Bochner-Weitzenböck formula and Li-Yau estimates on Finsler manifolds

Shin-ichi Ohta*& Karl-Theodor Sturm†

May 6, 2011

Abstract

We prove the Bochner-Weitzenböck formula for the (nonlinear) Laplacian on general Finsler manifolds and derive Li-Yau type gradient estimates as well as parabolic Harnack inequalities. Moreover, we deduce Bakry-Émery gradient estimates. All these estimates depend on lower bounds for the weighted flag Ricci tensor.

1 Introduction

Geometry and analysis on singular spaces is an important topic of current research. Besides Alexandrov spaces, Finsler manifolds constitute one of the most relevant classes of explicit examples of metric measure spaces. Finsler spaces quite often occur naturally via homogenization as scaling limits of discrete or Riemannian structures.

A *Finsler manifold* is a smooth manifold M equipped with a norm – or, more generally, a Minkowski norm – $F(x,\cdot)$ on each tangent space T_xM. The particular case of a Hilbert norm leads to the important subclasses of Riemannian manifolds. In our previous paper [OS1], we introduced and studied in detail the *heat flow* on a Finsler manifold. It can equivalently be defined

- either as gradient flow in $L^2(M,m)$ for the *energy* $\frac{1}{2}\int_M F(\boldsymbol{\nabla} u)^2\,dm$;
- or as gradient flow in the L^2-Wasserstein space $(\mathcal{P}_2(M), W_2)$ for the *relative entropy* $\int_M u\log u\,dm$.

For this *nonlinear* evolution semigroup, we proved $\mathcal{C}^{1,\alpha}$-regularity and L^p-contraction estimates as well as integrated upper Gaussian estimates à la Davies ([OS1, Sections 3, 4]).

Various fine properties of the heat flow are intimately linked with the *weighted flag Ricci tensor* Ric_N (with $N \in [\dim M, \infty]$) introduced in [Oh1]. In [Oh1], the first author

*Department of Mathematics, Kyoto University, Kyoto 606-8502, Japan (sohta@math.kyoto-u.ac.jp);
Supported in part by the Grant-in-Aid for Young Scientists (B) 20740036.
†Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany (sturm@uni-bonn.de)

proved that the curvature-dimension condition $\mathsf{CD}(K,N)$ in the sense of Lott-Sturm-Villani for a weighted Finsler space (M,F,m) is equivalent to a lower bound K for Ric_N. In particular, the relative entropy – regarded as a function on the L^2-Wasserstein space – is K-convex if and only if Ric_∞ is bounded from below by K. In [OS1, Theorem 5.6], we proved that lower bounds for the weighted flag Ricci curvature imply point-wise comparison results à la Cheeger-Yau. Surprisingly enough, however, in [OS2] we observed that exponential contraction/expansion bounds in Wasserstein distance

$$W_2(P_t\mu, P_t\nu) \le e^{Ct} W_2(\mu,\nu) \tag{1.1}$$

hold true if and only if we are in a Riemannian setting.

The main goal of this paper now is to derive the *Bochner-Weitzenböck formula* (Theorems 3.3, 3.6)

$$\Delta^{\boldsymbol{\nabla} u}\left(\frac{F(\boldsymbol{\nabla} u)^2}{2}\right) - D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) = \mathrm{Ric}_\infty(\boldsymbol{\nabla} u) + \|\boldsymbol{\nabla}^2 u\|^2_{HS(\boldsymbol{\nabla} u)}. \tag{1.2}$$

Here $\Delta^{\boldsymbol{\nabla} u}$ denotes the linearization of the Laplacian $\boldsymbol{\Delta}$ on M in direction $\boldsymbol{\nabla} u$. The formula (1.2) a priori holds true only on the set $\{x \in M \,|\, \boldsymbol{\nabla} u(x) \neq 0\}$. It extends to an identity in distributional sense on entire M.

Formulation and proof of (1.2) – as well as applications of it – require lot of care due to the lack of regularity. For instance, solutions to the heat equation $\boldsymbol{\Delta} u = \partial u/\partial t$ on M will be $\mathcal{C}^2$ *only if* M is Riemannian ([OS1]). As a first striking application we deduce the *Bakry-Émery gradient estimate* (Theorem 4.1)

$$F\big(\boldsymbol{\nabla} P_t u(x)\big)^2 \le e^{-2Kt} P^{\boldsymbol{\nabla} u}_{0,t}\big(F(\boldsymbol{\nabla} u)^2\big)(x). \tag{1.3}$$

Here P_t denotes the (nonlinear) semigroup generated by the Laplacian $\boldsymbol{\Delta}$, whereas $P^{\boldsymbol{\nabla} u}_{0,t}$ denotes the (linear, symmetric) Markov transition operator on $L^2(M,m)$ with time-dependent generator L_s, $0<s<t$, obtained by linearization of $\boldsymbol{\Delta}$ in direction $\boldsymbol{\nabla} u(s,\cdot)$. An immediate consequence of (1.3) is the growth bound for Lipschitz constants

$$\mathrm{Lip}(P_t u) \le e^{-Kt}\,\mathrm{Lip}(u). \tag{1.4}$$

Note that in the Riemannian setting (or more abstract linear frameworks), according to the Kantorovich-Rubinstein duality, the bounds (1.1) and (1.4) are equivalent to each other (with $C=-K$).

Finally, we prove the *Li-Yau gradient estimate* (Theorem 4.4) for positive solutions to the heat equation on Finsler spaces which in the case $\mathrm{Ric}_N \ge 0$ simply reads as

$$F\big(\boldsymbol{\nabla}(\log u)(t,x)\big)^2 - \partial_t(\log u)(t,x) \le \frac{N}{2t},$$

and deduce as a corollary the famous *Li-Yau type Harnack inequality* (Theorem 4.5)

$$u(s,x) \le u(t,y)\cdot\left(\frac{t}{s}\right)^{N/2}\exp\left(\frac{d(x,y)^2}{4(t-s)}\right)$$

for any $0<s<t$ and $x,y\in M$.

For (1.3) and the subsequent estimates, M is required to be compact. It is an open question whether this compactness assumption can be replaced by completeness.

2 Finsler geometry

We briefly review the fundamentals of Finsler geometry, for which we refer to [BCS] and [Sh], as well as some results from [Oh1] and [OS1].

2.1 Finsler manifolds

Let M be a connected, n-dimensional $\mathcal{C}^\infty$-manifold without boundary. Given a local coordinate $(x^i)_{i=1}^n$ on an open set $U \subset M$, we will always use the coordinate $(x^i, v^i)_{i=1}^n$ of TU such that

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

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 $c \geq 0$.

(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 $v \in T_xM \setminus 0$.

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 the horizontal direction. For $x, y \in M$, we define the distance from x to y in a natural way by

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

where the infimum is taken over all $\mathcal{C}^1$-curves $\eta : [0, 1] \longrightarrow M$ such that $\eta(0) = x$ and $\eta(1) = y$. Note that $d(y, x) \neq d(x, y)$ may 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 (i.e., $F(\dot\eta)$ is constant). See (2.6) 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 (cf. [BCS, Theorem 6.6.1]).

For 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 via

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

This is regarded as the best Riemannian approximation of $F|_{T_xM}$ in the direction v. In fact, the unit sphere of g_v is tangent to that of $F|_{T_xM}$ at $v/F(v)$ up to the second order. In particular, we have $g_v(v,v) = F(v)^2$. The *Cartan tensor*

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

is a quantity appearing only in the Finsler context. Indeed, A_{ijk} vanishes everywhere on $TM \setminus 0$ if and only if F comes from a Riemannian metric. We will repeatedly use the next useful fact on homogeneous functions.

Theorem 2.2 (cf. [BCS, Theorem 1.2.1]) *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 $v \in \mathbb{R}^n \setminus 0$ (in other words, H is* positively r-homogeneous*). Then we have*

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

for all $v \in \mathbb{R}^n \setminus 0$.

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 $i, j, k = 1, 2, \ldots, n$. Define the *formal Christoffel symbol*

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

for $v \in TM \setminus 0$, where (g^{ij}) stands for the inverse matrix of (g_{ij}). We also introduce 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 $v \in TM \setminus 0$, and $G^i(0) = N^i_j(0) := 0$ by convention. Note that G^i is positively 2-homogeneous, so that Theorem 2.2 implies $\sum_{j=1}^{n} N^i_j(v)v^j = G^i(v)$.

By using the nonlinear connections N^i_j, the coefficients of the *Chern connection* is given by

$$\Gamma^i_{jk} := \gamma^i_{jk} - \sum_{l,m=1}^{n} \frac{g^{il}}{F}(A_{jlm}N^m_k + A_{lkm}N^m_j - A_{jkm}N^m_l) \tag{2.4}$$

on $TM \setminus 0$. Namely, the corresponding *covariant derivative* of a vector field $X = \sum_{i=1}^{n} X^i(\partial/\partial x^i)$ by $v \in T_xM$ with *reference vector* $w \in T_xM \setminus 0$ is defined by

$$D^w_v X(x) := \sum_{i,j=1}^{n} \left\{ v^j \frac{\partial X^i}{\partial x^j}(x) + \sum_{k=1}^{n} \Gamma^i_{jk}(w)v^j X^k(x) \right\} \frac{\partial}{\partial x^i}\Big|_x. \tag{2.5}$$

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

$$D^{\dot\eta}_{\dot\eta}\dot\eta(t) = \sum_{i=1}^n \big\{\ddot\eta^i(t) + G^i\big(\dot\eta(t)\big)\big\}\frac{\partial}{\partial x^i}\Big|_{\eta(t)} = 0. \tag{2.6}$$

2.2 Nonlinear Laplacian and the associated heat flow

Let us denote by $\mathcal{L}^* : T^*M \longrightarrow TM$ the *Legendre transform*. Precisely, $\mathcal{L}^*$ is sending $\alpha \in T^*_xM$ to the unique element $v \in T_xM$ such that $\alpha(v) = F^*(\alpha)^2$ and $F(v) = F^*(\alpha)$, where F^* stands for the dual norm of F. Note that $\mathcal{L}^*|_{T^*_xM}$ becomes a linear operator only when $F|_{T_xM}$ is an inner product. For a differentiable function $u : M \longrightarrow \mathbb{R}$, the *gradient vector* of u 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 $Du(x) = 0$, then clearly $\boldsymbol{\nabla} u(x) = 0$. If $Du(x) \neq 0$, then we can write in coordinates

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

We must be careful when $Du(x) = 0$, because $g_{ij}(\boldsymbol{\nabla} u(x))$ is not defined as well as the Legendre transform $\mathcal{L}^*$ being only continuous at the zero section. For later convenience we set in general

$$M_V := \{x \in M \,|\, V(x) \neq 0\} \tag{2.7}$$

for a vector field V on M, and $M_u := M_{\boldsymbol{\nabla} u}$. For a differentiable vector field V on M and $x \in M_V$, we define $\boldsymbol{\nabla} V(x) \in T^*_xM \otimes T_xM$ by using the covariant derivative (2.5) as

$$\boldsymbol{\nabla} V(v) := D^V_v V(x) \in T_xM, \qquad v \in T_xM. \tag{2.8}$$

We also set $\boldsymbol{\nabla}^2 u(x) := \boldsymbol{\nabla}(\boldsymbol{\nabla} u)(x)$ for a twice differentiable function $u : M \longrightarrow \mathbb{R}$ and $x \in M_u$.

Lemma 2.3 *$\boldsymbol{\nabla}^2 u(x)$ is symmetric in the sense that $g_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla}^2 u(v), w) = g_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla}^2 u(w), v)$ for all $v, w \in T_xM$.*

Proof. This can be checked by hand. Indeed, choosing a coordinate $\{(\partial/\partial x^i)|_x\}_{i=1}^n$ orthonormal with respect to $g_{\boldsymbol{\nabla} u(x)}$, we see by calculation

$$\begin{aligned}\boldsymbol{\nabla}^2 u\bigg(\frac{\partial}{\partial x^i}\bigg) &= D^{\boldsymbol{\nabla} u}_{\partial/\partial x^i}\bigg(\sum_{j,k=1}^n g^{jk}\big(\boldsymbol{\nabla} u(x)\big)\frac{\partial u}{\partial x^k}(x)\frac{\partial}{\partial x^j}\bigg)\\ &= \sum_{j=1}^n\bigg\{\frac{\partial^2 u}{\partial x^i \partial x^j}(x) - \sum_{k=1}^n \Gamma^k_{ij}\big(\boldsymbol{\nabla} u(x)\big)\frac{\partial u}{\partial x^k}(x)\bigg\}\frac{\partial}{\partial x^j}.\end{aligned}$$

□

From now on, we fix an arbitrary positive $\mathcal{C}^\infty$-measure m on M as our base measure. Define the *divergence* of a differentiable vector field V on M with respect to m by

$$\operatorname{div}_m V := \sum_{i=1}^n\bigg(\frac{\partial V_i}{\partial x^i} + V_i\frac{\partial \Phi}{\partial x^i}\bigg),$$

where we decompose m in coordinates as $dm = e^{\Phi}\, dx^1 dx^2 \cdots dx^n$. This can be rewritten (and extended to weakly differentiable vector fields) in the weak form as

$$\int_M \phi \operatorname{div}_m V \, dm = -\int_M D\phi(V)\, dm$$

for all $\phi \in \mathcal{C}_c^\infty(M)$. Then we define the distributional *Laplacian* of $u \in H^1_{\mathrm{loc}}(M)$ by $\boldsymbol{\Delta} u := \operatorname{div}_m(\boldsymbol{\nabla} u)$ in the weak sense that

$$\int_M \phi \boldsymbol{\Delta} u \, dm := -\int_M D\phi(\boldsymbol{\nabla} u)\, dm$$

for $\phi \in \mathcal{C}_c^\infty(M)$. We remark that $H^1_{\mathrm{loc}}(M)$ (and $L^2_{\mathrm{loc}}(M)$ etc.) is defined solely in terms of the differentiable structure of M.

Given a vector field V such that $V \neq 0$ on M_u, we can define the gradient vector and the Laplacian on the weighted Riemannian manifold (M, g_V, m) by

$$\nabla^V u := \begin{cases} \displaystyle\sum_{i,j=1}^n g^{ij}(V) \frac{\partial u}{\partial x^j} \frac{\partial}{\partial x^i} & \text{on } M_u, \\ 0 & \text{on } M \setminus M_u, \end{cases} \qquad \Delta^V u := \operatorname{div}_m(\nabla^V u),$$

where the latter should be read in the sense of distributions. Note that, in the definition of $\Delta^V u$, we used the divergence with respect to m rather than the volume form of g_V. It is not difficult to see $\nabla^{\boldsymbol{\nabla} u} u = \boldsymbol{\nabla} u$ and $\Delta^{\boldsymbol{\nabla} u} u = \boldsymbol{\Delta} u$ ([OS1, Lemma 2.4]). We also observe for later use that, given u, f_1 and f_2,

$$Df_2(\nabla^{\boldsymbol{\nabla} u} f_1) = g_{\boldsymbol{\nabla} u}(\nabla^{\boldsymbol{\nabla} u} f_1, \nabla^{\boldsymbol{\nabla} u} f_2) = Df_1(\nabla^{\boldsymbol{\nabla} u} f_2). \tag{2.9}$$

In [OS1], we have studied the associated *nonlinear heat equation* $\partial u/\partial t = \boldsymbol{\Delta} u$. The nonlinearity is inherited from the Legendre transform. Given an open set $\Omega \subset M$, define the *Dirichlet energy* of $u \in H^1_{\mathrm{loc}}(\Omega)$ by

$$\mathcal{E}_\Omega(u) := \frac{1}{2} \int_\Omega F(\boldsymbol{\nabla} u)^2 \, dm.$$

We will suppress Ω if $\Omega = M$, namely $\mathcal{E} = \mathcal{E}_M$. Set

$$H^1(\Omega) := \{u \in H^1_{\mathrm{loc}}(\Omega) \cap L^2(\Omega, m) \mid \mathcal{E}_\Omega(u) < \infty\},$$

and let $H^1_0(\Omega)$ be the closure of $\mathcal{C}_c^\infty(\Omega)$ with respect to the (Minkowski) norm $\|u\|_{H^1(\Omega)} := \|u\|_{L^2(\Omega)} + \mathcal{E}_\Omega(u)^{1/2}$.

Definition 2.4 (Global and local solutions) (1) For $T > 0$, we say that a function u on $[0,T] \times M$ is a *global solution* to the heat equation if $u \in L^2([0,T], H^1_0(M)) \cap H^1([0,T], H^{-1}(M))$ and if

$$\int_M \phi \frac{\partial u_t}{\partial t} \, dm = -\int_M D\phi(\boldsymbol{\nabla} u_t)\, dm$$

holds for all $t \in [0,T]$ and $\phi \in \mathcal{C}_c^\infty(M)$, where we set $u_t := u(t, \cdot)$.

(2) Given an open interval $I \subset \mathbb{R}$ and an open set $\Omega \subset M$, we say that a function u on $I \times \Omega$ is a *local solution* to the heat equation on $I \times \Omega$ if $u \in L^2_{\mathrm{loc}}(I \times \Omega)$, $F(\boldsymbol{\nabla} u) \in L^2_{\mathrm{loc}}(I \times \Omega)$ and if

$$\int_I \int_\Omega u \frac{\partial \phi}{\partial t}\, dmdt = \int_I \int_\Omega D\phi(\boldsymbol{\nabla} u)\, dmdt$$

holds for all $\phi \in \mathcal{C}^\infty_c(I \times \Omega)$.

Global solutions can be constructed as the gradient flow of the energy functional $\mathcal{E}$ in $L^2(M, m)$. We summarize the existence and regularity properties established in [OS1] in the next theorem.

Theorem 2.5 ([OS1, Sections 3,4])

(i) *For each $u_0 \in H^1_0(M)$ and $T > 0$, there exists a 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 m for all $t \in (0, T)$. If M is compact, then such a global solution is uniquely determined by its initial datum u_0.*

(ii) *Given an open set $\Omega \subset M$, the continuous version of any local solution u to the heat equation on Ω enjoys the H^2_{loc}-regularity in x as well as the $\mathcal{C}^{1,\alpha}$-regularity in both t and x. Furthermore, the distributional time derivative $\partial_t u$ lies in $H^1_{\mathrm{loc}}(\Omega) \cap \mathcal{C}(\Omega)$.*

We remark that the mild smoothness assumption [OS1, (4.4)] clearly holds true for our $\mathcal{C}^\infty$-smooth F and m.

2.3 Ricci curvature

The *Ricci curvature* (as the trace of the *flag curvature*) on a Finsler manifold is defined by using the Chern connection (and is in fact independent of the choice of connection). Instead of giving a precise definition in coordinates, here we explain a useful interpretation due to Shen [Sh, §6.2]. Given a unit vector $v \in T_x M$ (i.e., $F(v) = 1$), we extend it to a $\mathcal{C}^\infty$-vector field V on a neighborhood of x in such a way that every integral curve of V is geodesic, and consider the Riemannian structure g_V induced from (2.2). Then the Ricci curvature $\mathrm{Ric}(v)$ of v with respect to F coincides with the 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, the *weighted Ricci curvature* for (M, F, m) is introduced in [Oh1] as follows.

Definition 2.6 (Weighted Ricci curvature) Given a unit vector $v \in T_x M$, let $\eta : (-\varepsilon, \varepsilon) \longrightarrow M$ be the geodesic such that $\dot\eta(0) = v$. We decompose m as $m = e^{-\Psi}\, \mathrm{vol}_{\dot\eta}$ along η, where $\mathrm{vol}_{\dot\eta}$ is the volume form of $g_{\dot\eta}$. Define

$$(1)\ \mathrm{Ric}_n(v) := \begin{cases} \mathrm{Ric}(v) + (\Psi \circ \eta)''(0) & \text{if } (\Psi \circ \eta)'(0) = 0, \\ -\infty & \text{otherwise,} \end{cases}$$

$$(2)\ \mathrm{Ric}_N(v) := \mathrm{Ric}(v) + (\Psi \circ \eta)''(0) - \frac{(\Psi \circ \eta)'(0)^2}{N - n} \quad \text{for } N \in (n, \infty),$$

(3) $\mathrm{Ric}_\infty(v) := \mathrm{Ric}(v) + (\Psi \circ \eta)''(0)$.

For $c \geq 0$ and $N \in [n, \infty]$, we define $\mathrm{Ric}_N(cv) := c^2\, \mathrm{Ric}_N(v)$.

It is established in [Oh1, Theorem 1.2] that, for $K \in \mathbb{R}$, the bound $\mathrm{Ric}_N(v) \geq KF(v)^2$ is equivalent to Lott, Villani and the second author's *curvature-dimension condition* $\mathsf{CD}(K, N)$. This extends the corresponding result on (weighted) Riemannian manifolds (due to [vRS], [St1], [St2], [St3], [LV2], [LV3]), and has many analytic and geometric applications (see [Oh1], [OS1] or a survey [Oh2]).

Remark 2.7 (a) 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 $(M, g_{\boldsymbol{\nabla} u}, m)$. Indeed, on a Minkowski space $(\mathbb{R}^n, \|\cdot\|)$, it is easy to see that $(\mathbb{R}^n, g_{\boldsymbol{\nabla} u})$ is not flat for some function u.

(b) 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 $N \in [n, \infty]$. It is also known that, for Finsler manifolds of *Berwald type* (i.e., Γ^k_{ij} is constant on each T_xM), the *Busemann-Hausdorff measure* satisfies $(\Psi \circ \eta)' \equiv 0$ (in other words, Shen's **S**-curvature vanishes, see [Sh, §7.3]). In general, however, there may not exist any measure with vanishing **S**-curvature (see [Oh3] for such an example). This is a reason why we begin with an arbitrary measure m.

For later convenience, we introduce the following notations.

Definition 2.8 (Reverse Finsler structure) Define the *reverse Finsler structure* $\overleftarrow{F}$ of F by $\overleftarrow{F}(v) := F(-v)$. We say that F is *reversible* if $\overleftarrow{F} = F$. We will put an arrow $\leftarrow$ on those quantities associated with $\overleftarrow{F}$, for example, $\overleftarrow{d}(x, y) = d(y, x)$, $\overleftarrow{\boldsymbol{\nabla}} u = -\boldsymbol{\nabla}(-u)$ and $\overleftarrow{\mathrm{Ric}}_N(v) = \mathrm{Ric}_N(-v)$.

We say that (M, F) is *backward complete* if $(M, \overleftarrow{F})$ is forward complete. Compact Finsler manifolds are both forward and backward complete.

3 Bochner-Weitzenböck formula

In this section, we prove the Bochner-Weitzenböck formula. For the sake of simplicity, we first derive a general formula for vector fields V, and then apply it to gradient vector fields of functions. We start with a point-wise calculation on M_V (recall (2.7)) followed by an integrated formula.

3.1 Point-wise calculation

We follow the argumentation in [Vi, Chapter 14] in the Riemannian situation, with the help of [Oh1] and [OS1].

As in the previous section, let (M, F) be an n-dimensional $\mathcal{C}^\infty$-Finsler manifold and m be a positive $\mathcal{C}^\infty$-measure on M. Fix a $\mathcal{C}^\infty$-vector field V on M and $x \in M_V$. For

$t \in (0, \delta]$ with sufficiently small $\delta > 0$, we introduce the map $\mathcal{T}_t$ and the vector field V_t on a neighborhood of x by

$$\mathcal{T}_t(y) := \exp_y[tV(y)], \qquad V_t\big(\mathcal{T}_t(y)\big) := \frac{\partial \mathcal{T}_t}{\partial t}(y). \tag{3.1}$$

As the curve $\sigma(t) := \mathcal{T}_t(y)$ is geodesic, taking the covariant differentiation $D^{\dot\sigma}_{\sigma}$ of $t \longmapsto V_t(\sigma(t))$ yields the *pressureless Euler equation*

$$\frac{\partial V_t}{\partial t} + D^{V_t}_{V_t} V_t = 0, \tag{3.2}$$

where we abbreviated as

$$\frac{\partial V_t}{\partial t} = \sum_{i=1}^n \frac{\partial V_t^i}{\partial t}\frac{\partial}{\partial x^i}, \qquad \text{with } V_t = \sum_{i=1}^n V_t^i \frac{\partial}{\partial x^i}.$$

Now we put $\eta(t) := \mathcal{T}_t(x)$, take an orthonormal basis $\{e_i\}_{i=1}^n$ of (T_xM, g_V) such that $e_n = \dot\eta(0)/F(\dot\eta(0))$, and consider the vector field along η

$$E_i(t) := D(\mathcal{T}_t)_x(e_i) \in T_{\eta(t)}M, \qquad i = 1, 2, \dots, n.$$

Then each E_i is a Jacobi field along η. (Note that $E_n(t) \neq \dot\eta(t)/F(\dot\eta(t))$ in general, since $\mathcal{T}_t(\eta(s)) \neq \eta(s+t)$ for $s, t > 0$.) We set $E_i'(t) := D^{\dot\eta}_{\dot\eta} E_i(t)$ for simplicity. Define the $n \times n$ matrix-valued function $B(t)$ by $E_i'(t) = \sum_{j=1}^n b_{ij}(t) E_j(t)$. (This B corresponds to U in [Vi, Chapter 14].) Along the discussion in [Oh1, Lemma 7.2], we obtain the following *Riccati type equation*.

Lemma 3.1 (Riccati type equation) *For η and B as above, we have*

$$\frac{d(\operatorname{tr} B)}{dt}(t) + \operatorname{tr}\big(B(t)^2\big) + \operatorname{Ric}\big(\dot\eta(t)\big) = 0, \qquad t \geq 0.$$

Proof. We give only an outline of the proof, see [Oh1] for details. Consider the matrix-valued function $A(t) = (g_{\dot\eta}(E_i(t), E_j(t)))$ and observe $A' = BA + AB^T$ by definition (B^T is the transpose of B). Since each E_i is a Jacobi field, we have $(BA - AB^T)' \equiv 0$, so that

$$BA - AB^T \equiv B(0) - B(0)^T, \qquad A' = 2BA - B(0) + B(0)^T.$$

We also deduce from the Jacobi equation for E_i that

$$A'' = -2\operatorname{Ric}_{\dot\eta} + 2BAB^T = -2\operatorname{Ric}_{\dot\eta} + 2B^2A - 2B\big(B(0) - B(0)^T\big),$$

where $(\operatorname{Ric}_{\dot\eta})_{ij} := g_{\dot\eta}(R^{\dot\eta}(E_i, \dot\eta)\dot\eta, E_j)$ and $R^{\dot\eta}$ is the curvature tensor. Comparing this with $A'' = 2B'A + 2BA'$, we find $B' = -\operatorname{Ric}_{\dot\eta} A^{-1} - B^2$. Taking the trace with respect to $g_{\dot\eta}$ completes the proof. □

Lemma 3.2 (i) *We have $B(t) = \boldsymbol{\nabla} V_t(\eta(t))$ in the sense that, for each $i = 1, 2, \dots, n$, $\boldsymbol{\nabla} V_t(E_i(t)) = \sum_{j=1}^n b_{ij}(t) E_j(t)$.*

(ii) *It holds that* $\mathrm{tr}(B(t)) = \mathrm{div}_m V_t(\eta(t)) + D\Psi(\dot\eta(t))$, *where* $m = e^{-\Psi}\,\mathrm{vol}_{\dot\eta}$ *along* η *such that* $\mathrm{vol}_{\dot\eta}$ *denotes the Riemannian volume measure of* $g_{\dot\eta}$.

Proof. (i) By the definition (2.8), $\boldsymbol{\nabla} V_t(E_i(t)) = D^{V_t}_{E_i(t)} V_t(\eta(t))$. Setting

$$E_i(t) = \frac{\partial}{\partial\delta}\Big[\mathcal{T}_t(\exp_x \delta e_i)\Big]\Big|_{\delta=0} =: \sum_{k=1}^n E_i^k(t)\frac{\partial}{\partial x^k}\Big|_{\eta(t)},$$

we have

$$E_i'(t) = \sum_{k=1}^n \frac{dE_i^k}{dt}(t)\frac{\partial}{\partial x^k} + \sum_{j,k,l=1}^n \Gamma^l_{jk}\big(\dot\eta(t)\big)\dot\eta^j(t)E_i^k(t)\frac{\partial}{\partial x^l},$$

where $\dot\eta(t) = \sum_{j=1}^n \dot\eta^j(t)(\partial/\partial x^j)|_{\eta(t)}$. Exchanging the order of differentiations (by δ and t) in the first term, we deduce from (3.1) that

$$\frac{dE_i^k}{dt}(t) = \frac{\partial}{\partial\delta}\Big[V_t^k\big(\mathcal{T}_t(\exp_x \delta e_i)\big)\Big]\Big|_{\delta=0}.$$

Therefore we obtain $E_i'(t) = D^{V_t}_{E_i(t)} V_t(\eta(t)) = \boldsymbol{\nabla} V_t(E_i(t))$ and complete the proof.

(ii) Choose a coordinate $(x^i)_{i=1}^n$ around $\eta(t)$ such that $\{(\partial/\partial x^i)|_{\eta(t)}\}_{i=1}^n$ is an orthonormal basis of $(T_{\eta(t)}M, g_{V_t})$. We will suppress evaluations at $\eta(t)$. Recall first that

$$\boldsymbol{\nabla} V_t\bigg(\frac{\partial}{\partial x^i}\bigg) = D^{V_t}_{\partial/\partial x^i} V_t = \sum_{k=1}^n \bigg\{\frac{\partial V_t^k}{\partial x^i} + \sum_{j=1}^n \Gamma^k_{ij}(V_t)V_t^j\bigg\}\frac{\partial}{\partial x^k}.$$

Thus we have

$$\mathrm{tr}\big(B(t)\big) = \mathrm{tr}(\boldsymbol{\nabla} V_t) = \sum_{i=1}^n \bigg\{\frac{\partial V_t^i}{\partial x^i} + \sum_{j=1}^n \Gamma^i_{ij}(V_t)V_t^j\bigg\}.$$

We also find, by (2.4) and (2.3),

$$\sum_{j=1}^n \Gamma^i_{ij}(V_t)V_t^j = \sum_{j=1}^n \bigg\{\frac{1}{2}\frac{\partial g_{ii}}{\partial x^j}(V_t)V_t^j - A_{iij}(V_t)\frac{G^j(V_t)}{F(V_t)}\bigg\}.$$

Next we observe from $dm = e^{-\Psi}\sqrt{\det g_{V_t}}\,dx^1dx^2\cdots dx^n$ on η that

$$\mathrm{div}_m V_t\big(\eta(t)\big) + D\Psi\big(\dot\eta(t)\big) = \sum_{i=1}^n \bigg\{\frac{\partial V_t^i}{\partial x^i} + V_t^i\frac{\partial(\log\sqrt{\det g_{V_t}})}{\partial x^i}\bigg\}.$$

Note that

$$\begin{aligned}\frac{\partial(\log\sqrt{\det g_{V_t}})}{\partial x^i} &= \frac{1}{2}\,\mathrm{tr}\left(\frac{\partial g_{V_t}}{\partial x^i}\cdot g_{V_t}^{-1}\right)\\ &= \frac{1}{2}\sum_{j=1}^n\bigg\{\frac{\partial g_{jj}}{\partial x^i}(V_t) + \sum_{k=1}^n \frac{\partial g_{jj}}{\partial v^k}(V_t)\frac{\partial V_t^k}{\partial x^i}\bigg\}.\end{aligned}$$

This yields, together with (3.1),

$$\begin{aligned}\sum_{i=1}^n V_t^i \frac{\partial(\log\sqrt{\det g_{V_t}})}{\partial x^i} &= \frac{1}{2}\sum_{i,j=1}^n \left\{ \frac{\partial g_{jj}}{\partial x^i}(V_t)V_t^i + \frac{\partial g_{jj}}{\partial v^i}(V_t)\ddot{\eta}^i(t) \right\} \\ &= \sum_{i,j=1}^n \left\{ \frac{1}{2}\frac{\partial g_{jj}}{\partial x^i}(V_t)V_t^i - \frac{A_{jji}(V_t)}{F(V_t)}G^i(V_t) \right\} \\ &= \sum_{i,j=1}^n \Gamma_{ij}^i(V_t)V_t^j.\end{aligned}$$

We used the geodesic equation (2.6) of η in the second line. This completes the proof. □

We deduce from Lemmas 3.1, 3.2(ii) that

$$\frac{d}{dt}\Big|_{t=0+}\Big[\operatorname{div}_m V_t\big(\eta(t)\big)\Big] + \operatorname{tr}\big(B(0)^2\big) + \operatorname{Ric}_\infty\big(\dot\eta(0)\big) = 0.$$

Thanks to (3.2), we have

$$\begin{aligned}\frac{d}{dt}\Big|_{t=0+}\Big[\operatorname{div}_m V_t\big(\eta(t)\big)\Big] &= D(\operatorname{div}_m V)\big(\dot\eta(0)\big) + \operatorname{div}_m\left(\frac{\partial V_t}{\partial t}\Big|_{t=0+}\right)(x) \\ &= D(\operatorname{div}_m V)\big(\dot\eta(0)\big) - \operatorname{div}_m(D_V^V V)(x).\end{aligned}$$

Combining these, we obtain

$$\operatorname{div}_m(D_V^V V) - D(\operatorname{div}_m V)(V) = \operatorname{Ric}_\infty(V) + \operatorname{tr}\big(B(0)^2\big) \tag{3.3}$$

at x.

Now, we take $u \in \mathcal{C}^\infty(M)$ and $x \in M_u$, and apply (3.3) to $V = \boldsymbol{\nabla} u$. Then the symmetry of $\boldsymbol{\nabla}^2 u$ (Lemma 2.3) allows us to simplify (3.3) in two respects. First, Lemma 3.2(i) immediately yields $\operatorname{tr}(B(0)^2) = \|\boldsymbol{\nabla}^2 u(x)\|^2_{HS(\boldsymbol{\nabla} u)}$, where $\|\cdot\|_{HS(\boldsymbol{\nabla} u)}$ stands for the Hilbert-Schmidt norm with respect to $g_{\boldsymbol{\nabla} u}$. Second, for each $i = 1, 2, \dots, n$,

$$g_{\boldsymbol{\nabla} u}\left(D^{\boldsymbol{\nabla} u}_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u), \frac{\partial}{\partial x^i}\right) = g_{\boldsymbol{\nabla} u}\big(\boldsymbol{\nabla} u, D^{\boldsymbol{\nabla} u}_{\partial/\partial x^i}(\boldsymbol{\nabla} u)\big) = \frac{\partial}{\partial x^i}\left(\frac{g_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u)}{2}\right).$$

This implies

$$D^{\boldsymbol{\nabla} u}_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u) = \sum_{i,j=1}^n g^{ji}(\boldsymbol{\nabla} u) g_{\boldsymbol{\nabla} u}\left(D^{\boldsymbol{\nabla} u}_{\boldsymbol{\nabla} u}(\boldsymbol{\nabla} u), \frac{\partial}{\partial x^i}\right)\frac{\partial}{\partial x^j} = \nabla^{\boldsymbol{\nabla} u}\left(\frac{F(\boldsymbol{\nabla} u)^2}{2}\right).$$

Plugging these into (3.3), we obtain (3.4) below on M_u.

Theorem 3.3 (Point-wise Bochner-Weitzenböck formula) *Given $u \in \mathcal{C}^\infty(M)$, we have*

$$\Delta^{\boldsymbol{\nabla} u}\left(\frac{F(\boldsymbol{\nabla} u)^2}{2}\right) - D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) = \operatorname{Ric}_\infty(\boldsymbol{\nabla} u) + \|\boldsymbol{\nabla}^2 u\|^2_{HS(\boldsymbol{\nabla} u)} \tag{3.4}$$

as well as

$$\Delta^{\boldsymbol{\nabla} u}\left(\frac{F(\boldsymbol{\nabla} u)^2}{2}\right) - D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u) \geq \operatorname{Ric}_N(\boldsymbol{\nabla} u) + \frac{(\boldsymbol{\Delta} u)^2}{N} \tag{3.5}$$

for $N \in [n, \infty]$, point-wise on M_u.

Proof. We have seen that (3.4) holds on M_u. The second assertion is clear if $N = \infty$. For $N \in (n, \infty)$, we observe from Lemma 3.2 that, as $\{e_i\}_{i=1}^n$ is orthonormal so that $B(0)$ is symmetric,

$$\begin{aligned}\|\boldsymbol{\nabla}^2 u\|^2_{HS(\boldsymbol{\nabla} u)} = \operatorname{tr}\big(B(0)^2\big) &= \frac{(\operatorname{tr} B(0))^2}{n} + \left\| B(0) - \frac{\operatorname{tr} B(0)}{n} I_n \right\|^2_{HS} \\ &\geq \frac{(\boldsymbol{\Delta} u + D\Psi(\boldsymbol{\nabla} u))^2}{n}.\end{aligned}$$

Plugging $a = \boldsymbol{\Delta} u$ and $b = D\Psi(\boldsymbol{\nabla} u)$ into

$$\frac{(a+b)^2}{n} = \frac{a^2}{N} - \frac{b^2}{N-n} + \frac{N(N-n)}{n}\left(\frac{a}{N} + \frac{b}{N-n}\right)^2,$$

we obtain

$$\frac{(\boldsymbol{\Delta} u + D\Psi(\boldsymbol{\nabla} u))^2}{n} \geq \frac{(\boldsymbol{\Delta} u)^2}{N} - \frac{D\Psi(\boldsymbol{\nabla} u)^2}{N-n}.$$

Together with (3.4), this yields the desired estimate (3.5). The remaining case of $N = n$ is derived as the limit. □

Remark 3.4 Our Bochner-Weitzenböck formula (3.4) can not be simply derived from the formula of the weighted Riemannian manifold $(M, g_{\boldsymbol{\nabla} u}, m)$. This is because $\mathrm{Ric}_\infty(\boldsymbol{\nabla} u)$ and $\|\boldsymbol{\nabla}^2 u\|^2_{HS(\boldsymbol{\nabla} u)}$ are different from those for $g_{\boldsymbol{\nabla} u}$ (unless all integral curves of $\boldsymbol{\nabla} u$ are geodesic), while $D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u)$ and $\Delta^{\boldsymbol{\nabla} u}(F(\boldsymbol{\nabla} u)^2/2)$ are same for F and $g_{\boldsymbol{\nabla} u}$. Recall Remark 2.7(a) and note that $\boldsymbol{\nabla}^2 u(v) \neq (\nabla^{\boldsymbol{\nabla} u})^2 u(v)$ in general (for $v \neq \boldsymbol{\nabla} u$).

Let us close the subsection with additional comments on the situation of $V = \boldsymbol{\nabla} u$ with $u \in \mathcal{C}_c^\infty(M)$. A little knowledge of optimal transport theory is necessary for which we refer to [Vi] and [Oh1]. A potential function φ_t of the vector field V_t (i.e., $\boldsymbol{\nabla}\varphi_t = V_t$) is given by

$$\varphi_t(y) := \inf_{x \in M}\left\{\frac{d(x,y)^2}{2t} + u(x)\right\}. \tag{3.6}$$

This is the *Hopf-Lax formula* providing a (viscosity) solution to the *Hamilton-Jacobi equation*

$$\frac{\partial \varphi_t}{\partial t} + \frac{F(\boldsymbol{\nabla}\varphi_t)^2}{2} = 0 \tag{3.7}$$

that corresponds to (3.2). Indeed, we can apply the proof of [LV1, Theorem 2.5] verbatim to see (3.7), although our distance is nonsymmetric. In view of optimal transport theory, (3.6) is rewritten as

$$t\varphi_t(y) = \inf_{x \in M}\left\{\frac{d(x,y)^2}{2} - \big(-tu(x)\big)\right\} = (-tu)^c(y),$$

where $(-tu)^c$ stands for the *c-transform* of the function $-tu$ for the cost $c(x,y) = d(x,y)^2/2$. Thus, if we put $\eta(t) = \exp_x[t\boldsymbol{\nabla} u(x)]$, then it holds that

$$\dot\eta(t) = -\overleftarrow{\boldsymbol{\nabla}}(-\varphi_t)\big(\eta(t)\big) = \boldsymbol{\nabla}\varphi_t\big(\eta(t)\big).$$

Hence, in terms of Wasserstein geometry, the vector field $\boldsymbol{\nabla}\varphi_t$ is the tangent vector at time t of the geodesic $(\exp[t\boldsymbol{\nabla}u]_\sharp\mu)_{t\in[0,\delta]}$ starting from any absolutely continuous probability measure μ (whereas δ depends on u), where $\exp[t\boldsymbol{\nabla}u]_\sharp\mu$ denotes the push-forward of μ by the map $\exp[t\boldsymbol{\nabla}u]$.

3.2 Integrated formula

We need to be careful in extending the point-wise estimate (3.5) to a global one (in the weak sense), because some quantities are undefined on $M \setminus M_u$. We prove an auxiliary lemma to overcome this difficulty.

Lemma 3.5 *Let L be a locally uniformly elliptic, linear second order differential operator in divergence form on a Riemannian manifold. Let $\mathcal{E}^L$ denote the associated Dirichlet form with domain $H_0^1(M)$ and with intrinsic gradient denoted by ∇^L, i.e.,*

$$\mathcal{E}^L(h,h) = -\int_M h\cdot Lh\,dm = \int_M |\nabla^L h|^2\,dm, \qquad h\in H_0^1(M).$$

Then we have the following.

(i) *For each $h\in H_0^1(M)$, it holds $\nabla^L h = 0$ a.e. on $h^{-1}(0)$.*

(ii) *If $h\in H_0^1(M)\cap L^\infty(M)$, then $\nabla^L(h^2/2) = h\nabla^L h = 0$ a.e. on $h^{-1}(0)$.*

(iii) *The assertions in* (i) *and* (ii) *also hold true if h merely lies locally in the respective spaces.*

Proof. (i) is a particular case of the general fact that, for local regular Dirichlet forms with square field operator Γ,

$$\mathcal{E}^L(h,h) = \int_M \Gamma(h,h)\,dm = \int_{\{h\neq 0\}} \Gamma(h,h)\,dm.$$

(In the monograph [BH], this is called the *energy image density*.)

(ii) follows from the chain rule.

(iii) Given an arbitrary open relatively compact set $U\subset M$, choose a cut-off function $\phi\in H_c^1(M)\cap L^\infty(M)$ such that $\phi\equiv 1$ on U. Then, for each $h\in H_{\mathrm{loc}}^1(M)$, it holds $h\phi\in H_0^1(M)$ and we have $\nabla^L(h\phi) = \nabla^L\phi$ a.e. on $(h\phi)^{-1}(0)\cap U = \phi^{-1}(0)\cap U$. Thus we can reduce the claim to (i), and it is similar for (ii). □

Theorem 3.6 (Integrated Bochner-Weitzenböck formula) *Given $u\in H_{\mathrm{loc}}^2(M)\cap \mathcal{C}^1(M)$ with $\boldsymbol{\Delta}u\in H_{\mathrm{loc}}^1(M)$, we have*

$$\begin{aligned}&-\int_M D\phi\Big(\nabla^{\boldsymbol{\nabla}u}\Big(\frac{F(\boldsymbol{\nabla}u)^2}{2}\Big)\Big)\,dm\\ &=\int_M \phi\big\{D(\boldsymbol{\Delta}u)(\boldsymbol{\nabla}u) + \mathrm{Ric}_\infty(\boldsymbol{\nabla}u) + \|\boldsymbol{\nabla}^2 u\|^2_{HS(\boldsymbol{\nabla}u)}\big\}\,dm\end{aligned}$$

for all $\phi \in H^1_c(M) \cap L^\infty(M)$, *and*

$$-\int_M D\phi\Big(\nabla^{\boldsymbol{\nabla} u}\Big(\frac{F(\boldsymbol{\nabla} u)^2}{2}\Big)\Big)\,dm \geq \int_M \phi\Big\{D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u)+\mathrm{Ric}_N(\boldsymbol{\nabla} u)+\frac{(\boldsymbol{\Delta} u)^2}{N}\Big\}\,dm \quad (3.8)$$

for $N \in [n,\infty]$ *and all nonnegative functions* $\phi \in H^1_c(M) \cap L^\infty(M)$.

Proof. As the proofs are common, we consider only (3.8) with $N < \infty$.

(I) Let us first treat the case of $u \in \mathcal{C}^\infty(M)$. If $\phi \in H^1_c(M_u)$, then (3.8) follows from (3.5) via integration by parts.

For an arbitrary bounded nonnegative function $\phi \in H^1_c(M) \cap L^\infty(M)$, set

$$\phi_k := \min\big\{\phi, k^2 F(\boldsymbol{\nabla} u)^2\big\}, \qquad k \in \mathbb{N}.$$

Then $\phi_k \in H^1_c(M_u)$ and $\lim_{k\to\infty} \phi_k(x) = \phi(x)$ for all $x \in M_u$. Hence we have

$$-\int_M D\phi_k\Big(\nabla^{\boldsymbol{\nabla} u}\Big(\frac{F(\boldsymbol{\nabla} u)^2}{2}\Big)\Big)\,dm \geq \int_M \phi_k\Big\{D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u)+\mathrm{Ric}_N(\boldsymbol{\nabla} u)+\frac{(\boldsymbol{\Delta} u)^2}{N}\Big\}\,dm. \quad (3.9)$$

In the limit as k goes to infinity, the right hand side converges to

$$\int_{M_u} \phi\Big\{D(\boldsymbol{\Delta} u)(\boldsymbol{\nabla} u)+\mathrm{Ric}_N(\boldsymbol{\nabla} u)+\frac{(\boldsymbol{\Delta} u)^2}{N}\Big\}\,dm.$$

The first two terms of the integrand obviously vanish on the set $M \setminus M_u$. In order to see that also the third term vanishes, let us fix a local coordinate $(x^i)_{i=1}^n$ and apply Lemma 3.5 to the function $h := \partial u/\partial x^i \in H^1_{\mathrm{loc}}(M)$. It implies that $\partial h/\partial x^j = 0$ a.e. on $h^{-1}(0)$ for all $j = 1, 2, \dots, n$. In other words, $\partial^2 u/\partial x^i \partial x^j = 0$ for all $i, j = 1, 2, \dots, n$ on $\bigcap_{i=1}^n \{\partial u/\partial x^i = 0\}$. Hence, in particular, $\boldsymbol{\Delta} u = 0$ a.e. on $M \setminus M_u$.

To see the passage to the limit on the left hand side of (3.9), we observe from Lemma 3.5 that $\nabla^{\boldsymbol{\nabla} u}(F(\boldsymbol{\nabla} u)^2) = 2F(\boldsymbol{\nabla} u)\nabla^{\boldsymbol{\nabla} u}(F(\boldsymbol{\nabla} u))$ vanishes a.e. on $M \setminus M_u$. Thus, putting

$$\Omega_k := \big\{x \in M_u \,\big|\, \phi(x) > k^2 F\big(\boldsymbol{\nabla} u(x)\big)^2\big\} = \{x \in M_u \,|\, \phi(x) \neq \phi_k(x)\},$$

we find

$$\begin{aligned}
&\Big|\int_M D(\phi-\phi_k)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)^2\big)\Big)\,dm\Big| \\
&\leq \Big|\int_{\Omega_k} D\phi\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)^2\big)\Big)\,dm\Big| + k^2\int_{\Omega_k} D\big(F(\boldsymbol{\nabla} u)^2\big)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)^2\big)\Big)\,dm.
\end{aligned}$$

The first term of the right hand side tends to zero since Ω_k decreases to a null set as k goes to infinity. For the second term, note that

$$D\big(F(\boldsymbol{\nabla} u)^2\big)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)^2\big)\Big) = 4F(\boldsymbol{\nabla} u)^2 \cdot D\big(F(\boldsymbol{\nabla} u)\big)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)\big)\Big)$$

on M_u. Therefore, by the choice of Ω_k, we obtain

$$k^2 \int_{\Omega_k} D\big(F(\boldsymbol{\nabla} u)^2\big)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)^2\big)\Big)\, dm \le 4 \int_{\Omega_k} \phi \cdot D\big(F(\boldsymbol{\nabla} u)\big)\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} u)\big)\Big)\, dm \\ \to 0 \qquad (k \to \infty).$$

Hence, as k goes to infinity, the left hand side of (3.9) converges to

$$-\int_M D\phi\bigg(\nabla^{\boldsymbol{\nabla} u}\bigg(\frac{F(\boldsymbol{\nabla} u)^2}{2}\bigg)\bigg)\, dm.$$

This proves the claim for $u \in \mathcal{C}^\infty(M)$.

(II) Now let us consider the general case of $u \in H^2_{\mathrm{loc}}(M) \cap \mathcal{C}^1(M)$ with $\boldsymbol{\Delta} u \in H^1_{\mathrm{loc}}(M)$. Choose $u_k \in \mathcal{C}^\infty(M)$ with $u_k \to u$ locally in the H^2-norm as $k \to \infty$. Using integration by parts, we rewrite (3.8) for the function u_k in the form

$$\begin{aligned} &-\int_M D\phi\bigg(\nabla^{\boldsymbol{\nabla} u_k}\bigg(\frac{F(\boldsymbol{\nabla} u_k)^2}{2}\bigg)\bigg)\, dm \\ &\ge -\int_M \mathrm{div}_m(\phi \boldsymbol{\nabla} u_k)\boldsymbol{\Delta} u_k\, dm + \int_M \phi\bigg\{\, \mathrm{Ric}_N(\boldsymbol{\nabla} u_k) + \frac{(\boldsymbol{\Delta} u_k)^2}{N}\,\bigg\}\, dm. \end{aligned}$$

As k goes to infinity, each of the terms appearing in this inequality converges towards the repective term with u in the place of u_k. This finally proves the claim for general u. □

4 Applications

In linear semigroup theory, the estimates as in Theorems 3.3, 3.6 are also called the Γ_2-*inequality* or the *Bakry-Émery curvature-dimension condition* (after [BE]). Applying it to solutions to the heat equation, we obtain Bakry-Émery and Li-Yau type estimates (for $N = \infty$ and $N < \infty$, respectively). We remark that Theorem 2.5 ensures that solutions to the heat equation enjoy enough regularity for applying Theorem 3.6. Throughout the section, we assume the compactness of M. We will consider only global solutions $u : [0,T] \times M \longrightarrow \mathbb{R}$ to the heat equation (more precisely, their continuous versions on $(0,T] \times M$) for simplicity.

Given a global solution $u : [0,T] \times M \longrightarrow \mathbb{R}$ to the heat equation, let us fix a measurable one-parameter family of non-vanishing vector fields $\{V_r\}_{r \in [0,T]}$ on M such that $V_r = \boldsymbol{\nabla} u_r$ on M_{u_r}. For $0 \le s < t \le T$, denote by $P^{\boldsymbol{\nabla} u}_{s,t}$ the (linear, symmetric) Markov transition operator on $L^2(M,m)$ associated with the time-dependent generator Δ^V. Precisely, for each $h \in L^2(M,m)$, $h_t = P^{\boldsymbol{\nabla} u}_{s,t} h$ for $t \in [s,T]$ is the weak solution to $\partial h_t / \partial t = \Delta^{V_t} h_t$ with $h_s = h$.

Theorem 4.1 (Bakry-Émery gradient estimate) *Let (M,F,m) be compact and satisfy $\mathrm{Ric}_\infty \ge K$ for some $K \in \mathbb{R}$, and let $u : [0,T] \times M \longrightarrow \mathbb{R}$ be a global solution to the heat equation. Then we have*

$$F\big(\boldsymbol{\nabla} u_t(x)\big)^2 \le e^{-2K(t-s)} P^{\boldsymbol{\nabla} u}_{s,t}\big(F(\boldsymbol{\nabla} u_s)^2\big)(x) \tag{4.1}$$

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

Proof. For fixed $t \in (0,T]$, u as above and for an arbitrary nonnegative function $h \in \mathcal{C}(M)$, we set

$$H(s) := e^{2Ks} \int_M P_{s,t}^{\boldsymbol{\nabla} u} h \cdot \frac{F(\boldsymbol{\nabla} u_s)^2}{2}\, dm, \qquad 0 \le s \le t.$$

Then we deduce from the definition of $P_{s,t}^{\boldsymbol{\nabla} u}$ that

$$\begin{aligned} H'(s) &= e^{2Ks} \int_M D\left(\frac{F(\boldsymbol{\nabla} u_s)^2}{2}\right)\left(\nabla^{\boldsymbol{\nabla} u_s}(P_{s,t}^{\boldsymbol{\nabla} u} h)\right) dm \\ &\quad + e^{2Ks} \int_M P_{s,t}^{\boldsymbol{\nabla} u} h \cdot \frac{\partial}{\partial s}\left(\frac{F(\boldsymbol{\nabla} u_s)^2}{2}\right) dm + 2KH(s). \end{aligned}$$

Note that, thanks to Lemma 3.5, the integrals in the right hand side are well-defined. By (2.9), the first term in the right hand side coincides with

$$e^{2Ks} \int_M D(P_{s,t}^{\boldsymbol{\nabla} u} h)\left(\nabla^{\boldsymbol{\nabla} u_s}\left(\frac{F(\boldsymbol{\nabla} u_s)^2}{2}\right)\right) dm.$$

The other terms are calculated as, with the help of Theorem 2.2,

$$\begin{aligned} &e^{2Ks} \int_M P_{s,t}^{\boldsymbol{\nabla} u} h \left\{ D\left(\frac{\partial u_s}{\partial s}\right)(\boldsymbol{\nabla} u_s) + KF(\boldsymbol{\nabla} u_s)^2 \right\} dm \\ &= e^{2Ks} \int_M P_{s,t}^{\boldsymbol{\nabla} u} h \left\{ D(\boldsymbol{\Delta} u_s)(\boldsymbol{\nabla} u_s) + KF(\boldsymbol{\nabla} u_s)^2 \right\} dm. \end{aligned}$$

To be precise, we used the positive 0-homogeneity of g_{ij} to see, on M_{u_s},

$$\begin{aligned} \frac{\partial}{\partial s}\left(\frac{F(\boldsymbol{\nabla} u_s)^2}{2}\right) &= \frac{\partial}{\partial s}\left[\frac{1}{2}\sum_{i,j=1}^n g^{ij}(\boldsymbol{\nabla} u_s)\frac{\partial u_s}{\partial x^i}\frac{\partial u_s}{\partial x^j}\right] \\ &= \sum_{i,j=1}^n g^{ij}(\boldsymbol{\nabla} u_s)\frac{\partial u_s}{\partial x^i}\frac{\partial^2 u_s}{\partial s \partial x^j} + \frac{1}{2}\sum_{i,j,k=1}^n \frac{\partial g^{ij}}{\partial v^k}(\boldsymbol{\nabla} u_s)\frac{\partial (\boldsymbol{\nabla} u_s)_k}{\partial s}\frac{\partial u_s}{\partial x^i}\frac{\partial u_s}{\partial x^j} \\ &= D\left(\frac{\partial u_s}{\partial s}\right)(\boldsymbol{\nabla} u_s) - \frac{1}{2}\sum_{i,j,k}\frac{\partial g_{ij}}{\partial v^k}(\boldsymbol{\nabla} u_s)(\boldsymbol{\nabla} u_s)_i(\boldsymbol{\nabla} u_s)_j\frac{\partial (\boldsymbol{\nabla} u_s)_k}{\partial s} \\ &= D\left(\frac{\partial u_s}{\partial s}\right)(\boldsymbol{\nabla} u_s), \qquad \text{where } \boldsymbol{\nabla} u_s = \sum_{i=1}^n (\boldsymbol{\nabla} u)_i \frac{\partial}{\partial x^i}. \end{aligned} \tag{4.2}$$

Therefore we have $H'(s) \le 0$ by the weak formulation of the Bochner-Weitzenböck formula (3.8) (for $\phi = P_{s,t}^{\boldsymbol{\nabla} u} h$ and $N = \infty$) together with the hypothesis $\mathrm{Ric}_\infty \ge K$.

Thus H is non-increasing and, in particular,

$$\begin{aligned} e^{2Kt} \int_M h\, \frac{F(\boldsymbol{\nabla} u_t)^2}{2}\, dm &\le e^{2Ks} \int_M P_{s,t}^{\boldsymbol{\nabla} u} h \cdot \frac{F(\boldsymbol{\nabla} u_s)^2}{2}\, dm \\ &= e^{2Ks} \int_M h \cdot P_{s,t}^{\boldsymbol{\nabla} u}\left(\frac{F(\boldsymbol{\nabla} u_s)^2}{2}\right) dm. \end{aligned}$$

Since this holds true for any nonnegative h, we obtain the claimed inequality (4.1) for a.e. $x \in M$. The Hölder continuity in x of both sides of (4.1) finally allows to deduce it for every $x \in M$. □

We stress that it is used in (4.1) the mixture of the nonlinear operator $u_s \longmapsto u_t$ and the linear one $P_{s,t}^{\boldsymbol{\nabla} u}$. As an immediate corollary to Theorem 4.1, we obtain a growth bound for Lipschitz constants. For a continuous function $u \in \mathcal{C}(M)$, define

$$\mathrm{Lip}(u) := \sup_{x,y \in M} \frac{f(y) - f(x)}{d(x,y)}.$$

Corollary 4.2 *Assume that (M, F, m) is compact and satisfies $\mathrm{Ric}_\infty \geq K$ for some $K \in \mathbb{R}$, and let $u : [0,T] \times M \longrightarrow \mathbb{R}$ be a global solution to the heat equation. Then we have*

$$\|F(\boldsymbol{\nabla} u_t)\|_{L^\infty} \leq e^{-Kt} \|F(\boldsymbol{\nabla} u_0)\|_{L^\infty}$$

and, if $u_0 \in \mathcal{C}(M)$,

$$\mathrm{Lip}(u_t) \leq e^{-Kt} \mathrm{Lip}(u_0)$$

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

Remark 4.3 In the Riemannian case, each of the properties in Corollary 4.2 characterizes manifolds with $\mathrm{Ric}_\infty \geq K$ ([vRS, Theorem 1.3]), and $\mathrm{Ric}_\infty \geq K$ is also equivalent to the K-contraction property of heat flow in the L^2-Wasserstein space ([vRS, Corollary 1.4], see (1.1)). The equivalence between the gradient estimate and the contraction property is known to hold for more general linear semigroups by Kuwada's duality [Ku]. Comparing Theorem 4.1 and Corollary 4.2 with the non-contraction property of heat flow established in [OS2], we observe that a completely different phenomenon occurs for nonlinear semigroups.

Next we consider Li-Yau type estimates for $N < \infty$ (cf. [LY] and [Da]).

Theorem 4.4 (Li-Yau gradient estimate) *Let us assume that (M, F, m) is compact and satisfies $\mathrm{Ric}_N \geq K$ for some $N \in [n, \infty)$ and $K \in \mathbb{R}$, put $K' := \min\{K, 0\}$ and take a positive global solution $u : [0,T] \times M \longrightarrow \mathbb{R}$ to the heat equation. Then we have*

$$F\big(\boldsymbol{\nabla}(\log u)(t,x)\big)^2 - \theta \partial_t (\log u)(t,x) \leq N\theta^2 \left(\frac{1}{2t} - \frac{K'}{4(\theta - 1)} \right)$$

on $(0,T] \times M$ for any $\theta > 1$.

Proof. We follow the argumentation in [Da, Lemmas 5.3.2, 5.3.3] and focus on the modifications required because of nonlinearity and the lack of higher order regularity. Throughout the proof, we fix a measurable one-parameter family of non-vanishing vector fields $\{V_r\}_{r \in [0,T]}$ as described in the paragraph preceding to Theorem 4.1. We divide the proof into five steps. Note that it suffices to show the claim for $t = T$.

(I) First consider the function $f := \log u$ which is H^2 in space and $\mathcal{C}^{1,\alpha}$ in both space and time. As u solves the heat equation, f satisfies

$$\boldsymbol{\Delta} f_t + F(\boldsymbol{\nabla} f_t)^2 = \partial_t f_t \tag{4.3}$$

for every t in the weak sense that

$$\int_M \{-D\phi(\boldsymbol{\nabla} f_t) + \phi F(\boldsymbol{\nabla} f_t)^2\}\, dm = \int_M \phi\, \partial_t f_t\, dm$$

for each $\phi \in H^1(M)$. Note also that $g_{\boldsymbol{\nabla} f_t} = g_{\boldsymbol{\nabla} u_t}$ a.e. on M_{u_t} and $\boldsymbol{\Delta} f_t \in H^1(M)$ for each t.

(II) Next we verify that the function $\sigma(t,x) := t\partial_t f(t,x)$ ($\in H^1(M)$ by Theorem 2.5) satisfies

$$\Delta^V \sigma - \partial_t \sigma + \frac{\sigma}{t} + 2D\sigma(\boldsymbol{\nabla} f) = 0 \tag{4.4}$$

in the sense of distributions on $(0,T)\times M$. Indeed, for each $\phi \in H_0^1((0,T)\times M)$,

$$\begin{aligned}
&\int_0^T \int_M \Big\{ - D\phi(\nabla^V \sigma) + \partial_t \phi \cdot \sigma + \frac{\phi\sigma}{t} + 2\phi D\sigma(\boldsymbol{\nabla} f) \Big\}\, dmdt \\
&= \int_0^T \int_M \big\{ - D(t\phi)\big(\nabla^V(\partial_t f)\big) + \partial_t(t\phi)\partial_t f + 2t\phi D(\partial_t f)(\boldsymbol{\nabla} f) \big\}\, dmdt \\
&= \int_0^T \int_M \big\{ D\big(\partial_t(t\phi)\big)(\boldsymbol{\nabla} f) + \partial_t(t\phi)\big(\boldsymbol{\Delta} f + F(\boldsymbol{\nabla} f)^2\big) + t\phi\partial_t\big(F(\boldsymbol{\nabla} f)^2\big) \big\}\, dmdt \\
&= 0,
\end{aligned}$$

where we used (4.3) and (4.2) in the second equality. To be precise, a calculation similar to (4.2) ensures

$$\partial_t\big(D(t\phi)(\boldsymbol{\nabla} f)\big) = D\big(\partial_t(t\phi)\big)(\boldsymbol{\nabla} f) + D(t\phi)\big(\nabla^V(\partial_t f)\big).$$

(III) Now let us consider the function $\alpha(t,x) := t\{F(\boldsymbol{\nabla} f(t,x))^2 - \theta\partial_t f(t,x)\}$. It lies in $H^1(M)$ for each t and is Hölder continuous in both space and time. Using the previous identity (4.4) for σ as well as our version of the Bochner-Weitzenböck formula, we shall deduce that

$$\Delta^V \alpha + 2D\alpha(\boldsymbol{\nabla} f) - \partial_t \alpha \geq \beta \tag{4.5}$$

again in the distributional sense on $(0,T)\times M$, where β denotes the continuous function

$$\beta(t,x) := -\frac{\alpha}{t} + 2t\bigg\{ \frac{1}{N}\big(F(\boldsymbol{\nabla} f)^2 - \partial_t f\big)^2 + KF(\boldsymbol{\nabla} f)^2 \bigg\}.$$

Indeed, for each nonnegative $\phi \in H_0^1((0,T)\times M)$, (4.4), (4.2) and (4.3) show

$$\begin{aligned}
&\int_0^T \int_M \Big\{ - D\phi(\nabla^V \alpha) + 2\phi D\alpha(\boldsymbol{\nabla} f) + \partial_t \phi \cdot \alpha + \frac{\phi\alpha}{t} \Big\}\, dmdt \\
&= \int_0^T \int_M \Big\{ - tD\phi\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} f)^2\big)\Big) + 2t\phi D\big(F(\boldsymbol{\nabla} f)^2\big)(\boldsymbol{\nabla} f) \\
&\qquad - \phi\big\{F(\boldsymbol{\nabla} f)^2 + 2tD(\partial_t f)(\boldsymbol{\nabla} f)\big\} + \phi F(\boldsymbol{\nabla} f)^2 \Big\}\, dmdt \\
&= \int_0^T \int_M \Big\{ - tD\phi\Big(\nabla^{\boldsymbol{\nabla} u}\big(F(\boldsymbol{\nabla} f)^2\big)\Big) - 2t\phi D(\boldsymbol{\Delta} f)(\boldsymbol{\nabla} f) \Big\}\, dmdt.
\end{aligned}$$

Thanks to (3.8) and the hypothesis $\mathrm{Ric}_N \geq K$, this is estimated from below by

$$2t\int_0^T \int_M \phi\bigg\{ KF(\boldsymbol{\nabla} f)^2 + \frac{(\boldsymbol{\Delta} f)^2}{N} \bigg\}\, dmdt$$

and (4.3) completes the proof of (4.5).

(IV) Let (s,x) denote the maximizer of α on $[0,T]\times M$ (note that M is compact). Without restriction, $\alpha(s,x)>0$ and thus $s>0$ (otherwise, the assertion of the theorem is obvious). We shall claim that this implies $\beta(s,x)\le 0$. Assume the contrary, $\beta(s,x)>0$. It would imply $\beta>0$ on a neighborhood of (s,x). Hence, according to (4.5) on such a neighborhood, the function α would be a strict sub-solution to the linear parabolic operator

$$\mathrm{div}_m(\nabla^V\alpha)+2D\alpha(\boldsymbol{\nabla} f)-\partial_t\alpha.$$

Therefore, $\alpha(s,x)$ would be strictly less than the supremum of α on the boundary of any small parabolic cylinder $[s-\delta,s]\times B_\delta(x)$, where $B_\delta(x):=\{y\in M\,|\,d(x,y)<\delta\}$. In particular, α could not be maximal at (s,x).

(V) The conclusion of the previous step is that $\beta(s,x)\le 0$. In other words,

$$\alpha\ge 2s^2\left\{\frac{1}{N}\big(F(\boldsymbol{\nabla} f)^2-\partial_t f\big)^2+KF(\boldsymbol{\nabla} f)^2\right\}$$

at the maximum point (s,x) of α. Following the reasoning of [Da, Lemma 5.3.3], we conclude

$$\alpha(s,x)\le\frac{N\theta^2}{2}\left(1-\frac{K's}{2(\theta-1)}\right)$$

and thus, since $K'\le 0$,

$$\alpha(T,y)\le\alpha(s,x)\le\frac{N\theta^2}{2}\left(1-\frac{K'T}{2(\theta-1)}\right)$$

for all $y\in M$. This completes the proof. □

We proceed along the line of [Da, Theorem 5.3.5] and obtain the Harnack inequality.

Theorem 4.5 (Li-Yau Harnack inequality) *Assume that (M,F,m) is compact and satisfies $\mathrm{Ric}_N\ge K$ for some $N\in[n,\infty)$ and $K\in\mathbb{R}$, put $K':=\min\{K,0\}$ and take a nonnegative global solution $u:[0,T]\times M\longrightarrow\mathbb{R}$ to the heat equation. Then we have*

$$u(s,x)\le u(t,y)\cdot\left(\frac{t}{s}\right)^{\theta N/2}\exp\left(\frac{\theta d(x,y)^2}{4(t-s)}-\frac{\theta K'N(t-s)}{4(\theta-1)}\right)$$

for any $\theta>1$, $0<s<t\le T$ and $x,y\in M$.

Proof. Replacing u by $u+\varepsilon$ if necessary, we may assume without restriction that u is positive. Let $\eta(\tau)=\exp_x(\tau v)$ for $\tau\in[s,t]$ be the minimal geodesic from $x=\eta(s)$ to $y=\eta(t)$ with suitable $v\in T_xM$. Then obviously $F(\dot\eta(\tau))=d(x,y)/(t-s)$ for all τ. We also put $f:=\log u$,

$$\Theta:=\frac{\theta d(x,y)^2}{4(t-s)^2}-\frac{\theta K'N}{4(\theta-1)},\qquad \sigma(\tau):=f\big(\tau,\eta(\tau)\big)+\frac{\theta N}{2}\log\tau+\Theta\tau.$$

Then we have, by Theorem 4.4,

$$\begin{aligned}\partial_\tau \sigma(\tau) &= Df(\dot\eta) + \partial_t f + \frac{\theta N}{2\tau} + \Theta \\ &\geq -F(\boldsymbol{\nabla} f)F(\dot\eta) + \frac{F(\boldsymbol{\nabla} f)^2}{\theta} - N\theta\left(\frac{1}{2t} - \frac{K'}{4(\theta-1)}\right) + \frac{\theta N}{2\tau} + \Theta \\ &\geq -F(\boldsymbol{\nabla} f)\frac{d(x,y)}{t-s} + \frac{F(\boldsymbol{\nabla} f)^2}{\theta} + \frac{\theta d(x,y)^2}{4(t-s)^2} \geq 0.\end{aligned}$$

Hence we obtain

$$u(s,x)\cdot s^{\theta N/2}e^{\Theta s} = e^{\sigma(s)} \leq e^{\sigma(t)} = u(t,y)\cdot t^{\theta N/2}e^{\Theta t}$$

which proves the claim. □

In both Theorems 4.4 and 4.5, we can choose $\theta = 1$ if $K' = 0$. Thus we have the following.

Corollary 4.6 *Let (M, m, F) be compact with $\mathrm{Ric}_N \geq 0$ for some $N \in [n, \infty)$, and take a nonnegative global solution $u : [0,T] \times M \longrightarrow \mathbb{R}$ to the heat equation. Then we have*

$$u(s,x) \leq u(t,y)\cdot\left(\frac{t}{s}\right)^{N/2}\exp\left(\frac{d(x,y)^2}{4(t-s)}\right)$$

for any $0 < s < t \leq T$ and $x, y \in M$.

References

[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.

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

[BH] N. Bouleau and F. Hirsch, Dirichlet forms and analysis on Wiener space, Walter de Gruyter & Co., Berlin, 1991.

[Da] E. B. Davies, Heat kernels and spectral theory, Cambridge University Press, Cambridge, 1989.

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

[LY] P. Li and S.-T. Yau, On the parabolic kernel of the Schrödinger operator, Acta Math. **156** (1986), 153–201.

[LV1] J. Lott and C. Villani, Hamilton-Jacobi semigroup on length spaces and applications, J. Math. Pures Appl. **88** (2007), 219–229.

[LV2] J. Lott and C. Villani, Weak curvature conditions and functional inequalities, J. Funct. Anal. **245** (2007), 311–333.

[LV3] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. **169** (2009), 903–991.

[Oh1] S. Ohta, Finsler interpolation inequalities, Calc. Var. Partial Differential Equations **36** (2009), 211–249.

[Oh2] S. Ohta, Optimal transport and Ricci curvature in Finsler geometry, Adv. Stud. Pure Math. **57** (2010), 323–342.

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

[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, Preprint (2010). Available at arXiv:1009.2312

[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.

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

[St1] K.-T. Sturm, Convex functionals of probability measures and nonlinear diffusions on manifolds, J. Math. Pures Appl. **84** (2005), 149–168.

[St2] K.-T. Sturm, On the geometry of metric measure spaces. I, Acta Math. **196** (2006), 65–131.

[St3] K.-T. Sturm, On the geometry of metric measure spaces. II, Acta Math. **196** (2006), 133–177.

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