A semigroup approach to Finsler geometry: Bakry–Ledoux's isoperimetric inequality

Shin-ichi Ohta*†‡

September 30, 2023

Abstract

We develop the celebrated semigroup approach à la Bakry et al on Finsler manifolds, where natural Laplacian and heat semigroup are nonlinear, based on the Bochner–Weitzenböck formula established by Sturm and the author. We show the L^1-gradient estimate on Finsler manifolds (under some additional assumptions in the noncompact case), which is equivalent to a lower weighted Ricci curvature bound and the improved Bochner inequality. As a geometric application, we prove Bakry–Ledoux's Gaussian isoperimetric inequality, again under some additional assumptions in the noncompact case. This extends Cavalletti–Mondino's inequality on reversible Finsler manifolds to non-reversible metrics, and improves the author's previous estimate, both based on the localization (also called needle decomposition) method.

Mathematics Subject Classification (2010): 53C60, 58J35, 49Q20

Contents

*Department of Mathematics, Osaka University, Osaka 560-0043, Japan (s.ohta@math.sci.osaka-u.ac.jp)
†RIKEN Center for Advanced Intelligence Project (AIP), 1-4-1 Nihonbashi, Tokyo 103-0027, Japan
‡Supported in part by JSPS Grant-in-Aid for Scientific Research (KAKENHI) 15K04844, 19H01786.

1 Introduction

The aim of this article is to put forward the semigroup approach in geometric analysis on Finsler manifolds, based on the Bochner–Weitzenböck formula established in [OS3]. There are already a number of applications of the Bochner–Weitzenböck formula (including [WX, Xi, YH, Oh7]), and the machinery in this article would contribute to a further development. In addition, our treatment of a nonlinear generator and the associated nonlinear semigroup (Laplacian and heat semigroup) could be of independent interest from the analytic viewpoint.

The celebrated theory developed by Bakry, Émery, Ledoux et al (called the Γ-*calculus*) studies symmetric generators and the associated linear, symmetric diffusion semigroups under a kind of Bochner inequality (called the (*analytic*) *curvature-dimension condition*). Attributed to Bakry–Émery's original work [BE], this condition will be denoted by $\mathrm{BE}(K,N)$ in this introduction, where $K \in \mathbb{R}$ and $N \in (1,\infty]$ are parameters corresponding to 'curvature' and 'dimension', respectively. This technique is extremely powerful in studying various inequalities (log-Sobolev and Poincaré inequalities, gradient estimates, etc.) in a unified way, we refer to [BE] and the recent book [BGL] for more on this theory.

On a Riemannian manifold equipped with the Laplacian Δ, $\mathrm{BE}(K,N)$ means the following Bochner-type inequality:

$$\Delta\left[\frac{\|\nabla u\|^2}{2}\right] - \langle \nabla(\Delta u), \nabla u\rangle \ge K\|\nabla u\|^2 + \frac{(\Delta u)^2}{N}.$$

Thereby a Riemannian manifold with Ricci curvature not less than K and dimension not greater than N (more generally, a weighted Riemannian manifold of weighted Ricci curvature $\mathrm{Ric}_N \ge K$) is a fundamental example satisfying $\mathrm{BE}(K,N)$.

Later, inspired by [CMS, OV], Sturm [vRS, St1, St2] and Lott–Villani [LV] introduced the (*geometric*) *curvature-dimension condition* $\mathrm{CD}(K,N)$ for metric measure spaces in terms of optimal transport theory. The condition $\mathrm{CD}(K,N)$ characterizes $\mathrm{Ric} \ge K$ and $\dim \le N$ (or $\mathrm{Ric}_N \ge K$) for (weighted) Riemannian manifolds, and its formulation requires a lower regularity of spaces than $\mathrm{BE}(K,N)$. We refer to Villani's book [Vi] for more on this rapidly developing theory. It was shown in [Oh2] that $\mathrm{CD}(K,N)$ also holds and characterizes $\mathrm{Ric}_N \ge K$ for Finsler manifolds, where the natural Laplacian and the associated heat semigroup are nonlinear. For this reason, Ambrosio, Gigli and Savaré [AGS1] introduced a reinforced version $\mathrm{RCD}(K,\infty)$ called the *Riemannian curvature-dimension condition* as the combination of $\mathrm{CD}(K,\infty)$ and the linearity of heat semigroup, followed by the finite-dimensional analogue $\mathrm{RCD}^*(K,N)$ investigated by Erbar, Kuwada and Sturm [EKS] (see also [Gi1, Gi2]). It then turned out that $\mathrm{RCD}^*(K,N)$ is equivalent to $\mathrm{BE}(K,N)$ ([AGS2, EKS]), this equivalence justifies the term 'curvature-dimension condition' which actually came from the similarity to Bakry's theory.

In this article, we develop the theory of Bakry et al on Finsler manifolds. We consider a Finsler manifold M equipped with a Finsler metric $F : TM \longrightarrow [0,\infty)$ and a positive $\mathcal{C}^\infty$-measure $\mathfrak{m}$ on M. We will not assume that F is *reversible*, thereby $F(-v) \neq F(v)$ is allowed. The key ingredient, the *Bochner inequality* under $\mathrm{Ric}_N \ge K$, was established in [OS3] as follows:

$$\Delta^{\nabla u}\left[\frac{F^2(\nabla u)}{2}\right] - d(\Delta u)(\nabla u) \ge KF^2(\nabla u) + \frac{(\Delta u)^2}{N}. \tag{1.1}$$

This Bochner inequality has the same form as the Riemannian case by means of the mixture of the nonlinear Laplacian Δ and its linearization $\Delta^{\nabla u}$. Despite of this mixture, we could derive Bakry–Émery's L^2-gradient estimate as well as Li–Yau's estimates on compact manifolds (see [OS3, §4]). We proceed further in this direction and show the *improved Bochner inequality* under $\mathrm{Ric}_\infty \ge K$ (Proposition 3.5):

$$\Delta^{\nabla u}\left[\frac{F^2(\nabla u)}{2}\right] - d(\Delta u)(\nabla u) - KF^2(\nabla u) \ge d[F(\nabla u)]\big(\nabla^{\nabla u}[F(\nabla u)]\big). \tag{1.2}$$

The first application of (1.2) is the L^1-*gradient estimate* (Theorem 3.7), where we include also the noncompact case but with some additional (likely redundant) assumptions, see the theorem below where we assume the same conditions. We also see that the Bochner inequalities (1.1) (with $N=\infty$), (1.2) and the L^2- and L^1-gradient estimates are all equivalent to $\mathrm{Ric}_\infty \ge K$ (Theorem 3.9).

The second, geometric application of (1.2) is a generalization of *Bakry–Ledoux's Gaussian isoperimetric inequality* (Theorem 4.1):

Theorem (Bakry–Ledoux's isoperimetric inequality) *Let $(M,F,\mathfrak{m})$ be complete and satisfy $\mathrm{Ric}_\infty \ge K > 0$, $\mathfrak{m}(M)=1$, $\mathsf{C}_F < \infty$ and $\mathsf{S}_F < \infty$. We also assume that*

$$d[F(\nabla u_t)]\big(\nabla^{\nabla u_t}[F(\nabla u_t)]\big) \in L^1(M)$$

holds for any global solution $(u_t)_{t\ge 0}$ to the heat equation with $u_0 \in \mathcal{C}_c^\infty(M)$ and any $t>0$. Then we have

$$\mathcal{I}_{(M,F,\mathfrak{m})}(\theta) \ge \mathcal{I}_K(\theta) \tag{1.3}$$

for all $\theta \in [0,1]$, where

$$\mathcal{I}_K(\theta) := \sqrt{\frac{K}{2\pi}}\mathrm{e}^{-Kc^2(\theta)/2} \qquad \text{with} \quad \theta = \int_{-\infty}^{c(\theta)} \sqrt{\frac{K}{2\pi}}\mathrm{e}^{-Ka^2/2}\,da.$$

Here $\mathcal{I}_{(M,F,\mathfrak{m})} : [0,1] \longrightarrow [0,\infty)$ is the *isoperimetric profile* defined as the least boundary area of sets $A \subset M$ with $\mathfrak{m}(A) = \theta$ (see the beginning of Section 4), and C_F (resp. S_F) is the (2-)*uniform convexity* (resp. *smoothness*) *constant* which bounds the *reversibility*,

$$\Lambda_F := \sup_{v \in TM\setminus 0} \frac{F(v)}{F(-v)} \in [1,\infty], \tag{1.4}$$

as $\Lambda_F \le \min\{\sqrt{\mathsf{C}_F}, \sqrt{\mathsf{S}_F}\}$ (see Lemma 2.4). (In particular, the forward completeness is equivalent to the backward completeness, and we denoted it by the plain *completeness* in the theorem.) All the conditions $\mathsf{C}_F < \infty$, $\mathsf{S}_F < \infty$, and $d[F(\boldsymbol{\nabla} u_t)](\nabla^{\boldsymbol{\nabla} u_t}[F(\boldsymbol{\nabla} u_t)]) \in L^1(M)$ hold true in the compact case. In the noncompact case, however, there are technical difficulties and it is unclear how to remove them in this semigroup approach (see §3.4 for a discussion). We remark that, in [Oh8] based on the needle decomposition, we did not need those conditions.

The inequality (1.3) has the same form as the Riemannian case in [BL], and it is sharp and the model space is the real line $\mathbb{R}$ equipped with the normal (Gaussian) distribution $d\mathfrak{m} = \sqrt{K/2\pi}\, \mathrm{e}^{-Kx^2/2}\, dx$. See [BL] for the original work of Bakry and Ledoux on linear diffusion semigroups (influenced by Bobkov's works [Bob1, Bob2]), and [Bor, SC] for the classical Euclidean or Hilbert cases. We also refer to [AM] for the Gaussian isoperimetric inequality on $\mathrm{RCD}(K,\infty)$-spaces by a refinement of the Γ-calculus.

The above theorem extends Cavalletti–Mondino's isoperimetric inequality in [CM] to non-reversible Finsler manifolds. Precisely, they considered essentially non-branching metric measure spaces $(X, d, \mathfrak{m})$ satisfying $\mathrm{CD}(K,N)$ for $K \in \mathbb{R}$ and $N \in (1,\infty)$, and showed the sharp *Lévy–Gromov type isoperimetric inequality* of the form

$$\mathcal{I}_{(X,d,\mathfrak{m})}(\theta) \ge \mathcal{I}_{K,N,D}(\theta)$$

with $\operatorname{diam} X \le D$ $(\le \infty)$. The case of $N = \infty$ is not included in [CM] for technical reasons on the structure of $\mathrm{CD}(K,\infty)$-spaces, but the same argument gives (1.3) (corresponding to $N = D = \infty$) for reversible Finsler manifolds. The proof in [CM] is based on the *needle decomposition* (also called *localization*) inspired by Klartag's work [Kl] on Riemannian manifolds, extending the successful technique in convex geometry. Along the lines of [CM], in [Oh8] we have generalized the needle decomposition to non-reversible Finsler manifolds, however, then we obtain only a weaker isoperimetric inequality,

$$\mathcal{I}_{(M,F,\mathfrak{m})}(\theta) \ge \Lambda_F^{-1} \cdot \mathcal{I}_{K,N,D}(\theta), \tag{1.5}$$

with Λ_F in (1.4). The inequality (1.3) improves (1.5) in the case where $N = D = \infty$ and $K > 0$, and supports a conjecture that the sharp isoperimetric inequality in the non-reversible case is the same as the reversible case, namely Λ_F^{-1} in (1.5) would be removed.

The organization of this article is as follows: In Section 2 we review the basics of Finsler geometry, including the weighted Ricci curvature and the Bochner–Weitzenböck formula. Section 3 is devoted to a detailed study of the nonlinear heat semigroup and its linearizations, we improve the Bochner inequality under $\mathrm{Ric}_\infty \ge K$ and show the L^1-gradient estimate. We prove the isoperimetric inequality in Section 4.

Acknowledgements. I am grateful to Kazumasa Kuwada for his suggestion to consider this problem and for many valuable discussions. I also thank Karl-Theodor Sturm and Kohei Suzuki for stimulating discussions.

2 Geometry and analysis on Finsler manifolds

We review the basics of Finsler geometry (we refer to [BCS, Sh] for further reading), and introduce the weighted Ricci curvature and the nonlinear Laplacian studied in [Oh2, OS1] (see also [GS] for the latter).

Throughout the article, let M be a connected $\mathcal{C}^\infty$-manifold without boundary of dimension $n \ge 2$. We also fix an arbitrary positive $\mathcal{C}^\infty$-measure $\mathfrak{m}$ on M.

2.1 Finsler manifolds

Given local coordinates $(x^i)_{i=1}^n$ on an open set $U \subset M$, we will always use the fiber-wise linear coordinates $(x^i, v^j)_{i,j=1}^n$ of TU such that

$$v = \sum_{j=1}^n v^j \frac{\partial}{\partial x^j}\Big|_x \in T_xM, \qquad x \in U.$$

Definition 2.1 (Finsler structures) We say that a nonnegative function $F : TM \longrightarrow [0,\infty)$ is a $\mathcal{C}^\infty$-*Finsler structure* of M if the following three conditions hold:

(1) (*Regularity*) F is $\mathcal{C}^\infty$ on $TM \setminus 0$, where 0 stands for the zero section;

(2) (*Positive* 1-*homogeneity*) It holds $F(cv) = cF(v)$ for all $v \in TM$ and $c \ge 0$;

(3) (*Strong convexity*) The $n \times n$ matrix

$$\big(g_{ij}(v)\big)_{i,j=1}^n := \left(\frac{1}{2}\frac{\partial^2(F^2)}{\partial v^i \partial v^j}(v)\right)_{i,j=1}^n \tag{2.1}$$

is positive-definite for all $v \in TM \setminus 0$.

We call such a pair (M,F) a $\mathcal{C}^\infty$-*Finsler manifold*.

In other words, F provides a Minkowski norm on each tangent space which varies smoothly in horizontal directions. If $F(-v) = F(v)$ holds for all $v \in TM$, then we say that F is *reversible* or *absolutely homogeneous*. The strong convexity means that the unit sphere $T_xM \cap F^{-1}(1)$ (called the *indicatrix*) is 'positively curved' and implies the strict convexity: $F(v+w) \le F(v) + F(w)$ for all $v, w \in T_xM$ and equality holds only when $v = aw$ or $w = av$ for some $a \ge 0$.

In the coordinates $(x^i, \alpha_j)_{i,j=1}^n$ of T^*U given by $\alpha = \sum_{j=1}^n \alpha_j \, dx^j$, we will also consider

$$g^*_{ij}(\alpha) := \frac{1}{2}\frac{\partial^2[(F^*)^2]}{\partial \alpha_i \partial \alpha_j}(\alpha), \qquad i,j = 1,2,\ldots,n,$$

for $\alpha \in T^*U \setminus 0$. Here $F^* : T^*M \longrightarrow [0,\infty)$ is the *dual Minkowski norm* to F, namely

$$F^*(\alpha) := \sup_{v \in T_xM,\, F(v) \le 1} \alpha(v) = \sup_{v \in T_xM,\, F(v)=1} \alpha(v)$$

for $\alpha \in T_x^*M$. It is clear by definition that $\alpha(v) \le F^*(\alpha)F(v)$, and hence

$$\alpha(v) \ge -F^*(\alpha)F(-v), \qquad \alpha(v) \ge -F^*(-\alpha)F(v).$$

We remark that, however, $\alpha(v) \ge -F^*(\alpha)F(v)$ does not hold in general.

Let us denote by $\mathcal{L}^* : T^*M \longrightarrow TM$ the *Legendre transform*. Precisely, $\mathcal{L}^*$ is sending $\alpha \in T_x^*M$ to the unique element $v \in T_xM$ such that $F(v) = F^*(\alpha)$ and $\alpha(v) = F^*(\alpha)^2$. In coordinates we can write down

$$\mathcal{L}^*(\alpha) = \sum_{i,j=1}^{n} g_{ij}^*(\alpha)\alpha_i \frac{\partial}{\partial x^j}\Big|_x = \sum_{j=1}^{n} \frac{1}{2}\frac{\partial[(F^*)^2]}{\partial \alpha_j}(\alpha)\frac{\partial}{\partial x^j}\Big|_x$$

for $\alpha \in T_x^*M \backslash 0$ (the latter expression makes sense also at 0). Note that $g_{ij}^*(\alpha) = g^{ij}(\mathcal{L}^*(\alpha))$ for $\alpha \in T_x^*M \setminus 0$, where $(g^{ij}(v))$ denotes the inverse matrix of $(g_{ij}(v))$. The map $\mathcal{L}^*|_{T_x^*M}$ is being a linear operator only when $F|_{T_xM}$ comes from an inner product. We also define $\mathcal{L} := (\mathcal{L}^*)^{-1} : TM \longrightarrow T^*M$.

For $x, y \in M$, we define the (asymmetric) *distance* from x to y by

$$d(x, y) := \inf_{\eta} \int_0^1 F\big(\dot{\eta}(t)\big)\, dt,$$

where $\eta : [0, 1] \longrightarrow M$ runs over all $\mathcal{C}^1$-curves such that $\eta(0) = x$ and $\eta(1) = y$. Note that $d(y, x) \neq d(x, y)$ can happen since F is only positively homogeneous. A $\mathcal{C}^\infty$-curve η on M is called a *geodesic* if it is locally minimizing and has a constant speed with respect to d, similarly to Riemannian or metric geometry. See (2.7) below for the precise geodesic equation. For $v \in T_xM$, if there is a geodesic $\eta : [0, 1] \longrightarrow M$ with $\dot{\eta}(0) = v$, then we define the *exponential map* by $\exp_x(v) := \eta(1)$. We say that (M, F) is *forward complete* if the exponential map is defined on whole TM. Then the Hopf–Rinow theorem ensures that any pair of points is connected by a minimal geodesic (see [BCS, Theorem 6.6.1]).

Given each $v \in T_xM \setminus 0$, the positive-definite matrix $(g_{ij}(v))_{i,j=1}^n$ in (2.1) induces the Riemannian structure g_v of T_xM by

$$g_v\bigg(\sum_{i=1}^{n} a_i \frac{\partial}{\partial x^i}\Big|_x, \sum_{j=1}^{n} b_j \frac{\partial}{\partial x^j}\Big|_x\bigg) := \sum_{i,j=1}^{n} g_{ij}(v) a_i b_j. \tag{2.2}$$

Notice that this definition is coordinate-free and $g_v(v, v) = F^2(v)$ holds. One can regard g_v as the best Riemannian approximation of $F|_{T_xM}$ in the direction v. The *Cartan tensor*

$$A_{ijk}(v) := \frac{F(v)}{2}\frac{\partial g_{ij}}{\partial v^k}(v), \qquad v \in TM \setminus 0,$$

measures the variation of g_v in vertical directions, and vanishes everywhere on $TM \setminus 0$ if and only if F comes from a Riemannian metric.

The following useful fact on homogeneous functions (see [BCS, Theorem 1.2.1]) plays a fundamental role in our calculus.

Theorem 2.2 (Euler's theorem) *Suppose that a differentiable function $H : \mathbb{R}^n \setminus 0 \longrightarrow \mathbb{R}$ satisfies $H(cv) = c^r H(v)$ for some $r \in \mathbb{R}$ and all $c > 0$ and $v \in \mathbb{R}^n \backslash 0$ (that is,* positively r-homogeneous*). Then we have, for all $v \in \mathbb{R}^n \setminus 0$,*

$$\sum_{i=1}^{n} \frac{\partial H}{\partial v^i}(v) v^i = rH(v).$$

Observe that g_{ij} is positively 0-homogeneous on each T_xM, and hence

$$\sum_{i=1}^{n} A_{ijk}(v)v^i = \sum_{j=1}^{n} A_{ijk}(v)v^j = \sum_{k=1}^{n} A_{ijk}(v)v^k = 0 \tag{2.3}$$

for all $v \in TM \setminus 0$ and $i, j, k = 1, 2, \dots, n$. Define the *formal Christoffel symbol*

$$\gamma^i_{jk}(v) := \frac{1}{2}\sum_{l=1}^{n} g^{il}(v)\left\{\frac{\partial g_{lk}}{\partial x^j}(v) + \frac{\partial g_{jl}}{\partial x^k}(v) - \frac{\partial g_{jk}}{\partial x^l}(v)\right\} \tag{2.4}$$

for $v \in TM \setminus 0$, and the *geodesic spray coefficients* and the *nonlinear connection*

$$G^i(v) := \sum_{j,k=1}^{n} \gamma^i_{jk}(v)v^j v^k, \qquad N^i_j(v) := \frac{1}{2}\frac{\partial G^i}{\partial v^j}(v)$$

for $v \in TM \setminus 0$ ($G^i(0) = N^i_j(0) := 0$ by convention). Note that G^i is positively 2-homogeneous, hence Theorem 2.2 implies $\sum_{j=1}^{n} N^i_j(v)v^j = G^i(v)$.

By using N^i_j, the coefficients of the *Chern connection* are given by

$$\Gamma^i_{jk}(v) := \gamma^i_{jk}(v) - \sum_{l,m=1}^{n} \frac{g^{il}}{F}(A_{lkm}N^m_j + A_{jlm}N^m_k - A_{jkm}N^m_l)(v) \tag{2.5}$$

on $TM \setminus 0$. The corresponding *covariant derivative* of a vector field X by $v \in T_xM$ with *reference vector* $w \in T_xM \setminus 0$ is defined as

$$D^w_v X(x) := \sum_{i,j=1}^{n}\left\{v^j \frac{\partial X^i}{\partial x^j}(x) + \sum_{k=1}^{n} \Gamma^i_{jk}(w)v^j X^k(x)\right\}\frac{\partial}{\partial x^i}\Big|_x \in T_xM. \tag{2.6}$$

Then the *geodesic equation* is written as, with the help of (2.3),

$$D^{\dot\eta}_{\dot\eta}\dot\eta(t) = \sum_{i=1}^{n}\left\{\ddot\eta^i(t) + G^i\big(\dot\eta(t)\big)\right\}\frac{\partial}{\partial x^i}\Big|_{\eta(t)} = 0. \tag{2.7}$$

2.2 Uniform convexity and smoothness

We will need the following quantity associated with (M, F):

$$\mathsf{S}_F := \sup_{x\in M}\sup_{v,w\in T_xM\setminus 0} \frac{g_v(w,w)}{F^2(w)} \in [1, \infty].$$

Since $g_v(w,w) \le \mathsf{S}_F F^2(w)$ and g_v is the Hessian of $F^2/2$ at v, the constant S_F measures the (fiber-wise) concavity of F^2 and is called the (2-)*uniform smoothness constant* (see [Oh1]). We remark that $\mathsf{S}_F = 1$ holds if and only if (M, F) is Riemannian. The following lemma is a standard fact, we give a proof for thoroughness.

Lemma 2.3 *For any $x \in M$, $v \in T_xM \setminus 0$ and $\alpha := \mathcal{L}(v)$, we have*

$$\sup_{w \in T_xM \setminus 0} \frac{g_v(w,w)}{F^2(w)} = \sup_{\beta \in T_x^*M \setminus 0} \frac{F^*(\beta)^2}{g_\alpha^*(\beta,\beta)},$$

where g_α^* is the inner product of T_x^*M defined by

$$g_\alpha^*(\beta,\beta) := \sum_{i,j=1}^n g_{ij}^*(\alpha)\beta_i\beta_j, \qquad \beta = \sum_{i=1}^n \beta_i \, dx^i.$$

Proof. Choose local coordinates $(x^i)_{i=1}^n$ around x such that $g_{ij}(v) = \delta_{ij}$ and set

$$\mathbb{S}_x := \left\{ w = \sum_{i=1}^n w^i \frac{\partial}{\partial x^i} \in T_xM \,\middle|\, \sum_{i=1}^n (w^i)^2 = 1 \right\},$$
$$\mathbb{S}_x^* := \left\{ \beta = \sum_{i=1}^n \beta_i \, dx^i \in T_x^*M \,\middle|\, \sum_{i=1}^n (\beta_i)^2 = 1 \right\}.$$

First, given $w \in \mathbb{S}_x$, we take $\beta \in \mathbb{S}_x^*$ such that $\beta(w) = 1$. Then we have $1 = \beta(w) \le F^*(\beta)F(w)$ and hence

$$\frac{g_v(w,w)}{F^2(w)} = \frac{1}{F^2(w)} \le F^*(\beta)^2 = \frac{F^*(\beta)^2}{g_\alpha^*(\beta,\beta)}.$$

Next, for $\beta' \in \mathbb{S}_x^*$, take $w' \in \mathbb{S}_x$ with $\beta'(w') = F^*(\beta')F(w')$. Then $F^*(\beta')F(w') = \beta'(w') \le 1$ and hence $1/F^2(w') \ge F^*(\beta')^2$. This completes the proof. $\square$

One can in a similar manner introduce the (2-)*uniform convexity constant*:

$$\mathsf{C}_F := \sup_{x \in M} \sup_{v,w \in T_xM \setminus 0} \frac{F^2(w)}{g_v(w,w)} = \sup_{x \in M} \sup_{\alpha,\beta \in T_x^*M \setminus 0} \frac{g_\alpha^*(\beta,\beta)}{F^*(\beta)^2} \in [1,\infty]. \tag{2.8}$$

Again, $\mathsf{C}_F = 1$ holds if and only if (M,F) is Riemannian. We remark that S_F and C_F control the *reversibility constant* Λ_F defined in (1.4) as follows.

Lemma 2.4 *We have*

$$\Lambda_F \le \min\{\sqrt{\mathsf{S}_F}, \sqrt{\mathsf{C}_F}\}.$$

Proof. For any $v \in TM \setminus 0$, we observe

$$\frac{F^2(v)}{F^2(-v)} = \frac{g_v(v,v)}{F^2(-v)} = \frac{g_v(-v,-v)}{F^2(-v)} \le \mathsf{S}_F,$$

and similarly

$$\frac{F^2(v)}{F^2(-v)} = \frac{F^2(v)}{g_{-v}(v,v)} \le \mathsf{C}_F.$$

$\square$

2.3 Weighted Ricci curvature

The *Ricci curvature* (as the trace of the *flag curvature*) on a Finsler manifold is defined by using some connection. Instead of giving a precise definition in coordinates (for which we refer to [BCS]), here we explain a useful interpretation in [Sh, §6.2] going back to (at least) [Au]. Given a unit vector $v \in T_xM \cap F^{-1}(1)$, we extend it to a $\mathcal{C}^\infty$-vector field V on a neighborhood U of x in such a way that every integral curve of V is geodesic, and consider the Riemannian structure g_V of U induced from (2.2). Then the *Finsler* Ricci curvature $\mathrm{Ric}(v)$ of v with respect to F coincides with the *Riemannian* Ricci curvature of v with respect to g_V (in particular, it is independent of the choice of V).

Inspired by the above interpretation of the Ricci curvature as well as the theory of weighted Ricci curvature (also called the *Bakry–Émery–Ricci curvature*) of Riemannian manifolds, the *weighted Ricci curvature* for $(M, F, \mathfrak{m})$ was introduced in [Oh2] as follows. Recall that $\mathfrak{m}$ is a positive $\mathcal{C}^\infty$-measure on M, from here on it comes into play.

Definition 2.5 (Weighted Ricci curvature) Given a unit vector $v \in T_xM$, let V be a $\mathcal{C}^\infty$-vector field on a neighborhood U of x as above. We decompose $\mathfrak{m}$ as $\mathfrak{m} = \mathrm{e}^{-\Psi} \mathrm{vol}_{g_V}$ on U, where $\Psi \in \mathcal{C}^\infty(U)$ and vol_{g_V} is the volume form of g_V. Denote by $\eta : (-\varepsilon, \varepsilon) \longrightarrow M$ the geodesic such that $\dot\eta(0) = v$. Then, for $N \in (-\infty, 0) \cup (n, \infty)$, define

$$\mathrm{Ric}_N(v) := \mathrm{Ric}(v) + (\Psi \circ \eta)''(0) - \frac{(\Psi \circ \eta)'(0)^2}{N - n}.$$

We also define as the limits:

$$\mathrm{Ric}_\infty(v) := \mathrm{Ric}(v) + (\Psi \circ \eta)''(0), \qquad \mathrm{Ric}_n(v) := \lim_{N \downarrow n} \mathrm{Ric}_N(v).$$

For $c \geq 0$, we set $\mathrm{Ric}_N(cv) := c^2 \mathrm{Ric}_N(v)$.

We will denote by $\mathrm{Ric}_N \geq K$, $K \in \mathbb{R}$, the condition $\mathrm{Ric}_N(v) \geq KF^2(v)$ for all $v \in TM$. In the Riemannian case, the study of Ric_∞ goes back to Lichnerowicz [Li], he showed a Cheeger–Gromoll type splitting theorem (see [Oh5] for a Finsler counterpart). The range $N \in (n, \infty)$ has been well studied by Bakry [Ba, §6], Qian [Qi] and many others. The study of the range $N \in (-\infty, 0)$ is more recent; see [Mi2] for isoperimetric inequalities, [Oh6] for the curvature-dimension condition, and [Wy] for splitting theorems (for $N \in (-\infty, 1]$).

It was established in [Oh2] (and [Oh6] for $N < 0$, [Oh8] for $N = 0$) that, for $K \in \mathbb{R}$, the bound $\mathrm{Ric}_N \geq K$ is equivalent to Lott, Sturm and Villani's *curvature-dimension condition* $\mathrm{CD}(K, N)$. This extends the corresponding result on weighted Riemannian manifolds and has many geometric and analytic applications (see [Oh2, OS1] among others).

Remark 2.6 (S-curvature) For a Riemannian manifold (M, g, vol_g) endowed with the Riemannian volume measure, clearly we have $\Psi \equiv 0$ and hence $\mathrm{Ric}_N = \mathrm{Ric}$ for all N. It is also known that, for Finsler manifolds of *Berwald type* (i.e., Γ^k_{ij} is constant on each $T_xM \setminus 0$), the *Busemann–Hausdorff measure* satisfies $(\Psi \circ \eta)' \equiv 0$ (in other words, Shen's **S**-*curvature* vanishes, see [Sh, §7.3]). For a general Finsler manifold, however, there may not exist any measure with vanishing **S**-curvature (see [Oh3] for such an example). This is a reason why we chose to begin with an arbitrary measure $\mathfrak{m}$.

For later convenience, we introduce the following notations.

Definition 2.7 (Reverse Finsler structures) We define the *reverse Finsler structure* $\overleftarrow{F}$ of F by $\overleftarrow{F}(v) := F(-v)$.

We will put an arrow $\leftarrow$ on those quantities associated with $\overleftarrow{F}$, we have for example $\overleftarrow{d}(x,y) = d(y,x)$, $\overleftarrow{\mathrm{Ric}}_N(v) = \mathrm{Ric}_N(-v)$ and $\overleftarrow{\boldsymbol{\nabla}} u = -\boldsymbol{\nabla}(-u)$. We say that (M,F) is *backward complete* if $(M, \overleftarrow{F})$ is forward complete. If $\Lambda_F < \infty$, then these completenesses are mutually equivalent, and we may call it simply *completeness*.

2.4 Nonlinear Laplacian and heat flow

For a differentiable function $u : M \longrightarrow \mathbb{R}$, the *gradient vector* at x is defined as the Legendre transform of the derivative of u: $\boldsymbol{\nabla} u(x) := \mathcal{L}^*(du(x)) \in T_xM$. If $du(x) \neq 0$, then we can write down in coordinates as

$$\boldsymbol{\nabla} u = \sum_{i,j=1}^{n} g^*_{ij}(du) \frac{\partial u}{\partial x^j} \frac{\partial}{\partial x^i}.$$

We need to be careful when $du(x) = 0$, because $g^*_{ij}(du(x))$ is not defined as well as the Legendre transform $\mathcal{L}^*$ is only continuous at the zero section. Therefore we set

$$M_u := \{x \in M \,|\, du(x) \neq 0\}.$$

For a twice differentiable function $u : M \longrightarrow \mathbb{R}$ and $x \in M_u$, we define a kind of *Hessian* $\boldsymbol{\nabla}^2 u(x) \in T^*_xM \otimes T_xM$ by using the covariant derivative (2.6) as

$$\boldsymbol{\nabla}^2 u(v) := D_v^{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u)(x) \in T_xM, \qquad v \in T_xM.$$

The operator $\boldsymbol{\nabla}^2 u(x)$ is symmetric in the sense that

$$g_{\boldsymbol{\nabla} u}\big(\boldsymbol{\nabla}^2 u(v), w\big) = g_{\boldsymbol{\nabla} u}\big(v, \boldsymbol{\nabla}^2 u(w)\big)$$

for all $v, w \in T_xM$ with $x \in M_u$ (see, for example, [OS3, Lemma 2.3]).

Define the *divergence* of a differentiable vector field V on M with respect to the measure $\mathfrak{m}$ by

$$\mathrm{div}_{\mathfrak{m}} V := \sum_{i=1}^{n} \left(\frac{\partial V^i}{\partial x^i} + V^i \frac{\partial \Phi}{\partial x^i} \right), \qquad V = \sum_{i=1}^{n} V^i \frac{\partial}{\partial x^i},$$

where we decomposed $\mathfrak{m}$ as $d\mathfrak{m} = \mathrm{e}^{\Phi}\, dx^1 dx^2 \cdots dx^n$. One can rewrite in the weak form as

$$\int_M \phi \,\mathrm{div}_{\mathfrak{m}} V \, d\mathfrak{m} = -\int_M d\phi(V)\, d\mathfrak{m} \qquad \text{for all } \phi \in \mathcal{C}_c^\infty(M),$$

that makes sense for measurable vector fields V with $F(V) \in L^1_{\mathrm{loc}}(M)$. Then we define the distributional *Laplacian* of $u \in H^1_{\mathrm{loc}}(M)$ by $\boldsymbol{\Delta} u := \mathrm{div}_{\mathfrak{m}}(\boldsymbol{\nabla} u)$ in the weak sense that

$$\int_M \phi \boldsymbol{\Delta} u \, d\mathfrak{m} := -\int_M d\phi(\boldsymbol{\nabla} u)\, d\mathfrak{m} \qquad \text{for all } \phi \in \mathcal{C}_c^\infty(M).$$

Notice that the space $H^1_{\rm loc}(M)$ is defined solely in terms of the differentiable structure of M. Since taking the gradient vector (more precisely, the Legendre transform) is a nonlinear operation, our Laplacian $\boldsymbol{\Delta}$ is a nonlinear operator unless F is Riemannian.

In [OS1, OS3], we have studied the associated *nonlinear heat equation* $\partial_t u = \boldsymbol{\Delta} u$. In order to recall some results in [OS1], we define the *Dirichlet energy* of $u \in H^1_{\rm loc}(M)$ by

$$\mathcal{E}(u) := \frac{1}{2}\int_M F^2(\boldsymbol{\nabla} u)\, d\mathfrak{m} = \frac{1}{2}\int_M F^*(du)^2\, d\mathfrak{m}.$$

We remark that $\mathcal{E}(u) < \infty$ does not necessarily imply $\mathcal{E}(-u) < \infty$. Define $H^1_0(M)$ as the closure of $\mathcal{C}^\infty_c(M)$ with respect to the (absolutely homogeneous) norm

$$\|u\|_{H^1} := \|u\|_{L^2} + \{\mathcal{E}(u) + \mathcal{E}(-u)\}^{1/2}.$$

Note that $(H^1_0(M), \|\cdot\|_{H^1})$ is a Banach space.

Definition 2.8 (Global solutions) We say that a function u on $[0,T] \times M$, $T > 0$, is a *global solution* to the heat equation $\partial_t u = \boldsymbol{\Delta} u$ if it satisfies the following:

(1) $u \in L^2\big([0,T], H^1_0(M)\big) \cap H^1\big([0,T], H^{-1}(M)\big)$;

(2) For every $\phi \in \mathcal{C}^\infty_c(M)$, we have

$$\int_M \phi \cdot \partial_t u_t\, d\mathfrak{m} = -\int_M d\phi(\boldsymbol{\nabla} u_t)\, d\mathfrak{m}$$

for almost all $t \in [0,T]$, where we set $u_t := u(t,\cdot)$.

We refer to [Ev] for the notations as in (1). Denoted by $H^{-1}(M)$ is the dual Banach space of $H^1_0(M)$ (so that $H^1_0(M) \subset L^2(M) \subset H^{-1}(M)$). By noticing

$$\begin{aligned}\int_M |(d\phi - d\bar{\phi})(\boldsymbol{\nabla} u_t)|\, d\mathfrak{m} &\le \int_M \max\big\{F^*\big(d(\phi-\bar{\phi})\big), F^*\big(d(\bar{\phi}-\phi)\big)\big\} F(\boldsymbol{\nabla} u_t)\, d\mathfrak{m} \\ &\le \{2\mathcal{E}(\phi-\bar{\phi}) + 2\mathcal{E}(\bar{\phi}-\phi)\}^{1/2} \cdot \{2\mathcal{E}(u_t)\}^{1/2},\end{aligned}$$

the test function ϕ can be taken from $H^1_0(M)$. Global solutions can be constructed as gradient curves of the energy functional $\mathcal{E}$ in the Hilbert space $L^2(M)$. We summarize the existence and regularity properties established in [OS1, §§3, 4] in the next theorem.

Theorem 2.9 *Assume $\Lambda_F < \infty$.*

(i) *For each initial datum $u_0 \in H^1_0(M)$ and $T > 0$, there exists a unique global solution u to the heat equation on $[0,T] \times M$, and the distributional Laplacian $\boldsymbol{\Delta} u_t$ is absolutely continuous with respect to $\mathfrak{m}$ for all $t \in (0,T)$.*

(ii) *One can take the continuous version of a global solution u, and it enjoys the $H^2_{\rm loc}$-regularity in x as well as the $\mathcal{C}^{1,\alpha}$-regularity for some α in both t and x. Moreover, $\partial_t u$ lies in $H^1_{\rm loc}(M) \cap \mathcal{C}(M)$, and further in $H^1_0(M)$ if $\mathsf{S}_F < \infty$.*

We remark that the usual elliptic regularity yields that u is $\mathcal{C}^\infty$ on $\bigcup_{t>0}(\{t\} \times M_{u_t})$. The proof of $\partial_t u \in H_0^1(M)$ under $\mathsf{S}_F < \infty$ can be found in [OS1, Appendix A]. The uniqueness in (i) is a consequence of the convexity of F^* (see [OS1, Proposition 3.5]).

We finally remark that, by the construction of heat flow as the gradient flow of $\mathcal{E}$, it is readily seen that:

$$\text{If } u_0 \ge 0 \text{ almost everywhere, then } u_t \ge 0 \text{ almost everywhere for all } t > 0. \tag{2.9}$$

Indeed, if $u_t < 0$ on a non-null set, then the curve $\bar{u}_t := \max\{u_t, 0\}$ will give a less energy with a less L^2-length, a contradiction.

2.5 Bochner–Weitzenböck formula

Given $f \in H^1_{\mathrm{loc}}(M)$ and a measurable vector field V such that $V \neq 0$ almost everywhere on $M_f = \{x \in M \,|\, df(x) \neq 0\}$, we can define the gradient vector field and the Laplacian on the weighted Riemannian manifold $(M, g_V, \mathfrak{m})$ by

$$\nabla^V f := \begin{cases} \displaystyle\sum_{i,j=1}^n g^{ij}(V) \frac{\partial f}{\partial x^j} \frac{\partial}{\partial x^i} & \text{on } M_f, \\ 0 & \text{on } M \setminus M_f, \end{cases} \qquad \Delta^V f := \mathrm{div}_{\mathfrak{m}}(\nabla^V f),$$

where the latter is in the sense of distribution. We have $\nabla^{\boldsymbol{\nabla} u} u = \boldsymbol{\nabla} u$ and $\Delta^{\boldsymbol{\nabla} u} u = \boldsymbol{\Delta} u$ for $u \in H^1_{\mathrm{loc}}(M)$ ([OS1, Lemma 2.4]). We also observe that, for $f_1, f_2 \in H^1_{\mathrm{loc}}(M)$ and V such that $V \neq 0$ almost everywhere,

$$df_2(\nabla^V f_1) = g_V(\nabla^V f_1, \nabla^V f_2) = df_1(\nabla^V f_2). \tag{2.10}$$

We established in [OS3, Theorem 3.3] the following key ingredient of the Γ-calculus.

Theorem 2.10 (Bochner–Weitzenböck formula) *Given $u \in \mathcal{C}^\infty(M)$, we have*

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) = \mathrm{Ric}_\infty(\boldsymbol{\nabla} u) + \|\boldsymbol{\nabla}^2 u\|^2_{\mathrm{HS}(\boldsymbol{\nabla} u)} \tag{2.11}$$

as well as

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) \ge \mathrm{Ric}_N(\boldsymbol{\nabla} u) + \frac{(\boldsymbol{\Delta} u)^2}{N}$$

for $N \in (-\infty, 0) \cup [n, \infty]$ point-wise on M_u, where $\|\cdot\|_{\mathrm{HS}(\boldsymbol{\nabla} u)}$ denotes the Hilbert–Schmidt norm with respect to $g_{\boldsymbol{\nabla} u}$.

In particular, if $\mathrm{Ric}_N \ge K$, then we have

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) \ge K F^2(\boldsymbol{\nabla} u) + \frac{(\boldsymbol{\Delta} u)^2}{N} \tag{2.12}$$

on M_u, that we will call the *Bochner inequality*. One can further generalize the Bochner–Weitzenböck formula to a more general class of Hamiltonian systems (by dropping the positive 1-homogeneity; see [Lee, Oh4]).

Remark 2.11 (**F versus $g_{\boldsymbol{\nabla} u}$**) In contrast to $\Delta^{\boldsymbol{\nabla} u} u = \boldsymbol{\Delta} u$, $\mathrm{Ric}_N(\boldsymbol{\nabla} u)$ may not coincide with the weighted Ricci curvature $\mathrm{Ric}_N^{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u)$ of the weighted Riemannian manifold $(M, g_{\boldsymbol{\nabla} u}, \mathfrak{m})$. It is compensated in (2.11) by the fact that $\boldsymbol{\nabla}^2 u$ does not necessarily coincide with the Hessian of u with respect to $g_{\boldsymbol{\nabla} u}$.

The integrated form was shown in [OS3, Theorem 3.6], with the help of the following fact to overcome the ill-posedness of $\boldsymbol{\nabla} u$ on $M \setminus M_u$ (see [Leo, Exercise 10.37(iv)], [Ma, Lemma 1.7.1] for example).

Lemma 2.12 *For each $f \in H^1_{\mathrm{loc}}(M)$, we have $df = 0$ almost everywhere on $f^{-1}(0)$. If $f \in H^1_{\mathrm{loc}}(M) \cap L^\infty_{\mathrm{loc}}(M)$, then $d(f^2/2) = f\,df = 0$ also holds almost everywhere on $f^{-1}(0)$.*

Theorem 2.13 (Integrated form) *Assume $\mathrm{Ric}_N \geq K$ for some $K \in \mathbb{R}$ and $N \in (-\infty, 0) \cup [n, \infty]$. Given $u \in H^2_{\mathrm{loc}}(M) \cap \mathcal{C}^1(M)$ such that $\boldsymbol{\Delta} u \in H^1_{\mathrm{loc}}(M)$, we have*

$$-\int_M d\phi \left(\nabla^{\boldsymbol{\nabla} u} \left[\frac{F^2(\boldsymbol{\nabla} u)}{2} \right] \right) d\mathfrak{m} \geq \int_M \phi \left\{ d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) + K F^2(\boldsymbol{\nabla} u) + \frac{(\boldsymbol{\Delta} u)^2}{N} \right\} d\mathfrak{m}$$

for all bounded nonnegative functions $\phi \in H^1_c(M) \cap L^\infty(M)$.

Recall from Theorem 2.9(ii) that global solutions to the heat equation always enjoy $u \in H^1_0(M) \cap H^2_{\mathrm{loc}}(M) \cap \mathcal{C}^1(M)$ and $\boldsymbol{\Delta} u \in H^1_{\mathrm{loc}}(M)$.

3 Linearized semigroups and gradient estimates

In the Bochner–Weitzenböck formula (Theorem 2.10) in the previous section, we used the linearized Laplacian $\Delta^{\boldsymbol{\nabla} u}$ induced from the Riemannian structure $g_{\boldsymbol{\nabla} u}$. In the same spirit, we can consider the linearized heat equation associated with a global solution to the heat equation. This technique turned out useful and we have obtained gradient estimates à la Bakry–Émery and Li–Yau in [OS3, §4]. In this section we discuss such a linearization in detail and improve the L^2-gradient estimate to an L^1-bound (Theorem 3.7).

3.1 Linearized heat semigroups and their adjoints

Let $(u_t)_{t \geq 0}$ be a global solution to the heat equation. We will fix a measurable one-parameter family of *non-vanishing* vector fields $(V_t)_{t \geq 0}$ such that $V_t = \boldsymbol{\nabla} u_t$ on M_{u_t} for each $t \geq 0$. Given $f \in H^1_0(M)$ and $s \geq 0$, let $(P^{\boldsymbol{\nabla} u}_{s,t}(f))_{t \geq s}$ be the weak solution to the *linearized heat equation*:

$$\partial_t [P^{\boldsymbol{\nabla} u}_{s,t}(f)] = \Delta^{V_t} [P^{\boldsymbol{\nabla} u}_{s,t}(f)], \qquad P^{\boldsymbol{\nabla} u}_{s,s}(f) = f. \tag{3.1}$$

The existence and other properties of the linearized semigroup $P^{\boldsymbol{\nabla} u}_{s,t}$ are summarized in the following proposition.

Proposition 3.1 (Properties of linearized semigroups) *Assume that $(M, F, \mathfrak{m})$ is complete and satisfies $\mathsf{C}_F < \infty$ and $\mathsf{S}_F < \infty$, and let $(u_t)_{t \geq 0}$ and $(V_t)_{t \geq 0}$ be as above.*

(i) *For each $s \geq 0$, $T > 0$ and $f \in H_0^1(M)$, there exists a unique weak solution $f_t = P_{s,t}^{\nabla u}(f)$, $t \in [s, s+T]$, to (3.1). Moreover, $(f_t)_{t\in[s,s+T]}$ lies in $L^2([s,s+T], H_0^1(M)) \cap H^1([s,s+T], H^{-1}(M))$ as well as $\mathcal{C}([s,s+T], L^2(M))$.*

(ii) *The solution $(f_t)_{t\in[s,s+T]}$ in* (i) *is Hölder continuous on $(s, s+T) \times M$.*

(iii) *Assume that either $\mathfrak{m}(M) < \infty$ or $\mathrm{Ric}_\infty \geq K$ for some $K \in \mathbb{R}$ holds. If $c \leq f \leq C$ for some $-\infty < c < C < \infty$, then we have $c \leq f_t \leq C$ almost everywhere for all $t \in (s, s+T]$.*

Proof. (i) Let $s = 0$ without loss of generality. This unique existence follows from Theorem 4.1 and Remark 4.3 in [LM, Chapter III] (see also [RR, Theorem 11.3], where $A(t)$ is assumed to be continuous in t but it is in fact unnecessary). Precisely, in the notations in [LM], we take $H = L^2(M)$, $V = H_0^1(M)$, and put $A_t := -\Delta^{V_t} : H_0^1(M) \longrightarrow H^{-1}(M)$. We deduce with the help of (2.8) that, for any $h, \bar{h} \in H_0^1(M)$,

$$\left| \int_M \bar{h} \Delta^{V_t} h \, d\mathfrak{m} \right| = \left| \int_M g^*_{\mathcal{L}(V_t)}(dh, d\bar{h}) \, d\mathfrak{m} \right| \leq 2\sqrt{\mathcal{E}^{V_t}(h)}\sqrt{\mathcal{E}^{V_t}(\bar{h})} \leq 2\mathsf{C}_F \sqrt{\mathcal{E}(h)}\sqrt{\mathcal{E}(\bar{h})}$$

and

$$-\int_M h \Delta^{V_t} h \, d\mathfrak{m} = 2\mathcal{E}^{V_t}(h) \geq \frac{2}{\mathsf{S}_F}\mathcal{E}(h),$$

where $\mathcal{E}^{V_t}(h) := (1/2)\int_M g^*_{\mathcal{L}(V_t)}(dh, dh)\, d\mathfrak{m}$ denotes the energy functional on $(M, g_{V_t}, \mathfrak{m})$. Since $\Lambda_F < \infty$ by $\mathsf{C}_F < \infty$ (or $\mathsf{S}_F < \infty$), $\|h\|_{L^2} + \sqrt{\mathcal{E}(h)}$ is comparable with $\|h\|_{H^1}$. Therefore we have a unique solution $(f_t)_{t\in[0,T]}$ to (3.1) with $f_0 = f$ lying in $L^2([0,T], H_0^1(M)) \cap H^1([0,T], H^{-1}(M))$, and also in $\mathcal{C}([0,T], L^2(M))$ (see [Ev, §5.9.2], [RR, Lemma 11.4]).

(ii) The Hölder continuity is a consequence of the local uniform ellipticity of Δ^{V_t} (see [OS1, Proposition 4.4]).

(iii) This is seen for example by using the fundamental solution $q(t,x;s,y)$ to the equation $\partial_t[P_{s,t}^{\nabla u}(f)] = \Delta^{V_t}[P_{s,t}^{\nabla u}(f)]$ (see [Sal, §6]). We have

$$f_t(x) = \int_M q(t,x;s,y) f(y)\, \mathfrak{m}(dy),$$

and $\int_M q(t,x;s,y)\, \mathfrak{m}(dy) = 1$ (by $1 \in H_0^1(M)$ when $\mathfrak{m}(M) < \infty$, or by [Sal, §7] since $\mathrm{Ric}_\infty \geq K$ implies the squared exponential volume bound as in [St1, Theorem 4.24]). This completes the proof. $\square$

The uniqueness in (i) above ensures that $u_t = P_{s,t}^{\nabla u}(u_s)$. It follows from the non-expansion property,

$$\frac{d}{dt}\left[\|f_t\|_{L^2}^2\right] = -4\mathcal{E}^{V_t}(f_t) \leq 0,$$

that $P_{s,t}^{\nabla u}$ uniquely extends to a linear contraction semigroup acting on $L^2(M)$. Notice also that f is $\mathcal{C}^\infty$ on $\bigcup_{s<t<s+T}(\{t\} \times M_{u_t})$.

The operator $P_{s,t}^{\nabla u}$ is linear but not symmetric (with respect to the L^2-inner product). Let us denote by $\widehat{P}_{s,t}^{\nabla u}$ the *adjoint operator* of $P_{s,t}^{\nabla u}$. That is to say, given $\phi \in H_0^1(M)$ and $t > 0$, we define $(\widehat{P}_{s,t}^{\nabla u}(\phi))_{s\in[0,t]}$ as the solution to the equation

$$\partial_s[\widehat{P}_{s,t}^{\nabla u}(\phi)] = -\Delta^{V_s}[\widehat{P}_{s,t}^{\nabla u}(\phi)], \qquad \widehat{P}_{t,t}^{\nabla u}(\phi) = \phi. \tag{3.2}$$

Note that

$$\int_M \phi \cdot P^{\boldsymbol{\nabla} u}_{s,t}(f)\, d\mathfrak{m} = \int_M \widehat{P}^{\boldsymbol{\nabla} u}_{s,t}(\phi) \cdot f\, d\mathfrak{m} \tag{3.3}$$

indeed holds, since for $r \in (0, t-s)$

$$\begin{aligned}
&\partial_r \bigg[\int_M \widehat{P}^{\boldsymbol{\nabla} u}_{s+r,t}(\phi) \cdot P^{\boldsymbol{\nabla} u}_{s,s+r}(f)\, d\mathfrak{m} \bigg] \\
&= -\int_M \Delta^{V_{s+r}}[\widehat{P}^{\boldsymbol{\nabla} u}_{s+r,t}(\phi)] \cdot P^{\boldsymbol{\nabla} u}_{s,s+r}(f)\, d\mathfrak{m} + \int_M \widehat{P}^{\boldsymbol{\nabla} u}_{s+r,t}(\phi) \cdot \Delta^{V_{s+r}}[P^{\boldsymbol{\nabla} u}_{s,s+r}(f)]\, d\mathfrak{m} \\
&= 0.
\end{aligned}$$

One may rewrite (3.2) as

$$\partial_\sigma [\widehat{P}^{\boldsymbol{\nabla} u}_{t-\sigma,t}(\phi)] = \Delta^{V_{t-\sigma}}[\widehat{P}^{\boldsymbol{\nabla} u}_{t-\sigma,t}(\phi)], \qquad \sigma \in [0,t],$$

to see that the adjoint heat semigroup solves the linearized heat equation *backward in time*. (This evolution is sometimes called the *conjugate heat semigroup*, especially in the Ricci flow theory; see for instance [Ch$_+$, Chapter 5].) Therefore we see in the same way as $P^{\boldsymbol{\nabla} u}_{s,t}$ that $\|\widehat{P}^{\boldsymbol{\nabla} u}_{t-\sigma,t}(\phi)\|_{L^2}$ is non-increasing in σ and that $\widehat{P}^{\boldsymbol{\nabla} u}_{t-\sigma,t}$ extends to a linear contraction semigroup acting on $L^2(M)$.

Remark 3.2 In general, the semigroups $P^{\boldsymbol{\nabla} u}_{s,t}$ and $\widehat{P}^{\boldsymbol{\nabla} u}_{s,t}$ depend on the choice of an auxiliary vector field $(V_t)_{t \ge 0}$. We will not discuss this issue, but carefully replace V_t with $\boldsymbol{\nabla} u_t$ as far as it is possible (with the help of Lemma 2.12).

By a well known technique based on the Bochner inequality (2.12) with $N = \infty$, we obtained in [OS3, Theorem 4.1] the L^2-*gradient estimate* of the following form.

Theorem 3.3 (L^2-gradient estimate, compact case) *Assume that $(M, F, \mathfrak{m})$ is compact and satisfies $\mathrm{Ric}_\infty \ge K$ for some $K \in \mathbb{R}$. Then, given any global solution $(u_t)_{t \ge 0}$ to the heat equation, we have*

$$F^2\big(\boldsymbol{\nabla} u_t(x)\big) \le \mathrm{e}^{-2K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F^2(\boldsymbol{\nabla} u_s)\big)(x)$$

for all $0 < s < t < \infty$ and $x \in M$.

We remark that, by Theorem 2.9, $F^2(\boldsymbol{\nabla} u_s) \in H^1(M)$ and both sides in Theorem 3.3 are Hölder continuous. Let us stress that we use the nonlinear semigroup $(u_s \to u_t)$ in the LHS, while in the RHS the linearized semigroup $P^{\boldsymbol{\nabla} u}_{s,t}$ is employed.

Remark 3.4 In the proof of [OS3, Theorem 4.1], we did not distinguish $P^{\boldsymbol{\nabla} u}_{s,t}$ and $\widehat{P}^{\boldsymbol{\nabla} u}_{s,t}$ and treated $P^{\boldsymbol{\nabla} u}_{s,t}$ as a symmetric operator. However, the proof is valid by replacing $P^{\boldsymbol{\nabla} u}_{s,t}(h)$ with $\widehat{P}^{\boldsymbol{\nabla} u}_{s,t}(h)$. See the proof of Theorem 3.7 below which is based on a similar calculation (with the sharper inequality in Proposition 3.5).

3.2 Improved Bochner inequality

We shall give an inequality improving the Bochner inequality (2.12) with $N = \infty$, that will be used to show the L^1-gradient estimate as well as the isoperimetric inequality. In the context of linear diffusion operators, such an inequality can be derived from (2.12) by a self-improvement argument (see [BGL, §C.6], and also [Sav] for an extension to $\mathrm{RCD}(K,\infty)$-spaces). Here we give a direct proof by calculations in coordinates.

Proposition 3.5 (Improved Bochner inequality) *Assume* $\mathrm{Ric}_\infty \ge K$ *for some* $K \in \mathbb{R}$*. Then we have, for any* $u \in \mathcal{C}^\infty(M)$*,*

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) - KF^2(\boldsymbol{\nabla} u) \ge d[F(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F(\boldsymbol{\nabla} u)]\big) \tag{3.4}$$

point-wise on M_u*.*

Proof. By comparing (2.12) with $N = \infty$ and (3.4), it suffices to show

$$4F^2(\boldsymbol{\nabla} u)\|\boldsymbol{\nabla}^2 u\|^2_{\mathrm{HS}(\boldsymbol{\nabla} u)} \ge d[F^2(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F^2(\boldsymbol{\nabla} u)]\big). \tag{3.5}$$

Fix $x \in M_u$ and choose local coordinates such that $g_{ij}(\boldsymbol{\nabla} u(x)) = \delta_{ij}$. We first calculate the RHS of (3.5) at x as

$$\begin{aligned}
&d[F^2(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F^2(\boldsymbol{\nabla} u)]\big) = \sum_{i=1}^n \left(\frac{\partial[F^2(\boldsymbol{\nabla} u)]}{\partial x^i}\right)^2 \\
&= \sum_{i=1}^n \left(\frac{\partial}{\partial x^i}\left[\sum_{j,k=1}^n g^*_{jk}(du)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k}\right]\right)^2 \\
&= \sum_{i=1}^n \left(2\sum_{j=1}^n \frac{\partial u}{\partial x^j}\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{j,k=1}^n \frac{\partial g^*_{jk}}{\partial x^i}(du)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k} + \sum_{j,k,l=1}^n \frac{\partial g^*_{jk}}{\partial \alpha_l}(du)\frac{\partial^2 u}{\partial x^i \partial x^l}\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k}\right)^2 \\
&= \sum_{i=1}^n \left(2\sum_{j=1}^n \frac{\partial u}{\partial x^j}\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{j,k=1}^n \frac{\partial g^*_{jk}}{\partial x^i}(du)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k}\right)^2,
\end{aligned}$$

where we used Euler's theorem (Theorem 2.2, similarly to (2.3)) in the last equality. Next we observe from (2.6) and (2.5) that, again at x,

$$\begin{aligned}
&\boldsymbol{\nabla}^2 u\left(\frac{\partial}{\partial x^j}\right) = D^{\boldsymbol{\nabla} u}_{\partial/\partial x^j}(\boldsymbol{\nabla} u) \\
&= \sum_{i=1}^n \left\{\frac{\partial}{\partial x^j}\left[\sum_{k=1}^n g^*_{ik}(du)\frac{\partial u}{\partial x^k}\right] + \sum_{k=1}^n \Gamma^i_{jk}(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^k}\right\}\frac{\partial}{\partial x^i} \\
&= \sum_{i=1}^n \left\{\frac{\partial^2 u}{\partial x^j \partial x^i} + \sum_{k=1}^n \frac{\partial g^*_{ik}}{\partial x^j}(du)\frac{\partial u}{\partial x^k} + \sum_{k=1}^n \gamma^i_{jk}(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^k} - \sum_{l=1}^n \frac{A_{ijl}(\boldsymbol{\nabla} u)}{F(\boldsymbol{\nabla} u)}G^l(\boldsymbol{\nabla} u)\right\}\frac{\partial}{\partial x^i} \\
&= \sum_{i=1}^n \left\{\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^n \left(\gamma^i_{jk} - \frac{\partial g_{ik}}{\partial x^j}\right)(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^k} - \sum_{k=1}^n \frac{A_{ijk}G^k}{F}(\boldsymbol{\nabla} u)\right\}\frac{\partial}{\partial x^i}.
\end{aligned}$$

In the last line we used

$$\frac{\partial g^*_{ik}}{\partial x^j}\big(du(x)\big) = -\frac{\partial g_{ik}}{\partial x^j}\big(\boldsymbol{\nabla} u(x)\big).$$

Hence we deduce from the Cauchy–Schwarz inequality, (2.3) and (2.4) that

$$\begin{aligned}
&F^2(\boldsymbol{\nabla} u)\|\boldsymbol{\nabla}^2 u\|^2_{\mathrm{HS}(\boldsymbol{\nabla} u)} \\
&= \sum_{j=1}^n \left(\frac{\partial u}{\partial x^j}\right)^2 \cdot \sum_{i,j=1}^n \left(\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^n \left(\gamma^i_{jk} - \frac{\partial g_{ik}}{\partial x^j}\right)(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^k} - \sum_{k=1}^n \frac{A_{ijk}G^k}{F}(\boldsymbol{\nabla} u)\right)^2 \\
&\ge \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial u}{\partial x^j}\left\{\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{k=1}^n \left(\gamma^i_{jk} - \frac{\partial g_{ik}}{\partial x^j}\right)(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^k} - \sum_{k=1}^n \frac{A_{ijk}G^k}{F}(\boldsymbol{\nabla} u)\right\}\right)^2 \\
&= \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial u}{\partial x^j}\frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{j,k=1}^n \left(\gamma^i_{jk} - \frac{\partial g_{ik}}{\partial x^j}\right)(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k}\right)^2 \\
&= \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial u}{\partial x^j}\frac{\partial u^2}{\partial x^i \partial x^j} - \frac{1}{2}\sum_{j,k=1}^n \frac{\partial g_{jk}}{\partial x^i}(\boldsymbol{\nabla} u)\frac{\partial u}{\partial x^j}\frac{\partial u}{\partial x^k}\right)^2.
\end{aligned}$$

This completes the proof of (3.5) as well as (3.4). □

The following integrated form can be shown in the same way as Theorem 2.13, we refer to [OS3, Theorem 3.6] for details.

Corollary 3.6 (Integrated form) *Assume* $\mathrm{Ric}_\infty \ge K$ *for some* $K \in \mathbb{R}$*. Given* $u \in H^2_{\mathrm{loc}}(M) \cap \mathcal{C}^1(M)$ *such that* $\boldsymbol{\Delta} u \in H^1_{\mathrm{loc}}(M)$*, we have*

$$\begin{aligned}
&-\int_M d\phi\left(\nabla^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right]\right) d\mathfrak{m} \\
&\ge \int_M \phi\Big\{d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) + KF^2(\boldsymbol{\nabla} u) + d[F(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F(\boldsymbol{\nabla} u)]\big)\Big\}\, d\mathfrak{m}
\end{aligned}$$

for all bounded nonnegative functions $\phi \in H^1_c(M) \cap L^\infty(M)$*.*

3.3 L^1-gradient estimate

The improved Bochner inequality (3.4) yields the following L^1*-gradient estimate*, under a technical (likely redundant) assumption that $d[F(\boldsymbol{\nabla} u_t)](\nabla^{\boldsymbol{\nabla} u_t}[F(\boldsymbol{\nabla} u_t)]) \in L^1(M)$ for all $t > 0$, which holds in the compact case thanks to the H^2_{loc}-regularity (recall Theorem 2.9).

Theorem 3.7 (L^1-gradient estimate) *Let* $(M, F, \mathfrak{m})$ *be complete and satisfy* $\mathrm{Ric}_\infty \ge K$*,* $\mathsf{C}_F < \infty$ *and* $\mathsf{S}_F < \infty$*, and* $(u_t)_{t \ge 0}$ *be a global solution to the heat equation with* $u_0 \in \mathcal{C}^\infty_c(M)$*. We further assume that*

$$d[F(\boldsymbol{\nabla} u_t)]\big(\nabla^{\boldsymbol{\nabla} u_t}[F(\boldsymbol{\nabla} u_t)]\big) \in L^1(M) \tag{3.6}$$

for all $t > 0$*. Then we have*

$$F\big(\boldsymbol{\nabla} u_t(x)\big) \le \mathrm{e}^{-K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F(\boldsymbol{\nabla} u_s)\big)(x)$$

for all $0 \le s < t < \infty$ *and* $x \in M$*.*

Proof. Notice first that $F(\boldsymbol{\nabla} u_0) \in H_0^1(M) \cap L^\infty(M)$ since $u_0 \in \mathcal{C}_c^\infty(M)$. Fix arbitrary $\varepsilon > 0$ and let us consider the function

$$\xi_\sigma := \sqrt{\mathrm{e}^{-2K\sigma} F^2(\boldsymbol{\nabla} u_{t-\sigma}) + \varepsilon}, \qquad 0 \le \sigma \le t-s.$$

Note from the proof of [OS3, Theorem 4.1] that

$$\frac{\partial}{\partial\sigma}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right] = -\frac{\partial}{\partial t}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right] = -d(\boldsymbol{\Delta} u_{t-\sigma})(\boldsymbol{\nabla} u_{t-\sigma}). \tag{3.7}$$

Hence we have, on one hand,

$$\partial_\sigma \xi_\sigma = -\frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma}\big\{K F^2(\boldsymbol{\nabla} u_{t-\sigma}) + d(\boldsymbol{\Delta} u_{t-\sigma})(\boldsymbol{\nabla} u_{t-\sigma})\big\}.$$

On the other hand, for any nonnegative function $\phi \in \mathcal{C}_c^\infty(M)$, we observe

$$\begin{aligned}
&\int_M d\phi(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}} \xi_\sigma)\, d\mathfrak{m} = \int_M \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma} d\phi\left(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\right) d\mathfrak{m} \\
&= \int_M \left(d\left(\phi \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma}\right) + \phi \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma^2} d\xi_\sigma\right)\left(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\right) d\mathfrak{m} \\
&= \int_M d\left(\phi \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma}\right)\left(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\right) d\mathfrak{m} \\
&\quad + \int_M \phi \frac{\mathrm{e}^{-4K\sigma}}{\xi_\sigma^3} d\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\left(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\right) d\mathfrak{m} \\
&\le \int_M d\left(\phi \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma}\right)\left(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\left[\frac{F^2(\boldsymbol{\nabla} u_{t-\sigma})}{2}\right]\right) d\mathfrak{m} \\
&\quad + \int_M \phi \frac{\mathrm{e}^{-2K\sigma}}{\xi_\sigma} d[F(\boldsymbol{\nabla} u_{t-\sigma})]\big(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}[F(\boldsymbol{\nabla} u_{t-\sigma})]\big)\, d\mathfrak{m},
\end{aligned}$$

where we used $F^2(\boldsymbol{\nabla} u_{t-\sigma}) \le \mathrm{e}^{2K\sigma}\xi_\sigma^2$ in the last inequality. Therefore the improved Bochner inequality (Corollary 3.6) shows that

$$\Delta^{\boldsymbol{\nabla} u_{t-\sigma}} \xi_\sigma + \partial_\sigma \xi_\sigma \ge 0 \tag{3.8}$$

in the weak sense. Notice that the test function ϕ can be in fact taken from $H_0^1(M) \cap L^\infty(M)$ thanks to the hypothesis (3.6) and $\xi_\sigma \ge \sqrt{\varepsilon}$.

For a nonnegative function $\varphi \in \mathcal{C}_c^\infty(M)$ and $\sigma \in (0, t-s)$, set

$$\Phi(\sigma) := \int_M \varphi \cdot P_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\xi_\sigma)\, d\mathfrak{m} = \int_M \widehat{P}_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\varphi) \cdot \xi_\sigma\, d\mathfrak{m}.$$

We deduce from (3.2) and (2.10) that

$$\begin{aligned}
\Phi'(\sigma) &= \int_M \widehat{P}_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\varphi) \cdot \partial_\sigma \xi_\sigma\, d\mathfrak{m} - \int_M d\xi_\sigma\big(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}\big[\widehat{P}_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\varphi)\big]\big)\, d\mathfrak{m} \\
&= \int_M \widehat{P}_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\varphi) \cdot \partial_\sigma \xi_\sigma\, d\mathfrak{m} - \int_M d[\widehat{P}_{t-\sigma,t}^{\boldsymbol{\nabla} u}(\varphi)](\nabla^{\boldsymbol{\nabla} u_{t-\sigma}} \xi_\sigma)\, d\mathfrak{m}.
\end{aligned}$$

Therefore we can apply (3.8) with the test function $\widehat{P}^{\boldsymbol{\nabla} u}_{t-\sigma,t}(\varphi)$ (thanks to Proposition 3.1) to obtain $\Phi'(\sigma) \ge 0$. This implies

$$\int_M \varphi \cdot \xi_0 \, d\mathfrak{m} \le \int_M \varphi \cdot P^{\boldsymbol{\nabla} u}_{s,t}(\xi_{t-s}) \, d\mathfrak{m}.$$

By the arbitrariness of φ and ε, we have

$$F(\boldsymbol{\nabla} u_t) \le \mathrm{e}^{-K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F(\boldsymbol{\nabla} u_s)\big)$$

almost everywhere. Since both sides are Hölder continuous (Proposition 3.1(ii)), this completes the proof. □

It is a standard fact that the L^1-gradient estimate implies the L^2-bound.

Corollary 3.8 **(L^2-gradient estimate, noncompact case)** *Let $(M, F, \mathfrak{m})$ be complete and satisfy $\mathrm{Ric}_\infty \ge K$, $\mathsf{C}_F < \infty$ and $\mathsf{S}_F < \infty$, and $(u_t)_{t \ge 0}$ be a global solution to the heat equation with $u_0 \in \mathcal{C}_c^\infty(M)$ and satisfying (3.6) for all $t > 0$. Then we have*

$$F^2\big(\boldsymbol{\nabla} u_t(x)\big) \le \mathrm{e}^{-2K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F^2(\boldsymbol{\nabla} u_s)\big)(x)$$

for all $0 \le s < t < \infty$ and $x \in M$.

Proof. This is a consequence of a kind of Jensen's inequality:

$$P^{\boldsymbol{\nabla} u}_{s,t}(f)^2 \le P^{\boldsymbol{\nabla} u}_{s,t}(f^2)$$

for $f \in L^2(M) \cap L^\infty(M)$. For $\psi \in \mathcal{C}_c^\infty(M)$ with $0 \le \psi \le 1$ and $r \in \mathbb{R}$, we have

$$\begin{aligned} 0 \le P^{\boldsymbol{\nabla} u}_{s,t}\big((rf + \psi)^2\big) &= r^2 P^{\boldsymbol{\nabla} u}_{s,t}(f^2) + 2r P^{\boldsymbol{\nabla} u}_{s,t}(f\psi) + P^{\boldsymbol{\nabla} u}_{s,t}(\psi^2) \\ &\le r^2 P^{\boldsymbol{\nabla} u}_{s,t}(f^2) + 2r P^{\boldsymbol{\nabla} u}_{s,t}(f\psi) + 1. \end{aligned}$$

Letting $f\psi \to f$ in $L^2(M)$, we find $r^2 P^{\boldsymbol{\nabla} u}_{s,t}(f^2) + 2r P^{\boldsymbol{\nabla} u}_{s,t}(f) + 1 \ge 0$ for all $r \in \mathbb{R}$. Hence $P^{\boldsymbol{\nabla} u}_{s,t}(f)^2 - P^{\boldsymbol{\nabla} u}_{s,t}(f^2) \le 0$ as desired. □

3.4 On the hypothesis (3.6)

The hypothesis (3.6) seems redundant and indeed unnecessary for weighted Riemannian manifolds and RCD-spaces. Especially, when $K > 0$, the Gaussian decay of the measure ([St1, Theorem 4.26]) could imply (3.6). Let us give some more comments on (3.6).

3.4.1 Weighted Riemannian case

We essentially followed the proof of [BGL, Theorem 3.2.4] in Theorem 3.7. Then we have

$$d\xi_\sigma(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}} \xi_\sigma) \le \mathrm{e}^{-2K\sigma} d[F(\boldsymbol{\nabla} u_{t-\sigma})]\big(\nabla^{\boldsymbol{\nabla} u_{t-\sigma}}[F(\boldsymbol{\nabla} u_{t-\sigma})]\big),$$

and the improved Bochner inequality (Proposition 3.5) implies

$$\int_M d[F(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F(\boldsymbol{\nabla} u)]\big) \, d\mathfrak{m} \le \|\boldsymbol{\Delta} u\|_{L^2}^2 + |K| \cdot \|u\|_{L^2} \|\boldsymbol{\Delta} u\|_{L^2}$$

for $u \in \mathcal{C}_c^\infty(M)$. Now in [BGL], for a linear operator L, we make use of the density of $\mathcal{A}_0 = \mathcal{C}_c^\infty(M)$ in the domain $\mathcal{D}(\mathrm{L})$ with respect to the norm

$$\|f\|_{\mathcal{D}(\mathrm{L})} := \left(\|f\|_{L^2}^2 + \|\mathrm{L}f\|_{L^2}^2\right)^{1/2}$$

to extend the above estimate to $\mathcal{D}(\mathrm{L})$. This density is a consequence of the *hypo-ellipticity* (see [BGL, Proposition 3.2.1]), which is defined by the property that any solution to $\mathrm{L}^* f = \lambda f$ is smooth (see also [BGL, Definition 3.3.8], typically $\mathcal{A} = \mathcal{C}^\infty(M)$). This is not the case for operators with nonsmooth coefficients, thereby it is unclear if we can apply this method in the Finsler case (to the linearized Laplacian $\Delta^{\boldsymbol{\nabla} u}$).

3.4.2 RCD-case

In $\mathrm{RCD}(K,\infty)$-spaces, we obtain the *Wasserstein contraction estimate* of heat flow by the convexity of the relative entropy, and then the gradient estimates follow by the duality argument. Moreover, we can obtain the Bochner inequality by differentiating the gradient estimate (see [AGS1, AGS2, GKO, Sav] for details).

This method could avoid the use of the functional analytic argument involving $\mathcal{A}_0$ and $\mathcal{A}$, and what is important and interesting here is that the Bochner inequality derived from the gradient estimate is of the form:

$$\int \Delta\phi \cdot \frac{|\nabla u|^2}{2}\, d\mathfrak{m} \geq \int \phi\left\{d(\Delta u)(\nabla u) + K|\nabla u|^2\right\} d\mathfrak{m}$$

for $u \in \mathcal{D}(\Delta)$ with $\Delta u \in H^1$ and $\phi \in \mathcal{D}(\Delta) \cap L^\infty$ with $\Delta\phi \in L^\infty$. In the LHS, what we have directly from the point-wise Bochner inequality is

$$\int \phi \cdot \Delta\left[\frac{|\nabla u|^2}{2}\right] d\mathfrak{m},$$

and modifying this into the above LHS requires an approximation of ϕ by functions ϕ_k in $\mathcal{C}_c^\infty$ such that $\Delta\phi_k \to \Delta\phi$, namely the density of $\mathcal{C}_c^\infty$ in the $\mathcal{D}(\Delta)$-norm as in the approach of [BGL].

In the Finsler case, we know that the Wasserstein contraction fails (see Remark 3.10 below). Nonetheless, if one can show the Bochner inequality in the above form as well as $F(\boldsymbol{\nabla} u) \in L^\infty(M)$, then it follows from the argument along [Sav, Lemma 3.2] that $d[F(\boldsymbol{\nabla} u)](\nabla^{\boldsymbol{\nabla} u} F(\boldsymbol{\nabla} u)) \in L^1(M)$ and we obtain the gradient estimates.

3.5 Characterizations of lower Ricci curvature bounds

We close the section with several characterizations of the lower Ricci curvature bound $\mathrm{Ric}_\infty \geq K$.

Theorem 3.9 (Characterizations of $\mathrm{Ric}_\infty \geq K$) *Let $(M,F,\mathfrak{m})$ be complete and satisfy $\mathsf{C}_F < \infty$ and $\mathsf{S}_F < \infty$. We assume that (3.6) holds for all solutions $(u_t)_{t\geq 0}$ to the heat equation with $u_0 \in \mathcal{C}_c^\infty(M)$. Then, for each $K \in \mathbb{R}$, the following are equivalent:*

(I) $\mathrm{Ric}_\infty \geq K$.

(II) *The Bochner inequality*

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) \geq KF^2(\boldsymbol{\nabla} u)$$

holds on M_u for all $u \in \mathcal{C}^\infty(M)$.

(III) *The improved Bochner inequality*

$$\Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right] - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) - KF^2(\boldsymbol{\nabla} u) \geq d[F(\boldsymbol{\nabla} u)]\big(\nabla^{\boldsymbol{\nabla} u}[F(\boldsymbol{\nabla} u)]\big)$$

holds on M_u for all $u \in \mathcal{C}^\infty(M)$.

(IV) *The L^2-gradient estimate*

$$F^2(\boldsymbol{\nabla} u_t) \leq \mathrm{e}^{-2K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F^2(\boldsymbol{\nabla} u_s)\big), \qquad 0 \leq s < t < \infty,$$

holds for all global solutions $(u_t)_{t\geq 0}$ to the heat equation with $u_0 \in \mathcal{C}^\infty_c(M)$.

(V) *The L^1-gradient estimate*

$$F(\boldsymbol{\nabla} u_t) \leq \mathrm{e}^{-K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F(\boldsymbol{\nabla} u_s)\big), \qquad 0 \leq s < t < \infty,$$

holds for all global solutions $(u_t)_{t\geq 0}$ to the heat equation with $u_0 \in \mathcal{C}^\infty_c(M)$.

Proof. We have shown (I) $\Rightarrow$ (III) in Proposition 3.5, (III) $\Rightarrow$ (V) in Theorem 3.7, and (V) $\Rightarrow$ (IV) in Corollary 3.8. One can deduce (IV) $\Rightarrow$ (II) from the proof of [OS3, Theorem 4.1] or by differentiating $F^2(\boldsymbol{\nabla} u_t) \leq \mathrm{e}^{-2Kt} P^{\boldsymbol{\nabla} u}_{0,t}(F^2(\boldsymbol{\nabla} u_0))$ at $t = 0$ (recall (3.7), see also [GKO]). Let us finally prove (II) $\Rightarrow$ (I). Given $v_0 \in T_{x_0}M \setminus 0$, fix local coordinates $(x^i)_{i=1}^n$ around x_0 with $g_{ij}(v_0) = \delta_{ij}$ and $x^i(x_0) = 0$ for all i. Consider the function

$$u := \sum_{i=1}^n v_0^i x^i + \frac{1}{2}\sum_{i,j,k=1}^n \Gamma^k_{ij}(v_0) v_0^k x^i x^j$$

on a neighborhood of x_0, and observe that $\boldsymbol{\nabla} u(x_0) = v_0$ as well as $(\boldsymbol{\nabla}^2 u)|_{T_{x_0}M} = 0$ (see [OS3, Lemma 2.3] for the precise expression in coordinates of $\boldsymbol{\nabla}^2 u$). Then the Bochner–Weitzenböck formula (2.11) and (II) imply

$$\mathrm{Ric}_\infty(v_0) = \Delta^{\boldsymbol{\nabla} u}\left[\frac{F^2(\boldsymbol{\nabla} u)}{2}\right](x_0) - d(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u)(x_0) \geq KF^2(v_0).$$

This completes the proof. □

Remark 3.10 (The lack of contraction) In the Riemannian context, lower Ricci curvature bounds are also equivalent to contraction estimates of heat flow with respect to the Wasserstein distance (we refer to [vRS] for the Riemannian case, and [EKS] for the case of RCD-spaces). More generally, for linear semigroups, gradient estimates are directly equivalent to the corresponding contraction properties (see [Ku]). In our Finsler setting, however, the lack of the *commutativity* (see [OP]) prevents such a contraction estimate, at least in the same form (see [OS2] for details).

Remark 3.11 (Similarities to (super) Ricci flow theory) The methods in this section have connections with the Ricci flow theory. Ricci flow provides time-dependent Riemannian metrics obeying a kind of heat equation on the space of Riemannian metrics, while we considered the time-dependent (singular) Riemannian structures $g_{\boldsymbol{\nabla} u}$ for u solving the heat equation. More precisely, what corresponds to our lower Ricci curvature bound is *super Ricci flow* (super-solutions to the Ricci flow equation). We refer to [MT] for an inspiring work on a characterization of super Ricci flow in terms of the contraction of heat flow, and to [St3] for a recent investigation of super Ricci flow on time-dependent metric measure spaces including various characterizations related to Theorem 3.9. Then, again, what is missing in our Finsler setting is the contraction property, for which the Riemannian nature of the space is necessary.

4 Bakry–Ledoux's isoperimetric inequality

This section is devoted to the isoperimetric inequality, as a geometric application of the improved Bochner inequality (Proposition 3.5). We will assume $\mathrm{Ric}_\infty \ge K > 0$, then $\mathfrak{m}(M) < \infty$ holds (see [St1, Theorem 4.26]) and hence we can normalize $\mathfrak{m}$ as $\mathfrak{m}(M) = 1$ without changing Ric_∞ ($c\mathfrak{m}$ with $c > 0$ gives the same weighted Ricci curvature as $\mathfrak{m}$).

For a Borel set $A \subset M$, define the *Minkowski exterior boundary measure* as

$$\mathfrak{m}^+(A) := \liminf_{\varepsilon \downarrow 0} \frac{\mathfrak{m}(B^+(A,\varepsilon)) - \mathfrak{m}(A)}{\varepsilon},$$

where $B^+(A,\varepsilon) := \{y \in M \mid \inf_{x \in A} d(x,y) < \varepsilon\}$ is the forward ε-neighborhood of A. Then the (forward) *isoperimetric profile* $\mathcal{I}_{(M,F,\mathfrak{m})} : [0,1] \longrightarrow [0,\infty)$ of $(M,F,\mathfrak{m})$ is defined by

$$\mathcal{I}_{(M,F,\mathfrak{m})}(\theta) := \inf\{\mathfrak{m}^+(A) \mid A \subset M : \text{Borel set with } \mathfrak{m}(A) = \theta\}.$$

Clearly $\mathcal{I}_{(M,F,\mathfrak{m})}(0) = \mathcal{I}_{(M,F,\mathfrak{m})}(1) = 0$. The following is our main result (stated as Theorem in the introduction).

Theorem 4.1 (Bakry–Ledoux's isoperimetric inequality) *Let (M,F) be complete and satisfy $\mathrm{Ric}_\infty \ge K > 0$, $\mathfrak{m}(M) = 1$, $\mathsf{C}_F < \infty$ and $\mathsf{S}_F < \infty$. We assume that (3.6) holds for all solutions $(u_t)_{t \ge 0}$ to the heat equation with $u_0 \in \mathcal{C}_c^\infty(M)$. Then we have*

$$\mathcal{I}_{(M,F,\mathfrak{m})}(\theta) \ge \mathcal{I}_K(\theta) \tag{4.1}$$

for all $\theta \in [0,1]$, where

$$\mathcal{I}_K(\theta) := \sqrt{\frac{K}{2\pi}} \mathrm{e}^{-Kc^2(\theta)/2} \qquad \text{with } \theta = \int_{-\infty}^{c(\theta)} \sqrt{\frac{K}{2\pi}} \mathrm{e}^{-Ka^2/2} \, da.$$

Recall that, under $\mathsf{C}_F < \infty$ or $\mathsf{S}_F < \infty$, the forward completeness is equivalent to the backward completeness by Lemma 2.4. In the Riemannian case, the inequality (4.1) is due to Bakry and Ledoux [BL] (see also [BGL, §8.5.2]) and can be regarded as the dimension-free version of *Lévy–Gromov's isoperimetric inequality* (see [Lé1, Lé2, Gr]). Lévy–Gromov's classical isoperimetric inequality asserts that the isoperimetric profile of an n-dimensional Riemannian manifold (M,g) with $\mathrm{Ric} \ge n-1$ is bounded below by the profile of the unit sphere $\mathbb{S}^n$ (both spaces are equipped with the normalized volume measures). In (4.1), the role of the unit sphere is played by the real line $\mathbb{R}$ equipped with the Gaussian measure $\sqrt{K/2\pi}\,\mathrm{e}^{-Kx^2/2}\,dx$, thereby (4.1) is also called the *Gaussian isoperimetric inequality*.

In [Oh8], generalizing Cavalletti and Mondino's localization technique in [CM], we showed the slightly weaker inequality (recall the introduction)

$$\mathcal{I}_{(M,F,\mathfrak{m})}(\theta) \ge \Lambda_F^{-1} \cdot \mathcal{I}_K(\theta)$$

under the finite reversibility $\Lambda_F < \infty$ (but without $\mathsf{C}_F < \infty$ nor $\mathsf{S}_F < \infty$). In fact we have treated in [Oh8] the general curvature-dimension-diameter bound $\mathrm{Ric}_N \ge K$ and $\operatorname{diam} M \le D$ (in accordance with [Mi1]). Theorem 4.1 sharpens the estimate in [Oh8] in the special case of $N = D = \infty$ and $K > 0$.

4.1 Ergodicity

We begin with some properties induced from our hypothesis $\mathrm{Ric}_\infty \ge K > 0$.

Lemma 4.2 (Global Poincaré inequality) *Suppose that $(M,F,\mathfrak{m})$ is forward or backward complete, $\mathrm{Ric}_\infty \ge K > 0$ and $\mathfrak{m}(M) = 1$. Then we have, for any locally Lipschitz function $f \in H_0^1(M)$,*

$$\int_M f^2\,d\mathfrak{m} - \left(\int_M f\,d\mathfrak{m}\right)^2 \le \frac{1}{K}\int_M F^*(df)^2\,d\mathfrak{m}. \tag{4.2}$$

Proof. It is well known that the curvature bound $\mathrm{Ric}_\infty \ge K$ (or $\mathrm{CD}(K,\infty)$) implies the *log-Sobolev inequality*,

$$\int_M \rho\log\rho\,d\mathfrak{m} \le \frac{1}{2K}\int_M \frac{F^*(d\rho)^2}{\rho}\,d\mathfrak{m} \tag{4.3}$$

for nonnegative locally Lipschitz functions ρ with $\int_M \rho\,d\mathfrak{m} = 1$, and that (4.2) follows from (4.3) (see [OV, LV, Vi, Oh2]). Here we explain the latter step for thoroughness.

By truncation, let us assume that f is bounded. Since

$$\int_M f^2\,d\mathfrak{m} - \left(\int_M f\,d\mathfrak{m}\right)^2 = \int_M \left(f - \int_M f\,d\mathfrak{m}\right)^2 d\mathfrak{m},$$

we can further assume that $\int_M f\,d\mathfrak{m} = 0$. There is nothing to prove if $f \equiv 0$, thereby assume $\|f\|_{L^\infty} > 0$. For $\varepsilon \in \mathbb{R}$ with $|\varepsilon| < \|f\|_{L^\infty}^{-1}$, we consider the probability measure $(1+\varepsilon f)\mathfrak{m}$. Then the log-Sobolev inequality for $\rho_\varepsilon := 1 + \varepsilon f$ under $\mathrm{Ric}_\infty \ge K$ implies

$$\int_M (1+\varepsilon f)\log(1+\varepsilon f)\,d\mathfrak{m} \le \frac{1}{2K}\int_M \frac{\varepsilon^2 F^*(df)^2}{1+\varepsilon f}\,d\mathfrak{m}.$$

Expanding the LHS at $\varepsilon = 0$ yields

$$\int_M \left\{\varepsilon f + \frac{1}{2}(\varepsilon f)^2 + O(\varepsilon^3)\right\} d\mathfrak{m} = \frac{\varepsilon^2}{2}\int_M f^2\,d\mathfrak{m} + O(\varepsilon^3),$$

where $O(\varepsilon^3)$ in the LHS is uniform in M thanks to the boundedness of f. Hence we have

$$\frac{\varepsilon^2}{2}\int_M f^2\,d\mathfrak{m} \le \frac{1}{1-\varepsilon\|f\|_{L^\infty}}\frac{\varepsilon^2}{2K}\int_M F^*(df)^2\,d\mathfrak{m} + O(\varepsilon^3).$$

Dividing both sides by ε^2 and letting $\varepsilon \to 0$ implies (4.2). $\square$

The LHS of (4.2) is the *variance* of f:

$$\mathrm{Var}_{\mathfrak{m}}(f) := \int_M f^2\,d\mathfrak{m} - \left(\int_M f\,d\mathfrak{m}\right)^2.$$

We next show that the Poincaré inequality (4.2) yields the exponential decay of the variance and a kind of *ergodicity* along heat flow (similarly to [BGL, §4.2]), which is one of the key ingredients in the proof of Theorem 4.1 (see the proof of Corollary 4.5). Given a global solution $(u_t)_{t\ge 0}$ to the heat equation, since the finiteness of the total mass together with $\Lambda_F < \infty$ and the completeness implies $1 \in H_0^1(M)$, we observe the mass conservation:

$$\int_M P_{s,t}^{\nabla u}(f)\,d\mathfrak{m} = \int_M f\,d\mathfrak{m} \tag{4.4}$$

for any $f \in H_0^1(M)$ and $0 \le s < t < \infty$.

Proposition 4.3 (Variance decay and ergodicity) *Assume that $(M, F, \mathfrak{m})$ is complete and satisfies $\mathsf{C}_F < \infty$, $\mathsf{S}_F < \infty$, $\mathrm{Ric}_\infty \ge K > 0$ and $\mathfrak{m}(M) = 1$. Then we have, given any global solution $(u_t)_{t\ge 0}$ to the heat equation and $f \in H_0^1(M)$,*

$$\mathrm{Var}_{\mathfrak{m}}\big(P_{s,t}^{\nabla u}(f)\big) \le \mathrm{e}^{-2K(t-s)/\mathsf{S}_F}\,\mathrm{Var}_{\mathfrak{m}}(f)$$

for all $0 \le s < t < \infty$. In particular, $P_{s,t}^{\nabla u}(f)$ converges to the constant function $\int_M f\,d\mathfrak{m}$ in $L^2(M)$ as $t \to \infty$.

Proof. Put $f_t := P_{s,t}^{\nabla u}(f)$, then $\int_M f_t\,d\mathfrak{m} = \int_M f\,d\mathfrak{m}$ holds by (4.4). It follows from Lemmas 2.3, 4.2 that

$$\begin{aligned}\frac{d}{dt}\big[\mathrm{Var}_{\mathfrak{m}}(f_t)\big] &= -2\int_M df_t(\nabla^{V_t} f_t)\,d\mathfrak{m} = -2\int_M g^*_{\mathcal{L}(V_t)}(df_t, df_t)\,d\mathfrak{m}\\ &\le -\frac{2}{\mathsf{S}_F}\int_M F^*(df_t)^2\,d\mathfrak{m} \le -\frac{2K}{\mathsf{S}_F}\mathrm{Var}_{\mathfrak{m}}(f_t).\end{aligned}$$

Hence $\mathrm{e}^{2Kt/\mathsf{S}_F}\,\mathrm{Var}_{\mathfrak{m}}(f_t)$ is non-increasing in t, this completes the proof of the first assertion. Then the second assertion is straightforward since

$$\mathrm{Var}_{\mathfrak{m}}(f_t) = \int_M \left(f_t - \int_M f\,d\mathfrak{m}\right)^2 d\mathfrak{m} \to 0 \quad (t \to \infty).$$

$\square$

4.2 Key estimate

We next prove a key estimate which would have further applications (see [BL]). Define

$$\varphi(c) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} \mathrm{e}^{-b^2/2}\, db, \quad c \in \mathbb{R},$$

$$\mathscr{N}(\theta) := \varphi'\big(\varphi^{-1}(\theta)\big) = \frac{\mathrm{e}^{-\varphi^{-1}(\theta)^2/2}}{\sqrt{2\pi}}, \quad \theta \in (0,1).$$

We set also $\mathscr{N}(0) = \mathscr{N}(1) := 0$. Observe that $\mathscr{N}' = -\varphi^{-1}$ and $\mathscr{N}'' = -1/\mathscr{N}$ on $(0,1)$.

Theorem 4.4 *Assume that $(M,F,\mathfrak{m})$ is complete and satisfies $\mathrm{Ric}_\infty \ge K$ for some $K \in \mathbb{R}$, $\mathsf{C}_F < \infty$, $\mathsf{S}_F < \infty$ and $\mathfrak{m}(M) < \infty$. Then, given a global solution $(u_t)_{t\ge 0}$ to the heat equation with $u_0 \in \mathcal{C}_c^\infty(M)$, $0 \le u_0 \le 1$ and satisfying (3.6), we have*

$$\sqrt{\mathscr{N}^2(u_t) + \alpha F^2(\boldsymbol{\nabla} u_t)} \le P_{0,t}^{\boldsymbol{\nabla} u}\Big(\sqrt{\mathscr{N}^2(u_0) + c_\alpha(t) F^2(\boldsymbol{\nabla} u_0)}\Big) \tag{4.5}$$

on M for all $\alpha \ge 0$ and $t > 0$, where

$$c_\alpha(t) := \frac{1-\mathrm{e}^{-2Kt}}{K} + \alpha \mathrm{e}^{-2Kt} > 0$$

and $c_\alpha(t) := 2t + \alpha$ when $K = 0$.

For simplicity, we suppressed the dependence of c_α on K.

Proof. By replacing u_0 with $(1-2\varepsilon)u_0 + \varepsilon$, we can assume $\varepsilon \le u_0 \le 1-\varepsilon$ for some $\varepsilon > 0$, and then we have $\varepsilon \le u_t \le 1-\varepsilon$ for all $t > 0$ (recall (2.9)). Fix $t > 0$ and put

$$\zeta_s := \sqrt{\mathscr{N}^2(u_s) + c_\alpha(t-s) F^2(\boldsymbol{\nabla} u_s)}, \qquad 0 \le s \le t$$

(compare this function with ξ_σ in the proof of Theorem 3.7). Then (4.5) is written as $\zeta_t \le P_{0,t}^{\boldsymbol{\nabla} u}(\zeta_0)$ and it suffices to show $\partial_s[P_{s,t}^{\boldsymbol{\nabla} u}(\zeta_s)] \le 0$ in the weak sense. Observe from (3.3) and (3.2) that, for any nonnegative $\phi \in \mathcal{C}_c^\infty((0,t)\times M)$,

$$\begin{aligned}
\int_0^t \int_M \partial_s \phi_s \cdot P_{s,t}^{\boldsymbol{\nabla} u}(\zeta_s)\, d\mathfrak{m}\, ds &= \int_0^t \int_M \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\partial_s \phi_s) \cdot \zeta_s\, d\mathfrak{m}\, ds \\
&= \int_0^t \int_M \big\{\partial_s[\widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_s)] + \Delta^{V_s}[\widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_s)]\big\} \cdot \zeta_s\, d\mathfrak{m}\, ds \\
&= \int_0^t \int_M \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_s) \cdot (\Delta^{\boldsymbol{\nabla} u_s} \zeta_s - \partial_s \zeta_s)\, d\mathfrak{m}\, ds,
\end{aligned} \tag{4.6}$$

where in the second equality we deduce from the linearity of $\widehat{P}_{s,t}^{\boldsymbol{\nabla} u}$ that

$$\begin{aligned}
&\int_M \partial_s[\widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_s)] \cdot \zeta_s\, d\mathfrak{m} \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_M \big\{\widehat{P}_{s+\varepsilon,t}^{\boldsymbol{\nabla} u}(\phi_{s+\varepsilon}) - \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_{s+\varepsilon})\big\} \cdot \zeta_s\, d\mathfrak{m} + \int_M \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\partial_s \phi_s) \cdot \zeta_s\, d\mathfrak{m} \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_s^{s+\varepsilon} \int_M d\zeta_s\big(\nabla^{\boldsymbol{\nabla} u_{s+r}}[\widehat{P}_{s+r,t}^{\boldsymbol{\nabla} u}(\phi_{s+\varepsilon})]\big)\, d\mathfrak{m}\, dr + \int_M \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\partial_s \phi_s) \cdot \zeta_s\, d\mathfrak{m} \\
&= \int_M d\zeta_s\big(\nabla^{\boldsymbol{\nabla} u_s}[\widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\phi_s)]\big)\, d\mathfrak{m}\, dr + \int_M \widehat{P}_{s,t}^{\boldsymbol{\nabla} u}(\partial_s \phi_s) \cdot \zeta_s\, d\mathfrak{m}
\end{aligned}$$

for almost every s. We shall show that the RHS of (4.6) is nonnegative.

We first calculate by using (3.7) and $c_\alpha' = 2(1 - Kc_\alpha)$ as

$$\partial_s \zeta_s = \frac{1}{\zeta_s}\Big\{ \mathscr{N}(u_s)\mathscr{N}'(u_s)\boldsymbol{\Delta} u_s + \big(Kc_\alpha(t-s) - 1\big)F^2(\boldsymbol{\nabla} u_s) + c_\alpha(t-s)d(\boldsymbol{\Delta} u_s)(\boldsymbol{\nabla} u_s)\Big\}.$$

Next, we have

$$\nabla^{\boldsymbol{\nabla} u_s}\zeta_s = \frac{1}{\zeta_s}\bigg\{ \mathscr{N}(u_s)\mathscr{N}'(u_s)\boldsymbol{\nabla} u_s + \frac{c_\alpha(t-s)}{2}\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\bigg\}.$$

Hence

$$\begin{aligned}
\Delta^{\boldsymbol{\nabla} u_s}\zeta_s &= \frac{\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s}\boldsymbol{\Delta} u_s + \frac{\mathscr{N}'(u_s)^2 - 1}{\zeta_s}F^2(\boldsymbol{\nabla} u_s) - \frac{\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s^2}d\zeta_s(\boldsymbol{\nabla} u_s) \\
&\quad + \frac{c_\alpha(t-s)}{2\zeta_s}\Delta^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)] - \frac{c_\alpha(t-s)}{2\zeta_s^2}d\zeta_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big),
\end{aligned}$$

where we used $\mathscr{N}'' = -1/\mathscr{N}$ and $\Delta^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]$ is understood in the weak sense.

Now we apply the improved Bochner inequality (Corollary 3.6) to obtain

$$\begin{aligned}
\Delta^{\boldsymbol{\nabla} u_s}\zeta_s - \partial_s\zeta_s &= \frac{\mathscr{N}'(u_s)^2 - Kc_\alpha(t-s)}{\zeta_s}F^2(\boldsymbol{\nabla} u_s) \\
&\quad + \frac{c_\alpha(t-s)}{\zeta_s}\bigg\{\Delta^{\boldsymbol{\nabla} u_s}\bigg[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\bigg] - d(\boldsymbol{\Delta} u_s)(\boldsymbol{\nabla} u_s)\bigg\} \\
&\quad - \frac{\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s^2}d\zeta_s(\boldsymbol{\nabla} u_s) - \frac{c_\alpha(t-s)}{2\zeta_s^2}d\zeta_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big) \\
&\ge \frac{\mathscr{N}'(u_s)^2}{\zeta_s}F^2(\boldsymbol{\nabla} u_s) + \frac{c_\alpha(t-s)}{\zeta_s F^2(\boldsymbol{\nabla} u_s)}d\bigg[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\bigg]\bigg(\nabla^{\boldsymbol{\nabla} u_s}\bigg[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\bigg]\bigg) \\
&\quad - \frac{\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s^2}d\zeta_s(\boldsymbol{\nabla} u_s) - \frac{c_\alpha(t-s)}{2\zeta_s^2}d\zeta_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)
\end{aligned}$$

in the weak sense. Substituting

$$d\zeta_s = \frac{1}{\zeta_s}\bigg\{\mathscr{N}(u_s)\mathscr{N}'(u_s)du_s + \frac{c_\alpha(t-s)}{2}d[F^2(\boldsymbol{\nabla} u_s)]\bigg\}$$

and recalling (2.10), we obtain

$$
\begin{aligned}
\Delta^{\boldsymbol{\nabla} u_s}\zeta_s - \partial_s\zeta_s \geq{}& \frac{\zeta_s^2\mathscr{N}'(u_s)^2 - \mathscr{N}^2(u_s)\mathscr{N}'(u_s)^2}{\zeta_s^3}F^2(\boldsymbol{\nabla} u_s)\\
&- \frac{c_\alpha(t-s)\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s^3}du_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)\\
&+ \frac{c_\alpha(t-s)}{\zeta_s^3}\left\{\frac{\zeta_s^2}{F^2(\boldsymbol{\nabla} u_s)} - c_\alpha(t-s)\right\}d\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\left(\nabla^{\boldsymbol{\nabla} u_s}\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\right)\\
={}& \frac{c_\alpha(t-s)\mathscr{N}'(u_s)^2}{\zeta_s^3}F^4(\boldsymbol{\nabla} u_s)\\
&- \frac{c_\alpha(t-s)\mathscr{N}(u_s)\mathscr{N}'(u_s)}{\zeta_s^3}du_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)\\
&+ \frac{c_\alpha(t-s)}{\zeta_s^3}\frac{\mathscr{N}^2(u_s)}{F^2(\boldsymbol{\nabla} u_s)}d\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\left(\nabla^{\boldsymbol{\nabla} u_s}\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\right).
\end{aligned}
$$

Since the Cauchy–Schwarz inequality for $g_{\boldsymbol{\nabla} u_s}$ yields

$$
\left|du_s\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)\right| \leq F(\boldsymbol{\nabla} u_s)\sqrt{d[F^2(\boldsymbol{\nabla} u_s)]\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)},
$$

we conclude that

$$
\begin{aligned}
&\Delta^{\boldsymbol{\nabla} u_s}\zeta_s - \partial_s\zeta_s\\
&\geq \frac{c_\alpha(t-s)\mathscr{N}'(u_s)^2}{\zeta_s^3}F^4(\boldsymbol{\nabla} u_s)\\
&\quad - \frac{c_\alpha(t-s)\mathscr{N}(u_s)|\mathscr{N}'(u_s)|}{\zeta_s^3}F(\boldsymbol{\nabla} u_s)\sqrt{d[F^2(\boldsymbol{\nabla} u_s)]\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)}\\
&\quad + \frac{c_\alpha(t-s)}{\zeta_s^3}\frac{\mathscr{N}^2(u_s)}{F^2(\boldsymbol{\nabla} u_s)}d\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\left(\nabla^{\boldsymbol{\nabla} u_s}\left[\frac{F^2(\boldsymbol{\nabla} u_s)}{2}\right]\right)\\
&= \frac{c_\alpha(t-s)}{\zeta_s^3}\left(|\mathscr{N}'(u_s)|F^2(\boldsymbol{\nabla} u_s) - \frac{\mathscr{N}(u_s)}{2F(\boldsymbol{\nabla} u_s)}\sqrt{d[F^2(\boldsymbol{\nabla} u_s)]\big(\nabla^{\boldsymbol{\nabla} u_s}[F^2(\boldsymbol{\nabla} u_s)]\big)}\right)^2\\
&\geq 0
\end{aligned}
$$

in the weak sense. Notice that, similarly to the proof of Theorem 3.7, we can take test functions from $H_0^1(M) \cap L^\infty(M)$ by virtue of (3.6). Therefore the RHS of (4.6) is nonnegative and this completes the proof. $\square$

When $K > 0$, choosing $\alpha = K^{-1}$ and letting $t \to \infty$ in (4.5) yields the following.

Corollary 4.5 *Assume that* $(M, F, \mathfrak{m})$ *is complete and satisfies* $\mathrm{Ric}_\infty \geq K > 0$, $\mathsf{C}_F < \infty$, $\mathsf{S}_F < \infty$ *and* $\mathfrak{m}(M) = 1$. *Then, for any* $u \in \mathcal{C}_c^\infty(M)$ *with* $0 \leq u \leq 1$ *and satisfying* (3.6), *we have*

$$
\sqrt{K}\mathscr{N}\left(\int_M u\, d\mathfrak{m}\right) \leq \int_M \sqrt{K\mathscr{N}^2(u) + F^2(\boldsymbol{\nabla} u)}\, d\mathfrak{m}. \tag{4.7}
$$

Proof. Let $(u_t)_{t\ge 0}$ be the global solution to the heat equation with $u_0 = u$. Taking $\alpha = K^{-1}$, we find $c_\alpha \equiv K^{-1}$ and hence by (4.5)

$$\sqrt{K\mathscr{N}^2(u_t)} \le \sqrt{K\mathscr{N}^2(u_t) + F^2(\boldsymbol{\nabla} u_t)} \le P_{0,t}^{\boldsymbol{\nabla} u}\Big(\sqrt{K\mathscr{N}^2(u) + F^2(\boldsymbol{\nabla} u)}\Big).$$

Letting $t \to \infty$, we deduce from the ergodicity (Proposition 4.3) that

$$u_t \to \int_M u\, d\mathfrak{m},$$
$$P_{0,t}^{\boldsymbol{\nabla} u}\Big(\sqrt{K\mathscr{N}^2(u) + F^2(\boldsymbol{\nabla} u)}\Big) \to \int_M \sqrt{K\mathscr{N}^2(u) + F^2(\boldsymbol{\nabla} u)}\, d\mathfrak{m}$$

in $L^2(M)$. Thereby we obtain (4.7). □

4.3 Proof of Theorem 4.1

Proof. Let $\theta \in (0,1)$. Fix a closed set $A \subset M$ with $\mathfrak{m}(A) = \theta$ and consider

$$u^\varepsilon(x) := \max\{1 - \varepsilon^{-1} d(x,A), 0\}, \qquad \varepsilon > 0.$$

Note that $F(\boldsymbol{\nabla} u^\varepsilon) = \varepsilon^{-1}$ on $B^-(A,\varepsilon) \setminus A$, where $B^-(A,\varepsilon) := \{x \in M \mid \inf_{y\in A} d(x,y) < \varepsilon\}$ is the backward ε-neighborhood of A. Applying (4.7) to (smooth approximations of) u^ε and letting $\varepsilon \downarrow 0$ implies, with the help of $\mathscr{N}(0) = \mathscr{N}(1) = 0$,

$$\sqrt{K}\mathscr{N}(\theta) \le \liminf_{\varepsilon\downarrow 0} \frac{\mathfrak{m}(B^-(A,\varepsilon)) - \mathfrak{m}(A)}{\varepsilon}.$$

This is the desired isoperimetric inequality for the reverse Finsler structure $\overleftarrow{F}$ (recall Definition 2.7) since, with $c := \varphi^{-1}(\theta)/\sqrt{K}$,

$$\sqrt{K}\mathscr{N}(\theta) = \sqrt{\frac{K}{2\pi}}\mathrm{e}^{-Kc^2/2}, \qquad \theta = \varphi(\sqrt{K}c) = \sqrt{\frac{K}{2\pi}}\int_{-\infty}^{c} \mathrm{e}^{-Ka^2/2}\, da.$$

Because the curvature bound $\mathrm{Ric}_\infty \ge K$ is common to F and $\overleftarrow{F}$, we also obtain (4.1). □

References

[AGS1] L. Ambrosio, N. Gigli and G. Savaré, Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163** (2014), 1405–1490.

[AGS2] L. Ambrosio, N. Gigli and G. Savaré, Bakry–Émery curvature-dimension condition and Riemannian Ricci curvature bounds. Ann. Probab. **43** (2015), 339–404.

[AM] L. Ambrosio and A. Mondino, Gaussian-type isoperimetric inequalities in $\mathrm{RCD}(K,\infty)$ probability spaces for positive K. Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. **27** (2016), 497–514.

[Au] L. Auslander, On curvature in Finsler geometry. Trans. Amer. Math. Soc. **79** (1955), 378–388.

[Ba] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes. (French) Lectures on probability theory (Saint-Flour, 1992), 1–114, Lecture Notes in Math., **1581**, Springer, Berlin, 1994.

[BE] D. Bakry and M. Émery, Diffusions hypercontractives. (French) Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., **1123**, Springer, Berlin, 1985.

[BGL] D. Bakry, I. Gentil and M. Ledoux, Analysis and geometry of Markov diffusion operators. Springer, Cham, 2014.

[BL] D. Bakry and M. Ledoux, Lévy–Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator. Invent. Math. **123** (1996), 259–281.

[BCS] D. Bao, S.-S. Chern and Z. Shen, An introduction to Riemann-Finsler geometry. Springer-Verlag, New York, 2000.

[Bob1] S. Bobkov, A functional form of the isoperimetric inequality for the Gaussian measure. J. Funct. Anal. **135** (1996), 39–49.

[Bob2] S. G. Bobkov, An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. Ann. Probab. **25** (1997), 206–214.

[Bor] C. Borell, The Brunn–Minkowski inequality in Gauss space. Invent. Math. **30** (1975), 207–216.

[CM] F. Cavalletti and A. Mondino, Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. Invent. Math. **208** (2017), 803–849.

[Ch_+] B. Chow, S.-C. Chu, D. Glickenstein, C. Guenther, J. Isenberg, T. Ivey, D. Knopf, P. Lu, F. Luo, L. Ni, The Ricci flow: techniques and applications. Part I. Geometric aspects. American Mathematical Society, Providence, RI, 2007.

[CMS] D. Cordero-Erausquin, R. J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. Invent. Math. **146** (2001), 219–257.

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

[Ev] L. C. Evans, Partial differential equations. American Mathematical Society, Providence, RI, 1998.

[GS] Y. Ge and Z. Shen, Eigenvalues and eigenfunctions of metric measure manifolds. Proc. London Math. Soc. (3) **82** (2001), 725–746.

[Gi1] N. Gigli, On the differential structure of metric measure spaces and applications. Mem. Amer. Math. Soc. **236** (2015), no. 1113.

[Gi2] N. Gigli, The splitting theorem in non-smooth context. Preprint (2013). Available at arXiv:1302.5555

[GKO] N. Gigli, K. Kuwada and S. Ohta, Heat flow on Alexandrov spaces, Comm. Pure Appl. Math. **66** (2013), 307–331.

[Gr] M. Gromov, Metric structures for Riemannian and non-Riemannian spaces. Based on the 1981 French original. With appendices by M. Katz, P. Pansu and S. Semmes. Translated from the French by Sean Michael Bates. Birkhäuser Boston, Inc., Boston, MA, 1999.

[Kl] B. Klartag, Needle decompositions in Riemannian geometry. Mem. Amer. Math. Soc. **249** (2017), no. 1180.

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

[Lee] P. W. Y. Lee, Displacement interpolations from a Hamiltonian point of view. J. Funct. Anal. **265** (2013), 3163–3203.

[Leo] G. Leoni, A first course in Sobolev spaces. American Mathematical Society, Providence, RI, 2009.

[Lé1] P. Lévy, Leçons d'analyse fonctionnelle. Gauthier-Villars, Paris, 1922.

[Lé2] P. Lévy, Problèmes concrets d'analyse fonctionnelle. Avec un complément sur les fonctionnelles analytiques par F. Pellegrino (French). 2d ed. Gauthier-Villars, Paris, 1951.

[Li] A. Lichnerowicz, Variétés riemanniennes à tenseur C non négatif (French). C. R. Acad. Sci. Paris Sér. A-B **271** (1970), A650–A653.

[LM] J.-L. Lions and E. Magenes, Non-homogeneous boundary value problems and applications. Vol. I. Translated from the French by P. Kenneth. Springer-Verlag, New York-Heidelberg, 1972.

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

[Ma] V. Maz'ya, Sobolev spaces with applications to elliptic partial differential equations. Second, revised and augmented edition. Springer, Heidelberg, 2011.

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

[Mi1] E. Milman, Sharp isoperimetric inequalities and model spaces for curvature-dimension-diameter condition. J. Eur. Math. Soc. (JEMS) **17** (2015), 1041–1078.

[Mi2] E. Milman, Beyond traditional curvature-dimension I: new model spaces for isoperimetric and concentration inequalities in negative dimension. Trans. Amer. Math. Soc. **369** (2017), 3605–3637.

[Oh1] S. Ohta, Uniform convexity and smoothness, and their applications in Finsler geometry. Math. Ann. **343** (2009), 669–699.

[Oh2] S. Ohta, Finsler interpolation inequalities. Calc. Var. Partial Differential Equations **36** (2009), 211–249.

[Oh3] S. Ohta, Vanishing S-curvature of Randers spaces. Differential Geom. Appl. **29** (2011), 174–178.

[Oh4] S. Ohta, On the curvature and heat flow on Hamiltonian systems. Anal. Geom. Metr. Spaces **2** (2014), 81–114.

[Oh5] S. Ohta, Splitting theorems for Finsler manifolds of nonnegative Ricci curvature. J. Reine Angew. Math. **700** (2015), 155–174.

[Oh6] S. Ohta, (K, N)-convexity and the curvature-dimension condition for negative N. J. Geom. Anal. **26** (2016), 2067–2096.

[Oh7] S. Ohta, Some functional inequalities on non-reversible Finsler manifolds. Proc. Indian Acad. Sci. Math. Sci. **127** (2017), 833–855; Correction in *Ibid.* **131** (2021), Paper No. 23.

[Oh8] S. Ohta, Needle decompositions and isoperimetric inequalities in Finsler geometry. J. Math. Soc. Japan **70** (2018), 651–693.

[OP] S. Ohta and M. Pálfia, Gradient flows and a Trotter–Kato formula of semi-convex functions on CAT(1)-spaces. Amer. J. Math. **139** (2017), 937–965.

[OS1] S. Ohta and K.-T. Sturm, Heat flow on Finsler manifolds. Comm. Pure Appl. Math. **62** (2009), 1386–1433.

[OS2] S. Ohta and K.-T. Sturm, Non-contraction of heat flow on Minkowski spaces. Arch. Ration. Mech. Anal. **204** (2012), 917–944.

[OS3] S. Ohta and K.-T. Sturm, Bochner–Weitzenböck formula and Li–Yau estimates on Finsler manifolds. Adv. Math. **252** (2014), 429–448.

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

[Qi] Z. Qian, Estimates for weighted volumes and applications. Quart. J. Math. Oxford Ser. (2) **48** (1997), 235–242.

[RR] M. Renardy and R. C. Rogers, An introduction to partial differential equations. Second edition. Springer-Verlag, New York, 2004.

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

[Sal] L. Saloff-Coste, Uniformly elliptic operators on Riemannian manifolds. J. Differential Geom. **36** (1992), 417–450.

[Sav] G. Savaré, Self-improvement of the Bakry–Émery condition and Wasserstein contraction of the heat flow in $\mathrm{RCD}(K, \infty)$ metric measure spaces. Discrete Contin. Dyn. Syst. **34** (2014), 1641–1661.

[Sh] Z. Shen, Lectures on Finsler geometry. World Scientific Publishing Co., Singapore, 2001.

[St1] K.-T. Sturm, On the geometry of metric measure spaces. I. Acta Math. **196** (2006), 65–131.

[St2] K.-T. Sturm, On the geometry of metric measure spaces. II. Acta Math. **196** (2006), 133–177.

[St3] K.-T. Sturm, Super-Ricci flows for metric measure spaces. J. Funct. Anal. **275** (2018), 3504–3569.

[SC] V. N. Sudakov and B. S. Cirel'son, Extremal properties of half-spaces for spherically invariant measures. (Russian) Problems in the theory of probability distributions, II. Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **41** (1974), 14–24, 165.

[Vi] C. Villani, Optimal transport, old and new. Springer-Verlag, Berlin, 2009.

[WX] G. Wang and C. Xia, A sharp lower bound for the first eigenvalue on Finsler manifolds. Ann. Inst. H. Poincaré Anal. Non Linéaire **30** (2013), 983–996.

[Wy] W. Wylie, A warped product version of the Cheeger–Gromoll splitting theorem. Trans. Amer. Math. Soc. **369** (2017), 6661–6681.

[Xi] C. Xia, Local gradient estimate for harmonic functions on Finsler manifolds. Calc. Var. Partial Differential Equations **51** (2014), 849–865.

[YH] S.-T. Yin and Q. He, The first eigenvalue of Finsler p-Laplacian. Differential Geom. Appl. **35** (2014), 30–49.