2.3}$$

Theorem 2.1. *Suppose that (X, d, m) satisfies a doubling property. Then $\mathrm{Lip}(X) \subset \mathbb{D}_{\mathrm{loc}}$ and for every $u \in \mathrm{Lip}(X)$, $\Gamma(u, u) \le (\mathrm{Lip}\, u)^2 m$, that is, $\Gamma(u, u)$ is absolutely continuous with respect to m and*

$$\frac{d}{dm}\Gamma(u, u) \le (\mathrm{Lip}\, u)^2$$

almost everywhere.

The proof of Theorem 2.1 relies on the following three auxiliary lemmas.

Lemma 2.1. *For $n \in \mathbb{N}$, $E = \{x_i\}_{i=1}^n \subset X$ and $A = \{a_i\}_{i=1}^n \subset \mathbb{R}$, set*

$$d_{A,E}(x) \equiv \max_{i=1,\dots,n} \{a_i - d(x_i, x)\}.$$

Then $\Gamma(d_{A,E}, d_{A,E}) \le m$.

Moreover, if $\Gamma(d_x, d_x) = m$ for every $x \in X$, then $\Gamma(d_{A,E}, d_{A,E}) = m$.

Proof. We prove this by induction. It is easy to see that if $n = 1$, then from $\Gamma(a_1, v) = 0$ for all $v \in \mathbb{D}_{\mathrm{loc}}$ and from $\Gamma(d_{x_1}, d_{x_1}) \le m$ proven in [44], we deduce that

$$\Gamma(a_1 - d_{x_1}, a_1 - d_{x_1}) = \Gamma(a_1, a_1 - d_{x_1}) - \Gamma(d_{x_1}, a_1 - d_{x_1}) = \Gamma(d_{x_1}, d_{x_1}) \le m. \tag{2.4}$$

Now assume that the claim holds for n. We are going to prove it for $n + 1$. To this end, let $E_{n+1} = \{x_i\}_{i=1}^{n+1} \subset X$ and $A_{n+1} \subset \{a_i\}_{i=1}^{n+1} \in \mathbb{R}$. Notice that

$$\begin{aligned} d_{A_{n+1},E_{n+1}} &= \max_{i=1,\dots,n+1} \left\{a_i - d_{x_i}\right\} \\ &= \max\left\{\max_{i=1,\dots,n} \left\{a_i - d_{x_i}\right\}, a_{n+1} - d_{x_{n+1}}\right\} \\ &= \max\left\{d_{A_n,E_n}, a_{n+1} - d_{x_{n+1}}\right\}, \end{aligned}$$

where $A_n = A_{n+1} \setminus \{a_{n+1}\}$ and $E_n = E_{n+1} \setminus \{x_{n+1}\}$. Recall that the following truncation property was proven in [44]:

$$\Gamma(u \wedge v, u \wedge v) = 1_{u<v}\Gamma(u, u) + 1_{u\ge v}\Gamma(v, v),$$

where $u \wedge v = \min\{u, v\}$, and 1_F refers to the characteristic function of F. Denote $u \vee v = \max\{u, v\}$. Then we have

$$\Gamma(u \vee v, u \vee v) = \Gamma((-u) \wedge (-v), (-u) \wedge (-v)) = 1_{u>v}\Gamma(u, u) + 1_{u\le v}\Gamma(v, v),$$

and moreover, if $\Gamma(u, u) \le m$ and $\Gamma(v, v) \le m$, then $\Gamma(u \vee v, u \vee v) \le m$. Now

$$\Gamma(d_{A_n,E_n}, d_{A_n,E_n}) \le m$$

by induction and

$$\Gamma(a_{n+1} - d_{x_{n+1}}, a_{n+1} - d_{x_{n+1}}) \le m$$

by (2.4). Hence, we have

$$\Gamma(d_{A_{n+1},E_{n+1}}, d_{A_{n+1},E_{n+1}}) \le m,$$

as desired.

Moreover, if $\Gamma(d_x, d_x) = m$ for every $x \in \mathscr{X}$, then (2.4) holds with $\le$ replaced by $=$. With this, by induction, we further obtain $\Gamma(d_{A,E}, d_{A,E}) = m$. □

Lemma 2.2. *Let $V \subset X$ be a bounded open set. Define $u(x) \equiv \sup_{z\in V}\{v(z) - d(z,x)\}$. If $v \in \mathbb{D}_{\mathrm{loc}}$, $1_V \Gamma(v,v) \le 1_V m$ and $\|v\|_{\mathrm{Lip}(V)} \le 1$, then $\Gamma(u,u) \le m$.*

Proof. For every $n \in \mathbb{N}$, choose a maximal finite set of V, $\{x_{n,i}\} \subset V$, such that $d(x_{n,i}, x_{n,j}) \ge \frac{1}{n}\operatorname{diam} V$, and for all $x \in X$, set

$$u_n \equiv \max_i\{v(x_{n,i}) - d(x_{n,i}, x)\}.$$

Then, by Lemma 2.1, $\Gamma(u_n, u_n) \le m$, which implies that $\{u_n\}_{n\in\mathbb{N}}$ is a locally bounded set and hence has a subsequence which converges weakly in $\mathbb{D}_{\mathrm{loc}}$ to some u_0. Without loss of generality, we still denote this subsequence by $\{u_n\}_{n\in\mathbb{N}}$. Now $\Gamma(u_0,u_0) \le \lim_{n\to\infty}\Gamma(u_n,u_n) \le m$.

It suffices to show that $u = u_0$. To see this, we first notice that $u_n(x) \le u(x)$ for all $x \in X$. On the other hand, obviously, for all $x \in V$, $u(x) = v(x)$ and for all i, $u(x_{n,i}) = v(x_{n,i}) = u_n(x_{n,i})$. For any $x \in X$, there exists $z \in V$ such that $u(x) \le u(z) - d(z,x) + \frac{1}{n}\operatorname{diam} V$. By the choice of $x_{n,i}$, we can find $x_{n,i} \in B(z, \frac{2}{n}\operatorname{diam} V)$. Since $\|v\|_{\mathrm{Lip}(V)} \le 1$, we have $|u(z) - u(x_{n,i})| = |v(z) - v(x_{n,i})| \le d(x, x_{n,i})$. Hence

$$\begin{aligned}
u(x) &\le u(z) - d(z,x) + \frac{1}{n}\operatorname{diam} V\\
&= u(x_{n,i}) - d(x_{n,i},x) + u(z) - u(x_{n,i}) - d(z,x) + d(x_{n,i},x) + \frac{1}{n}\operatorname{diam} V\\
&\le u_n(x) + 2d(z,x_{n,i}) + \frac{1}{n}\operatorname{diam} V\\
&\le u_n(x) + \frac{1}{n}\operatorname{diam} V.
\end{aligned}$$

So $u_n \to u$ uniformly. Thus $u_n \to u = u_0$ weakly in $\mathbb{D}_{\mathrm{loc}}$, which implies that

$$\Gamma(u,u) \le \liminf_{n\to\infty}\Gamma(u_n,u_n) \le m.$$

This finishes the proof Lemma 2.2. □

The following lemma was established in [7, Lemma 6.30]. Its proof uses the Lusin theorem and relies on decay property of a doubling measure on a length space observed in [8].

Lemma 2.3. *Suppose that (X, d, m) satisfies a doubling property. Then for every ball $B(x_0, r_0) \subset X$, there exists a constant $C_2 \ge 1$ such that for every $n \in \mathbb{N}$ and $u \in \mathrm{Lip}(B(x_0, r_0))$, there exists a finite collection $\{B(x_{n,j}, r_{n,j})\}$ of mutually disjoint balls with $x_{n,j} \in B(x_0, r_0)$ and $r_{n,j} \le r_0$ satisfying that*

$$\operatorname{dist}(B(x_{n,i}, r_{n,i}), B(x_{n,j}, r_{n,j})) \ge \frac{1}{2}(r_{n,i} + r_{n,j}), \tag{2.5}$$

$$m(B(x_0, r_0) \setminus \cup_j B(x_{n,j}, r_{n,j})) \le C_2\frac{1}{n}m(B(x_0, r_0)), \tag{2.6}$$

$$\int_{B(x_{n,j}\,3r_{n,j})} |\mathrm{Lip}\, u(x) - \mathrm{Lip}\, u(x_{n,j})|^2\, dm \le \frac{1}{n} m(B(x_{n,j}\,3r_{n,j})) \tag{2.7}$$

and so for all $x, y \in B(x_{n,j}, r_{n,j})$ *with* $d(x, y) \ge \frac{1}{n} r_{n,j}$,

$$\frac{|u(x) - u(y)|}{d(x, y)} < \mathrm{Lip}\, u(x_{n,j}) + \frac{1}{n}. \tag{2.8}$$

Proof of Theorem 2.1. Let $u \in \mathrm{Lip}(X)$. It suffices to prove that for every ball $B(x_0, r_0) \subset X$,

$$\int_X 1_{B(x_0,r_0)}\, d\Gamma(u, u) \le \int_X 1_{B(x_0,r_0)} (\mathrm{Lip}\, u)^2\, dm. \tag{2.9}$$

Indeed, by this and a covering argument, one can show that $\Gamma(u, u)$ is absolutely continuous with respect to m, and $\frac{d}{dm}\Gamma(u, u) \le (\mathrm{Lip}\, u)^2$ almost everywhere. We omit the details.

To prove (2.9), we need the following construction via the MacShane extension, which is a slight modification of that in [7]. For $n \in \mathbb{N}$, let $\{B(x_{n,j}, r_{n,j})\}$ be the covering provided by Lemma 2.3. For every j, we choose a maximal set $\{z_{n,j,k}\} \subset B(x_0, r_0)$ such that for $k \ne \ell$,

$$d(z_{n,j,k}, z_{n,j,\ell}) \ge \frac{1}{n} r_{n,j}.$$

Define a function u_n on $\cup_j B(x_{n,j}, r_{n,j})$ as follows: for $x \in B(x_{n,j}, r_{n,j})$, set

$$u_n(x) \equiv \max_k \{u(z_{n,j,k}) - L_j d(z_{n,j,k}, x)\},$$

where $L_{n,j} \equiv \mathrm{Lip}\, u(x_{n,j}) + \frac{1}{n}$, and for $x \in X \setminus \cup_j B(x_{n,j}, r_{n,j})$, set

$$u_n(x) \equiv \sup_{z \in \cup_j B(x_{n,j}, r_{n,j})} \left\{u_n(z) - \|u_n\|_{\mathrm{Lip}(\cup_j B(x_{n,j}, r_{n,j}))} d(z, x)\right\}.$$

Notice that for almost all $x \in B(x_{n,j}, r_{n,j})$, since

$$L_{n,j} \ge \max_{k \ne \ell} \left\{\frac{|u(z_{n,j,k}) - u(z_{n,j,\ell})|}{d(z_{n,j,k}, z_{n,j,\ell})}\right\}, \tag{2.10}$$

we have

$$\mathrm{Lip}\, u_n(x) = \|u_n\|_{\mathrm{Lip}(B(x_{n,j}, r_{n,j}))} = L_{n,j}. \tag{2.11}$$

Then by Lemma 2.1 and the strong locality of Γ,

$$\begin{aligned} 1_{B(x_{n,j},r_{n,j})} \Gamma\left(\frac{1}{L_{n,j}} u_n, \frac{1}{L_{n,j}} u_n\right) &= 1_{B(x_{n,j},r_{n,j})} \Gamma\left(\max_k \left\{\frac{1}{L_{n,j}} u(z_{n,j,k}) - d(z_{n,j,k}, \cdot)\right\},\right. \\ &\qquad \left. \max_k \left\{\frac{1}{L_{n,j}} u(z_{n,j,k}) - d(z_{n,j,k}, \cdot)\right\}\right) \\ &\le 1_{B(x_{n,j},r_{n,j})} m. \end{aligned} \tag{2.12}$$

Moreover, by Lemma 2.3,

$$\Gamma(u_n, u_n) \le \|u\|_{\mathrm{Lip}(\cup_j B(x_{n,j}\, r_{n,j}))} m = (\sup_j L_{n,j})^2 m \le (\|u\|_{\mathrm{Lip}(X)} + 1)^2 m, \tag{2.13}$$

which implies that $\int_X 1_{B(x_0,r_0)}\Gamma(u_n,u_n)$ is bounded in $\mathbb{D}$. So there is a subsequence of $\{1_{B(x_0,r_0)}u_n\}_{n\in\mathbb{N}}$ weakly converging to some $v\in\mathbb{D}$. Without loss of generality, we still denote the subsequence by the sequence itself, and hence

$$\int_X d\Gamma(v,v)\le \liminf_{n\to\infty}\int_X 1_{B(x_0,r_0)}\,d\Gamma(u_n,u_n).$$

On the other hand, by (2.10), we have $u_n(z_{n,j,k}) = u(z_{n,j,k})$ for all j and k. For every $x\in B(x_j,r_j)$, by the choice of $z_{n,j,k}$, there exists $z_{n,j,k}$ such that $d(x,z_{n,j,k})\le\frac{1}{m}$, and hence

$$\begin{aligned}|u(x)-u_n(x)| &\le |u(x)-u(z_{n,j,k})|+|u_n(x)-u_n(z_{n,j,k})|\\ &\le (\|u\|_{\mathrm{Lip}(X)}+L_{n,j})d(x,z_{n,j,k})\le \frac{1}{n}(2\|u\|_{\mathrm{Lip}(X)}+1).\end{aligned}$$

For $x\in B(x_0,r_0)\setminus\cup_j B(x_{n,j},r_{n,j})$, we have

$$|u_n(x)-u(x)|\le |u_n(x)-u_n(z_{n,j,k})|+|u(x)-u(z_{n,j,k})|\le 2(2\|u\|_{\mathrm{Lip}(X)}+1)r_0.$$

Thus we have

$$\begin{aligned}\|u_n-u\|^2_{L^2(B(x_0,r_0))} &\lesssim \|u-u_n\|_{L^2(\cup_j B(x_{n,j},r_{n,j}))}\\ &\quad +2(2\|u\|_{\mathrm{Lip}(X)}+1)r_0 m(B(x_0,r_0)\setminus\cup_j B(x_{n,j},r_{n,j}))\\ &\lesssim C(u,B(x_0,r_0))\frac{1}{n}m(B(x_0,r_0)),\end{aligned}\tag{2.14}$$

where $C(u,B(x_0,r_0))$ is a constant independent of n. This means that $\{1_B(x_0,r_0)u_n\}_{n\in\mathbb{N}}$ converges to $1_{B(x_0,r_0)}u$ in $L^2(X)$, and hence $v=1_{B(x_0,r_0)}u$, which together with the locality of Γ implies that

$$\int_X 1_{B(x_0,r_0)}\,d\Gamma(u,u)\le\liminf_{n\to\infty}\int_X 1_{B(x_0,r_0)}\,d\Gamma(u_n,u_n).\tag{2.15}$$

Now we estimate $\int_X 1_{B(x_0,r_0)}\Gamma(u_n,u_n)$ from above. Observe that by (2.12),

$$1_{B(x_{n,j},r_{n,j})}\Gamma(u_n,u_n)\le (L_j)^2 1_{B(x_{n,j},r_{n,j})}m = (\mathrm{Lip}\,u_n)^2 1_{B(x_{n,j},r_{n,j})}m\tag{2.16}$$

which yields

$$\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}\,d\Gamma(u_n,u_n)\le\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}(\mathrm{Lip}\,u_n)^2\,dm\tag{2.17}$$

Moreover, by the triangle inequality, Lemma 2.3 again, (2.11), (2.11) and the doubling property, we have

$$\begin{aligned}&\left|\left\{\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}(\mathrm{Lip}\,u)^2\,dm\right\}^{1/2}-\left\{\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}(\mathrm{Lip}\,u_n)^2\,dm\right\}^{1/2}\right|\\ &\quad\le\left\{\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}(\mathrm{Lip}\,u-\mathrm{Lip}\,u_n)^2\,dm\right\}^{1/2}\\ &\quad\le\left\{\sum_j\int_X 1_{B(x_{n,j},r_{n,j})}(\mathrm{Lip}\,u-\mathrm{Lip}\,u(x_{n,j}))^2\,dm\right\}^{1/2}\end{aligned}$$

$$
\begin{aligned}
&+\left\{\sum_j \int_X 1_{B(x_{n,j},r_{n,j})}(L_{n,j}-\operatorname{Lip} u(x_{n,j}))^2\,dm\right\}^{1/2}\\
&\lesssim \frac{1}{n}\left\{\sum_j m(B(x_{n,j},3r_{n,j}))\,dm\right\}^{1/2}+\frac{1}{n}[m(B(x_0,2r_0))]^{1/2}\\
&\lesssim \frac{1}{n}m(B(x_0,r_0)). \qquad (2.18)
\end{aligned}
$$

From this and (2.17), it follows that

$$
\begin{aligned}
&\left\{\sum_j \int_X 1_{B(x_{n,j},r_{n,j})}\,d\Gamma(u_n,u_n)\right\}^{1/2}\\
&\quad\le \left\{\sum_j \int_X 1_{B(x_{n,j},r_{n,j})}(\operatorname{Lip} u)^2\,dm\right\}^{1/2}+C\frac{1}{n}m(B(x_0,r_0))\\
&\quad\le \left\{\int_X 1_{B(x_0,r_0)}(\operatorname{Lip} u)^2\,dm\right\}^{1/2}+C\frac{1}{n}m(B(x_0,r_0))
\end{aligned}
$$

which together with (2.6) and (2.13) yields

$$
\begin{aligned}
\left\{\int_X 1_{B(x_0,r_0)}\,d\Gamma(u_n,u_n)\right\}^{1/2} &\le \left\{\sum_j \int_X 1_{B(x_{n,j},r_{n,j})}\,d\Gamma(u_n,u_n)\right\}^{1/2}+\frac{C}{n}m(B(x_0,r_0))\\
&\le \left\{\int_X 1_{B(x_0,r_0)}(\operatorname{Lip} u)^2\,dm\right\}^{1/2}+C\frac{1}{n}m(B(x_0,r_0)).
\end{aligned}
$$

Therefore, by (2.15), we obtain (2.9). □

Corollary 2.1. *Assume that (X,d,m) satisfies a doubling property. For every $u\in \operatorname{Lip}(X)$,*

$$
\|u\|_{\operatorname{Lip}(X)}=\sup_{x\in X}\operatorname{Lip} u(x)=\|\operatorname{Lip} u\|_{L^\infty(X)}=\left\|\frac{d}{dm}\Gamma(u,u)\right\|^{1/2}_{L^\infty(X)}.
$$

Proof. By the definition of $\operatorname{Lip} u(x)$, we easily have $\|u\|_{\operatorname{Lip}(X)}\ge \operatorname{Lip} u(x)$ for all $x\in X$. The inequality $\sup_{x\in X}\operatorname{Lip} u(x)\ge \|\operatorname{Lip} u\|_{L^\infty(X)}$ is trivial. By Theorem 2.1, we also have $\|\operatorname{Lip} u\|_{L^\infty(X)}\ge \|\frac{d}{dm}\Gamma(u,u)\|^{1/2}_{L^\infty(X)}$. Now, the proof of Corollary 2.1 is reduced to proving that $\|u\|_{\operatorname{Lip}(X)}\le \|\frac{d}{dm}\Gamma(u,u)\|^{1/2}_{L^\infty(X)}$.

Fix $u\in \operatorname{Lip}(X)$ with $\|\frac{d}{dm}\Gamma(u,u)\|_{L^\infty(X)}<\infty$ (by Theorem 2.1 this actually holds for each $u\in \operatorname{Lip}(X)$). Then, for $\epsilon>0$, we have $v_\epsilon\equiv u(\|\frac{d}{dm}\Gamma(u,u)\|_{L^\infty(X)}+\epsilon)^{-1/2}\in \mathbb{D}_{\mathrm{loc}}$ and $\Gamma(v_\epsilon,v_\epsilon)\le m$. By (2.2), we have that for all $x,y\in X$, $|v_\epsilon(x)-v_\epsilon(y)|\le d(x,y)$, which implies that

$$
|u(x)-u(y)|\le \left(\left\|\frac{d}{dm}\Gamma(u,u)\right\|_{L^\infty(X)}+\epsilon\right)^{1/2} d(x,y).
$$

This, together with the arbitrariness of $\epsilon>0$, implies that $\|u\|_{\operatorname{Lip}(X)}\le \|\frac{d}{dm}\Gamma(u,u)\|^{1/2}_{L^\infty(X)}$ as desired. □

Remark 2.1. (i) In the proof above, we used the result that $\Gamma(d_x, d_x) \le m$ from [44], but did not use the conclusion from [10] that this also holds for each 1-Lipschitz function u.

(ii) The doubling property in Theorem 2.1 can be relaxed to a local doubling property: for every $x_0 \in X$, there exist $r_{x_0} > 0$ and C_{x_0} such that for all $x \in B(x_0, r_{x_0})$ and $r \le r_{x_0}$, $m(B(x, 2r)) \le C_{x_0} m(B(x, r)) < \infty$. We would like to know if Theorem 2.1 holds for a general strongly local Dirichlet form.

Applying Theorem 2.1, we clarify the relations of two kinds of weak Poincaré inequalities on X with the aid of a quasi-Newtonian property.

Recall that $(X, \mathscr{E}, m)$ is said to support a *weak* $(1, p)$*-Poincaré inequality* with $p \in [1, \infty)$ if there exist constants $\lambda \ge 1$ and $C > 0$ such that for all $u \in \mathrm{Lip}(X)$, $x \in X$ and $r > 0$,

$$\fint_{B(x,r)} |u - u_{B(x,r)}|\, dm \le Cr \left\{ \fint_{B(x,\lambda r)} \left[\frac{d}{dm} \Gamma(u, u) \right]^{p/2} dm \right\}^{1/p}. \tag{2.19}$$

Similarly, (X, d, m) is said to support a *weak* $(1, p)$*-Poincaré inequality* if (2.19) holds with $\left[\frac{d}{dm}\Gamma(u, u)\right]^{p/2}$ replaced by $(\mathrm{Lip}\, u)^p$.

We say that $(X, \mathscr{E}, m)$ satisfies a K*-quasi-Newtonian property* if for every $u \in \mathrm{Lip}(X)$, there exists a Borel representative g of $\sqrt{\frac{d}{dm}\Gamma(u, u)}$ such that for all Lipschitz curves $\gamma : [0, 1] \to X$,

$$|u(\gamma(0)) - u(\gamma(1))| \le K \int_\gamma g\, ds.$$

Here g is called a *Borel representative* of a measurable function h if g is a Borel measurable function and satisfies that $g(x) \ge h(x)$ for all $x \in X$ and $g(x) = h(x)$ for almost all $x \in X$. If $K = 1$, we say that $(X, \mathscr{E}, m)$ satisfies the *Newtonian property*; otherwise we say that $(X, \mathscr{E}, m)$ satisfies a *quasi-Newtonian property*.

Proposition 2.1. *Suppose that* (X, d, m) *satisfies a doubling property. Then for every* $p \in [1, \infty)$, $(X, \mathscr{E}, m)$ *supports a weak* $(1, p)$*-Poincaré inequality if and only if* $(X, \mathscr{E}, m)$ *satisfies a quasi-Newtonian property and* (X, d, m) *supports a weak* $(1, p)$*-Poincaré inequality.*

To prove Proposition 2.1, we recall the notion of an upper gradient; see [14] and also [20,40]. Recall that a nonnegative Borel measurable function g is called a *p-weak upper gradient of u* with $p \in [1, \infty)$ if

$$|u(x) - u(y)| \le \int_\gamma g\, ds \tag{2.20}$$

for all $\gamma \in \Gamma_{\mathrm{rect}} \setminus \Gamma_o$, where x and y are the endpoints of γ, Γ_{rect} denotes the collection of non-constant compact rectifiable curves and Γ_o has p-modulus zero in the sense that

$$\inf \left\{ \|\rho\|^p_{L^p(X)} : \rho \text{ is non-negative, Borel measurable, } \int_\gamma \rho\, ds \ge 1 \text{ for all } \gamma \in \Gamma_o \right\} = 0.$$

We denote by $N^{1,p}(X)$ the collection of functions $u \in L^p(X)$ that have a p-weak upper gradient $g \in L^p(X)$, and moreover, $\|u\|_{N^{1,p}(X)} = \|u\|_{L^p(X)} + \inf_g \|g\|_{L^p(X)}$, where g is taken over all p-weak upper gradients of u. We denote by $N^{1,p}_{\mathrm{loc}}(X)$ the class of functions $u \in L^p_{\mathrm{loc}}(X)$ that have a p-weak upper gradient that belongs to $L^p(B)$ for each ball B.

For the following relations between the weak upper gradient and the (approximate) pointwise Lipschitz constant, see [7, Theorem 6.38] with a correction in [21, Remark 2.16] and also [40,20]. For a measurable function u, its *approximate pointwise Lipschitz constant* is defined as

$$\operatorname{apLip} u(x) \equiv \inf_A \limsup_{y\in A, y\to x} \frac{|u(x)-u(y)|}{d(x,y)}$$

for every $x \in X$, where the infimum is taken over all Borel sets $A \subset X$ with a point of density at x. Notice that if $u \in \operatorname{Lip}(X)$, then $\operatorname{apLip} u = \operatorname{Lip} u$ almost everywhere.

Lemma 2.4. *Suppose that (X, d, m) satisfies a doubling property and supports a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$. Then for every $u \in N^{1,p}_{\mathrm{loc}}(X)$, there exists a unique p-weak upper gradient g_u of u such that $g_u = \operatorname{apLip} u$ almost everywhere and $g_u \le g$ almost everywhere whenever g is a p-weak upper gradient of u. In particular, if $u \in \operatorname{Lip}(X)$, then $g_u = \operatorname{Lip} u$ almost everywhere.*

Proposition 2.1 follows from Theorem 2.1 and the following lemma.

Lemma 2.5. *Suppose that (X, d, m) satisfies a doubling property and $(X, \mathscr{E}, m)$ supports a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$. Then there exists a constant $C_1 \ge 1$ such that for all $u \in N^{1,p}_{\mathrm{loc}}(X)$,*

$$(\operatorname{apLip} u)^2 \le C_1 \frac{d}{dm}\Gamma(u,u)$$

almost everywhere.

Proof of Lemma 2.5. Let $u \in N^{1,p}_{\mathrm{loc}}(X)$ and let g_u be the p-weak upper gradient of u as in Lemma 2.4. Set

$$g_k(x) \equiv \sup_{j\ge k} \left\{ \fint_{B(x,\lambda 2^{-j})} \left[\frac{d}{dm}\Gamma(u,u)\right]^{p/2} dm \right\}^{1/p}.$$

Then g_k is Borel measurable; indeed, g_k is lower semicontinuous. Observe that if $g_k(x) < \infty$, then $\lim_{j\to\infty} u_{B(x,2^{-j})}$ exists. In fact, since for every j,

$$\fint_{B(x,2^{-j})} |u - u_{B(x,2^{-j})}|\, dm \lesssim 2^{-j} \left\{ \fint_{B(x,\lambda 2^{-j})} \left[\frac{d}{dm}\Gamma(u,u)\right]^{p/2} dm \right\}^{1/p},$$

by a telescope argument, we have

$$|u_{B(x,2^{-j})} - u_{B(x,2^{-\ell})}| \lesssim 2^{-\min\{j,\ell\}} g_k(x) \to 0$$

as $j, \ell \to \infty$. For such an x, we define $\widetilde{u}(x) \equiv \lim_{j\to\infty} u_{B(x,2^{-j})}$. Generally, for $x \in X$, if $\lim_{j\to\infty} u_{B(x,2^{-j})}$ exists, then we define $\widetilde{u}(x) \equiv \lim_{j\to\infty} u_{B(x,2^{-j})}$; otherwise, set $\widetilde{u}(x) \equiv 0$. Obviously, $u(x) = \widetilde{u}(x)$ for almost all $x \in X$, and hence u and $\widetilde{u}$ generate the same element of $N^{1,p}_{\mathrm{loc}}(X)$.

Now we are going to check that g_k is a p-weak upper gradient of $\widetilde{u}$. Observe that by a telescope argument again, for all $x, y \in X$ with $d(x, y) \le 2^{-k-2}$, we have

$$|\widetilde{u}(x) - \widetilde{u}(y)| \lesssim d(x,y)[g_k(x) + g_k(y)].$$

Recall that, by [40, Proposition 3.1], $\widetilde{u}$ is absolutely continuous on p-almost every curve, namely, $\widetilde{u} \circ \gamma$ is absolutely continuous on $[0, \ell(\gamma)]$ for all arc-length parameterized paths $\gamma \in \Gamma_{\mathrm{rect}} \setminus \Gamma$, where Γ has p-modulus zero. For $\gamma \in \Gamma_{\mathrm{rect}} \setminus \Gamma$, we are going to show that

$$|\widetilde{u}(x) - \widetilde{u}(y)| \lesssim \int_\gamma g_k\, ds. \tag{2.21}$$

To this end, by the absolute continuity of u on γ, it suffices to show that for j large enough,

$$2^j \left| \int_0^{2^{-j}} \widetilde{u} \circ \gamma(t)\, dt - \int_{\ell(\gamma)-2^{-j}}^{\ell(\gamma)} \widetilde{u} \circ \gamma(t)\, dt \right| \lesssim \int_0^{\ell(\gamma)} g_k \circ \gamma(t)\, dz.$$

But, for j large enough, we have that

$$\begin{aligned}
&2^j \left| \int_0^{2^{-j}} \widetilde{u} \circ \gamma(t)\, dt - \int_{\ell(\gamma)-2^{-j}}^{\ell(\gamma)} \widetilde{u} \circ \gamma(t)\, dt \right| \\
&\quad = 2^j \left| \int_0^{\ell(\gamma)-2^{-j}} [\widetilde{u} \circ \gamma(t + 2^{-j}) - \widetilde{u} \circ \gamma(t)]\, dt \right| \\
&\quad \le 2^j \int_0^{\ell(\gamma)-2^{-j}} \left| \widetilde{u} \circ \gamma(t + 2^{-j}) - \widetilde{u} \circ \gamma(t) \right| dt \\
&\quad \lesssim \int_0^{\ell(\gamma)-2^{-j}} \left[g_k \circ \gamma(t + 2^{-j}) + g_k \circ \gamma(t) \right] dt \\
&\quad \lesssim \int_0^{\ell(\gamma)} g_k \circ \gamma(t)\, dt.
\end{aligned}$$

This gives (2.21) and hence g_k is a p-weak upper gradient of $\widetilde{u}$. Notice that Lemma 2.4 gives that $\operatorname{apLip} u$ coincides with the unique minimal p-weak upper gradient of u and hence that of $\widetilde{u}$ almost everywhere. So $\operatorname{apLip} u \lesssim g_k$ almost everywhere and hence, by the Lebesgue differentiation theorem, for almost all $x \in X$,

$$\begin{aligned}
\operatorname{apLip} u(x) &\lesssim \liminf_{k\to\infty} g_k(x) \\
&\lesssim \lim_{k\to\infty} \sup_{j\ge k} \left\{ \fint_{B(x,\lambda 2^{-j})} \left[\frac{d}{dm} \Gamma(u,u) \right]^{p/2} dm \right\}^{1/p} \\
&\lesssim \left\{ \frac{d}{dm} \Gamma(u,u)(x) \right\}^{1/2}.
\end{aligned}$$

This finishes the proof of Lemma 2.5. □

Moreover, from Theorem 2.1 and Lemma 2.4 we conclude the following result.

Proposition 2.2. *Suppose that (X, d, m) satisfies a doubling property and supports a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$. Then $(X, \mathscr{E}, m)$ satisfies the Newtonian property if and only if for all $u \in \mathrm{Lip}(X)$,*

$$\frac{d}{dm} \Gamma(u,u) = (\operatorname{Lip} u)^2$$

almost everywhere.

When $p = 2$, we further have the following conclusion. Recall that, as proved by Sturm [46], (X, d, m) satisfies the doubling property and $(X, \mathscr{E}, m)$ supports a weak $(1, 2)$-Poincaré inequality if and only if a scale invariant Harnack inequality for the parabolic operator $\frac{\partial}{\partial t} - \Delta$ on $\mathbb{R} \times X$ holds true, with Δ corresponding to $\mathscr{E}$.

Theorem 2.2. *Suppose that (X, d, m) satisfies a doubling property and $(X, \mathscr{E}, m)$ supports a weak $(1, 2)$-Poincaré inequality. Then the following hold:*

(i) $\mathbb{D} = N^{1,2}(X)$ *with equivalent norms,* $\mathrm{Lip}(X) \cap \mathscr{C}_0(X)$ *is dense in* $\mathbb{D}$, *and* $\mathbb{D}_{\mathrm{loc}} = N^{1,2}_{\mathrm{loc}}(X)$*;*

(ii) *for all* $u \in \mathbb{D}_{\mathrm{loc}}$, $\Gamma(u, u)$ *is absolutely continuous with respect to* m*, and there exists a constant* $C_1 \geq 1$ *such that for all* $u \in \mathbb{D}_{\mathrm{loc}}$,

$$\frac{d}{dm}\Gamma(u, u) \leq (\mathrm{apLip}\, u)^2 \leq C_1 \frac{d}{dm}\Gamma(u, u) \tag{2.22}$$

almost everywhere, where $\mathrm{apLip}\, u = \mathrm{Lip}\, u$ *almost everywhere for* $u \in \mathrm{Lip}(X)$.

(iii) *If* $(X, \mathscr{E}, m)$ *satisfies the Newtonian property, then* $C_1 = 1$ *in* (2.22).

Proof. Recall that if (X, d, m) supports a weak $(1, 2)$-Poincaré inequality, then $\mathrm{Lip}(X) \cap \mathscr{C}_0(X)$ is dense in $N^{1,2}(X)$ as proved in [40,7]. Notice that Theorem 2.1 and Lemma 2.5 imply that $\frac{d}{dm}\Gamma(u, u) \sim (\mathrm{Lip}\, u)^2$ holds almost everywhere for all $u \in \mathrm{Lip}(X)$. Under this, it was proved in [25,41] that $\mathbb{D} = N^{1,2}(X)$ with equivalent norms. This implies that $\mathrm{Lip}(X) \cap \mathscr{C}_0(X)$ is dense in $\mathbb{D}$ and also that $\mathbb{D}_{\mathrm{loc}} = N^{1,2}_{\mathrm{loc}}(X)$. This gives (i).

Obviously, for $u \in \mathrm{Lip}(X)$, (ii) follows from Theorem 2.1, Lemma 2.5 and Proposition 2.2. For $u \in \mathbb{D}_{\mathrm{loc}}$, by $\mathbb{D}_{\mathrm{loc}} = N^{1,2}_{\mathrm{loc}}(X)$, we have that $\Gamma(u, u)$ is absolutely continuous with respect to m, and Lemma 2.5 gives $(\mathrm{apLip}\, u)^2 \leq C_1 \frac{d}{dm}\Gamma(u, u)$ almost everywhere. Finally, a density argument together with the closedness of $\mathscr{E}$ and the fact that (ii) holds for Lipschitz functions leads to $\frac{d}{dm}\Gamma(u, u) \leq (\mathrm{apLip}\, u)^2$ for all $u \in \mathbb{D}_{\mathrm{loc}}$, which completes the proof of Theorem 2.2. □

3. Dirichlet forms: $\frac{d}{dm}\Gamma(u, u) \neq (\mathrm{Lip}\, u)^2$

This section is a continuation of Section 2. By perturbing the classical Dirichlet energy form on $\mathbb{R}^2$, we construct an example that satisfies the doubling property and a weak Poincaré inequality but so that the intrinsic length structure does not coincide with the gradient structure; see Proposition 3.1. This shows that doubling and Poincaré are not enough to obtain $\frac{d}{dm}\Gamma(u, u) = (\mathrm{Lip}\, u)^2$ almost everywhere (this fact can also be deduced from [47]; see Remark 3.1). Moreover, the gradient (differential) structure of our perturbed Dirichlet form does not coincide with the distinguished gradient (differential) structure of Cheeger; see Proposition 3.2. Notice that the distinguished differential structure of Cheeger coincides with the gradient structure of Γ if $(X, \mathscr{E}, m)$ further satisfies the Newtonian property; see Corollary 3.1.

Our example is a perturbation of the classical Dirichlet energy form on $\mathbb{R}^2$ on a large Cantor set E. Denote by m the Lebesgue measure on $\mathbb{R}^2$, and denote by $| \cdot - \cdot |$ the Euclidean distance. The classical Dirichlet form $\mathscr{E}$ is defined by $\mathscr{E}(u, u) = \int_{\mathbb{R}^2} |\nabla u|^2 \, dm$ with the domain $\mathbb{D} = W^{1,2}(\mathbb{R}^2)$, where ∇ is the distributional gradient. Notice that the Euclidean distance gives the intrinsic distance associated to $\mathscr{E}$. Thus $(\mathbb{R}^2, \mathscr{E}, m, |\cdot - \cdot|)$ satisfies a doubling property, a weak $(1, 1)$-Poincaré inequality and the Newtonian property. Moreover, the length structure coincides with the gradient structure, that is, $|\nabla u| = \mathrm{apLip}\, u$ almost everywhere for all $u \in W^{1,2}(\mathbb{R}^2)$.

Let F be the Cantor set constructed as follows: I_i are the two closed intervals obtained by removing the middle open interval with length $1/10$ from $[0,1]$ and are ordered from left to right; when $n \geq 2$, $I_{i_1 \cdots i_n}$ are the two closed intervals obtained by removing the middle open interval with length $(1/10)^n$ from $I_{i_1 \cdots i_{n-1}}$, and are ordered from left to right; $F \equiv \cap_{n\in\mathbb{N}} \cup_{i_1,\dots,i_n} I_{i_1\cdots i_n}$. Notice that F has positive 1-dimensional Lebesgue measure. Set $E \equiv F \times F$. Then $\mathbb{R}^2 \setminus E$ is dense in $\mathbb{R}^2$ and by the Fubini theorem, $m(E) > 0$.

Now, for any $\delta \in (0,1)$, we define a perturbation $\mathscr{E}_\delta$ of $\mathscr{E}$ by setting

$$\mathscr{E}_\delta(u,u) \equiv \int_{\mathbb{R}^2} (1 - \delta 1_E)|\nabla u|^2 \, dm.$$

It is easy to see that $\mathscr{E}_\delta$ is a regular, strongly local Dirichlet form with the domain $\mathbb{D} = W^{1,2}(\mathbb{R}^2)$, and for $u \in \mathbb{D}_{\mathrm{loc}}$, $\Gamma_E(u,u) = (1 - \delta 1_E)|\nabla u|^2 m$. Moreover, let d_δ be the intrinsic distance defined as in (2.1). Then

$$(1-\delta)|\nabla u|^2 \leq \frac{d}{dm}\Gamma_\delta(u,u) \leq |\nabla u|^2$$

implies that

$$|x-y| \leq d_\delta(x,y) \leq \frac{1}{1-\delta}|x-y|.$$

From this, it is easy to see that $(\mathbb{R}^2, \mathscr{E}_\delta, d_\delta, m)$ satisfies the doubling property and a weak $(1,1)$-Poincaré inequality. However, the intrinsic length structure does not coincide with the gradient structure when δ is close to 1.

Proposition 3.1. *There exists $\delta_E \in (0,1)$ such that, for every $\delta \in (\delta_E, 1)$, the intrinsic length structure and the gradient structure of $(\mathbb{R}^2, \mathscr{E}_\delta, d_\delta, m)$ do not coincide, that is, there exists $u \in \mathbb{D}$ such that $\frac{d}{dm}\Gamma(u,u) < (\mathrm{apLip}_{d_\delta} u)^2$ on some set of positive measure.*

To prove this, we need the following crucial property.

Lemma 3.1. *There exists a positive constant $C_E > 1$ such that for every pair of $x, y \in \mathbb{R}^2$, we can find a rectifiable curve γ joining x and y and satisfying*

(i) *$\ell_{\mathbb{R}^2}(\gamma) \leq C_E|x-y|$, where $\ell_{\mathbb{R}^2}(\gamma)$ is the length of γ with respect to the Euclidean distance,*
(ii) *the set $\gamma \cap E$ contains at most 2 points.*

Proof. It suffices to consider all pairs of $x, y \in [0,1]^2 \equiv [0,1] \times [0,1]$. Indeed, if both x and y belong to $\mathbb{R}^2 \setminus (0,1)^2$, (i) and (ii) obviously hold; if only one of x, y belongs to $(0,1)^2$, say $y \in (0,1)^2$, taking z to be the intersection of the boundary of $(0,1)^2$ and the interval joining x and y, and gluing the interval joining x, z and the assumed curve joining y, z, we obtain the desired rectifiable curve.

We claim that for all pairs of $x, y \in [0,1]^2 \setminus E$, there exists γ joining x and y such that $\ell_{\mathbb{R}^2}(\gamma) \lesssim |x-y|$ and $\gamma \cap E = \emptyset$. Assume that this claim holds for the moment. If $x \in E$ and $y \notin E$, since $\mathbb{R}^2 \setminus E$ is dense in $\mathbb{R}^2$, there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset [0,1]^2 \setminus E$ of points such that $x_n \to x$ as $n \to \infty$ and $|x - x_n| \leq \frac{1}{2}|x - x_{n-1}|$ for all $n \geq 1$, where $x_0 = y$. Let γ_n be the assumed rectifiable curve joining x_{n-1}, x_n for $n \in \mathbb{N}$. Set $\gamma \equiv (\cup_{n\in\mathbb{N}} \gamma_n) \cup \{x, y\}$. Observing that

$$|x_n - x_{n-1}| \leq |x - x_{n-1}| + |x - x_n| \leq 2|x - x_{n-1}| \leq 2^{-n+1}|x - x_0|,$$

we have that

$$\ell_{\mathbb{R}^2}(\gamma) \le \sum_{n\in\mathbb{N}} \ell_{\mathbb{R}^2}(\gamma_n) \lesssim \sum_{n\in\mathbb{N}} |x_n - x_{n-1}| \lesssim \sum_{n\in\mathbb{N}} 2^{-n}|x - x_0| \lesssim |x - y|.$$

Obviously, $\gamma_n \cap E = \emptyset$ for each $n \in \mathbb{N}$ implies that $\gamma \cap E$ contains a single point. If $x, y \in E$, we pick a point z in the intersection of $\mathbb{R}^2 \setminus E$ and the interval joining x and y. This reduces us to the case $x \in E$ and $y \notin E$.

Now we prove the above claim. Let $x, y \in [0,1]^2 \setminus E$. Let $L_x^{(1)} \equiv \{x + t(x_1, 0) : \ t \in \mathbb{R}\}$ be the line parallel to x_1-axis and $L_x^{(2)} \equiv \{x + t(0, x_2) : \ t \in \mathbb{R}\}$ parallel to x_2-axis. Observe that at least one of $L_x^{(1)}$ and $L_x^{(2)}$ does not intersect E. Otherwise, if both $L_x^{(1)}$ and $L_x^{(2)}$ intersect E, then $(x_1 + t_1x_1, x_2), (x_1, x_2 + t_2x_2) \in E$ for some $t_1, t_2 \in \mathbb{R}$, and hence, $x_1, x_2 \in F$. Thus $x = (x_1, x_2) \in E$, which is a contradiction. Similarly, define $L_y^{(1)}$ and $L_y^{(2)}$ and then at least one of $L_y^{(1)}$ and $L_y^{(2)}$ does not intersect E. If $L_x^{(1)}$ and $L_y^{(2)}$ do not intersect E, since $L_x^{(1)} \cap L_y^{(2)} \ne \emptyset$, there exists a unique $z \in (L_x^{(1)} \cap L_y^{(2)}) \cap [0,1]^2$. Then we take γ as the union of the interval joining x and z and the interval joining y and z. Obviously, γ is as desired. We reason analogously if $L_x^{(2)}$ and $L_y^{(1)}$ do not intersect E. However, it may happen that only $L_x^{(1)}$ and $L_y^{(1)}$ do not intersect E. In this case, we take $z = (z_1, x_2) \in L_x^{(1)}$ such that $z_1 \in [0,1] \setminus F$ but $|z_1 - x_1| \le |x_1 - y_1|/2$. Notice that the fact that $L_x^{(1)}$ and $L_y^{(1)}$ do not intersect E implies that $x_2, y_2 \in [0,1] \setminus F$, and that $z_1, x_2 \in [0,1] \setminus F$ implies that $L_z^{(2)}$ does not intersect E. Hence $w \equiv (z_1, y_2) \in L_z^{(2)} \cap L_y^{(1)}$ does not belong to E. The desired rectifiable curve γ is given by the union of the interval joining x, z, the one joining z, w and the one joining w, y. Indeed, obviously, we have $\gamma \in [0,1] \setminus E$, and moreover,

$$\ell_{\mathbb{R}^2}(\gamma) \le |x - z| + |z - w| + |w - y| \le \frac{1}{2}|x_1 - y_1| + |x_2 - y_2| + |z_1 - y_1| \lesssim |x - y|.$$

We reason analogously if $L_x^{(2)}$ and $L_y^{(2)}$ do not intersect E. This finishes the proof of Lemma 3.1. □

Proof of Proposition 3.1. We first prove that for all $x, y \in \mathbb{R}^2$,

$$d_\delta(x, y) \le C_E|x - y|, \tag{3.1}$$

where C_E is the constant from Lemma 3.1. For every $u \in \mathbb{D}_{\mathrm{loc}}$ with $\Gamma_\delta(u, u) \le m$, we have $u \in \mathrm{Lip}(\mathbb{R}^2)$ (with respect to the Euclidean distance) and $|\nabla u(x)| \le 1$ for almost all $x \in \mathbb{R}^2 \setminus E$. Thus u is locally 1-Lipschitz outside of E. For a pair of points $x, y \in \mathbb{R}^2$, let γ be a curve as in Lemma 3.1 of length at most $C_E|x - y|$. We conclude that $|u(x) - u(y)| \le C_E|x - y|$.

Let u be a smooth function with compact support. Then $u \in \mathrm{Lip}_{d_\delta}(\mathbb{R}^2) \subset \mathbb{D}_{\mathrm{loc}}$. For every $x \in E \cap (0,1)^2$, since

$$\limsup_{r\to 0} \sup_{d_\delta(x,y)\le r} \frac{|u(x) - u(y) - \nabla u(x)\cdot(x - y)|}{d_\delta(x, y)} = 0,$$

we have that

$$\mathrm{Lip}_{d_\delta} u(x) \ge \liminf_{r\to 0} \sup_{d_\delta(x,y)\le r} \frac{|u(x) - u(y)|}{d_\delta(x, y)} \ge \liminf_{r\to 0} \sup_{d_\delta(x,y)\le r} \frac{|\nabla u(x)\cdot(x - y)|}{d_\delta(x, y)}.$$

Assume that $\nabla u(x) \neq 0$. Observe that there exists a sequence $\{y_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^2 \setminus E$ such that $y_i = x + \epsilon_i \nabla u(x)$ and $\epsilon_i \to 0$. Choose $\delta_E \equiv 1 - \frac{1}{C_E}$. Then by (3.1) and $\delta \in (1 - \frac{1}{C_E}, 1)$,

$$\begin{aligned}\operatorname{Lip}_{d_\delta} u(x) &\geq |\nabla u(x)| \liminf_{i\to\infty} \frac{|x - y_i|}{d_\delta(x, y_i)} \geq \frac{1}{C_E} |\nabla u(x)| \\ &> (1 - \delta)|\nabla u(x)| = \frac{d}{dm}\Gamma(u, u)(x).\end{aligned}$$

Since $E \cap (0, 1)^2$ has positive measure, if u has non-vanishing gradient on this set, then we have $\frac{d}{dm}\Gamma(u, u) < (\operatorname{Lip}_{d_\delta} u)^2$ on $E \cap (0, 1)^2$ as desired. □

Remark 3.1. It can be also deduced from [47] that doubling and Poincaré are not sufficient to guarantee that $\frac{d}{dm}\Gamma(u, u) \equiv (\operatorname{Lip} u)^2$ almost everywhere. Indeed, Sturm [47, Theorem 2] constructed a Dirichlet form

$$\mathscr{E}_a(u, u) = \int_{\mathbb{R}^2} a(x)|\nabla u(x)|^2\, dm(x),$$

where $a(x)$ satisfies $0 < c \leq a(x) < 1$ for all $x \in \mathbb{R}^2$, for which the intrinsic distance d_a is exactly the Euclidean distance. Our construction is motivated by the Cheeger differential structure below.

Now we recall the distinguished differential structure constructed by Cheeger [7]. Assume that (X, d, m) satisfies the doubling property and a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$. Cheeger [7, Theorem 4.38] proved the existence of an atlas that consists of a countable collection $\{(U_\alpha, y^\alpha, k(\alpha))\}_{\alpha\in\mathcal{A}}$ of *charts*, where

(i) U_α's are measurable sets and $m(X \setminus \cup_\alpha U_\alpha) = 0$;
(ii) $k(\alpha) \in \mathbb{N}$, $\sup_{\alpha\in\mathcal{A}} k(\alpha) < \infty$ and if $m(U_\alpha \cap U_\beta) > 0$; then $k(\alpha) = k(\beta)$,
(iii) $y^\alpha \equiv (y_1^\alpha, \dots, y_{k(\alpha)}^\alpha) : U_\alpha \to \mathbb{R}^{k(\alpha)}$ is Lipschitz;
(iv) for every $\alpha \in \mathcal{A}$ and $u \in \operatorname{Lip}(X)$, there exist $V_\alpha(u) \subset U_\alpha$ and a collection

$$\left\{\frac{\partial u}{\partial y_i^\alpha} : U_\alpha \to \mathbb{R}\right\}_{1\leq i\leq k(\alpha)}$$

of bounded Borel measurable functions uniquely determined almost everywhere such that $m(U_\alpha \setminus V_\alpha(u)) = 0$ and for all $z \in V_\alpha(u)$,

$$u(w) = u(z) + \sum_{j=1}^{k(\alpha)} \frac{\partial u}{\partial y_i^\alpha}(y_j^\alpha(w) - y_j^\alpha(z)) + o(d(w, z)); \tag{3.2}$$

(v) if $m(U_\alpha \cap U_\beta) > 0$, then the matrix of $\frac{\partial y^\alpha}{\partial y^\beta}$ is invertible almost everywhere in $U_\alpha \cap U_\beta$.

The above atlas yields a bi-Lipschitz invariant measurable tangent bundle TX and cotangent bundle T^*X. In fact, for every Lipschitz function $u : X \to \mathbb{R}$, its differential du is defined as $(\frac{\partial u}{\partial y_1^\alpha}, \dots, \frac{\partial u}{\partial y_{k(\alpha)}^\alpha})$ and its derivative as $Du = \sum_{i=1}^{k(\alpha)} \frac{\partial u}{\partial y_i^\alpha} \frac{\partial}{\partial y_i^\alpha}$ on each U_α. Notice that TX is the dual of T^*X. For a Lipschitz function u, its derivative $Du : TX \to \mathbb{R}$ coincides with its differential du in the sense that

$$\langle Du(z), v\rangle_z \equiv \sum_{i=1}^{k(\alpha)} \frac{\partial u}{\partial y_i^\alpha}(z) v_i = du(v)(z)$$

for every $z \in V_\alpha(u) \subset U_\alpha$ and $v = \sum_{i=1}^{k(\alpha)} v_i \frac{\partial}{\partial y_i^\alpha} \in T_z X$. Moreover, for each $z \in V_\alpha(u) \subset U_\alpha$, a natural norm $\|\cdot\|_{T_z X}$ on $T_z X$ is defined by setting $\|v\|_{T_z X} \equiv \mathrm{Lip}_d\left(\sum_{i=1}^{k(\alpha)} v_i y_i^\alpha\right)$ for $v = \sum_{i=1}^{k(\alpha)} v_i \frac{\partial}{\partial y_i^\alpha} \in T_z X$ and hence $\|Du\|_{T_z X} = \mathrm{Lip}_d\left(\sum_{i=1}^{k(\alpha)} \frac{\partial u}{\partial y_i^\alpha} y_i^\alpha\right)$ for every Lipschitz function u. Generally, $\|v\|_{T_z X}$ is not Hilbertian. Cheeger [7, p. 460] introduced a distinguished inner product norm $|||v|||_{T_z X}$ associated to it as follows.

Let V be a k-dimensional vector space and $\|\cdot\|$ be a norm on V. Denote by V^* the dual space of V, endowed with the norm $\|\cdot\|^*$ induced by $\|\cdot\|$. Then a distinguished inner product norm $|||\cdot|||^*$ on V^* is obtained by identifying the functions of V^* with their restriction to the unit ball $B_{\|\cdot\|}(0,1)$(with respect to $\|\cdot\|$) and regarding the functions so obtained as elements of $L^2(B_{\|\cdot\|}(0,1), (k+1)\frac{\mathrm{Vol}(k)}{\mathrm{Vol}(k+2)} H^k_{\|\cdot\|})$. Here $\mathrm{Vol}(n)$ denotes the volume of the Euclidean unit ball of $\mathbb{R}^n$ and $H^k_{\|\cdot\|}$ is the k-dimensional Hausdorff measure associated to the metric induced by $\|\cdot\|$. In other words, for $v^* \in V^*$, we define

$$|||v^*|||^* \equiv \left((k+1)\frac{\mathrm{Vol}(k)}{\mathrm{Vol}(k+2)} \int_{B_{\|\cdot\|}(0,1)} |v^*(v)|^2 \, dH^k_{\|\cdot\|}(v)\right)^{1/2}. \tag{3.3}$$

Then the inner product norm $|||\cdot|||$ on V is defined by $|||v||| = \sup_{|||v^*|||^* \le 1} v^*(v)$. Notice that if $\|\cdot\|$ is an inner product norm, then $|||\cdot|||^* = \|\cdot\|^*$ and $|||\cdot||| = \|\cdot\|$.

Now we have two differential (gradient) structures on $(X, \mathscr{E}, d, m)$: the original one of Γ induced by $\mathscr{E}$ and the distinguished one $|||\cdot|||_{TX}$ induced from the intrinsic distance in the sense of Cheeger. Under some reasonable assumptions, they coincide as a corollary to Theorem 2.2.

Corollary 3.1. *Suppose that (X, d, m) satisfies a doubling property, and that $(X, \mathscr{E}, m)$ supports a weak $(1,2)$-Poincaré inequality and the Newtonian property. Then for all $u \in \mathbb{D}$, $\frac{d}{dm}\Gamma(u,u) = |||Du|||^2_{TX} = (\mathrm{apLip}_d u)^2$ almost everywhere.*

However, generally, $\frac{d}{dm}\Gamma$ and $|||\cdot|||^2_{TX}$ do not necessarily coincide. This will be illustrated by the above example in Proposition 3.2. To this end, notice that since $(\mathbb{R}^2, d_\delta, m)$ satisfies the doubling property and a weak $(1,2)$-Poincaré inequality, by [7], there exists some atlas. Up to some change of the coordinate functions, we can take the atlas with a single chart $U \subset \mathbb{R}^2$ with $m(\mathbb{R}^2 \setminus U) = 0$, and naturally, choose $x = (x_1, x_2)$ as the coordinate. Indeed, let $\{(U_\alpha, y^\alpha, k(\alpha))\}_{\alpha \in \mathcal{A}}$ be the atlas determined by [7] as above. We will compare it with the usual coordinates via the classical Rademacher theorem: for every $u \in \mathrm{Lip}(\mathbb{R}^2)$ and almost all $z \in \mathbb{R}^2$,

$$u(w) = u(z) + \sum_{i=1}^{2} \frac{\partial u(z)}{\partial x_i}(w_i - z_i) + o(|w - z|). \tag{3.4}$$

Notice that this formula also holds when $|w - z|$ is replaced by $d_\delta(w, z)$ since the two distances are equivalent. Now, for $\alpha \in \mathcal{A}$, applying (3.4) to $(y_i^\alpha)_{i=1}^{k(\alpha)}$, we get a Jacobian matrix $\frac{\partial y^\alpha}{\partial x} \equiv (\frac{\partial y_i^\alpha}{\partial x_j})_{i=1,\dots,k(\alpha);\ j=1,2}$ almost everywhere; while applying (3.2) to (x_1, x_2), we get $\frac{\partial x}{\partial y^\alpha} \equiv (\frac{\partial x_i}{\partial y_j^\alpha})_{i=1,2;\ j=1,\dots,k(\alpha)}$ on $V_\alpha(x_1) \cap V_\alpha(x_2)$. Then for almost all $z \in U_\alpha$, we have

$$w = z + \frac{\partial x}{\partial y^\alpha}(z)\frac{\partial y^\alpha}{\partial x}(z)(w - z) + o(d_\delta(w, z)),$$

which implies that $\frac{\partial x}{\partial y^\alpha}\frac{\partial y^\alpha}{\partial x} = Id_2$ almost everywhere in U_α. Similarly, $\frac{\partial y^\alpha}{\partial x}\frac{\partial x}{\partial y^\alpha} = Id_{k(\alpha)}$ almost everywhere in U_α. This implies that $k(\alpha) = 2$, and $\frac{\partial x}{\partial y^\alpha} = (\frac{\partial y^\alpha}{\partial x})^{-1}$ almost everywhere. Therefore on U_α, and hence on $\cup_{\alpha\in\mathcal{A}} U_\alpha$, we can use the uniform coordinate function x.

Under the above atlas $\{(U, x, 2)\}$, from the above argument, we also see that the Cheeger derivative $D_\delta u$ coincides with ∇u for all Lipschitz functions u, namely, $D_\delta u = \frac{\partial u}{\partial x_1}\frac{\partial}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial}{\partial x_2}$. For almost all $z \in U$ and $v = v_1\frac{\partial}{\partial x_1} + v_2\frac{\partial}{\partial x_2} \in T_z\mathbb{R}^2$, $\|v\|_{T_z\mathbb{R}^2} = \mathrm{Lip}_{d_\delta}(v_1x_1 + v_2x_2)$. But when δ is close to 1, the following result shows that $|||D_\delta u|||^2_{T\mathbb{R}^2}$ does not coincide with the squared gradient $\frac{d}{dm}\Gamma_\delta(u, u)$ and hence, the distinguished differential structure of Cheeger does not coincide with the original differential structure on $(\mathbb{R}^2, \mathscr{E}_\delta, d_\delta, m)$.

Proposition 3.2. *There exists a $\widetilde{\delta}_E \in (0, 1)$ such that for every $\delta \in (\widetilde{\delta}_E, 1)$, we can find a function $u \in \mathrm{Lip}(\mathbb{R}^2)$ such that $\frac{d}{dm}\Gamma_\delta(u, u) < |||D_\delta u|||^2_{T\mathbb{R}^2}$ on some set with positive measure.*

To this end, we need the following result, whose notation is that of the paragraph containing formula (3.3).

Lemma 3.2. *Assume that $\|\cdot\|$ and $\|\cdot\|_o$ are two norms on V that satisfy $M^{-1}\|\cdot\| \le \|\cdot\|_o \le M\|\cdot\|$ for some $M \ge 1$. Then there exists a positive constant $C(M, k)$ such that*

$$\frac{1}{C(M,k)}|||\cdot||| \le |||\cdot|||_o \le C(M,k)|||\cdot|||.$$

Proof. We first notice that $H^k_{\|\cdot\|_o}$ (and also $H^k_{\|\cdot\|}$) is a constant multiple of the Lebesgue measure on $\mathbb{R}^k$ due to its translation invariance. Then $H^k_{\|\cdot\|_o} = c_o H^k_{\|\cdot\|}$ for some $c_o > 0$. We claim that $M^{-k} \le c_o \le M^k$. Indeed, recall that for any set F, its k-dimensional Hausdorff measure with respect to the norm $\|\cdot\|_o$ is defined by $H^k_{\|\cdot\|_o}(F) \equiv \lim_{\epsilon\to 0+} H^k_{\|\cdot\|_o,\epsilon}(F)$ with

$$H^k_{\|\cdot\|_o,\epsilon}(F) \equiv \inf\left\{\sum_i (\mathrm{diam}_{\|\cdot\|_o} U_i)^k\right\},$$

where the infimum is taken over all covers $\{U_i\}_i$ of F with $\mathrm{diam}_{\|\cdot\|_o} U_i \le \epsilon$. Notice that $M^{-1}\|\cdot\| \le \|\cdot\|_o \le M\|\cdot\|$ implies that

$$M^{-1}\mathrm{diam}_{\|\cdot\|}U \le \mathrm{diam}_{\|\cdot\|_o}U \le M\mathrm{diam}_{\|\cdot\|}U.$$

Then it follows that $H^k_{\|\cdot\|_o,\epsilon}(F) \le M^k H^k_{\|\cdot\|,M\epsilon}(F)$ and hence $H^k_{\|\cdot\|_o}(F) \le M^k H^k_{\|\cdot\|}(F)$. Similarly, $H^k_{\|\cdot\|}(F) \le M^k H^k_{\|\cdot\|_o}(F)$ as desired.

Moreover, observing that $B_{\|\cdot\|_o}(0, 1) \subset B_{\|\cdot\|}(0, M)$, by the scaling property of the Lebesgue measure and hence of $H^k_{\|\cdot\|}$, we obtain

$$\begin{aligned}
|||v^*|||^*_o &\le \left(c_o(k+1)\frac{\mathrm{Vol}(k)}{\mathrm{Vol}(k+2)}\int_{B_{\|\cdot\|}(0,M)} |v^*(v)|^2\, dH^k_{\|\cdot\|}(v)\right)^{1/2}\\
&\le \left(c_o(k+1)M^{k+2}\frac{\mathrm{Vol}(k)}{\mathrm{Vol}(k+2)}\int_{B_{\|\cdot\|}(0,1)} |v^*(v)|^2\, dH^k_{\|\cdot\|}(v)\right)^{1/2}\\
&\le \left(c_o M^{k+2}\right)^{1/2} |||v^*|||^*,
\end{aligned}$$

which implies that $\|\!|v|\!\| \le \left(c_o M^{k+2}\right)^{1/2} \|\!|v|\!\|_o$. Similarly, we have

$$\|\!|v|\!\|_o \le \left(\frac{1}{c_o} M^{k+2}\right)^{1/2} \|\!|v|\!\|,$$

which finishes the proof of Lemma 3.2. $\square$

Proof of Proposition 3.2. For almost every $z \in U$ and $v = v_1 \frac{\partial}{\partial x_1} + v_2 \frac{\partial}{\partial x_2} \in T_z\mathbb{R}^2$, following Cheeger's definition, we have $\|v\|_{T_z\mathbb{R}^2} = \mathrm{Lip}_{d_\delta}(v_1x_1 + v_2x_2)(z)$. Set

$$\|v\| \equiv |\nabla(v_1x_1 + v_2x_2)(z)| = (v_1^2 + v_2^2)^{1/2}.$$

Since (3.1) implies that $\frac{1}{C_E}|\nabla u(z)| \le \mathrm{Lip}_{d_\delta} u(z) \le |\nabla u(z)|$ for $u \in \mathrm{Lip}(\mathbb{R}^2)$, we have that $\frac{1}{C_E}\|v\| \le \|v\|_{T_z\mathbb{R}^2} \le \|v\|$, which together with Lemma 3.2 leads to

$$\frac{1}{C(C_E, 2)} \|\!|v|\!\| \le \|\!|v|\!\|_{T_z\mathbb{R}^2} \le C(C_E, 2)\|\!|v|\!\|.$$

Notice that $\|\cdot\| = \|\!|\cdot|\!\|$. We choose $\widetilde{\delta}_E \equiv 1 - \frac{1}{C(C_E,2)}$. For every $\delta \in (\widetilde{\delta}_E, 1)$ and $z \in E \cap U$, we have

$$\begin{aligned}\frac{d}{dm}\Gamma_\delta(v_1x_1 + v_2x_2, v_1x_1 + v_2x_2)(z) &= (1-\delta)^2|\nabla(v_1x_1 + v_2x_2)(z)|^2 = (1-\delta)^2\|v\|^2 \\ &< \|\!|v|\!\|^2_{T_z\mathbb{R}^2} = (\mathrm{Lip}_{d_\delta}(v_1x_1 + v_2x_2)(z))^2.\end{aligned}$$

Since the set $E \cap U$ has positive measure, this finishes the proof of Proposition 3.2. $\square$

4. A Sierpinski gasket with $\frac{d}{dm}\Gamma(u, u) = (\mathrm{Lip}_d\, u)^2$

In this section, for the standard (resistance) Dirichlet form on the Sierpinski gasket equipped with the Kusuoka measure, we will identify the intrinsic length structure with the measurable Riemannian and the gradient structures; see Theorems 4.1–4.3. We begin with the definition of the Sierpinski gasket K.

Definition 4.1. Let $V_0 \equiv \{p_1, p_2, p_3\} \in \mathbb{R}^2$ be the set of the three vertices of an equilateral triangle, and for $p_i \in V_0$, define $F_i(x) \equiv (x + p_i)/2$ for all $x \in \mathbb{R}^2$. The *Sierpinski gasket* K is defined as the self-similar set associated with the family of contractions $\{F_i\}_{i=1}^3$, namely, K is the unique non-empty compact set satisfying $K = \cup_{i=1}^3 F_i(K)$.

On the Sierpinski gasket K, there is a standard resistance form $(\mathscr{E}, \mathbb{F})$. Before defining it, we recall the following standard notation and notions.

(i) Let $S \equiv \{1, 2, 3\}$. Set $W_0 \equiv S^0 = \{\emptyset\}$ and for $n \in \mathbb{N}$, $W_n \equiv S^n = \{i_1i_2\cdots i_n | i_j \in S\}$. Let $W_* \equiv \cup_{n\in\mathbb{N}\cup\{0\}} W_n$. Set $\Sigma \equiv S^{\mathbb{N}} = \{i_1i_2\cdots | i_n \in S\}$ and for $w \in W_n$, $\Sigma_w \equiv \{v \in \Sigma | v_1\cdots v_n = w\}$.

(ii) For $w = w_1\cdots w_n \in W_n$ with $n \in \mathbb{N}\cup\{0\}$, define $|w| = n$, write $F_w \equiv F_{w_1} \circ \cdots \circ F_{w_n}$ if $n \ne 1$ and $F_w = Id_K$ if $n = 0$, and set $K_w \equiv F_w(K)$. For $n \in \mathbb{N}\cup\{0\}$, $V_n \equiv \cup_{w\in W_n} F_w(V_0)$. $V_* = \cup_{n\in\mathbb{N}\cup\{0\}} V_n$.

(iii) For $w = w_1w_2\cdots \in \Sigma$, define $\pi(w) = \cap_{n\ge1} K_{w_1\cdots w_n}$. Then $\pi : \Sigma \to K$ is continuous, surjective, and $\sharp(\pi^{-1}(x)) = 2$ if $x \in \cup_{n=1}^\infty V_n$ and $\sharp\pi^{-1}(x) = 1$ otherwise. For $w \in W_*$, $\pi(\Sigma_w) = K_w$ and $\Sigma_w = \pi^{-1}(K_w)$.

For $n \in \mathbb{N} \cup \{0\}$ and each pair of $u, v \in \mathscr{C}(K)$, define

$$\mathscr{E}_n(u, v) \equiv \frac{1}{8}\left(\frac{5}{3}\right)^n \sum_{p,q \in V_n, p\sim q} [u(p) - u(q)][v(p) - v(q)],$$

where $p \sim q$ if and only if $p, q \in F_w(V_0)$ for some $w \in W_n$. Notice that $\mathscr{E}_n$ is a non-negative definite symmetric quadratic form on $\mathscr{C}(K)$ and $\mathscr{E}_n(u, u) \le \mathscr{E}_{n+1}(u, u)$ for all $u \in \mathscr{C}(K)$. Then the following resistance form $(\mathscr{E}, \mathbb{F})$ is well defined.

Definition 4.2. Let $\mathbb{F} \equiv \{u \in \mathscr{C}(K) \mid \lim_{n\to\infty} \mathscr{E}_n(u, u) < \infty\}$ and define $\mathscr{E}(u, v) \equiv \lim_{n\to\infty} \mathscr{E}_n(u, v)$ for all $u, v \in \mathbb{F}$.

Observe that, associated to the above resistance form $(\mathscr{E}, \mathbb{F})$, the square gradient $\Gamma(u, v)$ is well defined by (2.1) as a signed Radon measure on K for each pair of $u, v \in \mathbb{F}$.

Kusuoka [26] endowed $(K, \mathscr{E}, \mathbb{F})$ with a "Riemannian volume" measure. Here we recall its definition via the harmonic embedding Φ of K into $\mathbb{R}^2$; see [22,19]. We say that $h \in \mathbb{F}$ is an *E-harmonic function* for some compact set $E \subset K$ if $\mathscr{E}(h, u) = 0$ for all $u \in \mathbb{F}$ with $u = 0$ on E. Let $h_1, h_2 \in \mathbb{F}$ be V_0-harmonic functions satisfying

$$h_1(p_1) = h_2(p_1) = 0, \qquad h_1(p_2) = h_1(p_3) = 1, \quad \text{and} \quad -h_2(p_2) = h_2(p_3) = \frac{1}{\sqrt{3}}.$$

For the existence of such functions see, for example, [23, Section 3.2]. Observe that by [22, Theorem 3.6], the harmonic embedding $\Phi \equiv (h_1, h_2)$ actually induces a homeomorphism between K and $\Phi(K)$. Due to this, $\Phi(K)$ is called the harmonic Sierpinski gasket.

Definition 4.3. The *Kusuoka measure m* on K is defined by $m \equiv \Gamma(h_1, h_1) + \Gamma(h_2, h_2)$.

Notice that the Kusuoka measure m is non-atomic and satisfies $m(U) > 0$ for all open nonempty sets $U \subset K$; see [26,22] and below. Then by [23, Theorem 3.4.6], $(K, \mathscr{E}, m)$ is a strongly local, regular Dirichlet form on $L^2(K, m)$ with domain $\mathbb{D} = \mathbb{F}$. The strong locality obviously follows from the definition of $\mathscr{E}$. The intrinsic distance d associated to $(K, \mathscr{E}, m)$ is then defined as in (2.2). To distinguish it from the Euclidean distance on $\mathbb{R}^2$, we let $B_d(x, r) = \{y \in K \mid d(x, y) < r\}$ for $x \in K$ and $r > 0$, and denote by $\mathrm{Lip}_d(K)$ the space of Lipschitz functions and by $\mathrm{Lip}_d\, u$ (resp. $\mathrm{apLip}_d\, u$) the (resp. approximative) pointwise Lipschitz constant with respect to d.

The following result identifies the intrinsic length structure with the gradient structure on $(K, \mathscr{E}, d, m)$.

Theorem 4.1. *For every $u \in \mathbb{D}$, the energy measure $\Gamma(u, u)$ is absolutely continuous with respect to the Kusuoka measure m and $\frac{d}{dm}\Gamma(u, u) = (\mathrm{Lip}_d\, u)^2$ almost everywhere.*

To prove this, we first recall the following properties of $(K, \mathscr{E}, d, m)$.

Proposition 4.1. *The topology induced by d coincides with the original topology on K inherited from $\mathbb{R}^2$, (K, d, m) satisfies a doubling property, and $(K, \mathscr{E}, m)$ supports a weak $(1, 2)$-Poincaré inequality.*

Proposition 4.1 has been proved in [24, Theorem 6.2] and [19, Lemma 3.7 and Proposition 3.20] with the aid of the dual formula: for all $x, y \in K$,

$$d(x, y) = \inf\{\ell_{\mathbb{R}^2}(\Phi \circ \gamma) \mid \gamma : [0, 1] \to K, \gamma \text{ is continuous}, \gamma(0) = x, \gamma(1) = y\}, \tag{4.1}$$

where $\ell_{\mathbb{R}^2}(\Phi \circ \gamma)$ denotes the length of $\Phi \circ \gamma : [0,1] \to \mathbb{R}^2$ with respect to the Euclidean distance. Recall that (4.1) is proved in [19, Theorem 4.2], and the right hand side of (4.1) is first introduced in [24] as the harmonic geodesic metric. From (4.1), it easily follows that

$$d_\Phi(x, y) \equiv |\Phi(x) - \Phi(y)| \le d(x, y). \tag{4.2}$$

But, as pointed out in [24, p. 800], d_Φ is not comparable to d; indeed, there exists a double sequence $\{x_n, y_n\}_{n\in\mathbb{N}} \subset K$ such that $d_\Phi(x_n, y_n)/d(x_n, y_n) \to 0$ as $n \to \infty$.

We also need the following Rademacher theorem on K, which is a corollary to Theorem 4.2.

Proposition 4.2. *For every $u \in \mathbb{D}$, there exists a unique measurable vector field $\widetilde{\nabla} u$ such that $\frac{d}{dm}\Gamma(u,u) = |\widetilde{\nabla} u|^2$ almost everywhere, and for almost all $x \in K$ and all $y \in K$,*

$$|u(y) - u(x) - \widetilde{\nabla} u(x) \cdot (\Phi(y) - \Phi(x))| = o(d(x, y)). \tag{4.3}$$

With the help of the results in Section 2 and the above Propositions 4.1, 4.2 and (4.2), we now prove Theorem 4.1.

Proof of Theorem 4.1. Combining Proposition 4.1 and Theorem 2.2, we have that for all $u \in \mathbb{D}$, $\Gamma(u,u)$ is absolutely continuous with respect to m and (2.22) holds with some $C_1 \ge 1$. By Theorem 2.2 and Proposition 2.2 or some density arguments, the proof of Theorem 4.1 is reduced to verifying that we may take $C_1 = 1$. For almost all $x \in K$ satisfying (4.3) for all y, applying Proposition 4.2 and (4.2), we have

$$\begin{aligned}
\frac{|u(y) - u(x)|}{d(x,y)} &\le \frac{|u(y) - u(x) - \widetilde{\nabla} u(x) \cdot (\Phi(y) - \Phi(x))|}{d(x,y)} + \frac{|\widetilde{\nabla} u(x) \cdot (\Phi(y) - \Phi(x))|}{d(x,y)} \\
&\le o(1) + \left|\widetilde{\nabla} u(x)\right| \frac{|\Phi(y) - \Phi(x)|}{d(x,y)} \\
&\le o(1) + \left|\widetilde{\nabla} u(x)\right|,
\end{aligned}$$

which implies that $\mathrm{Lip}_d u(x) \le |\widetilde{\nabla} u(x)|$. This finishes the proof of Theorem 4.1. □

To prove Proposition 4.2, we need several geometric properties of $(K, \mathscr{E}, d, m)$. We first recall the geometric description of $\Phi(K)$; see [26,22] and also [24,19]. Let $\{T_i\}_{i=1}^3$ be the linear transformation on $\mathbb{R}^2$ with the matrix representations:

$$T_1 \equiv \begin{pmatrix} 3/5 & 0 \\ 0 & 1/5 \end{pmatrix}, \qquad T_2 \equiv \begin{pmatrix} 3/10 & -\sqrt{3}/10 \\ -\sqrt{3}/10 & 1/2 \end{pmatrix} \quad \text{and}$$

$$T_3 \equiv \begin{pmatrix} 3/10 & \sqrt{3}/10 \\ \sqrt{3}/10 & 1/2 \end{pmatrix}.$$

Define $H_i(x) = \Phi(p_i) + T_i(x - \Phi(p_i))$ for all $x \in \mathbb{R}^2$, $i = 1, 2, 3$. Then $\Phi(K)$ is exactly the self-similar set determined by the system $\{H_i\}_{i=1}^3$, namely, $\Phi(K) = \cup_{i=1}^3 H_i(\Phi(K))$. Moreover, $H_i \circ \Phi = \Phi \circ H_i$ and $\Phi : K \to \Phi(K)$ is a homeomorphism.

We recall the "Riemannian volume" m on Σ introduced in [26], namely, the Kusuoka measure via geometric description. There exists a unique Borel regular probability measure m_Σ on Σ such that for all $w = w_1 \cdots w_n \in W_*$, $m_\Sigma(\Sigma_w) = (5/3)^n \|T_w\|^2$, where $T_w = T_{w_1} \circ \cdots \circ T_{w_n}$ and

$\|T_w\|$ denotes its Hilbert–Schmidt norm; see [26]. The pushforward measure $\pi_* m_\Sigma = m_\Sigma \circ \pi^{-1}$ is exactly the Kusuoka measure m as in Definition 4.3. Indeed, for $w = w_1 \cdots w_n \in W_*$,

$$\pi_* m_\Sigma(K_w) = m_\Sigma(\pi^{-1}(K_w)) = m_\Sigma(\Sigma_w) = \left(\frac{5}{3}\right)^n \|T_w\|^2 = m(K_w); \tag{4.4}$$

see [22] and also [19, Proposition 2.14].

Now we collect some further properties, which will be used later. See [26,22,24,19] for their proofs or details.

Lemma 4.1. (i) *If $u, v \in \mathbb{D}$, then $u \circ F_i, v \circ F_i \in \mathbb{D}$ for $i = 1, 2, 3$ and*

$$\mathscr{E}(u, v) = \frac{5}{3} \sum_{i=1}^{3} \mathscr{E}(u \circ F_i, v \circ F_i).$$

(ii) *There exists a constant $C_2 \geq 1$ such that for $u \in \mathbb{D}$,*

$$\operatorname*{osc}_K u \leq C_2 \sqrt{\mathscr{E}(u, u)},$$

where $\operatorname{osc}_E u \equiv \sup_{x \in E} u(x) - \inf_{x \in E} u(x)$ for any set E.

For $s \in (0, 1]$, denote by $\Lambda(s)$ the collection of all $w = w_1 \cdots w_n \in W_*$ such that $\|T_w\| \leq s < \|T_{w_1} \circ \cdots \circ T_{w_{n-1}}\|$ when $n \geq 2$ and $\|T_w\| \leq s$ when $n = 1$. For $x \in K$ and $s \in (0, 1]$, set

$$K(x, s) \equiv \bigcup_{w \in \Lambda(s), x \in K_w} K_w, \quad \text{and} \quad U(x, s) \equiv \bigcup_{w \in \Lambda(s), K_w \cap K(x,s) \neq \emptyset} K_w.$$

Then we have the following results; see [24,19].

Lemma 4.2. (i) *For all $x \in K$, $s \in (0, 1]$ and $w \in \Lambda(s)$,*

$$\sharp\{v \in \Lambda(s) | K_v \cap K(x, s) \neq \emptyset\} \leq 6 \quad \text{and} \quad \sharp\{v \in \Lambda(s) | K_v \cap K_w \neq \emptyset\} \leq 4, \tag{4.5}$$

and that

$$B_d(x, \sqrt{2}s/50) \subset U(x, s) \subset B_d(x, 10s). \tag{4.6}$$

(ii) *There exists a positive constant $C_3 \geq 1$ such that if $w, v \in \Lambda(s)$ and $K_w \cap K_v \neq \emptyset$, then*

$$C_3^{-1} m(K_v) \leq m(K_w) \leq C_3 m(K_v). \tag{4.7}$$

(iii) *There exists $C_4 \geq 1$ such that for all $w \in W_*$ and $i \in \{1, 2, 3\}$,*

$$C_4^{-1} m(K_w) \leq m(K_{wi}) \leq m(K_w). \tag{4.8}$$

Applying the properties above, we obtain the following results.

Lemma 4.3. (i) *There exists a positive integer N such that for all $x \in K$ and $s \in (0, 1)$, if $w, v \in \Lambda(s)$ satisfy $x \in K_w$ and $K_v \cap K(x, s) \neq \emptyset$, then* $\max\{|w| - N, 0\} \leq |v| \leq |w| + N$.

(ii) *There exists a constant $C_6 \geq 1$ such that for all $x \in K$ and $s \in (0, 1)$, if $w \in \Lambda(s)$ and $x \in K_w$, then $m(B_d(x, s)) \leq C_6 m(K_w)$.*

Proof. (i) Without loss of generality, we may assume that $K_w \cap K_v \neq \emptyset$. Indeed, there must exist $\sigma \in \Lambda(s)$ such that $K_\sigma \cap K_w \neq \emptyset$ and $K_\sigma \cap K_v \neq \emptyset$. By (4.4), $v \in \Lambda(s)$ and (4.8), we have

$$C_4^{-1}\left(\frac{5}{3}\right)^{|v|-1} s^2 \le C_4^{-1}\left(\frac{5}{3}\right)^{|v|-1} \|T_{v_1\cdots v_{n-1}}\|^2 \le \left(\frac{5}{3}\right)^{|v|} \|T_v\|^2 \le \left(\frac{5}{3}\right)^{|v|} s^2,$$

and the same inequality also holds with v replaced by w. From this, it is easy to see that $\|T_w\| \le s < \|T_{w_1\cdots w_{n-1}}\| \lesssim s$. Moreover, by (4.4), $w, v \in \Lambda(s)$ and (4.7), we also have

$$C_3^{-1}\left(\frac{5}{3}\right)^{|v|} \|T_v\|^2 \le \left(\frac{5}{3}\right)^{|w|} \|T_w\|^2 \le C_3\left(\frac{5}{3}\right)^{|v|} \|T_v\|^2,$$

which gives that $\|T_w\|^2 \sim \left(\frac{5}{3}\right)^{|v|-|w|} s^2$. Hence, $\left(\frac{5}{3}\right)^{|v|-|w|} \sim 1$, which yields (i).

(ii) By the doubling property and (4.6),

$$m(B_d(x,s)) \lesssim m(B_d(x, \sqrt{2}s/25)) \lesssim m(U(x,s)).$$

Then, by (4.5), it suffices to show that for all $v \in \Lambda(s)$ such that $K_v \cap K(x,s) \neq \emptyset$, we have $m(K_v) \lesssim m(K_w)$. But this follows from (4.7) and the fact that there must exist $\sigma \in \Lambda(s)$ such that $K_\sigma \cap K_w \neq \emptyset$ and $K_\sigma \cap K_v \neq \emptyset$. This finishes the proof of Lemma 4.3. □

Lemma 4.4. *Let $h \equiv h_1$ or $h \equiv h_2$. Then for all $u, v, f, g \in \mathbb{D}$ and for almost all $x \in K$,*

$$\frac{d\Gamma(u,v)}{d\Gamma(h,h)}(x)\frac{d\Gamma(f,g)}{d\Gamma(h,h)}(x) = \frac{d\Gamma(u,g)}{d\Gamma(h,h)}(x)\frac{d\Gamma(f,v)}{d\Gamma(h,h)}(x), \tag{4.9}$$

and for all $c \in \mathbb{R}$,

$$\frac{d\Gamma(cu+v,g)}{d\Gamma(h,h)}(x) = c\frac{d\Gamma(u,g)}{d\Gamma(h,h)}(x) + \frac{d\Gamma(v,g)}{d\Gamma(h,h)}(x). \tag{4.10}$$

Moreover, (4.9) and (4.10) also hold with $\Gamma(h,h)$ replaced by m.

Proof. By [15, Theorem 5.6], for every $u \in \mathbb{D}$, $\Gamma(u,u)$ is absolutely continuous with respect to $\Gamma(h,h)$. Thus $\Gamma(h,h)$ and m are mutually absolutely continuous, and moreover, for every $u \in \mathbb{F}$, by (2.1) and the Cauchy–Schwarz inequality, both $\Gamma(u,h)$ and $\Gamma(u,u)$ are absolutely continuous with respect to m and $\Gamma(h,h)$. Let $\{f_i\}_{i\in\mathbb{N}}$ be an arbitrary complete orthonormal system of $\mathbb{D}$. By [15, Proposition 2.12] and the fact that the index of $(K, \mathscr{E})$ is 1 (see [15, Section 4]), there exists a sequence of functions, $\{\zeta^i\}_{i\in\mathbb{N}}$, such that for all $i, j \in \mathbb{N}$, $\frac{d\Gamma(f_i,f_j)}{d\Gamma(h,h)} = \zeta^i\zeta^j$ almost everywhere. Recall that u has a unique representation $u = \sum_{i\in\mathbb{N}} a_i(u) f_i$ with $\sum_{i\in\mathbb{N}} a_i(u)^2 < \infty$. Then $\gamma(u) \equiv \sum_{i\in\mathbb{N}} a_i(u)\zeta^i \in L^2(K, \Gamma(h,h))$ is well defined. Indeed, for $u, v \in \mathbb{D}$,

$$\left(\sum_{i=1}^N a_i(u)\zeta^i\right)\left(\sum_{j=1}^N a_j(v)\zeta^j\right) = \sum_{i,j=1}^N a_i(u)a_j(v)\frac{d\Gamma(f_i,f_j)}{d\Gamma(h,h)}$$

$$= \frac{d\Gamma\left(\sum_{i=1}^N a_i(u)f_i, \sum_{j=1}^N a_j(v)f_j\right)}{d\Gamma(h,h)} \to \frac{d\Gamma(u,v)}{d\Gamma(h,h)}$$

as $N \to \infty$ in $L^2(K, \Gamma(h,h))$ and hence almost everywhere. Moreover, from this, we deduce that $\gamma(u)\gamma(v) = \frac{d\Gamma(u,v)}{d\Gamma(h,h)}$ almost everywhere, which implies (4.9).

The linearity property $a_i(cu+v) = ca_i(u)+a_i(v)$ implies linearity of γ and hence also (4.10) by repeating the above argument.

Since $\Gamma(h,h)$ and m are mutually absolutely continuous, (4.9) (resp. (4.10)) with $\Gamma(h,h)$ replaced by m follows from the Radon–Nikodym theorem and (4.9) (resp. (4.10)). Indeed, for almost all $x \in X$,

$$\begin{aligned}\frac{d\Gamma(u,v)}{dm}(x) &= \lim_{r\to 0}\frac{\int_X 1_{B(x,r)}\,d\Gamma(u,v)}{\int_X 1_{B(x,r)}\,dm}\\ &= \lim_{r\to 0}\frac{\int_X 1_{B(x,r)}d\,\Gamma(u,v)}{\int_X 1_{B(x,r)}\,d\Gamma(h,h)}\cdot\lim_{r\to 0}\frac{\int_X 1_{B(x,r)}d\,\Gamma(h,h)}{\int_X 1_{B(x,r)}\,dm}\\ &= \frac{d\Gamma(u,v)}{d\Gamma(h,h)}(x)\frac{d\Gamma(h,h)}{dm}(x).\end{aligned}$$

This finishes the proof of Lemma 4.4. □

Remark 4.1. The set of points $x \in K$ for which (4.10) holds is independent of $c \in \mathbb{R}$.

The following Rademacher theorem is an improvement on [15, Theorem 5.4] and [19, Theorem 2.17(ii)].

Proposition 4.3. *Let $h \equiv h_1$ or $h \equiv h_2$. For every $u \in \mathbb{D}$, there exists a unique measurable function $\frac{du}{dh}$ such that for almost all $x \in K$ and all $y \in K$,*

$$\left|u(y)-u(x)-\frac{du(x)}{dh}(h(y)-h(x))\right| = o(d(x,y)). \tag{4.11}$$

Moreover, $\frac{d\Gamma(u,u)}{d\Gamma(h,h)} = |\frac{du}{dh}|^2$ almost everywhere.

Proof of Proposition 4.3. Set

$$\frac{du}{dh} \equiv \frac{d\Gamma(u,h)}{d\Gamma(h,h)},$$

and

$$R_x(y) \equiv u(y)-u(x)-\frac{du(x)}{dh}(h(y)-h(x))$$

for all $y \in K$ whenever $\frac{du(x)}{dh}$ exists. Then $R_x(\cdot) \in \mathbb{D}$ and (4.9) implies that

$$\frac{d\Gamma(u,u)}{d\Gamma(h,h)}(x) = \frac{d\Gamma(u,u)}{d\Gamma(h,h)}(x)\frac{d\Gamma(h,h)}{d\Gamma(h,h)}(x) = \left(\frac{d\Gamma(u,h)}{d\Gamma(h,h)}(x)\right)^2 = \left(\frac{du(x)}{dh}\right)^2. \tag{4.12}$$

Now it suffices to prove that for almost all $x \in K$ and all $s \in (0,1)$,

$$\sup_{y\in B_d(x,s)}\left|u(y)-u(x)-\frac{du(x)}{dh}(h(y)-h(x))\right| = o(s). \tag{4.13}$$

Recall that for all $x \in K$ and $s \in (0,1/10]$, $B_d(x,s) \subset U(x,10s)$; see (4.6). Therefore, for each $y \in B_d(x,s)$, there exist $w, v \in \Lambda(10s)$ such that $x \in K_w,\ K_w \cap K_v \neq \emptyset$ and $y \in K_v$, where w and v may be equal. Taking $y_* \in K_w \cap K_v = F_w(V_0) \cap F_v(V_0)$ and using $R_x(x) = 0$, we obtain

$$\begin{aligned}|R_x(y)| &\le |R_x(y)-R_x(y_*)| + |R_x(y_*)-R_x(x)|\\ &\le \operatorname*{osc}_{K_w} R_x + \operatorname*{osc}_{K_v} R_x = \operatorname*{osc}_{K} R_x\circ F_w + \operatorname*{osc}_{K} R_x\circ F_v.\end{aligned}$$

By Lemma 4.1, we have

$$|R_x(y)| \lesssim \sqrt{\Gamma(R_x \circ F_w, R_x \circ F_w)(K)} + \sqrt{\Gamma(R_x \circ F_v, R_x \circ F_v)(K)}$$
$$\lesssim \sqrt{\left(\frac{3}{5}\right)^{|w|} \Gamma(R_x, R_x)(K_w)} + \sqrt{\left(\frac{3}{5}\right)^{|v|} \Gamma(R_x, R_x)(K_v)}.$$

Noticing that $|v| \sim |w|$ by Lemma 4.3 and $U(x, 10s) \subset B_d(x, 500s/\sqrt{2})$ by (4.6), for $s \in (0, 1/500]$, we have

$$\sup_{y \in B_d(x,s)} |R_x(y)| \lesssim \sqrt{\left(\frac{3}{5}\right)^{|w|} \Gamma(R_x, R_x)(B_d(x, 500s/\sqrt{2}))}.$$

On the other hand, for $s \in (0, 1)$,

$$\frac{\Gamma(R_x, R_x)(B_d(x, s))}{\Gamma(h, h)(B_d(x, s))} = \frac{\Gamma\left(u - \frac{du(x)}{dh}h, u - \frac{du(x)}{dh}h\right)(B_d(x, s))}{\Gamma(h, h)(B_d(x, s))}$$
$$= \frac{\Gamma(u, u)(B_d(x, s))}{\Gamma(h, h)(B_d(x, s))} - 2\frac{du(x)}{dh}\frac{\Gamma(u, h)(B_d(x, s))}{\Gamma(h, h)(B_d(x, s))} + \left(\frac{du(x)}{dh}\right)^2.$$

By this, (4.12), the definition of $\frac{du}{dh}$ and the Radon–Nikodym theorem, we conclude that

$$\lim_{s \to 0} \frac{\Gamma(R_x, R_x)(B_d(x, s))}{\Gamma(h, h)(B_d(x, s))} = \frac{d\Gamma(u, u)(x)}{d\Gamma(h, h)} - 2\frac{du(x)}{dh}\frac{d\Gamma(u, h)(x)}{d\Gamma(h, h)} + \left(\frac{du(x)}{dh}\right)^2 = 0$$

for almost all $x \in K$. Therefore,

$$\sup_{y \in B_d(x,s)} |R_x(y)| = o\left(\sqrt{\left(\frac{3}{5}\right)^{|w|} \Gamma(h, h)(B_d(x, 500s/\sqrt{2}))}\right). \tag{4.14}$$

Observe that, by the doubling property, definition of m and Lemma 4.1(ii) and Lemma 4.3, we have

$$\Gamma(h, h)(B_d(x, 500s/\sqrt{2})) \le m(B_d(x, 500s/\sqrt{2})) \lesssim m(B_d(x, 10s)) \lesssim m(K_w).$$

Thus by (4.4) and $w \in \Lambda(10s)$, we arrive at

$$\sup_{y \in B_d(x,s)} |R_x(y)| = o\left(\sqrt{\left(\frac{3}{5}\right)^{|w|} m(K_w)}\right) = o(\|T_w\|) = o(s),$$

as desired.

To see the uniqueness, assume that a is a measurable function such that (4.13) holds with $\frac{du(x)}{dh}$ replaced by $a(x)$, for almost all $x \in K$. Then

$$\sup_{y \in B_d(x,s)} \left|\frac{du(x)}{dh} - a(x)\right| |h(y) - h(x)| = o(s). \tag{4.15}$$

Take $x = \pi^{-1}(w) \in K_w$ such that the above holds, and for $s \in (0, 1)$, and $w_1 \cdots w_{n_s} \in \Lambda(\sqrt{s}/25)$. Observe that (4.6) gives $K_{w_1 \cdots w_{n_s}} \subset U(x, \sqrt{2}s/25) \subset B_d(x, s)$. Since

$$\sup_{y \in B_d(x,s)} |h(y) - h(x)| \ge \frac{1}{2} \operatorname*{osc}_{B_d(x,s)} h$$

$$\geq \frac{1}{2} \operatorname*{osc}_{K_{w_1\cdots w_{n_s}}} h \geq \frac{1}{2}\mathscr{E}(h\circ F_{w_1\cdots w_{n_s}}, h\circ F_{w_1\cdots w_{n_s}}),$$

it suffices to show that $\mathscr{E}(h\circ F_{w_1\cdots w_{n_s}}, h\circ F_{w_1\cdots w_{n_s}}) \gtrsim s$. Indeed, this would imply that $\frac{du(x)}{dh} - a(x) = 0$.

By the martingale convergence theorem and the mutual absolute continuity of m and $\Gamma(h,h)$, for almost all $x \in K$, we have

$$\begin{aligned}\lim_{s\to 0}\frac{\mathscr{E}(h\circ F_{w_1\cdots w_{n_s}}, h\circ F_{w_1\cdots w_{n_s}})}{m(K_{w_1\cdots w_{n_s}})} &= \lim_{s\to 0}\frac{\Gamma(h\circ F_{w_1\cdots w_{n_s}}, h\circ F_{w_1\cdots w_{n_s}})(K)}{m(K_{w_1\cdots w_{n_s}})}\\ &= \lim_{s\to 0}\frac{\Gamma(h,h)(K_{w_1\cdots w_{n_s}})}{m(K_{w_1\cdots w_{n_s}})} = \frac{d\Gamma(h,h)}{dm}(x) > 0,\end{aligned}$$

which together with (4.8) and $w_1\cdots w_{n_s} \in \Lambda(\sqrt{2}s/25)$ implies that

$$\mathscr{E}(h\circ F_{w_1\cdots w_{n_s}}, h\circ F_{w_1\cdots w_{n_s}}) \gtrsim m(K_{w_1\cdots w_{n_s}}) \gtrsim m(K_{w_1\cdots w_{n_s-1}}) \gtrsim s.$$

This finishes the proof of Proposition 4.3. □

Now, we are going to identify the intrinsic length structure with the "measurable Riemannian structure" of Kusuoka [26] in Theorem 4.2.

Recall that Kusuoka further introduced the "measurable Riemannian metric" Z on K. Indeed, for $w \in W_*$, set $Z_m(w) \equiv T_w T_w^*/\|T_w\|^2$. Kusuoka [26] proved that

$$Z(w) \equiv \lim_{n\to\infty} Z_m(w_1\cdots w_n) \tag{4.16}$$

exists for almost all $w = w_1w_2\cdots \in \Sigma$, and moreover, rank $Z(w) = 1$ and Z is the orthogonal projection onto its image for almost all $w \in \Sigma$. Since $m_\Sigma(V_*) = 0$, the pushforward mapping $\pi_*Z = Z\circ\pi^{-1}$, which is still denote by Z by abuse of notation, is well defined on K.

The above measurable Riemannian structure is identified with the gradient structure in the following sense; see [26,22] and [24, Theorem 4.8]. Let $\mathscr{C}^1(K) \equiv \{v\circ\Phi | v \in \mathscr{C}^1(\mathbb{R}^2)\}$. Then $\mathscr{C}^1(K)$ is dense in $(\mathbb{D}, \|\cdot\|_{\mathbb{D}})$. Moreover, for every $u \in \mathscr{C}^1(K)$, define $\nabla u \equiv (\nabla v)\circ\Phi$, which is well defined since it is independent of the choice of $v \in \mathscr{C}^1(\mathbb{R}^2)$ with $u = v\circ\Phi$; see [22, Section 4]. Here $\nabla v(x) \equiv (\frac{\partial v(x)}{\partial x_1}, \frac{\partial v(x)}{\partial x_2})$ denotes the usual gradient for $v \in \mathscr{C}^1(\mathbb{R}^2)$.

Proposition 4.4. *For every $u \in \mathbb{D}$, there exists a measurable vector field $Y(u) \in \mathbb{R}^2$ such that $Y(u) \in \operatorname{Im} Z$, $\frac{d}{dm}\Gamma(u,u) = |Y(u)|^2$ almost everywhere, and hence $\mathscr{E}(u,u) = \int_K |Y(u)(x)|^2\, dm(x)$. Moreover, if $u \in \mathscr{C}^1(K)$, then $Y(u) = Z\nabla u$.*

Applying Propositions 4.3 and 4.4, we have a formula for the projection Z via the harmonic embedding (or coordinate) Φ.

Lemma 4.5. *The pushforward π_*Z to K of the projection Z on Σ as in (4.16) is given by*

$$\pi_*Z = \begin{pmatrix} \dfrac{1}{1+a^2} & \dfrac{a}{1+a^2} \\ \dfrac{a}{1+a^2} & \dfrac{a^2}{1+a^2}\end{pmatrix}$$

almost everywhere, where $a = \frac{dh_2}{dh_1} = \frac{d\Gamma(h_2,h_1)}{d\Gamma(h_1,h_1)}$*. The eigenvalues of* $\pi_* Z$ *are* $\lambda_1 = 0$ *and* $\lambda_2 = a^2 + 1$*, the corresponding eigenvectors are* $\xi_1 = (-\frac{a}{1+a^2}, \frac{1}{1+a^2})$ *and* $\xi_2 = (\frac{1}{1+a^2}, \frac{a}{1+a^2})$*. The projection space is* $\mathrm{Im} Z = (\frac{1}{1+a^2}, \frac{a}{1+a^2})\mathbb{R}$*.*

Proof. By the Radon–Nikodym theorem and Proposition 4.3, we have

$$a^2 = \frac{d\Gamma(h_2,h_1)}{d\Gamma(h_1,h_1)}\frac{d\Gamma(h_2,h_1)}{d\Gamma(h_1,h_1)} = \frac{d\Gamma(h_1,h_1)}{d\Gamma(h_1,h_1)}\frac{d\Gamma(h_2,h_2)}{d\Gamma(h_1,h_1)} = \frac{d\Gamma(h_2,h_2)}{d\Gamma(h_1,h_1)}$$

and

$$\frac{dm}{d\Gamma(h_1,h_1)} = \frac{d\Gamma(h_1,h_1)}{d\Gamma(h_1,h_1)} + \frac{d\Gamma(h_2,h_2)}{d\Gamma(h_1,h_1)} = 1 + a^2.$$

Hence by Proposition 4.4 and the Radon–Nikodym theorem,

$$Ze_i \cdot e_j = Z\nabla h_i \cdot \nabla h_j = \frac{d\Gamma(h_1,h_1)}{dm} = \frac{d\Gamma(h_i,h_j)}{d\Gamma(h_1,h_1)}\frac{d\Gamma(h_1,h_1)}{dm} = \frac{1}{1+a^2}\frac{d\Gamma(h_i,h_j)}{d\Gamma(h_1,h_1)},$$

which implies that

$$Ze_1 \cdot e_1 = \frac{1}{1+a^2}, \qquad Ze_1 \cdot e_2 = Ze_2 \cdot e_1 = \frac{a}{1+a^2} \quad \text{and} \quad Ze_2 \cdot e_2 = \frac{a^2}{1+a^2}.$$

The other conclusions follow from this by standard computations. This finishes the proof of Lemma 4.5. □

Now we improve Proposition 4.4 and [27, Theorem 2.17(i)] as follows. Notice that Proposition 4.2 follows from Theorem 4.2.

Theorem 4.2. *For every* $u \in \mathbb{D}$*, there exists a unique measurable vector field* $\widetilde{\nabla} u$ *such that for almost all* $x \in K$*, and for all* $s \in (0,1)$*,*

$$\sup_{y \in B_d(x,s)} |u(y) - u(x) - \widetilde{\nabla} u(x) \cdot (\Phi(y) - \Phi(x))| = o(s). \tag{4.17}$$

Moreover, $\widetilde{\nabla} u = Y(u) \in \mathrm{Im}\, Z$ *and* $\frac{d}{dm}\Gamma(u,u) = |Y(u)|^2 = |\widetilde{\nabla} u|^2$ *almost everywhere; in particular, if* $u \in \mathscr{C}^1(K)$*, then* $\widetilde{\nabla} u = Y(u) = Z\nabla u$ *almost everywhere.*

Proof. We first observe that, by Proposition 4.4, $\nabla h_1 = e_1$, and mutual absolute continuity of $\Gamma(h_1,h_1)$ and m, we obtain

$$|Z(x)e_1|^2 = |Z(x)\nabla h_1(x)|^2 = \frac{d\Gamma(h_1,h_1)}{dm}(x) > 0$$

for almost all $x \in K$. Similarly, we have that $|Z(x)e_2|^2 > 0$ for almost all $x \in K$. For such an $x \in K$, let $\zeta \equiv (\zeta_1, \zeta_2) = \zeta_1 e_1 + \zeta_2 e_2$ with $\zeta_1 \equiv |Ze_1|$ and $\zeta_2 \equiv Ze_1 \cdot e_2/|Ze_1|$. Take $\widetilde{\nabla} u \equiv \frac{du}{dh_1}\zeta_1\zeta$. Then for almost all $x \in K$, obviously, $\widetilde{\nabla} u(x) \in \mathrm{Im} Z(x)$. Since $|\zeta(x)|^2 = 1$ and $(\zeta_1(x))^2 = \frac{d\Gamma(h_1,h_1)}{dm}(x)$, applying Lemma 4.4 and (4.12), we further have

$$|\widetilde{\nabla} u(x)|^2 = \left(\frac{du(x)}{dh_1}\right)^2 \frac{d\Gamma(h_1,h_1)}{dm}(x) = \frac{d\Gamma(u,u)}{d\Gamma(h_1,h_1)}(x)\frac{d\Gamma(h_1,h_1)}{dm}(x) = \frac{d\Gamma(u,u)}{dm}(x).$$

Whenever $\widetilde{\nabla} u(x)$ exists, write

$$\begin{aligned}
&u(y) - u(x) - \widetilde{\nabla} u(x) \cdot (\Phi(y) - \Phi(x)) \\
&\quad = (\zeta_1(x))^2 \left[u(y) - u(x) - \frac{du(x)}{dh_1}(h_1(y) - h_1(x)) \right] \\
&\qquad + \left[(1 - (\zeta_1(x))^2)(u(y) - u(x)) - \zeta_1(x)\zeta_2(x)\frac{du(x)}{dh_1}(h_2(y) - h_2(x)) \right] \\
&\quad \equiv \widetilde{R}_x^{(1)}(y) + \widetilde{R}_x^{(2)}(y)
\end{aligned}$$

for all $y \in K$. Observe that Proposition 4.3 implies that $\sup_{y \in B_d(x,s)} |\widetilde{R}_x^{(1)}(y)| = o(s)$. Then (4.17) is reduced to proving $\sup_{y \in B_d(x,s)} |\widetilde{R}_x^{(2)}(y)| = o(s)$. To this end, observe that by Proposition 4.4 and the Radon–Nikodym theorem,

$$1 - (\zeta_1(x))^2 = 1 - \frac{d\Gamma(h_1, h_1)}{dm}(x) = \frac{d\Gamma(h_2, h_2)}{dm}(x)$$

and

$$\zeta_1(x)\zeta_2(x) = Z\nabla h_1(x) \cdot \nabla h_2(x) = \frac{d\Gamma(h_1, h_2)}{dm}(x) = \frac{d\Gamma(h_1, h_1)}{dm}(x)\frac{d\Gamma(h_1, h_2)}{d\Gamma(h_1, h_1)}(x)$$

for almost all $x \in K$. Also by Proposition 4.3 and the Radon–Nikodym theorem, for almost all $x \in K$,

$$\begin{aligned}
\frac{du}{dh_1}(x) &= \frac{d\Gamma(u, h_1)}{d\Gamma(h_1, h_1)}(x) \\
&= \frac{d\Gamma(u, h_1)}{d\Gamma(h_2, h_2)}(x)\frac{d\Gamma(h_2, h_2)}{d\Gamma(h_2, h_2)}(x)\frac{d\Gamma(h_2, h_2)}{d\Gamma(h_1, h_1)}(x) \\
&= \frac{d\Gamma(u, h_2)}{d\Gamma(h_2, h_2)}(x)\frac{d\Gamma(h_1, h_2)}{d\Gamma(h_2, h_2)}(x)\frac{d\Gamma(h_2, h_2)}{d\Gamma(h_1, h_1)}(x) \\
&= \frac{d\Gamma(u, h_2)}{d\Gamma(h_2, h_2)}(x)\frac{d\Gamma(h_1, h_2)}{d\Gamma(h_1, h_1)}(x) \\
&= \frac{du}{dh_2}(x)\frac{dh_2}{dh_1}(x).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\zeta_1(x)\zeta_2(x)\frac{du}{dh_1}(x) &= \frac{d\Gamma(h_1, h_1)}{dm}(x)\frac{d\Gamma(h_1, h_2)}{d\Gamma(h_1, h_1)}(x)\frac{d\Gamma(h_1, h_2)}{d\Gamma(h_1, h_1)}(x)\frac{d\Gamma(u, h_2)}{d\Gamma(h_2, h_2)}(x) \\
&= \frac{d\Gamma(h_1, h_1)}{dm}(x)\frac{d\Gamma(h_2, h_2)}{d\Gamma(h_1, h_1)}(x)\frac{d\Gamma(u, h_2)}{d\Gamma(h_2, h_2)}(x) \\
&= \frac{d\Gamma(h_2, h_2)}{dm}(x)\frac{d\Gamma(u, h_2)}{d\Gamma(h_2, h_2)}(x) \\
&= (1 - (\zeta_1(x))^2)\frac{du}{dh_2}(x).
\end{aligned}$$

Thus

$$\widetilde{R}_x^{(2)}(y) = (1 - (\zeta_1(x))^2)\left[u(y) - u(x) - \frac{du(x)}{dh_2}(h_2(y) - h_2(x)) \right]$$

and hence by Proposition 4.3, $\sup_{y \in B_d(x,s)} |\widetilde{R}_x^{(2)}(y)| = o(s)$ as desired.

The uniqueness of $\widetilde{\nabla} u$ follows from exactly the same argument that was used in the proof of Proposition 4.3. $\square$

We also extend Theorems 4.1 and 4.2 to the case of an energy measure $\Gamma(h, h)$, where h is a nontrivial V_0-harmonic function with $\mathscr{E}(h, h) = 1$. Notice that $\Gamma(h, h)$ and m are mutually absolutely continuous. Recall that $(K, \mathscr{E}, \Gamma(h, h))$ is a strongly local Dirichlet form on $L^2(K, \Gamma(h, h))$ with domain $\mathbb{D} = \mathbb{F}$. Denote by d_h the associated intrinsic distance. Propositions 4.1 and 4.3 still hold for $(K, \mathscr{E}, \Gamma(h, h), d_h)$ and a dual formula similar to (4.1) is still available. For the above, see [22,24,19]. The following result identifies the length structure with the length of the gradient on $(K, \mathscr{E}, \Gamma(h, h), d_h)$.

Theorem 4.3. *For every $u \in \mathbb{D}$, the energy measure $\Gamma(u, u)$ is absolutely continuous with respect to the Kusuoka measure $\Gamma(h, h)$ and the square of the length of the gradient satisfies $\frac{d\Gamma(u,u)}{d\Gamma(h,h)} = (\mathrm{Lip}_{d_h} u)^2$ almost everywhere. Moreover, for almost all $x \in K$ and all $y \in K$,*

$$\left| u(y) - u(x) - \frac{du(x)}{dh}(h(y) - h(x)) \right| = o(d_h(x, y))$$

where $\frac{du}{dh} = \frac{d\Gamma(u,h)}{d\Gamma(h,h)}$ and $|\frac{du}{dh}|^2 = \frac{d\Gamma(u,u)}{d\Gamma(h,h)}$.

We point out that Theorem 4.3 can be proved by repeating the above arguments as in Theorems 4.1 and 4.2, and all the properties needed in the arguments are available by [22,24,19]. We omit the details.

5. Heat flow, gradient flow and $\frac{d}{dm}\Gamma(u, u) = (\mathrm{Lip}\, u)^2$

In this section, under the Ricci curvature bounds of Lott–Sturm–Villani, we will clarify the relations between the coincidence of the intrinsic length structure and the gradient structure and the identification of the heat flow of $\mathscr{E}$ and the gradient flow of entropy; see Theorems 5.1 and 5.2. We begin with the definition of Wasserstein distance.

On a given metric space (X, d), denote by $\mathscr{P}(X)$ the *collection of all Borel probability measures* on X and endow it with weak $*$-topology, that is, $\mu_i \to \mu$ if and only if for all $f \in \mathscr{C}(X)$, $\int_X f\, d\mu_i \to \int_X f\, d\mu$. For $p \in [1, \infty)$, denote by $\mathscr{P}_p(X)$ the collection of all measures $\mu \in \mathscr{P}(X)$ such that $\int_X d^p(x_1, x)\, d\mu(x) < \infty$. Moreover, for every pair of $\mu, \nu \in \mathscr{P}(X)$, define the *$L^p$-Wasserstein distance* as

$$W_p(\mu, \nu) \equiv \inf_\pi \left(\int_{X \times X} [d(x, y)]^p\, d\pi(x, y) \right)^{1/p}, \tag{5.1}$$

where the infimum is taken over all couplings π of μ and ν. Recall that a *coupling* π of μ and ν is a probability measure $\pi \in \mathscr{P}(X \times X)$ with the property that for all measurable sets $A \subset X, \pi(A \times X) = \mu(A)$ and $\pi(X \times A) = \nu(A)$. There always exists (at least) one optimal coupling, and so the above infimum can be replaced by minimum; see for example [51, Proposition 2.1].

In the rest of this section, we always assume that X is compact, $\mathscr{E}$ is a regular, strongly local Dirichlet form on X and m is a probability measure, namely, $m(X) = 1$. Let d be the associated intrinsic distance as in (2.2) and assume that the topology induced by d coincides with the original topology on X. Then (X, d) is a compact length space by [44,48], and hence, $\mathscr{P}_2(X) = \mathscr{P}(X)$ equipped with the distance W_2 is a compact length space (hence a geodesic space); see [29]. Notice that the topology induced by W_2 coincides with the above weak $*$-topology (see for example [51]).

Let $U : [0, \infty) \to [0, \infty)$ be a continuous convex function with $U(0) = 0$ and define the associated *functional* $\mathscr{U} : \mathscr{P}_2(X) \to \mathbb{R} \cup \{+\infty\}$ by setting

$$\mathscr{U}(\mu) \equiv \int_X U\left(\frac{d\mu}{dm}\right) dm + U'(\infty)\mu_{\mathrm{sing}}(X),$$

where μ_{sing} is the singular part of the Lebesgue decomposition of μ with respect to m, and $U'(\infty) \equiv \lim_{r\to\infty} \frac{1}{r}U(r)$. If $U'(\infty) = \infty$, then $\mathscr{U}(\mu) < \infty$ means that μ is absolutely continuous with respect to m, namely, $\mu_{\mathrm{sing}} = 0$. If $U'(\infty) < \infty$, this is not necessarily the case.

Definition 5.1. Let U be a continuous convex function with $U(0) = 0$ and $\lambda \in \mathbb{R}$. Then $\mathscr{U}$ is called *weakly λ-displacement convex* if for all $\mu_0, \mu_1 \in \mathscr{P}_2(X)$, there exists some Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ along which

$$\mathscr{U}(\mu_t) \le t\mathscr{U}(\mu_1) + (1-t)\mathscr{U}(\mu_0) - \frac{1}{2}\lambda t(1-t)W_2(\mu_0, \mu_1)^2. \tag{5.2}$$

Remark 5.1. If for every pair of $\mu_0, \mu_1 \in \mathscr{P}_2(X)$ that are absolutely continuous with respect to m and have continuous densities, there exists some Wasserstein geodesic $\{\mu_t\}_{t\in[0,1]}$ along which (5.2) holds, then as shown in [29, Lemma 3.24], $\mathscr{U}$ is weakly λ-displacement convex.

A curve $\{\mu_t\}_{t\in I} \subset \mathscr{P}_2(X)$ on an interval $I \subset \mathbb{R}$ is *absolutely continuous* if there exists a function $f \in L^1(I)$ such that

$$W_2(\mu_t, \mu_s) \le \int_t^s f(r)\, dr \tag{5.3}$$

for all $s, t \in I$ with $t \le s$. Obviously, an absolutely continuous curve is continuous. For an absolutely continuous curve $\{\mu_t\}_{t\in I} \subset \mathscr{P}_2(X)$, its *metric derivative*

$$|\dot{\mu}_t| \equiv \lim_{s\to t} \frac{W_2(\mu_t, \mu_s)}{|t-s|}$$

is well defined for almost all $t \in I$; see [1, Theorem 1.1.2]. Moreover, $|\dot{\mu}_t| \in L^1(I)$, and it is the minimal function such that (5.3) holds. For $\mu \in \mathscr{P}_2(X)$, define the *local slope* of $\mathscr{U}$ at μ as

$$|\nabla^-\mathscr{U}|(\mu) \equiv \limsup_{\nu\to\mu, \nu\ne\mu} \frac{[\mathscr{U}(\mu) - \mathscr{U}(\nu)]_+}{W_2(\mu, \nu)},$$

where $a_+ = \max\{a, 0\}$.

Now we recall the definition of a gradient flow of a weakly λ-displacement convex functional $\mathscr{U}$, via the energy dissipation identity.

Definition 5.2. Let $\mathscr{U} : \mathscr{P}_2(X) \to \mathbb{R} \cup \{\infty\}$ be weakly λ-displacement convex for some $\lambda \in \mathbb{R}$. An absolutely continuous curve $\{\mu_t\}_{t\in[0,\infty)} \subset \mathscr{P}_2(X)$ is called a *gradient flow* of $\mathscr{U}$ if $\mathscr{U}(\mu_t) < \infty$ for all $t \ge 0$, and for all $0 \le t < s$,

$$\mathscr{U}(\mu_t) = \mathscr{U}(\mu_s) + \frac{1}{2}\int_t^s |\dot{\mu}_r|^2\, dr + \frac{1}{2}\int_t^s |\nabla^-\mathscr{U}|^2(\mu_r)\, dr. \tag{5.4}$$

Associated with the convex function $U_\infty : [0,\infty) \to [0,\infty)$ defined by $U_\infty(r) = r \log r$ for $r > 0$ and $U(0) = 0$, we have the functional $\mathscr{U}_\infty : \mathscr{P}_2(X) \to \mathbb{R} \cup \{+\infty\}$, which is well-defined and lower semicontinuous on $\mathscr{P}_2(X)$; see, for example, [29, Theorem B.33]. Denote by $\mathscr{P}_2^*(X)$ the collection of $\mu \in \mathscr{P}(X)$ such that $\mathscr{U}_\infty(\mu) < \infty$. Since $U'_\infty(\infty) = \infty$, $\mu \in \mathscr{P}_2^*(X)$ implies that μ is absolutely continuous with respect to m and $U_\infty(\frac{d\mu}{dm}) \in L^1(X)$. Recall that if $\mathscr{U}_\infty$ is weakly λ-displacement convex on $\mathscr{P}_2(X)$ for some $\lambda \in \mathbb{R}$, then (X, d, m) is said to have *Ricci curvature bounded from below* in the sense of Lott–Sturm–Villani [49,50,29].

Recently, under the compactness of X and weak λ-displacement convexity of $\mathscr{U}_\infty$, Gigli [12] has obtained the existence, uniqueness and stability of the gradient flow of $\mathscr{U}_\infty$; for the basics of the theory of gradient flows, see [1]. Under some further additional conditions, we are going to prove in Theorem 5.1 that this gradient flow is actually given by the heat flow. Recall that the heat flow is the unique gradient flow of the Dirichlet energy functional $\mathscr{E}$ on the Hilbert space $L^2(X)$. Moreover, the heat flow can be represented by the strongly continuous group $\{T_t\}_{t\ge 0}$ on $L^2(X)$ generated by the unique selfadjoint operator Δ, which is determined by

$$-\int_X u \Delta v \, dm = \mathscr{E}(u, v) = \int_X d\Gamma(u, v)$$

for all $u, v \in \mathbb{D}$. Indeed, for every $\mu \in \mathscr{P}(X)$, the *heat flow* $\{T_t\mu\}_{t\in[0,\infty)}$ is given by $T_0\mu = \mu$ and when $t > 0$, $T_t\mu$ is defined as the unique nonnegative Borel regular measure satisfying

$$\int_X \phi \, dT_t\mu = \int_X \int_X \phi(x) T_t(x, y) \, d\mu(y) \, dm(x).$$

Then by $T_t 1 = 1$, we have $T_t\mu \in \mathscr{P}(X)$, and hence $\{T_t\mu\}_{t\in[0,\infty)}$ is a curve in $\mathscr{P}_2(X)$. Notice that if $\mu = fm$, then $T_t\mu = (T_t f)m$ for all $t \ge 0$.

Theorem 5.1. *Assume that (X, d, m) is compact and satisfies a doubling property, and that $(X, \mathscr{E}, m)$ satisfies the Newtonian property. If $\mathscr{U}_\infty$ is weakly λ-displacement convex for some $\lambda \in \mathbb{R}$, then for every $\mu \in \mathscr{P}_2^*(X)$, the heat flow $\{T_t\mu\}_{t\in[0,\infty)}$ gives the unique gradient flow of $\mathscr{U}_\infty$ with initial value μ.*

We follow the procedure of [13] to prove Theorem 5.1. Let $\{\mu_t\}_{t\in[0,\infty)} \subset \mathscr{P}_2(X)$ be an absolutely continuous curve that satisfies $\mathscr{U}_\infty(\mu_t) < \infty$ for all $t \ge 0$. To prove that $\{\mu_t\}_{t\in[0,\infty)}$ is a gradient flow of $\mathscr{U}_\infty$, we observe that since $\mathscr{U}_\infty$ is weakly λ-displacement convex and lower semicontinuous, by [1, Corollary 2.4.10], for all $s > t \ge 0$, we have

$$|\mathscr{U}_\infty(\mu_t) - \mathscr{U}_\infty(\mu_s)| \le \int_t^s |\nabla^- \mathscr{U}_\infty|(\mu_r) |\dot{\mu}_r| \, dr, \tag{5.5}$$

which implies, by Young's inequality, that

$$\mathscr{U}_\infty(\mu_t) \le \mathscr{U}_\infty(\mu_s) + \frac{1}{2}\int_t^s |\dot{\mu}_r|^2 \, dr + \frac{1}{2}\int_t^s |\nabla^- \mathscr{U}_\infty|^2(\mu_r) \, dr. \tag{5.6}$$

So it suffices to check that for all $s > t \ge 0$,

$$\mathscr{U}_\infty(\mu_s) + \frac{1}{2}\int_t^s |\dot{\mu}_r|^2 \, dr + \frac{1}{2}\int_t^s |\nabla^- \mathscr{U}_\infty|^2(\mu_r) \, dr \le \mathscr{U}_\infty(\mu_t),$$

which is further reduced to proving

$$\frac{1}{2}\,|\dot{\mu}_r|^2 + \frac{1}{2}|\nabla^- \mathscr{U}_\infty|^2(\mu_r) \le -\frac{d}{dr}\mathscr{U}_\infty(\mu_r) \tag{5.7}$$

for almost all $r \ge 0$.

Therefore, with the aid of [36,37], Theorem 5.1 will follow from Lemma 5.1, Propositions 5.2 and 5.3.

Lemma 5.1. *Assume that (X, d, m) is compact and satisfies a doubling property, and that $(X, \mathscr{E}, m)$ supports a weak $(1, 2)$-Poincaré inequality. If $\mathscr{U}_\infty$ is weakly λ-displacement convex for some $\lambda \in \mathbb{R}$, then there exists a constant $C_6 \ge 1$ such that for $\mu = fm \in \mathscr{P}_2^*(X)$,*

$$|\nabla^- \mathscr{U}_\infty|^2(\mu) \le 4C_6 \int_X d\Gamma(\sqrt{f}, \sqrt{f}). \tag{5.8}$$

Moreover, if $(X, \mathscr{E}, m)$ satisfies the Newtonian property, then $C_6 = 1$.

Lemma 5.1 follows from the following result; see [51, Theorem 20.1].

Proposition 5.1. *Let U be a continuous convex function on $[0, \infty)$. Let $\{\mu_t\}_{t\in[0,1]} \subset \mathscr{P}_2(X)$ be an absolutely continuous geodesic with density $\{\rho_t\}_{t\in[0,1]}$, and $U(\rho_t) \in L^1(X)$ for all $t \in [0, 1]$. Further assume that $\rho_0 \in \mathrm{Lip}(X)$, $U(\rho_0)$, $\rho_0 U'(\rho_0) \in L^1(X)$ and U' is Lipschitz on $\rho_0(X)$. Then*

$$\liminf_{t\to 0}\left[\frac{\mathscr{U}(\mu_t) - \mathscr{U}(\mu_0)}{t}\right] \ge -\int_{X\times X} U''(\rho_0(x_0))|\nabla^-\rho_0|(x_0)d(x_0, x_1)\, d\pi(x_0, x_1),$$

where π is an optimal coupling of μ_0 and μ_1.

In Proposition 5.1 and below, for a measurable function f on X, set

$$|\nabla^- f|(x) \equiv \limsup_{y\to x} \frac{[f(x) - f(y)]_+}{d(x, y)}.$$

Obviously, $|\nabla^- f|(x) \le \mathrm{Lip}\, f(x)$ for all $x \in X$. However, if (X, d, m) satisfies a doubling property and supports a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$, then $|\nabla^- f| = \mathrm{Lip}\, f$ almost everywhere. See [28, Remark 2.27].

Proof of Lemma 5.1. We first assume that $f \in \mathrm{Lip}(X)$ and f is bounded away from zero. By the definition of $|\nabla^- \mathscr{U}_\infty|(\mu)$, it suffices to consider $\nu \in \mathscr{P}_2(X)$ with $\mathscr{U}_\infty(\nu) < \mathscr{U}_\infty(\mu)$. Since $\mathscr{U}_\infty(\nu) < \infty$, we have $\nu \in \mathscr{P}_2^*(X)$. By the convexity of $\mathscr{U}_\infty$, there exists a curve $\{\mu_t\}_{t\in[0,1]} \subset \mathscr{P}_2(X)$ such that $\mu_0 = \mu$ and $\mu_1 = \nu$, along which (5.2) holds. Moreover, by (5.2), we have $\mathscr{U}(\mu_t) < \infty$ for all $t \in [0, 1]$, which further means that μ_t is absolutely continuous with respect to m. Denote the density by ρ_t.

Notice that U_∞ and $\{\mu_t\}_{t\in[0,1]}$ fulfill all the conditions required in Proposition 5.1. So by $U''(s) = \frac{1}{s}$, optimality of π and the Hölder inequality, we have

$$\begin{aligned}\liminf_{t\to 0}\left[\frac{\mathscr{U}_\infty(\mu_t) - \mathscr{U}_\infty(\mu_0)}{t}\right] &\ge -\int_{X\times X} \frac{1}{[\rho_0(x_0)]^2}|\nabla^-\rho_0|(x_0)d(x_0, x_1)\, d\pi(x_0, x_1)\\ &\ge -\left\{\int_{X\times X} \frac{1}{[\rho_0(x_0)]^2}|\nabla^-\rho_0|^2(x_0)\, d\pi(x_0, x_1)\right\}^{1/2}\end{aligned}$$

$$\times \left\{ \int_{X\times X} d(x_0, x_1)^2 \, d\pi(x_0, x_1) \right\}^{1/2}$$
$$= -W_2(\mu_0, \mu_1) \left\{ \int_X \frac{1}{[\rho_0(x_0)]^2} |\nabla^- \rho_0|^2(x_0) \, d\mu_0(x_0) \right\}^{1/2}$$
$$= -W_2(\mu_0, \mu_1) \left\{ \int_X \frac{1}{f} |\nabla^- f|^2 \, dm \right\}^{1/2},$$

which together with $|\nabla^- f|^2 \le |\mathrm{Lip}\, f|^2 \le C_1^2 \Gamma(f, f)$ almost everywhere implies that

$$\limsup_{t\to 0} \left[\frac{\mathscr{U}_\infty(\mu_0) - \mathscr{U}_\infty(\mu_t)}{t W_2(\mu_0, \mu_1)} \right] \le C_1 \left\{ \int_X \frac{1}{f} \, d\Gamma(f, f) \right\}^{1/2}. \tag{5.9}$$

On the other hand, by the weak displacement convexity of $\mathscr{U}_\infty$, we have

$$\left[\frac{\mathscr{U}_\infty(\mu_0) - \mathscr{U}_\infty(\mu_1)}{W_2(\mu_0, \mu_1)} \right] \le \left[\frac{\mathscr{U}_\infty(\mu_0) - \mathscr{U}_\infty(\mu_t)}{t W_2(\mu_0, \mu_1)} \right] - \frac{1}{2} \lambda (1 - t) W_2(\mu_0, \mu_1),$$

which together with (5.9) yields that

$$\left[\frac{\mathscr{U}_\infty(\mu_0) - \mathscr{U}_\infty(\mu_1)}{W_2(\mu_0, \mu_1)} \right] \le C_1 \left\{ \int_X \frac{1}{f} \, d\Gamma(f, f) \right\}^{1/2} + \frac{1}{2} |\lambda| W_2(\mu_0, \mu_1),$$

and hence, letting $\mu_1 \to \mu_0$ with respect to W_2,

$$|\nabla^- \mathscr{U}_\infty|(\mu_0) \le C_1 \left\{ \int_X \frac{1}{f} \, d\Gamma(f, f) \right\}^{1/2} = 2C_1 \left\{ \int_X d\Gamma(\sqrt{f}, \sqrt{f}) \right\}^{1/2}.$$

This is as desired.

For $f \in \mathrm{Lip}(X)$ with $\sqrt{f} \in \mathbb{D}$, letting $f_n = (f \vee \frac{1}{n}) \wedge n$, we have $f_n \in \mathrm{Lip}(X)$ and $\frac{f_n m}{\|f_n\|_{L^1(X)}} \in \mathscr{P}_2(X)$. Moreover, since $f_n \ge \frac{1}{n}$, by the above argument, we have

$$|\nabla^- \mathscr{U}_\infty| \left(\frac{f_n m}{\|f_n\|_{L^1(X)}} \right) \le 2C_1 \frac{1}{\|f_n\|_{L^1(X)}} \left\{ \int_X d\Gamma(\sqrt{f_n}, \sqrt{f_n}) \right\}^{1/2}.$$

Moreover, recall that the lower semicontinuity of $\mathscr{U}_\infty$ implies that of $|\nabla^- \mathscr{U}_\infty|$; see [1, Corollary 2.4.11]. Since $\sqrt{f_n} \to \sqrt{f}$ in $\mathbb{D}$, $\|f_n\|_{L^1(X)} \to 1$ and $\frac{f_n m}{\|f_n\|_{L^1(X)}} \to fm$ in $\mathscr{P}(X)$, we have

$$|\nabla^- \mathscr{U}_\infty|(fm) \le \liminf_{n\to\infty} |\nabla^- \mathscr{U}_\infty| \left(\frac{f_n m}{\|f_n\|_{L^1(X)}} \right) \le 2C_1 \left\{ \int_X d\Gamma(\sqrt{f}, \sqrt{f}) \right\}^{1/2}.$$

Generally, for $\mu = fm \in \mathscr{P}_2^*(X)$, we may assume $\sqrt{f} \in \mathbb{D}$ without loss of generality. By Theorem 2.2, we know that $\mathrm{Lip}(X)$ is dense in $\mathbb{D}$. So there exists a sequence $\{g_n\}_{n\in\mathbb{N}} \subset \mathrm{Lip}(X)$ such that $g_n \to \sqrt{f}$ in $\mathbb{D}$. Since $f \ge 0$ almost everywhere, we still have $|g_n| \to \sqrt{f}$ in $\mathbb{D}$. Notice that $|g_n|^2 \in \mathrm{Lip}(X)$ and $\|g_n\|_{L^2(X)} \to 1$. By the lower semicontinuity of $\mathscr{U}_\infty$ again and the above result for Lipschitz functions, we have

$$|\nabla^- \mathscr{U}_\infty|(\mu) \le \liminf_{n\to\infty} |\nabla^- \mathscr{U}_\infty| \left(\frac{|g_n|^2 m}{\|g_n\|_{L^2(X)}} \right)$$

$$\leq 2C_1 \liminf_{n\to\infty} \frac{1}{\|g_n\|_{L^2(X)}} \left\{\int_X d\Gamma(|g_n|, |g_n|)\right\}^{1/2}$$
$$= 2C_1 \left\{\int_X d\Gamma(\sqrt{f}, \sqrt{f})\right\}^{1/2}.$$

This finishes the proof of Lemma 5.1. □

Proposition 5.2. *Assume that (X, d, m) is compact and satisfies a doubling property, and that $(X, \mathscr{E}, m)$ supports a weak $(1,2)$-Poincaré inequality. Let $\mu = fm \in \mathscr{P}_2^*(X)$. Then $\{T_t\mu\}_{t\in[0,\infty)} \subset \mathscr{P}_2^*(X)$, $\{\sqrt{T_t f}\}_{t\in[0,\infty)} \subset \mathbb{D}$ with a locally uniform bound on $(0,\infty)$, and $\mathscr{U}_\infty(T_t\mu)$ is locally Lipschitz on $(0,\infty)$ and for almost all $t \in (0,\infty)$,*

$$\frac{d}{dt}\mathscr{U}_\infty(T_t\mu) = -\int_X \frac{1}{T_t f}\, d\Gamma(T_t f, T_t f). \tag{5.10}$$

Proof. Recall that, under the assumptions of Proposition 5.2, it was proved in [45] that for every $t > 0$, the kernel T_t is locally Hölder continuous in each variable and satisfies

$$C^{-1}\frac{1}{m(B(x,\sqrt{t}))} e^{-\frac{d^2(x,y)}{c_1 t}} \leq T_t(x,y) \leq C\frac{1}{m(B(x,\sqrt{t}))} e^{-\frac{d^2(x,y)}{c_2 t}}. \tag{5.11}$$

So, for every $t \geq \frac{1}{n}$, $T_t f$ is continuous, and moreover, $0 < C(n)^{-1} \leq T_t(x,y) \leq C(n)$ implies that $C(n)^{-1} \leq T_t f(x) \leq C(n)$ for all $x \in X$. From this, it is easy to see that $0 \leq \mathscr{U}_\infty(T_t\mu) \leq C(n)\log C(n) < \infty$ for all $t \geq 1/n$. For $t \geq 1/n$, by $T_{1/n} f \in \mathbb{D}$ and the fact that $\mathscr{E}(T_t f, T_t f)$ is decreasing in t (both of these facts follow by functional calculus), we have

$$\int_X \frac{1}{T_t f}\, d\Gamma(T_t f, T_t f) \leq C(n)\int_X d\Gamma(T_t f, T_t f) \leq C(n)\int_X d\Gamma(T_{1/n} f, T_{1/n} f),$$

which together with the chain rule implies that $\sqrt{T_t f} \in \mathbb{D}$ with locally uniform bound on $(0,\infty)$.

Observe that the function $U_\infty(s) = s\log s$ is smooth on the interval $(\frac{1}{n}, n)$ for all n, and that $T_t f$, as $L^2(X)$-valued function in $(0,\infty)$, is locally Lipschitz on $(0,\infty)$. So $\mathscr{U}_\infty(T_t\mu)$ is locally Lipschitz in $(0,\infty)$. Therefore, by the chain rule for Γ and the fact that $\Gamma(1,h) = 0$ for all $h \in \mathbb{D}$, we have

$$\begin{aligned}\frac{d}{dt}\mathscr{U}_\infty(T_t\mu) &= \int_X U'_\infty(T_t f)\Delta T_t f\, dm = \int_X (\log T_t f + 1)\Delta T_t f\, dm\\ &= -\int_X d\Gamma(\log T_t f + 1, T_t f) = -\int_X d\Gamma(\log T_t f, T_t f)\\ &= -\int_X \frac{1}{T_t f}\, d\Gamma(T_t f, T_t f).\end{aligned}$$

This is as desired. □

The following result is essentially proved in [13, Proposition 3.7]. We point out that, comparing with the assumptions of Theorem 5.1, we can get rid of the Newtonian property here since in the proof, instead of $\frac{d}{dm}\Gamma(u,u) = (\mathrm{Lip}\, u)^2$, it is enough to use $\frac{d}{dm}\Gamma(u,u) \leq (\mathrm{Lip}\, u)^2$ (after writing the first version of this paper, we learned that this was also realized in [2, Lemma 6.1]). For completeness, we give its proof.

Proposition 5.3. *Assume that (X, d, m) is compact and satisfies a doubling property, and that $(X, \mathscr{E}, m)$ supports a weak $(1, 2)$-Poincaré inequality. For every $\mu = fm \in \mathscr{P}_2^*(X)$, $\{T_t\mu\}_{t\in[0,\infty)}$ is an absolutely continuous curve in $\mathscr{P}_2(X)$ and for almost all $t \in [0, \infty)$,*

$$|\dot{T}_t\mu|^2 \le \int_X \frac{1}{T_t f}\, d\Gamma(T_t f, T_t f). \tag{5.12}$$

To prove this, we recall the following result about the Hamilton–Jacobi semigroup established in [28,5]. For $\phi \in \mathrm{Lip}(X)$, set $Q_0 f = f$ and for $t \ge 0$, define

$$Q_t\phi(x) \equiv \inf_{y\in X}\left[\phi(y) + \frac{1}{2t}d^2(x, y)\right].$$

Proposition 5.4. *Assume that (X, d, m) satisfies a doubling property and supports a weak $(1, p)$-Poincaré inequality for some $p \in [1, \infty)$. Then the following hold:*

(i) *for all $t, s \ge 0$ and all $x \in X$, $Q_t Q_s\phi(x) = Q_{t+s}\phi(x)$;*
(ii) *for all $t \ge 0$, $Q_t\phi \in \mathrm{Lip}(X)$;*
(iii) *for all $t \in (0, \infty)$ and almost all $x \in X$,*

$$\frac{d}{dt}Q_t\phi(x) + \frac{1}{2}|\nabla^- Q_t\phi|^2(x) = 0.$$

Proof of Proposition 5.3. Let $t, s > 0$. By the Kantorovich duality,

$$\frac{1}{2}W_2^2(T_t\mu, T_{t+s}\mu) = \left[\int_X Q_1\phi\, dT_{t+s}\mu - \int_X \phi\, dT_t\mu\right];$$

for some $\phi \in L^1(X)$; see, for example, [51, Theorem 5.10] and [1, Section 6]. Moreover, by checking the proof (see, for example, [51, p. 66]), we know that $|\phi|$ is bounded and for all $x \in X$,

$$\phi(x) = \sup_{y\in X}\left[Q_1\phi(y) - \frac{1}{2}d^2(x, y)\right].$$

Since X is compact and hence bounded, we further have that $\phi, Q_1\phi \in \mathrm{Lip}(X)$; we omit the details. Observe that, by Proposition 5.4, $Q_r\phi$ as an $L^2(X)$-valued function of r is Lipschitz on $[0, 1]$ and hence is differentiable almost everywhere. Similarly, $T_{t+rs} f$ as an $L^2(X)$-valued function of r is Lipschitz on $[0, 1]$ and hence is differentiable almost everywhere. Therefore, $(Q_r\phi)T_{t+rs} f$ as an $L^1(X)$-valued function of r is Lipschitz in $[0, 1]$ and hence is differentiable almost everywhere. Thus

$$\begin{aligned}\frac{1}{2}W_2^2(T_t\mu, T_{t+s}\mu) &= \int_0^1 \frac{d}{dr}\int_X (Q_r\phi)(T_{t+rs} f)\, dm\, dr\\ &= \int_0^1\int_X\left[-\frac{1}{2}|\nabla^- Q_r\phi|^2(T_{t+rs} f) + s(Q_r\phi)\Delta T_{t+rs} f\right] dm\, dr.\end{aligned}$$

Since

$$\frac{d}{dm}\Gamma(Q_r\phi, Q_r\phi) \le (\mathrm{Lip}\, Q_r\phi)^2 = |\nabla^- Q_r\phi|^2$$

almost everywhere as given by Theorem 2.1 and [28, Remark 2.27], we have

$$\frac{1}{2}W_2^2(T_t\mu, T_{t+s}\mu) \le -\frac{1}{2}\int_0^1\int_X T_{t+rs}f\, d\Gamma(Q_r\phi, Q_r\phi)\, dr + s\int_0^1\int_X (Q_r\phi)\Delta T_{t+rs}f\, dm\, dr.$$

Moreover, by the Cauchy–Schwarz inequality for Dirichlet forms, we have

$$\begin{aligned}\int_X (Q_r\phi)\Delta T_{t+rs}f\, dm &= \int_X d\Gamma(Q_r\phi, T_{t+rs}f)\\ &= \int_X (T_{t+rs}f)^{1/2}\cdot\frac{1}{(T_{t+rs}f)^{1/2}}\, d\Gamma(Q_r\phi, T_{t+rs}f)\\ &\le \frac{1}{2s}\int_X T_{t+rs}f\, d\Gamma(Q_r\phi, Q_r\phi)\\ &\quad+\frac{s}{2}\int_X \frac{1}{T_{t+rs}f}\, d\Gamma(T_{t+rs}f, T_{t+rs}f),\end{aligned}$$

which implies that

$$W_2^2(T_t\mu, T_{t+s}\mu) \le s^2\int_0^1\int_X \frac{1}{T_{t+rs}f}\, d\Gamma(T_{t+rs}f, T_{t+rs}f)\, dr. \tag{5.13}$$

Since $T_{1/n}f \in \mathbb{D}$, by an argument as in the proof of Proposition 5.2, we have for $t \ge 1/n$,

$$\begin{aligned}W_2^2(T_t\mu, T_{t+s}\mu) &\le s^2C(n)\int_0^1\int_X d\Gamma(T_{t+rs}f, T_{t+rs}f)\, dr\\ &\le s^2C(n)\int_X d\Gamma(T_{1/n}f, T_{1/n}f),\end{aligned}$$

which implies that $T_t\mu$ is locally Lipschitz continuous and hence, $\{T_t\mu\}_{t\ge 0}$ is an absolutely continuous curve in $\mathscr{P}_2(X)$. Moreover, (5.13) also implies that

$$|\dot{T}_t\mu|^2 \le \int_X \frac{1}{T_tf}\, d\Gamma(T_tf, T_tf),$$

which is as desired. $\square$

For the proof of Theorem 5.1, we still need a very recent result of Rajala [36,37]: λ-displacement convexity of $\mathscr{U}_\infty$ implies that (X, d, m) supports a weak $(1, 1)$- and hence a weak $(1, 2)$-Poincaré inequality. This combined with Proposition 2.2 allows us to use Lemma 5.1 and Propositions 5.2 and 5.3 in the proof of Theorem 5.1.

Proof of Theorem 5.1. For $\mu = fm \in \mathscr{P}_2^*(X)$, it was proved in Proposition 5.3 that $\{T_t\mu\}_{t\in[0,\infty)}$ is an absolutely continuous curve in $\mathscr{P}_2^*(X)$. To prove that $\{T_t\mu\}_{t\in[0,\infty)}$ is a gradient flow of $\mathscr{U}_\infty$, since (5.6) follows from the displacement convexity of $\mathscr{U}_\infty$, it suffices to check the reverse inequality which is further reduced to (5.7). But (5.7) follows from (5.8) with $C_3 = 1$, and (5.10) and (5.12). $\square$

Under the assumptions of Proposition 5.2, for every $\mu \in \mathscr{P}_2(X)$ and $t > 0$, $T_t\mu$ is absolutely continuous with respect to m and its density is continuous and bounded away from zero. Indeed, for $t > 0$ and every nonnegative $\phi \in \mathscr{C}(X)$, by $\mu(X) = 1$ and (5.11), we have

$$\int_X \phi\, dT_t\mu = \int_X \int_X \phi(x) T_t(x, y)\, d\mu(y)\, dm(x) \le C(t) \int_X \phi(x)\, dm(x),$$

which implies the absolute continuity of $T_t\mu$. Let $f_t = \frac{d}{dm} T_t\mu$ for $t > 0$. Then $f_t \in L^1(X)$, and moreover, by the semigroup property, $f_t = T_{t/2} f_{t/2}$, which together with (5.11) and the continuity of the kernel of $T_{t/2}$ implies that f_t is continuous and bounded away from zero. Relying on the observations above and Theorem 5.1, we conclude the following result.

Corollary 5.1. *Let all the assumptions be as in Theorem 5.1. For every $\mu \in \mathscr{P}(X)$, the curve $\{T_t\mu\}_{t\in(0,\infty)} \subset \mathscr{P}^*(X)$ is absolutely continuous on each $[\epsilon, \infty)$ for all $\epsilon > 0$, and $\mathscr{U}_\infty(T_t\mu) < \infty$ for all $t > 0$ and (5.4) holds for all $s > t > 0$.*

Proof. Let $\mu \in \mathscr{P}(X)$. For every $n \in \mathbb{N}$, by the argument before Corollary 5.1, we see that $T_{1/n}\mu$ is absolutely continuous with respect to m and that the Radon–Nikodym derivative $f_{1/n} = \frac{d}{dm} T_{1/n}\mu$ belongs to $L^1(X)$, that is, $T_{1/n}\mu = f_{1/n} m \in \mathscr{P}_2^*(X)$. Hence Proposition 5.3 ensures that $\{T_t(f_{1/n} m)\}_{t\in[0,\infty)}$ is an absolutely continuous curve in $\mathscr{P}_2^*(X)$. By Theorem 5.1, $\{T_t(f_{1/n} m)\}_{t\in[0,\infty)}$ gives the unique heat flow with initial value $T_{1/n}\mu = f_{1/n} m$, which means that for all $s > t \ge 0$, (5.4) holds with $\mu_r = (T_r f_{1/n})m$ when $t \le r \le s$. Observe that $\{T_t\mu\}_{t\in[\frac{1}{n},\infty)} = \{T_t(T_{1/n}\mu)\}_{t\in[0,\infty)} = \{T_t(f_{1/n} m)\}_{t\in[0,\infty)}$. We further obtain $\mathscr{U}_\infty(T_t\mu) < \infty$ for all $t \ge \frac{1}{n}$, $\{T_t\mu\}_{t\in[1/n,\infty)}$ is an absolutely continuous curve in $\mathscr{P}^*(X)$, and for all $\frac{1}{n} \le t < s$, (5.4) holds with $\mu_r = T_r\mu$ when $t \le r \le s$. By the arbitrariness of n, we finally have that $\{T_t\mu\}_{t\ge(0,\infty)}$ is a locally absolutely continuous curve in $\mathscr{P}^*(X)$, $\mathscr{U}_\infty(T_t\mu) < \infty$ for all $t > 0$ and for all $0 < t < s$, (5.4) holds with $\mu_r = T_r\mu$ when $t \le r \le s$. $\square$

Furthermore, for $1 < N < \infty$, associated to the convex function $U_N \equiv Nr - Nr^{1-1/N}$ for $r \ge 0$, we have the functional $\mathscr{U}_N : \mathscr{P}_2(X) \to \mathbb{R} \cup \{+\infty\}$, which is well defined; see for example [29]. Assume that $\mathscr{U}_N$ is weakly λ-displacement convex for some $N \in [1, \infty)$ and $\lambda \ge 0$. Then, as proved in [29, Theorem 5.31], (X, d, m) satisfies a doubling property.

Corollary 5.2. *Assume that $(X, \mathscr{E}, m)$ satisfies the Newtonian property. If $\mathscr{U}_N$ is weakly 0-displacement convex for some $N \in (1, \infty)$, and $\mathscr{U}_\infty$ is weakly λ-displacement convex for some $\lambda \in \mathbb{R}$, then the heat flow gives the unique gradient flow of $\mathscr{U}_\infty$.*

Remark 5.2. As pointed out by the referee, the weak 0-displacement convexity assumption on $\mathscr{U}_N$ and the weak λ-displacement convexity assumption on $\mathscr{U}_\infty$ in Corollary 5.2 can be replaced by a weaker condition, the curvature-dimension condition $CD(\lambda, N)$ (see [50, Definition 1.3]). Indeed, under the condition $CD(\lambda, N)$, by [50, Proposition 1.6], $\mathscr{U}_\infty$ is weakly λ-displacement convex. By the compactness of X and [50, Corollary 2.4], m satisfies the doubling property, and by [36], (X, d, m) supports a weak $(1, 1)$-Poincaré inequality. With these, the conclusion of Corollary 5.2 follows from Theorem 5.1.

By the linearity and symmetry of heat flows, we also have the following property of the gradient flow of $\mathscr{U}_\infty$.

Corollary 5.3. *Let all the assumptions be as in Theorem 5.1. For every $\nu \in \mathscr{P}_2^*(X)$, let $\{\mu_t^\nu\}_{t\ge0}$ be the gradient flow of $\mathscr{U}_\infty$ with $\mu_0^\nu = \nu$*

Then

(i) *for all* $\nu_0, \nu_1 \in \mathscr{P}_2^*(X)$, $t \geq 0$ *and* $\lambda \in [0, 1]$,

$$\mu_t^{(1-\lambda)\nu_0+\lambda\nu_1} = (1-\lambda)\mu_t^{\nu_0} + \lambda\mu_t^{\nu_1};$$

(ii) *for all nonnegative* $f, g \in L^1(X)$ *with* $\|f\|_{L^1(X)} = \|g\|_{L^1(X)} = 1$ *and* $t \geq 0$,

$$\int_X f\, d\mu_t^{gm} = \int_X g\, d\mu_t^{fm}.$$

Recall that, under a non-branching condition, additional semiconcavity and local angle conditions, the linearity property in Corollary 5.3(i) was proved in [39]. For the definitions of *K-semiconcavity* and *local angle condition* introduced in [39], see Section 6. We do not know if these conditions hold under the assumptions of Theorem 5.1. Also recall that the linearity property fails on Finsler manifolds as pointed out in [33].

After we obtained Theorem 5.1, we learned about a related result established in [2, Theorem 9.3]. Indeed, instead of the Dirichlet energy form $\mathscr{E}$, Ambrosio, Gigli and Savaré [2] considered the Cheeger energy functional **Ch** on $L^2(X)$, which is not necessarily Hilbertian. They showed that, under the convexity of $\mathscr{U}_\infty$ and very few assumptions on X, the gradient flow of **Ch** coincides with the gradient flow of the entropy $\mathscr{U}_\infty$; see [2, Theorem 9.3]. Their proof also relies on the procedure outlined in [13] but works at a high level of generality.

Assume that X satisfies a doubling property and supports a weak (1, 2)-Poincaré inequality. Then, with the aid of Lemma 2.5, **Ch** can be written as

$$\mathbf{Ch}(f) = \int_X (\operatorname{apLip} f)^2\, dm.$$

Moreover, by Theorem 2.2(ii),

$$\mathscr{E}(f, f) \leq \mathbf{Ch}(f) \leq C_1\mathscr{E}(f, f), \tag{5.14}$$

while $C_1 = 1$ if we further assume that $(X, \mathscr{E}, m)$ supports a Newtonian property. We point out that, under the assumptions of Theorem 5.1, the conclusion of Theorem 5.1 follows from (5.14) with $C_1 = 1$ and [2, Theorem 9.3].

Recall that in Theorem 5.1, we showed that the Newtonian property is a sufficient condition to identify the heat flow and the gradient flow of entropy. Combining Theorem 2.2, Proposition 5.3 and [2, Theorem 9.3], we will show that the Newtonian property is also necessary in the following sense.

Theorem 5.2. *Assume that* (X, d, m) *is compact and satisfies a doubling property, and that* $\mathscr{U}_\infty$ *is weakly* λ*-displacement convex for some* $\lambda \in \mathbb{R}$*. Then the following are equivalent:*

(i) *For every* $\mu \in \mathscr{P}_2^*(X)$*, the heat flow* $\{T_t\mu\}_{t\in[0,\infty)}$ *gives the unique gradient flow of* $\mathscr{U}_\infty$ *with initial value* μ*.*
(ii) $(X, \mathscr{E}, m)$ *satisfies the Newtonian property.*
(iii) *For all* $u \in \mathbb{D}$, $\frac{d}{dm}\Gamma(u, u) = (\operatorname{apLip} u)^2$ *almost everywhere.*

Proof. We recall again that, by [36,37], the λ-displacement convexity of $\mathscr{U}_\infty$ implies that (X, d, m) supports a weak (1, 2)-Poincaré inequality. Then the equivalence of (ii) and (iii) follows from Theorem 2.2 and Lemma 2.4. If (ii) holds, then by Theorem 5.1, we have (i). Now assume that (i) holds. Let $f \in \mathrm{Lip}(X)$ be a positive function and set $\mu = fm \in \mathscr{P}_2^*(X)$. By Proposition 5.3, we have that for almost all $t \in [0, \infty)$,

$$|\dot{T}_t\mu|^2 \le \int_X \frac{1}{T_t f}\, d\Gamma(T_t f, T_t f). \tag{5.15}$$

Moreover the assumption (i) says that $\{T_t\mu\}_{t\in[0,\infty)}$ is the gradient flow of $\mathscr{U}_\infty$. By this, the convexity of $\mathscr{U}_\infty$, and Theorems 9.3 and 8.5 of [2], we know that $T_t\mu$ coincides with the gradient flow of **Ch** and satisfies that for almost all $t \in (0, \infty)$,

$$|\dot{T}_t\mu|^2 = \int_X \frac{1}{T_t f}(\mathrm{apLip}\, T_t f)^2\, dm.$$

This and (5.15), with the aid of $\Gamma(T_t f, T_t f) \le (\mathrm{apLip}\, T_t f)^2 m$ given in (2.22), further give $\Gamma(T_t f, T_t f) = (\mathrm{apLip}\, T_t f)^2 m$ for almost all $t \in (0, \infty)$. Therefore,

$$\mathscr{E}(f, f) = \lim_{t\to\infty} \mathscr{E}(T_t f, T_t f) \ge \liminf_{t\to 0} \int_X (\mathrm{apLip}\, T_t f)^2\, dm.$$

Observing that

$$\begin{aligned}\int_X (\mathrm{apLip}\, f - \mathrm{apLip}\, T_t f)^2\, dm &\le \int_X [\mathrm{apLip}(f - T_t f)]^2\, dm \\ &\lesssim \mathscr{E}(f - T_t f, f - T_t f) \to 0,\end{aligned}$$

we obtain $\mathscr{E}(f, f) = \int_X (\mathrm{apLip}\, f)^2\, dm$. With the help of $\Gamma(f, f) \le (\mathrm{apLip}\, f)^2 m$ given in (2.22) again, we have $\Gamma(f, f) = (\mathrm{apLip}\, f)^2 m$ as desired. Then a density argument yields (iii). □

6. Applications to (coarse) Ricci curvatures

In this section, we apply Corollary 2.1 to a variant of the dual formula of Kuwada [27] and the coarse Ricci curvature of Ollivier in Theorem 6.1 and Corollary 6.1, and then apply Theorem 5.1 to the Ricci curvatures of Bakry–Emery and Lott–Sturm–Villani in Corollary 6.2. We always let $\mathscr{E}$ be a regular, strongly local Dirichlet form on $L^2(X, m)$, assume that the topology induced by the intrinsic d coincides with the original topology on X and that (X, d, m) satisfies the doubling property.

We begin with a variant of the dual formula established in [27, Theorem 2.2], which is closely related to the coarse Ricci curvature of Ollivier [32]. Let $\{P_x\}_{x\in X} \subset \mathscr{P}(X)$ be a family of probability measures on X, so that the map $x \to P_x$ from X to $\mathscr{P}(X)$ is continuous. Then $\{P_x\}_{x\in X}$ defines a bounded linear operator P on $\mathscr{C}(X)$ by $Pf(x) = \int_X f(y)\, dP_x(y)$ and we denote its dual operator by $P^* : \mathscr{P}(X) \to \mathscr{P}(X)$. We also assume that P_x is absolutely continuous with respect to m with density $P_x(y)$ for all $x \in X$, and that $P_x(y)$ is a continuous function of x for almost all $y \in X$. Observe that we do not assume that $\{P_x\}_{x\in X}$ has any relation with the Dirichlet form $\mathscr{E}$.

By Corollary 2.1 and [27], we have the following result.

Theorem 6.1. *Assume that $\mathscr{E}$ is a regular, strongly local Dirichlet form on $L^2(X)$, the topology induced by the intrinsic distance is equivalent to the original topology on the locally compact space X, and (X, d, m) satisfies the doubling property. Let $K_1 \ge 0$ be a positive constant. Then the following are equivalent:*

(i) *For all $\mu, \nu \in \mathscr{P}(X)$, $W_1(P^*\mu, P^*\nu) \le K_1 W_1(\mu, \nu)$.*

(ii) *For all* $f \in \mathrm{Lip}(X) \cap L^\infty(X)$, $Pf \in \mathrm{Lip}(X)$ *and* $\|Pf\|_{\mathrm{Lip}(X)} \le K_1 \|f\|_{\mathrm{Lip}(X)}$.

(iii) *For all* $f \in \mathrm{Lip}(X) \cap L^\infty(X)$, $Pf \in \mathrm{Lip}(X)$ *and*

$$\|\mathrm{Lip}\, Pf\|_{L^\infty(X)} \le K_1 \|\mathrm{Lip}\, f\|_{L^\infty(X)}.$$

(iv) *For all* $f \in \mathrm{Lip}(X) \cap L^\infty(X)$, $Pf \in \mathrm{Lip}(X)$ *and*

$$\left\| \frac{d}{dm}\Gamma(Pf, Pf)\right\|_{L^\infty(X)} \le (K_1)^2 \left\| \frac{d}{dm}\Gamma(f, f)\right\|_{L^\infty(X)}.$$

Proof. For $K_1 > 0$, the equivalence between (i) and (ii) follows from [27, Theorem 2] (by taking $\widetilde{d} = K_1 d$ there). Notice that the proof of [27, Theorem 2] for the case $p = 1$ does not require any weak Poincaré inequality. The equivalence of (ii)–(iv) follows from

$$\|u\|^2_{\mathrm{Lip}(X)} = \|\mathrm{Lip}\, u\|^2_{L^\infty(X)} = \left\| \frac{d}{dm}\Gamma(u, u)\right\|_{L^\infty(X)}$$

with $u = f \in L^\infty(X)$ and $u = Pf \in L^\infty(X)$; see Corollary 2.1. For $K_1 = 0$, the equivalence of (i) through (iv) follows from the case $K_1 + \epsilon$ with $\epsilon > 0$ and an approximation argument. We omit the details. □

Associated to (X, d, P) with P as above, Ollivier [34] introduced the *coarse Ricci curvature* via

$$\kappa(x, y) = 1 - \frac{W_1(P^*\delta_x, P^*\delta_y)}{d(x, y)}.$$

(X, d, P) is said to have the *coarse Ricci curvature bounded from below by constant* K if $\kappa(x, y) \ge K$ for all $x, y \in X$. Obviously, $K \le 1$. Applying Theorem 6.1, we have the following result.

Corollary 6.1. *Under the assumptions of Theorem* 6.1, *the following are equivalent:*

(i) (X, d, P) *has the coarse Ricci curvature bounded from below by* $K \le 1$.

(ii) *For all* $\mu, \nu \in \mathscr{P}(X)$, $W_1(P^*\mu, P^*\nu) \le (1 - K) W_1(\mu, \nu)$.

(iii) *For all* $f \in \mathrm{Lip}(X) \cap L^\infty(X)$, $Pf \in \mathrm{Lip}(X)$ *and*

$$\left\| \frac{d}{dm}\Gamma(Pf, Pf)\right\|_{L^\infty(X)} \le (1 - K)^2 \left\| \frac{d}{dm}\Gamma(f, f)\right\|_{L^\infty(X)}.$$

Proof. By Theorem 6.1, (i) follows from (ii) or (iii). Conversely, if (i) holds, then (ii) holds with $\mu = \delta_x$ and $\nu = \delta_y$, which together with [27, Lemma 3.3] further yields that (ii) holds with μ and $\nu \in \mathscr{P}(X)$. □

On the other hand, combining [39, Theorem 1], [27], and Theorems 5.1 and 2.2 of our paper, and following the procedure of [13], we know that, under some additional conditions, a Ricci curvature bound from below in the sense of Lott–Sturm–Villani [29,49,50] implies that in the sense of Bakry–Emery [3,4]. Recall that $(X, \mathscr{E}, m)$ is said to have Ricci curvature bounded from below by $\lambda \in \mathbb{R}$ in the sense of Bakry–Emery if for all $f \in \mathbb{D}$ and $t \ge 0$, and for almost all $x \in X$,

$$\frac{d}{dm}\Gamma(T_t f, T_t f)(x) \le e^{-2\lambda t} T_t\left(\frac{d}{dm}\Gamma(f, f)\right)(x). \tag{6.1}$$

Indeed, Savaré [39] obtained the contraction property of the gradient flow of the entropy with the aid of the semiconcavity and local angle conditions. Recall that X is K*-semiconcave* if $K \ge 1$ and for every geodesic γ and $y \in X$,

$$d^2(\gamma(t), y) \ge (1-t)d^2(\gamma(0), y) + td^2(\gamma(1), y) - Kt(1-t)d^2(\gamma(0), \gamma(1)).$$

Moreover, X satisfies the *local angle condition* if for every triplet of geodesics γ_i, $i = 1, 2, 3$, emanating from the same initial point x_0, the corresponding angles $\angle(\gamma_i, \gamma_j) \in [0, \pi]$ satisfy

$$\angle(\gamma_1, \gamma_2) + \angle(\gamma_2, \gamma_3) + \angle(\gamma_3, \gamma_1) \le 2\pi,$$

where

$$\angle(\gamma_i, \gamma_j) \equiv \liminf_{s,t\to 0+} \frac{d^2(x_0, \gamma_i(s)) + d^2(x_0, \gamma_j(t)) - d^2(\gamma_i(s), \gamma_j(s))}{2d(x_0, \gamma_i(s))d(x_0, \gamma_j(t))}.$$

Kuwada established a dual relation between contraction of the gradient flow in Wasserstein distance and its pointwise Lipschitz constant estimate (see [27]). Under our assumptions, Theorem 2.2 identifies the pointwise Lipschitz constant with length of the gradient, while Theorem 5.1 identifies the heat flow and gradient flow.

Corollary 6.2. *Under the assumptions of Theorem 5.1, and further assuming that (X, d) is compact, (X, d, m) is K-semiconcave for some $K \ge 1$ and satisfies a local angle condition, if $\mathscr{U}_\infty$ is weakly λ-displacement convex for some $\lambda \in \mathbb{R}$, then the following hold:*

(i) *For all $\mu, \nu \in \mathscr{P}(X)$, $W_2(T_t\mu, T_t\nu) \le e^{-\lambda t} W_2(\mu, \nu)$,*
(ii) *For all $f \in \mathbb{D}$ and $t \ge 0$, $T_t f \in \mathrm{Lip}(X)$ and for all $x \in X$,*

$$[\mathrm{Lip}\, T_t f(x)]^2 \le e^{-2\lambda t} T_t (\mathrm{apLip}\, f)^2(x). \tag{6.2}$$

(iii) *For all $f \in \mathbb{D}$ and $t \ge 0$, (6.1) holds for almost all $x \in X$.*

Proof. To see (i), since the heat flow coincides with gradient flow of $\mathscr{U}_\infty$ as given in 5.1, it suffices to prove that for the gradient flows $\{\mu_t\}_{t\ge 0}$ and $\{\nu_t\}_{t\ge 0}$, $W_2(\mu_t, \nu_t) \le e^{-\lambda t} W_2(\mu_0, \nu_0)$. But this was already proved by Savaré [39] and hence we have (i).

Obviously, applying (ii) and Theorem 2.2(iii), we have (iii).

Moreover, (ii) follows from (i), [27, Theorem 2] and an approximation argument. Indeed, for $f \in \mathrm{Lip}(X)$, by [27, Theorem 2], (6.2) follows from (i). Generally, let $f \in \mathbb{D}$. By Theorem 2.2(i), $\mathrm{Lip}(X)$ is dense in $\mathbb{D}$. Thus, there exists a sequence $f_i \in \mathrm{Lip}(X)$ such that $f_i \to f$ in $\mathbb{D}$ as $i \to \infty$. For each $x \in X$,

$$\begin{aligned}|T_t f(x) - T_t f_i(x)| &\le \int_X T_t(x, y)|f(y) - f_i(y)|\,dy \\ &\le \|T_t(x, \cdot)\|_{L^2(X)} \|f - f_i\|_{L^2(X)} \le C(t)\|f - f_i\|_{L^2(X)},\end{aligned}$$

where $C(t) = \sup_{x\in X} \|T_t(x, \cdot)\|_{L^2(X)} < \infty$. Thus for each pair of $x, y \in X$,

$$|T_t f(x) - T_t f(y)| \le 2C(t)\|f - f_i\|_{L^2(X)} + |T_t f_i(x) - T_t f_i(y)|.$$

Notice that for every rectifiable curve γ,

$$|T_t f_i(x) - T_t f_i(y)| \le \int_\gamma \mathrm{Lip}\, T_t f_i\, ds \le e^{-\lambda t} \int_\gamma [T_t(\mathrm{Lip} f_i)^2]^{1/2}\, ds$$

and by Theorem 2.2(iii),

$$\begin{aligned}[T_t(\mathrm{Lip} f_i)^2]^{1/2} &\le [T_t(\mathrm{Lip} f_i - \mathrm{apLip}\, f)^2]^{1/2} + [T_t(\mathrm{apLip}\, f)^2]^{1/2} \\ &\le [T_t(\mathrm{apLip}(f_i - f))^2]^{1/2} + [T_t(\mathrm{apLip}\, f)^2]^{1/2} \\ &\le \widetilde{C}(t)\|\mathrm{apLip}(f - f_i)\|_{L^2(X)} + [T_t(\mathrm{apLip}\, f)^2]^{1/2} \\ &\le \widetilde{C}(t)\|f - f_i\|_{\mathbb{D}} + [T_t(\mathrm{apLip}\, f)^2]^{1/2},\end{aligned}$$

where $\widetilde{C}(t) = \sup_{x,y\in X} T_t(x, y) < \infty$. Then

$$|T_t f(x) - T_t f(y)| \le [2C(t) + \widetilde{C}(t)e^{-\lambda t}\ell(\gamma)]\|f - f_i\|_{\mathbb{D}} + e^{-\lambda t}\int_\gamma [T_t(\mathrm{apLip}\, f)^2]^{1/2}\, ds$$

and hence

$$|T_t f(x) - T_t f(y)| \le e^{-\lambda t}\int_\gamma [T_t(\mathrm{apLip}\, f)^2]^{1/2}\, ds \le \widetilde{C}(t)\ell(\gamma)\|f\|_{\mathbb{D}}.$$

Choosing γ to be a geodesic joining x and y, we see that $T_t f \in \mathrm{Lip}(X)$. Moreover, by the continuity of the heat kernel and hence of $e^{-\lambda t}[T_t(\mathrm{apLip}\, f)^2]^{1/2}$, we have that $\mathrm{Lip}\, T_t f(x) \le e^{-\lambda t}[T_t(\mathrm{apLip}\, f)^2(x)]^{1/2}$ for all $x \in X$. □

Remark 6.1. Notice that, by [39,32], compact Aleksandrov spaces with curvature bounded from below satisfy the assumptions of Corollary 6.2 (in particular, the K-semiconcavity and the local angle condition) and thus they have Ricci curvature bounded from below in the sense of Bakry–Emery. This conclusion can also be found in [13].

7. Asymptotics of the gradient of the heat kernel

We are going to give a characterization for the condition that $\Gamma(d_x, d_x) = m$ for all $x \in X$ via the short time asymptotics of the gradient of the heat semigroup; see Theorem 7.1.

Assume that X is compact and $(X, \mathscr{E}, m)$ has a spectral gap, that is, there exists a positive constant C_{spec} such that for all $u \in \mathbb{D}$,

$$\int_X \left| u - \fint_X u\, dm \right|^2 dm \le C_{\mathrm{spec}}\mathscr{E}(u, u).$$

Obviously, if $(X, \mathscr{E}, m)$ satisfies a weak Poincaré inequality in the sense of Section 2, then it has a spectral gap. Then the Varahdan asymptotic behavior of heat kernels was established in [38]: for all $x, y \in X$,

$$\lim_{t\to 0} -4t \log T_t(x, y) = d^2(x, y); \tag{7.1}$$

see [31] for Lipschitz manifolds and [16] for general local and conservative Dirichlet forms.

On the other hand, on a Riemannian manifold, Malliavin and Stroock [30] (see also [43]) proved that

$$\lim_{t\to 0} -4t[\nabla \log T_t(\cdot, y)](x) = [\nabla d^2(\cdot, y)](x), \tag{7.2}$$

for all $y \in M$ and all $x \in M$ outside the cut locus of y, where ∇ denotes the gradient on a Riemannian manifold. On $\mathbb{R}^n$, the Gaussian kernel $h_t(x) = c_n t^{n/2}\exp(-\frac{|x|^2}{4t})$ satisfies $|\nabla|x|^2| = 4t|\nabla \log h_t(x)|$ for all $t \in (0, \infty)$.

We show that a weak variant of (7.2) will reflect a connection between the length structure and gradient structure of Dirichlet forms.

Theorem 7.1. *Let $\mathscr{E}$ be a regular, strongly local Dirichlet form on $L^2(X,m)$. Assume that X is compact, the topology induced by d coincides with the original topology, and that $(X,\mathscr{E},m)$ has a spectral gap. Then the following are equivalent:*

(i) *For all $x\in X$, $\Gamma(d_x,d_x)=m$.*
(ii) *For every Borel measurable set A with positive measure and each $\varphi\in\mathbb{D}\cap\mathscr{C}_0(X)$,*

$$\int_X \varphi d\Gamma(t\log T_t 1_A, t\log T_t 1_A)\to\int_X \varphi d\Gamma(d_A^2/4, d_A^2/4). \tag{7.3}$$

Notice that (7.3) is a weak variant of (7.2) while (7.1) has a weak variant as established in [38, Theorem 3.10].

Proposition 7.1. *Under the assumptions of Theorem 7.1, for every Borel measurable set A with positive measure and each $\varphi\in\mathbb{D}\cap\mathscr{C}_0(X)$,*

$$\int_X(-t\log T_t 1_A)\varphi\, dm\to\int_X \varphi d_A^2/4\, dm. \tag{7.4}$$

Proof of Theorem 7.1. We first prove that (ii) implies (i). Let A be a Borel measurable set in X, $u_t=-t\log T_t 1_A$ for all $t>0$, and $u_0=d_A^2/4$. Then $u_t,u_0\in\mathbb{D}_{\mathrm{loc}}$. From $\Gamma(d_x,d_x)\le m$, it follows that $\Gamma(u_0,u_0)\le u_0\, m$. It suffices to prove the converse inequality.

Notice that the strong locality of $\mathscr{E}$ implies that Γ satisfies the Leibniz rule, namely, for $F\in C^1(\mathbb{R})$, every $\phi\in\mathbb{D}\cap L^\infty(X)$ and $\varphi\in\mathbb{D}\cap\mathscr{C}_0(X)$, we have

$$\int_X \varphi\, d\Gamma(F\circ\phi,F\circ\phi)=\int_X\varphi(F'\circ\phi)^2\, d\Gamma(\phi,\phi),$$

and if $F\in C^2(\mathbb{R})$,

$$\int_X\varphi\Delta(F\circ\phi)\, d\mu=\int_X\varphi(F'\circ\phi)\Delta\phi\, d\mu+\int_X\varphi(F''\circ\phi)\, d\Gamma(\phi,\phi).$$

Then for every $\varphi\in\mathbb{D}\cap\mathscr{C}_0(X)$, by

$$\int_X\varphi\left[\frac{d}{dt}T_t 1_A+\Delta T_t 1_A\right]d\mu=0,$$

we have

$$t\int_X\frac{du_t}{dt}\varphi\, d\mu-t\int_X\Delta u_t\varphi\, d\mu=\int_X u_t\varphi\, d\mu-\int_X\varphi\, d\Gamma(u_t,u_t),$$

from which it follows that

$$t\int_X\frac{du_t}{dt}\varphi\, d\mu+t\mathscr{E}(u_t,\varphi)=\int_X u_t\varphi\, d\mu-\int_X\varphi\, d\Gamma(u_t,u_t)$$

and that

$$\begin{aligned}\frac{1}{t}\int_0^t\int_X\varphi\, d\Gamma(u_s,u_s)\, ds&=\frac{1}{t}\int_0^t\int_X u_s\varphi\, d\mu\, ds\\&\quad-\frac{1}{t}\int_0^t\int_X s\frac{du_s}{ds}\varphi\, d\mu\, ds-\frac{1}{t}\int_0^t s\mathscr{E}(u_s,\varphi)\, ds.\end{aligned} \tag{7.5}$$

Let $\widetilde{\varphi} \in \mathbb{D} \cap \mathscr{C}_0(X)$ such that $\widetilde{\varphi} = 1$ on the support of φ. Notice that

$$|\mathscr{E}(u_s, \varphi)| = \left| \int_X \widetilde{\varphi}\, d\Gamma(u_s, \varphi) \right| \le \mathscr{E}(\varphi, \varphi) \int_X (\widetilde{\varphi})^2 \, d\Gamma(u_s, u_s).$$

Then, by (7.3), we know that $\mathscr{E}(u_s, \phi)$ is uniformly bounded with respect to s. Hence

$$\frac{1}{t} \int_0^t s\mathscr{E}(u_s, \varphi)\, ds \to 0 \tag{7.6}$$

as $t \to 0$.

By (7.4), for every $\varphi \in \mathbb{D} \cap \mathscr{C}_0(X)$,

$$\int_X u_0 \varphi \, d\mu = \lim_{s \to 0} \int_X u_s \varphi \, d\mu.$$

Hence

$$\frac{1}{t} \int_0^t \int_X u_s \varphi \, d\mu \, ds \to \int_X u_0 \varphi \, d\mu$$

and

$$\frac{1}{t} \int_0^t \int_X s \frac{du_s}{ds} \varphi \, d\mu = \lim_{\epsilon \to 0} \frac{s}{t} \int_X u_s \varphi \, d\mu \bigg|_{s=\epsilon}^{t} - \frac{1}{t} \int_0^t \int_X u_s \varphi \, d\mu \, ds \to 0$$

as $t \to 0$. From these two facts, (7.5) and (7.6), it follows that

$$\frac{1}{t} \int_0^t \int_X \varphi \, d\Gamma(u_s, u_s)\, ds \to \int_X u_0 \varphi \, d\mu$$

as $t \to 0$, which together with (7.3) implies

$$\int_X \varphi \, d\Gamma(u_0, u_0) = \int_X u_0 \varphi \, d\mu. \tag{7.7}$$

This gives $\Gamma(u_0, u_0) = u_0 \mu$. Moreover, if A is compact, then for all $\varphi \in \mathscr{C}_0(X)$ with $\operatorname{supp} \varphi \subset A^{\complement}$,

$$\int_X \varphi \, d\Gamma(d_A, d_A) = 4 \int_X \varphi \, d\Gamma(\sqrt{u_0}, \sqrt{u_0}) = \int_X \varphi \frac{1}{u_0} \, d\Gamma(u_0, u_0) = \int_X \varphi \, d\mu,$$

which means that $\Gamma(d_A, d_A) = \mu$. Since $d_{\overline{B(x_0,r)}} \to d_{x_0}$ in $\mathbb{D}_{\text{loc}}$ as $r \to 0$, we have (i).

Now we turn to prove that (i) implies (ii). It suffices to prove that $\mathscr{E}(u_0, u_0) \ge \|u_0\|_{L^1(X)}$. Indeed, from this and $\Gamma(u_0, u_0) \le u_0 m$, it follows that $u_0 - \frac{d}{dm}\Gamma(u_0, u_0) = 0$ almost everywhere, which further implies that $\frac{d}{dm}\Gamma(d_A, d_A) = 1$ almost everywhere on $A^{\complement}$, and hence gives (i). We first observe that, by our assumption (ii),

$$\mathscr{E}(u_0, u_0) = \lim_{s \to 0} \mathscr{E}(u_s, u_s) = \lim_{t \to 0} \frac{1}{t} \int_0^t \mathscr{E}(u_s, u_s)\, dt. \tag{7.8}$$

But, taking $\varphi = 1$, (7.5) yields that

$$\begin{aligned} \frac{1}{t} \int_0^t \mathscr{E}(u_s, u_s)\, dt &= \frac{1}{t} \int_0^t \|u_s\|_{L^1(X)}\, ds - \frac{1}{t} \int_0^t s \frac{d}{ds} \|u_s\|_{L^1(X)}\, ds \\ &= \frac{1}{t} \int_0^t \|u_s\|_{L^1(X)}\, ds - \frac{1}{t} (s \|u_s\|_{L^1(X)}) \bigg|_{s \to 0}^{s=t} + \frac{1}{t} \int_0^t \|u_s\|_{L^1(X)}\, ds \end{aligned}$$

$$= 2\frac{1}{t}\int_0^t \|u_s\|_{L^1(X)}\,ds - \|u_t\|_{L^1(X)}.$$

Then, by (7.4) as given in Proposition 7.1, taking $\varphi = 1$, we have $\|u_t\|_{L^1(X)} \to \|u_0\|_{L^1(X)}$ as t tends to 0, which yields that

$$\lim_{t\to 0}\frac{1}{t}\int_0^t \mathscr{E}(u_s, u_s)\,dt = \|u_0\|_{L^1(X)}.$$

Combining this with (7.8), we have $\mathscr{E}(u_0, u_0) \geq \|u_0\|_{L^1(X)}$ as desired. □

Remark 7.1. There exist a large variety of $(X, \mathscr{E}, m)$ satisfying $\Gamma(d_x, d_x) = m$ for all $x \in X$, including compact Riemannian manifolds, compact Alexandrov spaces, and the Sierpinski gasket considered in Section 3. Theorem 7.1(ii) then gives the short time asymptotics of the gradient of the heat kernel for them.

Acknowledgments

Pekka Koskela was supported by the Academy of Finland grant 120972. Yuan Zhou was supported by the Program for New Century Excellent Talents in University of Ministry of Education of China and National Natural Science Foundation of China (Grant No. 11201015).

References

[1] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, second ed., in: Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008, p. 334.
[2] L. Ambrosio, N. Gigli, G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below, arXiv:1106.2090.
[3] D. Bakry, On sobolev and logarithmic inequalities for Markov semigroups, in: New Trends in Stochastic Analysis, Charingworth, 1994, World Scientific Publishing, River Edge, NJ, 1997, pp. 43–75.
[4] D. Bakry, M. Emery, Diffusions hypercontractives, in: Seminaire de probabilities, vol. XIX, 1983–1984, pp. 177–206.
[5] Z.M. Balogh, A. Engoulatov, L. Hunziker, O.E. Maasalo, Functional inequalities and Hamilton–Jacobi equations in geodesic spaces, Potential Anal. 36 (2012) 317–337.
[6] A. Beurling, J. Deny, Dirichlet spaces, Proc. Natl. Acad. Sci. USA 45 (1959) 208–215.
[7] J. Cheeger, Differentiability of Lipschitz functions on metric measure spaces, Geom. Funct. Anal. 9 (1999) 428–517.
[8] T.H. Colding, W.P. Minicozzi, Liouville theorems for harmonic sections and applications, Comm. Pure Appl. Math. 51 (1998) 113–138.
[9] M. Erbar, The heat equation on manifolds as a gradient flow in the Wasserstein space, Ann. Inst. Henri Poincaré Probab. Statist. 46 (2010) 1–23.
[10] R.L. Frank, D. Lenz, D. Wingert, Intrinsic metrics for non-local symmetric Dirichlet forms and applications to spectral theory. arXiv:1012.5050.
[11] M. Fukushima, Y. Ōshima, M. Takeda, Dirichlet Forms and Symmetric Markov Processes, in: de Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 1994, p. 392.
[12] N. Gigli, On the heat flow on metric measure spaces: existence, uniqueness and stability, Calc. Var. Partial Differential Equations 39 (2010) 101–120.
[13] N. Gigli, K. Kuwada, S.-I. Ohta, Heat flow on Alexandrov spaces. Preprint.
[14] J. Heinonen, P. Koskela, Quasiconformal maps in metric spaces with controlled geometry, Acta Math. 181 (1998) 1–61.
[15] M. Hino, Energy measures and indices of Dirichlet forms, with applications to derivatives on some fractals, Proc. Lond. Math. Soc. 100 (3) (2010) 269–302.
[16] M. Hino, J.A. Ramírez, Small-time Gaussian behavior of symmetric diffusion semigroups, Ann. Probab. 31 (2003) 1254–1295.

[17] R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker–Planck equation, SIAM J. Math. Anal. 29 (1998) 1–17.

[18] N. Juillet, Diffusion by optimal transport in Heisenberg group. Preprint.

[19] N. Kajino, Heat kernel asymptotics for measurable Riemannian structure on the Sierpinski gasket, Potential Anal. 36 (2012) 67–115.

[20] S. Keith, Modulus and the Poincaré inequality on metric measure spaces, Math. Z. 245 (2003) 255–292.

[21] S. Keith, Measurable differentiable structures and the Poincaré inequality, Indiana Univ. Math. J. 53 (2004) 1127–1150.

[22] J. Kigami, Harmonic metric and Dirichlet form on the Sierpiński gasket, in: Asymptotic Problems in Probability Theory: Stochastic Models and Diffusions on Fractals, Sanda, Kyoto, 1990, in: Pitman Res. Notes Math. Ser., vol. 283, Longman Sci. Tech., Harlow, 1993, pp. 201–218.

[23] J. Kigami, Analysis on Fractals, in: Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001.

[24] J. Kigami, Measurable Riemannian geometry on the Sierpínski gasket: the Kusuoka measure and the Gaussian heat kernel estimate, Math. Ann. 340 (2008) 781–804.

[25] P. Koskela, N. Shanmugalingam, J.T. Tyson, Dirichlet forms, Poincaré inequalities, and the Sobolev spaces of Korevaar and Schoen, Potential Anal. 21 (2004) 241–262.

[26] S. Kusuoka, Dirichlet forms on fractals and products of random matrices, Publ. Res. Inst. Math. Sci. 25 (1989) 659–680.

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

[28] J. Lott, C. Villani, Hamilton–Jacobi semigroup on length spaces and applications, J. Math. Pures Appl. 88 (9) (2007) 219–229.

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

[30] P. Malliavin, D.W. Stroock, Short time behavior of the heat kernel and its logarithmic derivatives, J. Differential Geom. 44 (1996) 550–570.

[31] J.R. Norris, Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds, Acta Math. 179 (1997) 79–103.

[32] S.-I. Ohta, Gradient flows on Wasserstein spaces over compact Alexandrov spaces, Amer. J. Math. 131 (2009) 475–516.

[33] S.-I. Ohta, K.-T. Sturm, Heat flow on Finsler manifolds, Comm. Pure Appl. Math. 62 (2009) 1386–1433.

[34] Y. Ollivier, Ricci curvature of Markov chains on metric spaces, J. Funct. Anal. 256 (2009) 810–864.

[35] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations 26 (2001) 101–174.

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

[37] T. Rajala, Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm. Preprint.

[38] J.A. Ramírez, Short-time asymptotics in Dirichlet spaces, Comm. Pure Appl. Math. 54 (2001) 259–293.

[39] G. Savaré, Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds, C. R. Math. Acad. Sci. Paris 345 (2007) 151–154.

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

[41] N. Shanmugalingam, A universality property of Sobolev spaces in metric measure spaces, in: Sobolev Spaces in Mathematics. I, in: Int. Math. Ser. (N. Y.), vol. 8, Springer, New York, 2009, pp. 345–359.

[42] P. Stollmann, A dual characterization of length spaces with application to Dirichlet metric spaces, Studia Math. 198 (2010) 221–233.

[43] D.W. Stroock, J. Turetsky, Short time behavior of logarithmic derivatives of the heat kernel, Asian J. Math. 1 (1997) 17–33.

[44] K.-T. Sturm, Analysis on local Dirichlet spaces. I. Recurrence, conservativeness and L^p-Liouville properties, J. Reine Angew. Math. 456 (1994) 173–196.

[45] K.-T. Sturm, Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations, Osaka J. Math. 32 (1995) 275–312.

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

[47] K.T. Sturm, Is a diffusion process determined by its intrinsic metric? Chaos Solitons Fractals 8 (1997) 1855–1860.

[48] K.T. Sturm, The geometric aspect of Dirichlet forms, in: New Directions in Dirichlet Forms, in: AMS/IP Stud. Adv. Math., vol. 8, Amer. Math. Soc., Providence, RI, 1998, pp. 233–277.
[49] K.-T. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (2006) 65–131.
[50] K.-T. Sturm, On the geometry of metric measure spaces. II, Acta Math. 196 (2006) 133–177.
[51] C. Villani, Optimal Transport: Old and New, in: Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 338, Springer-Verlag, Berlin, 2009, p. 973.