ESAIM: COCV 23 (2017) 827–850
DOI: 10.1051/cocv/2016015

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

GENERALIZED LI−YAU ESTIMATES AND HUISKEN'S MONOTONICITY FORMULA

Paul W.Y. Lee[1]

Abstract. We prove a generalization of the Li−Yau estimate for a broad class of second order linear parabolic equations. As a consequence, we obtain a new Cheeger−Yau inequality and a new Harnack inequality for these equations. We also prove a Hamilton−Li−Yau estimate, which is a matrix version of the Li−Yau estimate, for these equations. This results in a generalization of Huisken's monotonicity formula for a family of evolving hypersurfaces. Finally, we also show that all these generalizations are sharp in the sense that the inequalities become equality for a family of fundamental solutions, which however different from the Gaussian heat kernels on which the equality was achieved in the classical case.

Mathematics Subject Classification. 58J35.

Received August 20, 2015. Accepted March 25, 2016.

1. Introduction

The Harnack inequality is one of the most fundamental results in the regularity theory of non-linear elliptic and parabolic equations. In the case of linear parabolic equations in divergence form, this inequality was first done in [26]. A sharp version of this inequality which takes into account the geometry of the underlying manifold was first done in [24]. In fact, the key result in [24] is a sharp gradient estimate, now known as the Li−Yau estimate, for linear parabolic equations on Riemannian manifolds with a lower bound on the Ricci curvature. The sharp Harnack inequality can be obtained by integrating this estimate along geodesics. Because of this, this estimate and its generalizations are called differential Harnack inequalities.

There are numerous generalizations of the Li−Yau estimate. In the case of geometric evolution equations, this includes the evolution equations for hypersurfaces [2,14,21], the Yamabe flow [13], the Ricci flow [6,19] and its Kähler analogue [8,28]. For a more detail account of these generalizations as well as further developments, see [27].

In the case of the heat equation $\dot\rho_t = \Delta\rho_t$ on a Riemannian manifold of dimension n with non-negative Ricci curvature, the Li−Yau estimate is the following inequality for any positive solution ρ_t

$$\Delta \log \rho_t \geq -\frac{n}{2t} \tag{1.1}$$

for any time $t > 0$.

Keywords and phrases. Differential Harnack inequality, monotonicity formula.

[1] Room 216, Lady Shaw Building, The Chinese University of Hong Kong, Shatin, Hong Kong, P.R. China.
wylee@math.cuhk.edu.hk

This is sharp in the sense that the equality case of (1.1) is satisfied by the following solution of the heat equation on the Euclidean space:

$$\rho_t(x) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|x|^2}{4t}\right),$$

where $|x|$ is the Euclidean norm of x in $\mathbb{R}^n$.

On the other hand, there are also generalizations of the inequality (1.1) to other second order linear parabolic equations under the, so called, curvature-dimension conditions (see for instance [4]). They are estimates of the form

$$L \log \rho_t \geq -\frac{n}{2t}, \tag{1.2}$$

where L is a linear differential operator without constant term and ρ_t is a solution of the equation $\dot{\rho}_t = L\rho_t$.

However, the following

$$\rho_t(x) = \left(\frac{2\pi(\exp(2tk)-1)}{k\exp(2tk)}\right)^{-n/2} \exp\left(\frac{-k|x|^2}{2(\exp(2tk)-1)}\right). \tag{1.3}$$

is a solution of the equation

$$\dot{\rho}_t = \Delta\rho_t - k\langle x, \nabla\rho_t\rangle = \Delta\rho_t - \left\langle \nabla\left(\frac{k}{2}|x|^2\right), \nabla\rho_t\right\rangle \tag{1.4}$$

where $k > 0$ is a constant.

The solutions (1.3) never satisfy the equality case of (1.2). Motivated by this observation, we prove the following generalization of (1.1).

Theorem 1.1 (Generalized Li$-$Yau estimate)**.** *Assume that the Ricci curvature of the underlying Riemannian manifold M is non-negative. Let $U_1, U_2 : M \to \mathbb{R}$ be two smooth functions and let*

$$V := \Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2.$$

Assume that $|\nabla V|$ is bounded and $\Delta V \leq nk^2$. Let ρ_t be a positive solution of the equation

$$\dot{\rho}_t = \Delta\rho_t + \langle\nabla U_1, \nabla\rho_t\rangle + U_2\rho_t.$$

Then ρ_t satisfies

$$\Delta \log\rho_t + \frac{1}{2}\Delta U_1 \geq -\frac{nk}{2}\coth(kt). \tag{1.5}$$

By letting $U_1 \equiv 0$, $U_2 \equiv 0$ and k goes to 0, we recover the estimate (1.1). Note also that the solution (1.3) achieves the equality case of (1.5) and the assumptions of Theorem 1.1 with $U_1 = -\frac{k}{2}|x|^2$ and $U_2 \equiv 0$.

Recall that we can obtain the Harnack inequality by integrating (1.1) along geodesics. An analogue of this fact also holds true in our setting. However, instead of integrating along geodesics, the correct paths in this case are the minimizers of the following functional:

$$c_{s,t}(x,y) = \inf_{\gamma(s)=x, \gamma(t)=y} \int_s^t \frac{1}{2}|\dot{\gamma}(\tau)|^2 + V(\gamma(\tau))\mathrm{d}\tau, \tag{1.6}$$

where the infimum is taken over all paths $\gamma : [s,t] \to M$ joining x and y and $V = \Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2$. The idea of considering functionals of the form (1.6) already appeared in [24]. In the case of the Ricci flow, a version of the cost function (1.6), called L-distance, appeared in [29].

Theorem 1.2 (Generalized Harnack inequality)**.** *Under the same assumptions as in Theorem 1.1 and that V is bounded below, the following estimate holds:*

$$\frac{\rho_t(y)}{\rho_s(x)} \geq \left(\frac{\sinh(kt)}{\sinh(ks)}\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\left(c_{s,t}(x,y) + U_1(y) - U_1(x)\right)\right).$$

By letting $U_1 \equiv 0$, $U_2 \equiv 0$, and k goes to 0, we recover the following Harnack estimate.

Corollary 1.3 (The Harnack inequality [24,26])**.** *Assume that the Ricci curvature of the underlying Riemannian manifold M is non-negative. Then any positive solution ρ_t of the equation $\dot{\rho}_t = \Delta\rho_t$ satisfies the following estimate:*

$$\frac{\rho_t(y)}{\rho_s(x)} \geq \left(\frac{t}{s}\right)^{-n/2} \mathrm{e}^{-\frac{d^2(x,y)}{4(t-s)}}.$$

It is also known that Corollary 1.3 recovers the heat kernel comparison theorem of Cheeger–Yau [11] if we let ρ_t be the heat kernel and letting s goes to 0. The same principle also works for Theorem 1.2.

Theorem 1.4 (Generalized Cheeger–Yau comparison theorem)**.** *Under the same assumptions as in Theorem 1.2, the following estimate holds for the fundamental solution p_t of the equation $\dot{\rho}_t = \Delta\rho_t + \langle\nabla U_1, \nabla\rho_t\rangle + U_2\rho_t$:*

$$p_t(x,y) \geq \left(\frac{k}{4\pi\sinh(kt)}\right)^{\frac{n}{2}} \exp\left(-\frac{1}{2}\left(c_{0,t}(x,y) + U_1(y) - U_1(x)\right)\right). \tag{1.7}$$

In the case $U_1 = -\frac{k|x|^2}{2}$ and $U_2 \equiv 0$, the cost function is given by

$$c_{0,t}(0,y) = \frac{k|y|^2\coth(kt)}{2} - knt$$

(see the proof of Thm. 3.4) and right hand side of (1.7) becomes the fundamental solution (1.3). Therefore, all inequalities in Theorem 1.4 become equalities in this case.

Again, by setting $U_1 \equiv 0$, $U_2 \equiv 0$ and letting k goes to 0, we recover the Cheeger–Yau estimate.

Corollary 1.5. [11] (The Cheeger–Yau heat kernel comparison)**.** *Assume that the Ricci curvature of the underlying Riemannian manifold M is non-negative. Then the heat kernel $p_t(x,y)$ of the equation $\dot{\rho}_t = \Delta\rho_t$ satisfy the following estimate:*

$$p_t(x,y) \geq \frac{1}{(4\pi t)^{n/2}}\mathrm{e}^{-\frac{d^2(x,y)}{4t}}.$$

As another consequence of Theorem 1.1, we obtain the following Liouville type theorem.

Corollary 1.6 (A Liouville type theorem)**.** *Assume that the Ricci curvature of the underlying Riemannian manifold M is non-negative. Suppose that $|\nabla V|$ is bounded and $\Delta V \leq nk^2$. Then any positive solution ρ of the equation*

$$\Delta\rho + \langle\nabla U_1, \nabla\rho\rangle + U_2\rho = 0 \tag{1.8}$$

satisfies

$$\left|\nabla\log\rho + \frac{1}{2}\nabla U_1\right|^2 \leq \frac{1}{2}V.$$

In particular, if $V(x) < 0$ at some point x in M, then equation (1.8) does not admit any positive solution. If $V \equiv 0$, then there is a positive constant C such that

$$\rho = C\mathrm{e}^{-\frac{1}{2}U_1}.$$

As a special case of Corollary 1.7, we recover the following result in [33].

Corollary 1.7 (The Liouville theorem). *Assume that the Ricci curvature of the underlying Riemannian manifold M is non-negative. Then any non-negative harmonic function is a constant.*

In [18], Hamilton proved a matrix version of (1.1) for the heat equation, called the Hamilton–Li–Yau estimate (see also a Kähler analogue in [27]). Another matrix version of the differential Harnack inequality also appeared in [19] which is one of the most fundamental result in the theory of the Ricci flow (see also an interesting generalization in [6] and a Kähler analogue in [9]). The following is a matrix version of (1.1).

Theorem 1.8 (Generalized Hamilton–Li–Yau estimate). *Assume that the sectional curvature of the underlying compact Riemannian manifold M is non-negative and the Ricci curvature is parallel. Let $U_1, U_2 : M \to \mathbb{R}$ be two smooth functions satisfying the following condition for some non-negative constant k:*

$$\nabla^2 \left(\Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2 \right) \le k^2 I.$$

Then any positive solution ρ_t of the equation $\dot{\rho}_t = \Delta \rho_t + \langle \nabla U_1, \nabla \rho_t \rangle + U_2 \rho_t$ satisfies the following estimate:

$$\nabla^2 \log \rho_t + \frac{1}{2}\nabla^2 U_1 \ge -\frac{k \coth(kt)}{2} I,$$

where ∇^2 denotes the Hessian operator.

Once again, if the underlying manifold is $\mathbb{R}^n$, $U_1(x) = -\frac{k}{2}|x|^2$, and $U_2 \equiv 0$, then

$$\nabla^2 \left(\Delta U_1 + \frac{1}{2}|\nabla U_1|^2 \right) = k^2 I,$$

and

$$\nabla^2 \log \rho_t + \frac{1}{2}\nabla^2 U_1 = -\frac{k}{2}\coth(kt) I.$$

Therefore, the inequalities in Theorem 1.8 are equalities in this case.

By setting $U_1 \equiv 0$, $U_2 \equiv 0$, and letting $k \to 0$, we recover

Theorem 1.9 (The Hamilton–Li–Yau estimate [18]). *Assume that the sectional curvature of the underlying compact Riemannian manifold M is non-negative and the Ricci curvature is parallel. Then any positive solution ρ_t of the equation $\dot{\rho}_t = \Delta \rho_t$ satisfies the following estimate:*

$$\nabla^2 \log \rho_t \ge -\frac{1}{2t} I.$$

In [20], Theorem 1.9 was used to prove a generalization of Huisken's monotonicity formula for the mean curvature flow [22]. More precisely, let M be a m-dimensional sub-manifold of a n-dimensional Riemannian manifold N. Let $\varphi_t : M \to N$ be a family of immersions evolved according to the following equation

$$\dot{\varphi}_t = \mathfrak{H}_t(\varphi_t), \tag{1.9}$$

where $\mathfrak{H}_t$ is the mean curvature vector of the sub-manifold $M_t := \varphi_t(M)$.

Theorem 1.10 (Huisken's monotonicity formula [20,22]). *Assume that the sectional curvature of the underlying compact Riemannian manifold N is non-negative and the Ricci curvature is parallel. Let φ_t be a solution of* (1.9) *and let ρ_t be a positive solution of the heat equation $\dot{\rho}_t = \bar{\Delta}\rho_t$ on N. Here $\bar{\Delta}$ denotes the Laplacian operator on N. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left((T-t)^{\frac{n-m}{2}}\int_{\varphi_t(M)}\rho_{T-t}\,\mathrm{d}\mu_t\right) \le -(T-t)^{\frac{n-m}{2}}\int_{\varphi_t(M)}\rho_{T-t}\left(\left|\frac{\nabla_t^{\perp}u_t}{u_t}-\boldsymbol{\mathfrak{H}}_t\right|^2\right)\mathrm{d}\mu_t,$$

where μ_t is the Riemannian volume of M_t, $\bar{\nabla}u$ is the gradient of u on N, and $\nabla_t^{\perp}u_t$ is the projection of $\bar{\nabla}u$ onto the normal bundle of M_t.

In particular, the quantity $(T-t)^{\frac{n-m}{2}}\int_{\varphi_t(M)}\rho_{T-t}\,\mathrm{d}\mu_t$ is monotone.

There is an analogue of this monotonicity formula in the setting of Theorem 1.8. In this case, the evolving hypersurfaces M_t satisfy the following equation instead

$$\dot{\varphi}_t = \boldsymbol{\mathfrak{H}}_t(\varphi_t) + \nabla_t^{\perp}U. \tag{1.10}$$

We remark that the term $\boldsymbol{\mathfrak{H}}_t(\varphi_t) + \nabla_t^{\perp}U$ is a generalization of mean curvature first appeared in [17]. In particular, equation (1.10) is the gradient flow of the weighted volume functional

$$\int_{\varphi(M)} \mathrm{e}^{-U}\mathrm{d}\nu,$$

where ν is the Riemannian volume on $\varphi(M)$ induced by the one on N.

Special cases of the equation was also studied in [5, 30].

Theorem 1.11 (Generalized Huisken's monotonicity formula)**.** *Assume that the sectional curvature of the underlying compact Riemannian manifold N is non-negative and the Ricci curvature is parallel. Let $U\colon M \to \mathbb{R}$ be a smooth function satisfying the following condition for some positive constant k:*

$$\bar{\nabla}^2\left(-\Delta U + \frac{1}{2}|\nabla U|^2\right) \le k^2 I,$$

where $\bar{\nabla}^2$ is the Hessian operator on N. Let φ_t be a solution of (1.10) and let ρ_t be a positive solution of the equation

$$\dot{\rho}_t = \bar{\Delta}\rho_t + \langle\bar{\nabla}U, \bar{\nabla}\rho_t\rangle + \rho_t\bar{\Delta}U$$

on N. Then

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\left(\sinh^{\frac{n-m}{2}}(k(T-t))\int_{\varphi_t(M)}\rho_{T-t}\,\mathrm{d}\mu_t\right)\\ &\le -\sinh^{\frac{n-m}{2}}(k(T-t))\int_{\varphi_t(M)}\rho_{T-t}\left(\frac{1}{2}\Delta_t^{\perp}U + \left|\frac{\bar{\nabla}^{\perp}u_t}{u_t}-\boldsymbol{\mathfrak{H}}_t\right|^2\right)\mathrm{d}\mu_t,\end{aligned}$$

where $\Delta_t^{\perp}U$ is defined by $\Delta_t^{\perp}U = \sum_k \langle\bar{\nabla}_{\boldsymbol{n}_k(t)}U, \boldsymbol{n}_k(t)\rangle$.

As an immediate consequence, we have

Corollary 1.12. *Assume that the sectional curvature of the underlying compact Riemannian manifold N is non-negative and the Ricci curvature is parallel. Let $U\colon M \to \mathbb{R}$ be a smooth function satisfying the following condition for some constants K and k with $k > 0$:*

$$\bar{\nabla}^2\left(-\Delta U + \frac{1}{2}|\nabla U|^2\right) \le k^2 I \quad \textit{and} \quad \bar{\nabla}^2 U \ge KI.$$

Let φ_t be a solution of (1.10) *and let ρ_t be a positive solution of equation*

$$\dot{\rho}_t = \bar{\Delta}\rho_t + \langle \bar{\nabla} U, \bar{\nabla}\rho_t \rangle + \rho_t \bar{\Delta} U$$

on N. Then

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{-\frac{K(n-m)(T-t)}{2}} \sinh^{\frac{n-m}{2}}(k(T-t)) \int_{\varphi_t(M)} \rho_{T-t}\,\mathrm{d}\mu_t\right) \\ &\le -\mathrm{e}^{-\frac{K(n-m)(T-t)}{2}} \sinh^{\frac{n-m}{2}}(k(T-t)) \int_{\varphi_t(M)} \rho_{T-t}\left(\left|\frac{\nabla_t \perp u_t}{u_t} - \mathfrak{H}_t\right|^2\right)\mathrm{d}\mu_t. \end{aligned}$$

In particular, $\mathrm{e}^{-\frac{K(n-m)(T-t)}{2}} \sinh^{\frac{n-m}{2}}(k(T-t)) \int_{\varphi_t(M)} \rho_{T-t}\,\mathrm{d}\mu_t$ *is monotone.*

Remarkably, Corollary 1.12 is also sharp. In this case, we set $M = \mathbb{R}^n$, $U = -\frac{k}{2}|x|^2$, and $K = -k$. Then

$$\rho_t(x) = \left(\frac{2\pi(\exp(2tk) - 1)}{k\exp(2tk)}\right)^{-n/2} \exp\left(\frac{-k|x|^2}{2(\exp(2tk) - 1)}\right)\exp(knt)$$

is a solution of the equation

$$\dot{\rho}_t = \bar{\Delta}\rho_t + \langle \bar{\nabla} U, \bar{\nabla}\rho_t \rangle + \rho_t \bar{\Delta} U = \bar{\Delta}\rho_t - k\langle x, \bar{\nabla}\rho_t\rangle - kn\rho_t.$$

It follows from the proof of Corollary 1.12 that all inequalities in the corollary are equalities in this case.

Assuming that the underlying manifold M is compact, Theorem 1.1 can be proved using the Bochner formula and the maximum principle. However, instead of the Bochner formula, we will prove a general result (Thms. 2.1 and 2.3) using a moving frame argument motivated by the theory of optimal transportation (see [32]). This allows a more unified treatment for Theorems 1.1 and 1.8 under the compactness assumption. In Sections 3 and 4, we show that the above generalization of the Li–Yau estimate and its matrix analogue are simple consequences of Theorems 2.1 and 2.3. In Section 5, we give the proof of the generalized Huisken's monotonicity formula.

The Aronzon–Bénilan estimate is a differential Harnack inequality for the porous medium equation

$$\dot{\rho}_t = \Delta(\rho_t^m).$$

In Section 5, we will prove a generalization of Aronzon–Bénilan estimate using Theorem 2.1 and 2.3. We will prove sharp Laplace and Hessian type comparison theorems for the cost function (7.1) in Section 6. In Section 7, a semigroup proof, in the spirit of [4], of the generalized Li–Yau estimates will be discussed (again assuming M is compact). In Section 8, we give a proof of Theorem 1.1 without any compactness assumption.

2. Preliminaries

In this section, we state and prove general results which will be used in the next few sections. For this, we will introduce some notations. Let M be a n-dimensional compact manifold without boundary equipped with a Riemannian metric denoted by $\langle\cdot,\cdot\rangle$ or g. The corresponding Riemann curvature tensor is denoted by $\mathbf{Rm}$. Let F be a function on the space of all $n \times n$ matrices. We assume that F is invariant under orthogonal changes of variables (*i.e.* $F(O^T AO) = F(A)$ for each orthogonal matrix O). For each linear map $W\colon T_xM \to T_xM$ of the tangent space T_xM at a point x, we set $F(W) = F(\mathcal{W})$, where $\mathcal{W}$ is the matrix with ijth entry equal to $\langle W(v_i), v_j\rangle$ and $\{v_1, \ldots, v_n\}$ is an orthonormal frame at x. This is well-defined since F is invariant under orthogonal changes of variables. Note that this condition is not needed or can be relaxed when the tangent bundle TM of M is parallelizable. For instance, when the manifold is the flat torus, this condition can be completely removed. Finally, if u, v, and w are tangent vectors, then $u \otimes v$ denotes the linear map defined by $u \otimes v(w) = \langle v, w\rangle\, u$.

The following is a generalization of the Li–Yau estimate [24].

Theorem 2.1. *Assume that there is a non-negative function $b_t : M \to \mathbb{R}$, a time dependent vector field Y_t on a compact manifold M, and a fibre-preserving bundle homomorphism $W_t : TM \to TM$ of the tangent bundle TM such that*

(1) $F'(A)(B^2) \geq k_1 F(B)^2$ *for some non-negative constant* k_1*;*
(2) $F'(\nabla X_t)(W_t + \boldsymbol{Rm}(\cdot, X_t)X_t) \geq k_3$ *for some constant* k_3*;*
(3) $F'(\nabla X_t)(\nabla(\dot{X}_t + \nabla_{X_t} X_t) + W_t) + k_2 F(\nabla X_t)^2 \leq F'(\nabla X_t)(b_t \nabla^2(F(\nabla X_t)) + \nabla(F(\nabla X_t)) \otimes Y_t)$*;*
(4) $k_1 + k_2 > 0$.

Then

$$F(\nabla X_t) \leq \frac{1}{k_1 + k_2} a_{(k_1+k_2)k_3}(t),$$

where

$$a_K(t) = \begin{cases} \sqrt{K}\cot(\sqrt{K}\,t) & \text{if } K > 0 \\ \frac{1}{t} & \text{if } K = 0 \\ \sqrt{-K}\coth(\sqrt{-K}\,t) & \text{if } K < 0. \end{cases}$$

Remark 2.2. Note that the above theorem can be further generalized to include situation considered in [4] if F is allowed to depend on X_t, not just ∇X_t. However, we will not pursue this here.

A matrix version of Li–Yau estimate was done by Hamilton [18]. The following is the corresponding matrix version of Theorem 2.1.

Theorem 2.3. *Assume that there is a non-negative function $b_t : M \to \mathbb{R}$, a time dependent vector field Y_t on a compact manifold M, and a fibre-preserving bundle homomorphism $W_t : TM \to TM$ of the tangent bundle TM such that*

(1) $w \mapsto \langle X_t, w\rangle$ *is a closed 1-form;*
(2) $W_t + \boldsymbol{Rm}(\cdot, X_t)X_t \geq k_3 I$ *for some constant* k_3*;*
(3) $\left\langle \nabla_v(\dot{X}_t + \nabla X_t(X_t)), v\right\rangle + k_2 \langle \nabla X_t(\nabla X_t(v)), v\rangle + \langle W_t v, v\rangle \leq b_t \langle \Delta\nabla X_t(v), v\rangle + \langle \nabla_{Y_t}\nabla_v X_t, v\rangle$ *for each eigenvector of the linear map* $w \mapsto \nabla_w X_t$ *with the largest eigenvalue;*
(4) $1 + k_2 > 0$.

Then

$$\nabla X_t \leq \frac{1}{1 + k_2} a_{(1+k_2)k_3}(t) I.$$

As a consequence, we obtain the following estimate on the volume growth of a set under the flow of the vector field X_t if $F = \mathbf{tr}$.

Corollary 2.4. *Under the assumptions of Theorem 2.1 with $F = \boldsymbol{tr}$,*

$$(b_{(k_1+k_2)k_3}(t))^{-\frac{1}{k_1+k_2}} \boldsymbol{vol}(\varphi_t(D))$$

is a decreasing function of time t, where

$$b_K(t) = \begin{cases} \frac{1}{\sqrt{K}}\sin(\sqrt{K}t) & \text{if } K > 0 \\ t & \text{if } K = 0 \\ \frac{1}{\sqrt{-K}}\sinh(\sqrt{-K}t) & \text{if } K < 0. \end{cases}$$

The rest of this section is devoted to the proof of the above mentioned results.

Proof of Theorem 2.1. Let φ_t be the one-parameter family of diffeomorphisms defined by the vector field X_t: $\dot\varphi_t = X_t(\varphi_t)$ and $\varphi_0(x) = x$. Let $\gamma(s)$ be a curve which start from x with initial velocity v: $\gamma(0) = x$ and $\gamma'(0) = v$. Then

$$\frac{D}{\mathrm{d}t}\mathrm{d}\varphi_t(v) = \frac{D}{\mathrm{d}s}\frac{D}{\mathrm{d}t}\varphi_t(\gamma(s))\Big|_{s=0} = \nabla_{\mathrm{d}\varphi_t(v)}X_t.$$

Let $v_1(0),\dots,v_n(0)$ be an orthonormal frame at a point x and let $v_1(t),\dots,v_n(t)$ be the parallel transport of $v_1(0),\dots,v_n(0)$ along the path $\varphi_t(x)$. Let $A(t)$ be the matrix defined by

$$\mathrm{d}\varphi_t(v_i(0)) = \sum_{j=1}^n A_{ij}(t)v_j(t).$$

It follows that

$$\sum_{j=1}^n \dot A_{ij}(t)v_j(t) = \sum_{j=1}^n A_{ij}(t)\nabla_{v_j(t)}X_t.$$

Therefore, if $S_{ij}(t) = \left\langle \nabla_{v_i(t)}X_t, v_j(t)\right\rangle$, then $S(t) = A(t)^{-1}\dot A(t)$ and we have

$$\begin{aligned}\dot S(t) &= -A(t)^{-1}\dot A(t)A(t)^{-1}\dot A(t) + A(t)^{-1}\ddot A(t)\\ &= -S(t)^2 + A(t)^{-1}\ddot A(t).\end{aligned} \tag{2.1}$$

On the other hand, if we differentiate the equation $\dot\varphi_t = X_t(\varphi_t)$, then we get

$$\frac{D}{\mathrm{d}t}\dot\varphi_t = \dot X_t(\varphi_t) + \nabla_{X_t}X_t(\varphi_t)$$

and

$$\frac{D}{\mathrm{d}s}\frac{D}{\mathrm{d}t}\dot\varphi_t(\gamma(s))\Big|_{s=0} = \nabla_{\mathrm{d}\varphi_t(v)}\left(\dot X_t + \nabla_{X_t}X_t\right).$$

By the definition of the Riemann curvature tensor $\mathbf{Rm}$, it follows that

$$\frac{D^2}{\mathrm{d}t^2}\mathrm{d}\varphi_t(v) + \mathbf{Rm}(\mathrm{d}\varphi_t(v), X_t(\varphi_t))X_t(\varphi_t) = \nabla_{\mathrm{d}\varphi_t(v)}\left(\dot X_t + \nabla_{X_t}X_t\right).$$

Therefore, by the definition of the matrix $A(t)$, the following holds

$$\ddot A(t) + A(t)(R(t) - M(t)) = 0,$$

where

$$R_{ij}(t) = \langle \mathbf{Rm}(v_i(t), X_t(\varphi_t(x)))X_t(\varphi_t(x)), v_j(t)\rangle$$

and

$$M_{ij}(t) = \left\langle \nabla_{v_i(t)}\left(\dot X_t + \nabla_{X_t}X_t\right), v_j(t)\right\rangle_{\varphi_t(x)}.$$

By combining this with (2.1), we obtain

$$\dot S(t) + S(t)^2 + R(t) = M(t). \tag{2.2}$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}F(S(t)) + F'(S(t))(S(t)^2 + R(t)) = F'(S(t))(M(t)). \tag{2.3}$$

Let t_0 be the first time where $F(\nabla X_{t_0}(\varphi_{t_0}(x))) = ka_K(t_0)$ for some point x, where $k > 0$. By assumption, we have

$$F'(\nabla X_{t_0})(\nabla(\dot{X}_{t_0} + \nabla_{X_{t_0}} X_{t_0}) + W_t) + k_2 F(\nabla X_{t_0})^2 \leq 0 \tag{2.4}$$

at $\varphi_{t_0}(x)$.

In the matrix notation, we have

$$F'(S(t_0))(M(t_0) + \mathcal{W}(t_0)) + k_2 F(S(t_0))^2 \leq 0,$$

where $\mathcal{W}(t_0)$ be the matrix with ijth entry equal to $\langle W_t(v_i(t)), v_j(t)\rangle$.

By combining this with (2.3) and using the assumptions, we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} F(S(t_0)) + k_1 F(S(t_0))^2 + k_3 &\leq \frac{\mathrm{d}}{\mathrm{d}t} F(S(t_0)) + k_1 F(S(t_0))^2 + F'(S(t_0))(R(t_0) + \mathcal{W}(t_0)) \\
&\leq \frac{\mathrm{d}}{\mathrm{d}t} F(S(t_0)) + F'(S(t_0)^2) + F'(S(t_0))(R(t_0) + \mathcal{W}(t_0)) \\
&= F'(S(t_0))(M(t_0) + \mathcal{W}(t_0)) \\
&\leq -k_2 F(S(t_0))^2.
\end{aligned}$$

By the definition of t_0, we have $ka_K(t_0) = F(S(t_0))$ and $k\dot{a}_K(t_0) \leq \frac{\mathrm{d}}{\mathrm{d}t} F(S(t_0))$. Therefore, the above inequality becomes

$$k\dot{a}_K(t_0) + (k_1 + k_2)k^2 a_K(t_0)^2 + k_3 \leq 0.$$

Since a_K satisfies

$$\dot{a}_K + a_K^2 + K = 0, \tag{2.5}$$

it follows that

$$k((k_1 + k_2)k - 1)a(t_0)^2 + k_3 - kK \leq 0.$$

Therefore, we obtain a contradiction if $k = \frac{1}{k_1+k_2}$ and $K < (k_1 + k_2)k_3$. Hence

$$F(\nabla X_t) < \frac{1}{k_1 + k_2} a_K(t)$$

for all $K < (k_1 + k_2)k_3$. By letting $K \to (k_1 + k_2)k_3$, we obtain

$$F(\nabla X_t) \leq \frac{1}{k_1 + k_2} a_{(k_1+k_2)k_3}(t). \qquad \square$$

Proof of Theorem 2.3. Here, we use the same notations as in the proof of Theorem 2.1. By assumption the one-form $v \mapsto \langle X_t, v\rangle$ is closed. This is equivalent to $\langle \nabla_v X_t, w\rangle = \langle v, \nabla_w X_t\rangle$. It follows that the matrices $S(t)$ are all symmetric. Let t_0 be the first time such that there is a point x and a unit tangent vector v in the tangent space $T_{\varphi_t(x)}M$ at $\varphi_t(x)$ such that $\langle \nabla_v X_{t_0}, v\rangle = \langle S(t_0)v, v\rangle = ka_K(t_0)$. Here v denotes both the vector v and its matrix representation with respect to the orthonormal frame $v_1(t), \ldots, v_n(t)$. In particular, $ka_K(t_0)$ is the largest eigenvalue of $S(t_0)$ with eigenvector v. By parallel translating along geodesics, we extend v to a vector field still denoted by v. It follows that $\nabla v = 0$ and $\Delta v = 0$. Therefore, the following holds by assumption

$$\begin{aligned}
&\left\langle \nabla_v(\dot{X}_{t_0} + \nabla X_{t_0}(X_{t_0})), v \right\rangle + k_2 \langle \nabla X_{t_0}(\nabla X_{t_0}(v)), v\rangle + \langle W_{t_0} v, v\rangle \\
&\leq b_{t_0} \langle \Delta \nabla X_{t_0}(v), v\rangle + \left\langle \nabla_{Y_{t_0}} \nabla_v X_{t_0}, v \right\rangle \\
&\leq b_{t_0} \Delta \langle \nabla X_{t_0}(v), v\rangle + \nabla_{Y_{t_0}} \langle \nabla_v X_{t_0}, v\rangle \leq 0.
\end{aligned}$$

In terms of the matrix notations, the above inequality becomes

$$\left\langle (M(t_0)+k_2S(t_0)^2+\mathcal{W}(t_0))v,v\right\rangle \le 0.$$

This, together with (2.2) and (2.5), gives

$$\begin{aligned}0 &\le \frac{\mathrm{d}}{\mathrm{d}t}\left(\langle S(t)v,v\rangle - ka_K(t)\right)\Big|_{t=t_0}\\ &= -\left\langle S(t_0)^2v,v\right\rangle + \langle (M(t_0)-R(t_0))v,v\rangle + ka_K(t_0)^2+kK\\ &\le -(1+k_2)\left\langle S(t_0)^2v,v\right\rangle - \langle(\mathcal{W}(t_0)+R(t_0))v,v\rangle + ka_K(t_0)^2+kK.\end{aligned}$$

By assumption, $\mathcal{W}(t)+R(t)\ge k_3I$. It follows that

$$k(1-(1+k_2)k)a_K(t_0)^2+kK\ge k_3.$$

Therefore, we obtain a contradiction if $k=\frac{1}{1+k_2}$ and $K<k_3(1+k_2)$. It follows that

$$\nabla X_t \le \frac{a_{k_3(1+k_2)}(t)}{1+k_2}I. \qquad \square$$

Proof of Corollary 2.4. If $F(\nabla X)\ge \mathbf{tr}(\nabla X)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\det A(t)\le F\left(\nabla X_{\varphi_t(x)}\right).$$

It follows that

$$\frac{\det(\mathrm{d}\varphi_{t_1})}{\det(\mathrm{d}\varphi_{t_0})}\le \exp\left(\int_{t_0}^{t_1}F(\nabla X_{\varphi_t(x)})\mathrm{d}t\right)\le \frac{b_{(k_1+k_2)k_3}^{\frac{1}{k_1+k_2}}(t_1)}{b_{(k_1+k_2)k_3}^{\frac{1}{k_1+k_2}}(t_0)},$$

where

$$b_K(t)=\begin{cases}\frac{1}{\sqrt{K}}\sin(\sqrt{K}t) & \text{if } K>0\\ t & \text{if } K=0\\ \frac{1}{\sqrt{-K}}\sinh(\sqrt{-K}t) & \text{if } K<0.\end{cases} \qquad \square$$

3. A generalization of the Li–Yau estimate: The case on compact manifolds

In this section, we prove the following generalization of the Li–Yau estimate.

Theorem 3.1. *Assume that the Ricci curvature of the underlying compact Riemannian manifold M is non-negative. Let U_1 and U_2 be two functions on M satisfying*

$$\Delta\left(-\Delta U_1-\frac{1}{2}|\nabla U_1|^2+2U_2\right)\ge k_3.$$

Then any positive solution ρ_t of the equation

$$\dot{\rho}_t=\Delta\rho_t+\langle\nabla\rho_t,\nabla U_1\rangle+U_2\rho_t. \tag{3.1}$$

satisfies

$$2\Delta\log\rho_t+\Delta U_1\ge -na_{\frac{k_3}{n}}(t).$$

By integrating the above generalization of Li–Yau estimate, one obtains a Harnack inequality. For this, we need to consider the following functional

$$\int_{s_0}^{s_1} \frac{1}{2}|\dot\gamma(\tau)|^2 + V(\gamma(\tau))\mathrm{d}\tau,$$

where $\gamma : [s_0, s_1] \to M$ and $V = \Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2$.

Let c_{s_0,s_1} be the corresponding cost function defined by

$$c_{s_0,s_1}(x,y) = \inf \int_{s_0}^{s_1} \frac{1}{2}|\dot\gamma(\tau)|^2 + V(\gamma(\tau))\mathrm{d}\tau, \tag{3.2}$$

where the infimum is taken over all paths γ satisfying $\gamma(s_0) = x$ and $\gamma(s_1) = y$.

Corollary 3.2. *Under the assumptions of Theorem 3.1, the following holds*

$$\frac{\rho_{s_1}(y)}{\rho_{s_0}(x)} \geq \left(\frac{b_{\frac{k_3}{n}}(s_1)}{b_{\frac{k_3}{n}}(s_0)}\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\left(c_{s_0,s_1}(x,y) + U_1(y) - U_1(x)\right)\right).$$

If we let ρ_t be the fundamental solution $p_t(x,y)$ of equation (3.1) and let $s \to 0$ in Corollary 3.2, then we obtain the following generalization of Cheeger–Yau estimate [11].

Corollary 3.3. *Let p_t be the fundamental solution of equation (3.1). Under the assumptions of Theorem 3.1, the following holds*

$$p_t(x,y) \geq \left(4\pi b_{\frac{k_3}{n}}(t)\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\left(c_{0,t}(x,y) + U_1(y) - U_1(x)\right)\right).$$

Finally, we will show that the equality case in Corollary 3.3 is achieved by (1.3). More precisely,

Theorem 3.4. *Let ρ_t be defined by (1.3), $U_1(x) = -\frac{k}{2}|x|^2$, and $U_2 \equiv 0$. Then*

$$p_t(0,x) = \exp\left(-\frac{1}{2}\left(c_{0,t}(0,x) + U_1(x) - U_1(0)\right)\right)\left(4\pi b_{-k^2}(t)\right)^{-\frac{n}{2}}.$$

Proof of Theorem 3.1. If we specialize Theorem 2.1 to the case where $F = \mathbf{tr}$ and $X_t = \nabla h_t$, then the assumptions of Theorem 2.1 are satisfied if $k_1 = \frac{1}{n}$, $\mathbf{tr}(W_t) + \mathbf{Rc}(X_t, X_t) \geq k_3$, and

$$\Delta\left(\dot h_t + \frac{1}{2}|\nabla h_t|^2\right) + k_2(\Delta h_t)^2 + \mathbf{tr}(W_t) \leq b_t \Delta\Delta h_t + \langle \nabla \Delta h_t, Y_t\rangle. \tag{3.3}$$

Let $h_t = -2\log\rho_t - U_1$. Then the following holds

$$\dot h_t + \frac{1}{2}|\nabla h_t|^2 = \Delta h_t + \Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2.$$

Therefore, under the assumptions of the theorem, (3.3) holds with $k_2 = 0$ and $b_t \equiv 1$. Hence, the result follows from Theorem 2.1. □

Proof of Corollary 3.2. Let γ be a minimizer of (3.2) which satisfies $\gamma(s_0) = x_0$ and $\gamma(s_1) = x_1$. Using the notations in the proof of Theorem 3.1, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}h_t(\gamma(t)) - \frac{1}{2}|\dot\gamma(t)|^2 &\leq \frac{\mathrm{d}}{\mathrm{d}t}h_t(\gamma(t)) - \langle \nabla h_t(\gamma(t)), \dot\gamma(t)\rangle + \frac{1}{2}|\nabla h_t|^2_{\gamma(t)} \\ &= \Delta h_t(\gamma(t)) + \Delta U_1(\gamma(t)) + \frac{1}{2}|\nabla U_1|^2_{\gamma(t)} - 2U_2(\gamma(t)) \\ &\leq n a_{\frac{k_3}{n}}(t) + V(\gamma(t)).\end{aligned}$$

In the last inequality above, we have used Theorem 3.1.

By integrating the above inequality and noting that $\dot{b}_K = b_K a_K$, we obtain

$$h_{s_1}(x_1) - h_{s_0}(x_0) \le c_{s_0,s_1}(x_0,x_1) + n\log(b_{\frac{k_3}{n}}(s_1)) - n\log(b_{\frac{k_3}{n}}(s_0)).$$

By taking exponential of the above inequality, the result follows. □

Proof of Corollary 3.3. By Corollary 3.2, we have

$$\frac{p_t(x,y)}{p_s(x,x)} \ge \exp\left(-\frac{1}{2}\left(c_{s,t}(x,y) + U_1(y) - U_1(x)\right)\right)\left(\frac{b_{\frac{k_3}{n}}(t)}{b_{\frac{k_3}{n}}(s)}\right)^{-\frac{n}{2}}.$$

Since $\lim_{s\to 0}(4\pi s)^{n/2}p_s(x,x) = 1$ (see [16,31]), the above inequality gives

$$p_t(x,y) \ge \exp\left(-\frac{1}{2}\left(c_{0,t}(x,y) + U_1(y) - U_1(x)\right)\right)\left(4\pi b_{\frac{k_3}{n}}(t)\right)^{-\frac{n}{2}}$$

as claimed. □

Proof of Theorem 3.4. In this special case, the cost function (3.2) is given by

$$c_{0,t}(0,y) = \inf \int_0^t \frac{1}{2}|\dot{\gamma}(s)|^2 + V(\gamma(s))\mathrm{d}s, \tag{3.4}$$

where $V(x) = -kn + \frac{1}{2}k^2|x|^2$ and the infimum is taken over all paths γ satisfying $\gamma(0) = 0$ and $\gamma(t) = y$.

If $x(\cdot)$ is a minimizer of the above infimum, then it satisfies the following equations (see [15])

$$\dot{x} = p, \qquad \dot{p} = k^2 x.$$

Since $x(0) = 0$ and $x(t) = y$, it follows that

$$x(s) = \frac{\sinh(ks)}{\sinh(kt)}x(t).$$

If we substitute this back into (3.4), then we obtain

$$c_{0,t}(0,y) = \frac{k|y|^2\coth(kt)}{2} - knt.$$

A computation shows that

$$\begin{aligned} p_t(0,y) &= \exp\left(-\frac{1}{2}\left(c_{0,t}(0,y) - \frac{k}{2}|y|^2\right)\right)(4\pi b_{-k^2}(t))^{-\frac{n}{2}} \\ &= \exp\left(\frac{-k|y|^2}{2(\exp(2kt)-1)}\right)\left(\frac{2\pi(\exp(2tk)-1)}{k\exp(2kt)}\right)^{-\frac{n}{2}} \end{aligned}$$

as claimed. □

4. A generalization of Hamilton's matrix Li–Yau estimate

In this section, we show that the following generalization of Hamilton–Li–Yau estimate is a consequence of Theorem 2.3.

Theorem 4.1. *Assume that the sectional curvature of the underlying compact Riemannian manifold M is non-negative and the Ricci curvature is parallel. Let U_1 and U_2 be two functions on M satisfying*

$$-\nabla^2\left(\Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2\right) \geq k_3 I.$$

Then any solution ρ_t of Equation (3.1) satisfies

$$-2\nabla^2 \log \rho_t - \nabla^2 U_1 \leq a_{k_3}(t) I.$$

Proof. We need the following lemma.

Lemma 4.2. *Assume that the sectional curvature of a Riemannian manifold is non-negative at a point x and the Ricci curvature $\boldsymbol{Rc}$ satisfies $\nabla \boldsymbol{Rc}_x = 0$. Then, for any smooth function f, the following holds*

$$\Delta(\nabla_v \mathrm{d}f(v))(x) \geq \langle \nabla_v \nabla \Delta f, v\rangle_x .$$

Here we consider the Hessian $\nabla \mathrm{d}f$ of f as a self-adjoint operator on T_xM. The vector field v is defined as an eigenvector of the operator $\nabla \mathrm{d}f$ at x corresponding to the largest eigenvalue and it is extended to a neighborhood of x by parallel translation along geodesics.

Proof. Let $e_1, \ldots, e_n$ be an orthonormal frame at x and let us extend them to vector fields defined locally near x by parallel translation along geodesics. It follows that $\nabla v(x) = 0$ and $\nabla_{e_i}\nabla_{e_i} v(x) = 0$ (throughout this proof we sum over repeated indices without mentioning). Therefore,

$$\Delta(\nabla_v \mathrm{d}f(v)) = \nabla_{e_i}\nabla_{e_i}\nabla_v \mathrm{d}f(v).$$

Let α be a $(0,1)$-tensor and β be a $(0,2)$-tensor. By Ricci identity, we have

(1) $\nabla_{v_1}\nabla_{v_2}\alpha(v_3) = \nabla_{v_2}\nabla_{v_1}\alpha(v_3) - \alpha(\mathbf{Rm}(v_1,v_2)v_3)$;

(2) $\nabla_{v_4}\nabla_{v_1}\nabla_{v_2}\alpha(v_3)$
$= \nabla_{v_4}\nabla_{v_2}\nabla_{v_1}\alpha(v_3) - \nabla_{v_4}\alpha(\mathbf{Rm}(v_1,v_2)v_3) - \alpha(\nabla_{v_4}\mathbf{Rm}(v_1,v_2)v_3)$;

(3) $\nabla_{v_1}\nabla_{v_2}\beta(v_3,v_4)$
$= \nabla_{v_2}\nabla_{v_1}\beta(v_3,v_4) - \beta(\mathbf{Rm}(v_1,v_2)v_3,v_4) - \beta(v_3,\mathbf{Rm}(v_1,v_2)v_4)$.

Here, for instance, $\nabla_{v_4}\nabla_{v_1}\nabla_{v_2}\alpha(v_3)$ denotes

$$\nabla(\nabla(\nabla\alpha))(v_4,v_1,v_2,v_3).$$

It follows that

$$\begin{aligned}\Delta(\nabla_v \mathrm{d}f(v)) &= \nabla_{e_i}(\nabla_v\nabla_{e_i}\mathrm{d}f(v) - \mathrm{d}f(\mathbf{Rm}(e_i,v)v))\\ &= \nabla_{e_i}\nabla_v\nabla_v\mathrm{d}f(e_i) - \nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \mathrm{d}f(\nabla_{e_i}\mathbf{Rm}(e_i,v)v)\\ &= \nabla_v\nabla_{e_i}\nabla_v\mathrm{d}f(e_i) - \nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \nabla_v\mathrm{d}f(\mathbf{Rm}(e_i,v)e_i) - \nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \mathrm{d}f(\nabla_{e_i}\mathbf{Rm}(e_i,v)v)\\ &= \nabla_v\nabla_v\nabla_{e_i}\mathrm{d}f(e_i) - \nabla_v\mathrm{d}f(\mathbf{Rm}(e_i,v)e_i)\\ &\quad - \mathrm{d}f(\nabla_v\mathbf{Rm}(e_i,v)e_i) - \nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \nabla_v\mathrm{d}f(\mathbf{Rm}(e_i,v)e_i) - \nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \mathrm{d}f(\nabla_{e_i}\mathbf{Rm}(e_i,v)v)\\ &= \langle\nabla_v\nabla\Delta f, v\rangle - 2\nabla_v\mathrm{d}f(\mathbf{Rm}(e_i,v)e_i)\\ &\quad - \mathrm{d}f(\nabla_v\mathbf{Rm}(e_i,v)e_i) - 2\nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v)\\ &\quad - \mathrm{d}f(\nabla_{e_i}\mathbf{Rm}(e_i,v)v).\end{aligned}$$

Since the Ricci curvature is parallel, we have, by the contracted Bianchi identity,

$$\Delta(\nabla_v\mathrm{d}f(v)) = \langle\nabla_v\nabla\Delta f, v\rangle - 2\nabla_v\mathrm{d}f(\mathbf{Rm}(e_i,v)e_i) - 2\nabla_{e_i}\mathrm{d}f(\mathbf{Rm}(e_i,v)v).$$

If e_i is an eigenvector of the hessian of f with eigenvalue λ_i and v is an eigenvector of the hessian of f with the largest eigenvalue λ, then

$$\begin{aligned}\Delta(\nabla_v\mathrm{d}f(v)) &= \langle\nabla_v\nabla\Delta f, v\rangle + 2\lambda\mathbf{Rc}(v,v) - 2\lambda_i\langle e_i, \mathbf{Rm}(e_i,v)v\rangle\\ &\ge \langle\nabla_v\nabla\Delta f, v\rangle.\end{aligned}$$

Here we use the assumption that the sectional curvature is non-negative.

When $X_t = \nabla h_t$, the conditions become $\mathcal{W}(t) + R(t) \ge k_3 I$ and

$$\left\langle\nabla_v\nabla\left(\dot{h}_t + \frac{1}{2}|\nabla h_t|^2\right), v\right\rangle + k_2\left\langle(\nabla^2 h_t)^2 v, v\right\rangle + \langle W_t v, v\rangle \le b_t\left\langle\Delta\nabla^2 h_t(v), v\right\rangle + \left\langle\nabla_{Y_t}\nabla^2 h_t(v), v\right\rangle$$

for each eigenvector v of the symmetric operator $\nabla^2 h_t$ with the largest eigenvalue.

Recall that if ρ_t is a positive solution of the equation

$$\dot{\rho}_t = \Delta\rho_t + \langle\nabla\rho_t, \nabla U_1\rangle + U_2\rho_t,$$

then $h_t = -2\log\rho_t - U_1$ satisfies

$$\dot{h}_t + \frac{1}{2}|\nabla h_t|^2 = \Delta h_t + \Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2.$$

It follows that

$$\nabla^2\left(\dot{h}_t + \frac{1}{2}|\nabla h_t|^2\right) + W_t = \nabla^2\Delta h_t,$$

where $W_t = -\nabla^2\left(\Delta U_1 + \frac{1}{2}|\nabla U_1|^2 - 2U_2\right)$.

Therefore, if we assume that the Ricci curvature is parallel, the sectional curvature is non-negative, and $W_t \ge k_3 I$, then

$$\left\langle\nabla^2\left(\dot{h}_t + \frac{1}{2}|\nabla h_t|^2\right)(v), v\right\rangle + \langle W_t(v), v\rangle \le \left\langle\Delta\nabla^2 h_t(v), v\right\rangle.$$

It follows that

$$\nabla^2 h_t = -2\nabla^2\log\rho_t - \nabla^2 U_1 \le a_{k_3}(t)I. \qquad \square$$

5. A generalization of Huisken's monotonicity formula

This section is devoted to the proof of Theorem 1.1. First, let us recall the notations used. Let M be a submanifold of dimension m in a Riemannian manifold N of dimension n. The mean curvature flow is a family of immersions $\varphi_t : M \to N$ which satisfy

$$\dot{\varphi}_t = \mathfrak{H}_t(\varphi_t) + \nabla_t^{\perp} U(\varphi_t),$$

where $\mathfrak{H}_t$ is the mean curvature vector of $M_t := \varphi_t(M)$, $\bar{\nabla} U$ denotes the gradient of U with respect to the Riemannian metric on N, and $\nabla_t^{\perp} U$ is the projection of $\bar{\nabla}_t U$ onto the normal bundle of M_t. We also introduce the following notation for the part of the Laplacian in the normal bundle $\Delta_t^{\perp} U = \sum_k \left\langle \bar{\nabla}_{\mathbf{n}_k} \bar{\nabla} U, \mathbf{n}_k \right\rangle$.

Theorem 5.1. *Assume that the sectional curvature of the underlying compact Riemannian manifold N is non-negative and the Ricci curvature is parallel. Let $U : M \to \mathbb{R}$ be a smooth function satisfying the following condition for some positive constant k:*

$$\nabla^2 \left(\Delta U - \frac{1}{2} |\nabla U|^2 \right) \geq k_3 I,$$

Let φ_t be a solution of (1.10) and let ρ_t be a positive solution of the equation

$$\dot{\rho}_t = \bar{\Delta} \rho_t + \left\langle \bar{\nabla} U, \bar{\nabla} \rho_t \right\rangle + \rho_t \bar{\Delta} U$$

on N. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(b_{k_3} (T-t)^{\frac{n-m}{2}} \int_{\varphi_t(M)} u_t \, \mathrm{d}\mu_t \right) \leq -b_{k_3} (T-t)^{\frac{n-m}{2}} \int_{\varphi_t(M)} u_t \left(\frac{1}{2} \Delta_t^{\perp} U + \left| \frac{\nabla_t^{\perp} u_t}{u_t} - \mathfrak{H}_t \right|^2 \right) \mathrm{d}\mu_t.$$

The rest of this section is devoted to the proof of the above theorem. Next, we pick a convenient moving frame along φ_t.

Lemma 5.2. *Let $\sigma(\cdot)$ be a path in N such that $\sigma(t)$ is contained in $M_t := \varphi_t(M)$. Then there is a family of orthonormal frames*

$$\boldsymbol{n}_1(\psi_t), \ldots, \boldsymbol{n}_{n-m}(\psi_t), v_1(t), \ldots, v_m(t)$$

defined along $\sigma(\cdot)$ such that

(1) *$v_1(t), \ldots, v_m(t)$ are contained in the tangent bundle TM_t of M_t;*
(2) *$\boldsymbol{n}_1(t), \ldots, \boldsymbol{n}_{n-m}(t)$ are in the normal bundle $TM_t^{\perp}$ of M_t;*
(3) *$\dot{v}_1(t), \ldots, \dot{v}_m(t)$ are in $TM_t^{\perp}$;*
(4) *$\dot{\boldsymbol{n}}_1(t), \ldots, \dot{\boldsymbol{n}}_{n-m}(t)$ are in TM_t.*

Here $\dot{v}_i(t)$ denotes the covariant derivative of $v_i(t)$ with respect to the Riemannian metric $\langle \cdot, \cdot \rangle$ of N.

Moreover, if $\tilde{\boldsymbol{n}}_1(t), \ldots, \tilde{\boldsymbol{n}}_{n-m}(t), \tilde{v}_1(t), \ldots, \tilde{v}_m(t)$ is another such family, then there are orthogonal matrices $O^{(1)}$ and $O^{(2)}$ (independent of time) of size $(n-m) \times (n-m)$ and $m \times m$, respectively, such that

$$\tilde{\boldsymbol{n}}_i(t) = \sum_{j=1}^{n-m} O_{ij}^{(1)} \boldsymbol{n}_j(t) \text{ and } \tilde{v}_i(t) = \sum_{j=1}^{m} O_{ij}^{(2)} v_j(t).$$

The proof of Lemma 5.2 and that of ([23], Lem. 3.1) is completely analogous and is therefore omitted. From now on, we call any orthonormal moving frame which satisfies the conditions in Lemma 5.2 a parallel adapted frame along $\sigma(\cdot)$.

Let $\mathbf{n}_t$ be a normal vector in $TM_t^\perp$ and let $\mathfrak{S}_t^{\mathbf{n}_t} : TM_t \to TM_t$ be the shape operator of the submanifold M_t defined by

$$\langle \mathfrak{S}_t^{\mathbf{n}_t}(v_1), v_2\rangle = -\left\langle \bar{\nabla}_{v_1}\mathbf{n}_t, v_2\right\rangle.$$

Here $\bar{\nabla}$ denotes the Levi−Civita connection on N.

Recall that the mean curvature vector $\boldsymbol{\mathfrak{H}}_t$ of M_t is given by

$$\boldsymbol{\mathfrak{H}}_t = \sum_{i,j}\left\langle \mathfrak{S}_t^{\mathbf{n}_i(t)}(v_j(t)), v_j(t)\right\rangle \mathbf{n}_i(t).$$

Lemma 5.3. *Let $\boldsymbol{n}_1(t), \ldots, \boldsymbol{n}_{n-m}(t), v_1(t), \ldots, v_m(t)$ be a parallel adapted frame along $\varphi_t(x)$, where φ_t satisfies the following equation*

$$\dot{\varphi}_t = \sum_i F_i(t, \varphi_t)\boldsymbol{n}_i(t).$$

Let $A(t)$ and $G^k(t)$ be families of matrices defined by

$$\mathrm{d}\varphi_t(v_i(0)) = \sum_j A_{ij}(t)v_j(t) \text{ and } G^k_{ij}(t) = \left\langle \mathfrak{S}_t^{\boldsymbol{n}_k(t)}(v_i(t)), v_j(t)\right\rangle,$$

respectively. Then

$$\dot{A}(t) = -\sum_k F_k(t, \varphi_t)A(t)G^k(t),$$

where ∇_t is the gradient with respect to the induced metric on M_t.

Proof. Let $\gamma(s)$ be a curve in M such that $\frac{\mathrm{d}}{\mathrm{d}s}\gamma(s)\Big|_{s=0} = v_i(0)$. Then

$$\frac{\overline{D}}{\mathrm{d}t}\mathrm{d}\varphi_t(v_i(0)) = \sum_j \left(\dot{A}_{ij}(t)v_j(t) + A_{ij}(t)\dot{v}_j(t)\right).$$

On the other hand, we have

$$\frac{\overline{D}}{\mathrm{d}t}\mathrm{d}\varphi_t(v_i(0)) = \frac{\overline{D}}{\mathrm{d}s}\dot{\varphi}_t(\gamma(s))\Big|_{s=0} = \sum_k \left(\langle\nabla_t F_k(t,\varphi_t), \mathrm{d}\varphi_t(v_i(0))\rangle\, \mathbf{n}_k(t) + F_k(t,\varphi_t)\bar{\nabla}_{\mathrm{d}\varphi_t(v_i(0))}\mathbf{n}_k(t)\right).$$

It follows that

$$\dot{A}_{ij}(t) = -\sum_{l,k} A_{il}(t)F_k(t,\varphi_t)\left\langle \mathfrak{S}_t^{\mathbf{n}_k(t)}v_l(t), v_j(t)\right\rangle. \qquad \square$$

Proof of Theorem 5.1. Let ρ_t be the density of $\varphi_t^*\mu_t$ with respect to μ_0: $\rho_t\mu_0 = \varphi_t^*\mu_t$. Let $\mathbf{n}_1(t), \ldots, \mathbf{n}_{n-m}(t), v_1(t), \ldots, v_m(t)$ be a parallel adapted frame along the path $\varphi_t(x)$ and let $A(t)$ be the family of matrices defined by

$$\mathrm{d}\varphi_t(v_i(0)) = \sum_{j=1}^n A_{ij}(t)v_j(t).$$

Then $\rho_t = \det A(t)$ and we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\varphi_t(M)} u_t\,\mathrm{d}\mu_t = \frac{\mathrm{d}}{\mathrm{d}t}\int_M u_t(\varphi_t)\det A(t)\,\mathrm{d}\mu_0 &= \int_M \left[\dot{u}_t(\varphi_t) + u_t(\varphi_t)\mathbf{tr}(A(t)^{-1}\dot{A}(t)) + \left\langle\bar{\nabla}u_t, \dot{\varphi}_t\right\rangle_{\varphi_t}\right]\det A(t)\,\mathrm{d}\mu_0\\ &= \int_{\varphi_t(M)}\left(\dot{u}_t + \sum_k F_k\left\langle\bar{\nabla}u_t, \mathbf{n}_k(t)\right\rangle - u_t\sum_k F_k\mathbf{tr}(G^k(t))\right)\mathrm{d}\mu_t.\end{aligned}$$

Let $u_t = \rho_{T-t}$. Then we have, by assumptions, $F_k(t,\cdot) = \mathbf{tr}(G^k(t)) + \langle \bar{\nabla} U, \mathbf{n}_k(t)\rangle$ and $\dot{u}_t = -\bar{\Delta} u_t - \langle \bar{\nabla} U, \bar{\nabla} u_t\rangle - (\bar{\Delta} U)u_t$. Then the above equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\varphi_t(M)} u_t\,\mathrm{d}\mu_t = \int_{\varphi_t(M)} \Big(-\bar{\Delta} u_t - \langle \nabla U, \nabla u_t\rangle - (\bar{\Delta} U)u_t - u_t|\mathfrak{H}_t|^2 + \langle \nabla_t^{\perp} u_t, \mathfrak{H}_t\rangle - u_t\langle \nabla_t^{\perp} U, \mathfrak{H}_t\rangle \Big)\mathrm{d}\mu_t, \quad (5.1)$$

where $\nabla_t^{\perp} u$ is the projection of $\bar{\nabla} u$ onto the normal bundle of M_t.

A simple calculation shows that

$$\begin{aligned} \Delta u &= \sum_{i=1}^{n} \left\langle \bar{\nabla}_{v_i}(\bar{\nabla} u - \sum_k \langle \mathbf{n}_k, \bar{\nabla} u\rangle \mathbf{n}_k), v_i \right\rangle \\ &= \bar{\Delta} u - \sum_k \langle \bar{\nabla}_{\mathbf{n}_k}\bar{\nabla} u, \mathbf{n}_k\rangle + \sum_k \langle \mathbf{n}_k, \bar{\nabla} u\rangle \mathbf{tr}(G^k(t)) \\ &= \bar{\Delta} u - \Delta_t^{\perp} u + \langle \mathfrak{H}, \nabla_t^{\perp} u\rangle . \end{aligned}$$

Therefore, we have

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\int_{\varphi_t(M)} u_t\,\mathrm{d}\mu_t = \int_{\varphi_t(M)} \Big(-\Delta u_t - \Delta_t^{\perp} u_t - u_t \Delta_t^{\perp} U \\ &- \langle \nabla U, \nabla u_t\rangle - \Delta U u_t - u_t|\mathfrak{H}_t|^2 + 2\langle \nabla_t^{\perp} u_t, \mathfrak{H}_t\rangle \Big)\mathrm{d}\mu_t \\ &= \int_{\varphi_t(M)} \Big(-\Delta_t^{\perp} u_t - u_t\Delta_t^{\perp} U - u_t|\mathfrak{H}_t|^2 \\ &+ 2u_t \left\langle \frac{\nabla_t^{\perp} u_t}{u_t}, \mathfrak{H}\right\rangle - u_t\left|\frac{\nabla_t^{\perp} u_t}{u_t}\right|^2 + u_t\left|\frac{\nabla_t^{\perp} u_t}{u_t}\right|^2 \Big)\mathrm{d}\mu_t \\ &= -\int_{\varphi_t(M)} u_t \left(\Delta_t^{\perp}\log u_t + \Delta_t^{\perp} U + \left|\frac{\nabla_t^{\perp} u_t}{u_t} - \mathfrak{H}_t\right|^2 \right)\mathrm{d}\mu_t. \end{aligned}$$

$$-\Delta_t^{\perp}\log \rho_{T-t} - \frac{1}{2}\Delta_t^{\perp} U_1 \le \frac{n-m}{2} a_{k_3}(T-t).$$

By Theorem 4.1,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\varphi_t(M)} u_t\,\mathrm{d}\mu_t - \frac{n-m}{2}a_{k_3}(T-t)\int_{\varphi_t(M)} u_t\,\mathrm{d}\mu_t \le \int_{\varphi_t(M)} u_t\left(-\frac{1}{2}\Delta_t^{\perp} U - \left|\frac{\nabla_t^{\perp} u_t}{u_t} - \mathfrak{H}_t\right|^2 \right)\mathrm{d}\mu_t.$$

Since $\dot{b}_k = a_k b_k$, the result follows. ☐

6. A generalization of the Aronzon–Bénilan estimate

The Aronzon–Bénilan estimate [3] is a differential Harnack inequality for the porous medium equation

$$\dot{\rho}_t = \Delta(\rho_t^m).$$

In this section, we apply Theorem 2.1 and prove the following generalization of the Aronzon–Bénilan estimate.

Theorem 6.1. *Assume that the Ricci curvature of the underlying compact Riemannian manifold M is non-negative. Let U be a function on M satisfying*

$$\Delta U \ge \frac{k_3}{2m}$$

where $m > 1$. Then any smooth positive solution ρ_t of the equation

$$\dot{\rho}_t = \Delta(\rho_t^m) + U\rho_t^{2-m}.$$

satisfies

$$\frac{2m}{m-1}\Delta(\rho_t^{m-1}) \le \frac{2n}{2+n(m-1)}a_{\frac{k_3(2+n(m-1))}{2n}}(t).$$

Proof. A computation shows that $h_t = \frac{2m}{1-m}\rho_t^{m-1}$ satisfies

$$\dot{h}_t + \frac{1}{2}|\nabla h_t|^2 = \frac{1}{2}(1-m)h_t\Delta h_t - 2mU.$$

It follows that

$$\Delta\left(\dot{h}_t + \frac{1}{2}|\nabla h_t|^2\right) + \frac{1}{2}(m-1)(\Delta h_t)^2 + 2m\Delta U = (1-m)\left\langle \nabla h_t, \nabla \Delta h_t\right\rangle + m\rho_t^{m-1}\Delta\Delta h_t.$$

The rest follows from the assumptions and Theorem 2.1. □

7. On Laplacian and Hessian comparison type theorems

In this section, we prove versions of Laplacian and Hessian type comparison theorems for the following cost function

$$c_{s,t}(x,y) = \inf_{\gamma(s)=x,\gamma(t)=y}\int_s^t \frac{1}{2}|\dot{\gamma}(\tau) - \nabla U_1(\gamma(\tau))|^2 - U_2(\gamma(\tau))\mathrm{d}\tau. \tag{7.1}$$

More precisely,

Theorem 7.1. *Assume that*

(1) *the Ricci curvature of the underlying manifold M is non-negative;*
(2) $\Delta\left(U_2 - \frac{1}{2}|\nabla U_1|^2\right) \ge k_3$ *for some negative constant k_3.*

Then the cost function $c_{0,t}$ defined by (1.6) satisfies

$$\Delta_x c_{0,t}(x_0,x) \le \sqrt{-k_3 n}\coth\left(\sqrt{-\frac{k_3}{n}}\,t\right),$$

wherever $c_{0,t}(x_0,\cdot)$ is twice differentiable.

Theorem 7.2. *Assume that*

(1) *the sectional curvature of the underlying manifold M is non-negative;*
(2) $\nabla^2\left(U_2 - \frac{1}{2}|\nabla U_1|^2\right) \ge k_3 I$ *for some negative constant k_3.*

Then the cost function $c_{0,t}$ defined by (1.6) satisfies

$$\nabla_x^2 c_{0,t}(x_0,x) \le \sqrt{-k_3}\coth\left(\sqrt{-k_3}\,t\right)I$$

wherever $c_{0,t}(x_0,\cdot)$ is twice differentiable.

Remark 7.3. The function $x \mapsto c_{0,t}(x_0,x)$ is locally semi-concave. In particular, it is twice differentiable Lebesgue almost everywhere by Alexandrov's theorem. Therefore, the conclusions in Theorems 7.1 and 7.2 hold Lebesgue almost everywhere (see [32] for the definitions and the results).

Remark 7.4. We can see that the above theorems are sharp by looking at the case $M = \mathbb{R}^n$, $U_1 \equiv 0$, and $U_2(x) = -\frac{k^2}{2}|x|^2$. We have

$$\nabla^2 \left(U_2 - \frac{1}{2}|\nabla U_1|^2\right) = -k^2 I, \quad \Delta \left(U_2 - \frac{1}{2}|\nabla U_1|^2\right) = -k^2 n$$

which is the equality case in the second conditions of Theorems 7.1 and 7.2. We also have

$$\nabla_x^2 c_{0,t}(0,x) = k\coth(kt)I, \quad \Delta_x c_{0,t}(0,x) = kn\coth(kt)$$

which gives the equality case in the conclusions of Theorems 7.1 and 7.2.

Remark 7.5. A Bishop-Gromov type volume comparison theorem follows from Corollary 2.4.

The proof of Theorem 7.2 is similar to that of Theorem 7.1 and will be omitted.

Proof of Theorem 7.1. If $c_{0,t}(x_0, x)$ is smooth, then the result follows from Theorems 2.1 and 2.3. Indeed, the Legendre transform of the Lagrangian

$$L(x,v) = \frac{1}{2}g_{ij}(x)(v^i - g^{il}(x)(U_1)_{x_l}(x))(v^j - g^{jk}(x)U_{x_k}(x)) - U_2(x)$$

is given by

$$\begin{aligned} H(x,p) &= \sup_{v \in T_x M} [p(v) - L(x,v)] \\ &= \frac{1}{2}g^{ij}(x)p_i p_j + g^{ij}(x)p_i (U_1)_{x_j}(x) + U_2(x). \end{aligned}$$

Here we sum over repeated indices.

The corresponding Hamilton–Jacobi equation is given by

$$\dot{f}_t + \frac{1}{2}|\nabla f_t|^2 + \langle \nabla U_1, \nabla f_t\rangle + U_2 = 0 \tag{7.2}$$

and $c_{0,t}(x_0, x)$ is a particular solution (see [1]).

If we set $X_t = \nabla\left(c_{0,t}(x_0,\cdot) + U_1\right)$, then

$$\mathbf{tr}\left(\nabla\left(\dot{X}_t + \nabla_{X_t}X_t\right) + \nabla^2\left(U_2 - \frac{1}{2}|\nabla U_1|^2\right)\right) = 0.$$

Therefore,

$$\Delta_x c_{0,t}(x_0,x) \le n a_{\frac{k_3}{n}}(t) = \sqrt{-k_3 n}\coth\left(\sqrt{-\frac{k_3}{n}}\,t\right)$$

by Theorem 2.1.

In general, if x is a point where $c_{0,t}(x_0,\cdot)$ is twice differentiable, then there is a unique minimizer γ to the infimum (1.6) joining x_0 and x. Moreover, $c_{0,s}(x_0,\cdot)$ is smooth at $\gamma(s)$ for each s in $(0,t)$ (see [7]). Therefore, the proof of Theorem 2.1 still applies. Note that, in this case, (2.4) is an equality. $\square$

8. On the semigroup approach

In this section, we give a semigroup proof of Theorems 3.1 and 6.1 which does not require any use of maximum principle. Such a proof was first given by [4], assuming that the equation

$$\dot{\rho}_t = L\rho_t \tag{8.1}$$

is given by an operator L without constant term which is self-adjoint with respect to a weighted L^2 inner-product.

In the case of the heat equation, the key idea is to consider expressions of the form

$$P_{T-t}(\rho_t|\nabla\rho_t|^2) \quad \text{and} \quad P_{T-t}\dot{\rho}_t,$$

where P_t is the heat semigroup.

Since the heat semi-group is symmetric, it is equivalent to consider the followings instead

$$\int_M \rho_t|\nabla\rho_t|^2\varrho_{T-t}\mathrm{d}\mathbf{vol} \quad \text{and} \quad \int_M \dot{\rho}_t\varrho_{T-t}\mathrm{d}\mathbf{vol}, \tag{8.2}$$

where ϱ_t ranges over all solutions of (8.1).

When L is not self-adjoint but still linear, we also consider the expressions in (8.2). However, in this case, ϱ_t ranges over solutions of the equation

$$\dot{\varrho}_t = L^*\varrho_t$$

instead, where L^* is the adjoint of L.

Proof of Theorem 3.1. Let ϱ_t be a positive solution of the equation

$$\dot{\varrho}_t = \Delta\varrho_t - \langle\nabla U_1, \nabla\varrho_t\rangle + (U_2 - \Delta U_1)\varrho_t.$$

Let k_t be a one-parameter family of smooth functions. A computation shows that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_M \rho_t\, k_t\, \varrho_{T-t}\,\mathrm{d}\mathbf{vol} &= \int_M \dot{\rho}_t\, k_t\, \varrho_{T-t} + \rho_t\, \dot{k}_t\, \varrho_{T-t} - \rho_t\, k_t\, \dot{\varrho}_{T-t}\,\mathrm{d}\mathbf{vol}\\ &= \int_M \rho_t\,(\dot{k}_t - \Delta k_t - 2\,\langle\nabla f_t, \nabla k_t\rangle)\,\varrho_{T-t}\,\mathrm{d}\mathbf{vol},\end{aligned}$$

where $f_t = \log\rho_t + \frac{1}{2}U_1$.

It follows that $c := \int_M \rho_t\varrho_{T-t}\,\mathrm{d}\mathbf{vol}$ is independent of t. By Bochner formula, we also have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_M \rho_t(\Delta f_t)\,\varrho_{T-t}\mathrm{d}\mathbf{vol} &= \int_M \rho_t\left(2|\nabla^2 f_t|^2 + 2\mathbf{Rc}(\nabla f_t, \nabla f_t) - \frac{1}{2}\Delta V\right)\varrho_{T-t}\mathrm{d}\mathbf{vol}\\ &\geq \int_M \rho_t\left(\frac{2}{n}(\Delta f_t)^2 - \frac{k^2 n}{2}\right)\varrho_{T-t}\mathrm{d}\mathbf{vol}\\ &\geq \int_M \rho_t\left(\frac{4a(t)\Delta f_t}{n} - \frac{2a(t)^2}{n} - \frac{k^2 n}{2}\right)\varrho_{T-t}\mathrm{d}\mathbf{vol}.\end{aligned}$$

So $b(t) := \int_M \rho_t(\Delta h_t)\,\varrho_{T-t}\mathrm{d}\mathbf{vol}$ satisfies

$$\dot{b} \geq \frac{4a(t)}{n}b(t) - \frac{2a(t)^2 c}{n} - \frac{k^2 nc}{2}.$$

If $a(t) = -\frac{kn}{2}\coth(kt)$, then

$$\dot{b}(t) \geq -2kb(t)\coth(kt) - \frac{k^2 n(\coth^2(kt)+1)c}{2}.$$

It follows that

$$\int_M \rho_t(\Delta f_t)\,\varrho_{T-t}\mathrm{d}\mathbf{vol} = b(t) \geq -\frac{kn}{2}\coth(kt)c$$
$$= -\frac{kn}{2}\coth(kt)\int_M \rho_t\varrho_{T-t}\,\mathrm{d}\mathbf{vol}.$$

By setting $t = T$, we obtain

$$\int_M \rho_T(\Delta f_T)\,\varrho_0\mathrm{d}\mathbf{vol} = -\frac{kn}{2}\coth(kt)\int_M \rho_T\varrho_0\,\mathrm{d}\mathbf{vol}.$$

Since ϱ_0 is arbitrary, we must have

$$\dot{f}_t - |\nabla f_t|^2 + \frac{1}{2}V = \Delta f_t \geq -\frac{kn}{2}\coth(kt). \qquad \square$$

9. The generalized Li–Yau estimate without compactness assumption

In this section, we give another prove of Theorem 1.1 without making any compactness assumption. The proof uses the standard localization argument as in [24].

Proof of Theorem 1.1. Let ρ_t be a positive solution of the equation $\dot{\rho}_t = \Delta\rho_t + \langle\nabla U_1\nabla\rho_t\rangle + U_2\rho_t$ and let $f_t = \log\rho_t + \frac{1}{2}U_1$. Then $\dot{f}_t = |\nabla f_t|^2 + U_2 - \frac{1}{4}|\nabla U_1|^2 - \frac{1}{2}\Delta U_1 + \Delta f_t = |\nabla f_t|^2 + \Delta f_t - \frac{1}{2}V$, where $V = -2U_2 + \frac{1}{2}|\nabla U_1|^2 + \Delta U_1$. It follows that

$$\ddot{f}_t - \Delta\dot{f}_t - 2\left\langle\nabla f_t, \nabla\dot{f}_t\right\rangle = 0$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}|\nabla f_t|^2 - \Delta|\nabla f_t|^2 - 2\left\langle\nabla f_t, \nabla|\nabla f_t|^2\right\rangle \leq -\left\langle\nabla f_t, \nabla V\right\rangle - \frac{2}{n}\left(\dot{f}_t - |\nabla f_t|^2 + \frac{1}{2}V\right)^2.$$

Let $F_t = a_1\dot{f}_t + a_2|\nabla f_t|^2 + a_3V + a_4$, where

$$a_2 = \frac{(\exp(2tk)-1)^2}{\exp(2tk)} = (\exp(tk) - \exp(-tk))^2,$$

$a_1 = -\alpha a_2$, $a_3 = -\frac{\alpha}{2}a_2$, $\alpha > 1$. a_4 is a function of time t to be determined.

Then

$$\begin{aligned}\dot{F}_t - \Delta F_t - 2\left\langle\nabla f_t, \nabla F_t\right\rangle \leq{}& -a_2\left\langle\nabla f_t, \nabla V\right\rangle - \frac{2a_2}{n}\left(\dot{f}_t - |\nabla f_t|^2 + \frac{1}{2}V\right)^2\\ &- a_3\Delta V - 2a_3\left\langle\nabla f_t, \nabla V\right\rangle - \alpha\dot{a}_2\dot{f}_t + \dot{a}_2|\nabla f_t|^2 + \dot{a}_3V + \dot{a}_4\\ \leq{}& -\frac{2a_2}{n}\left(\frac{1}{\alpha a_2}F_t + \left(1 - \frac{1}{\alpha}\right)|\nabla f_t|^2 - \left(\frac{a_3}{\alpha a_2} + \frac{1}{2}\right)V - \frac{a_4}{\alpha a_2}\right)^2\\ &- a_3\Delta V - (a_2 + 2a_3)\left\langle\nabla f_t, \nabla V\right\rangle + \frac{\dot{a}_2}{a_2}F_t + \left(\dot{a}_3 - \frac{\dot{a}_2a_3}{a_2}\right)V + \dot{a}_4 - \frac{\dot{a}_2a_4}{a_2}.\end{aligned}$$

Let $G_t = \eta F_t$, where η is a cut off function. Let us fix a time t. At a maximum point of G_t, we have

$$\nabla F_t = -\frac{F_t}{\eta}\nabla\eta, \quad \Delta F_t \leq \frac{2F_t}{\eta^2}|\nabla\eta|^2 - \frac{F_t}{\eta}\Delta\eta, \quad \dot{F}_t \geq 0.$$

Therefore,

$$-\frac{2F_t}{\eta^2}|\nabla\eta|^2+\frac{F_t}{\eta}\Delta\eta+2\frac{F_t}{\eta}\langle\nabla f_t,\nabla\eta\rangle\le-\frac{2a_2}{n}\left(\frac{1}{\alpha a_2}F_t+\left(1-\frac{1}{\alpha}\right)|\nabla f_t|^2-\left(\frac{a_3}{\alpha a_2}+\frac{1}{2}\right)V-\frac{a_4}{\alpha a_2}\right)^2$$
$$-a_3\Delta V-(a_2+2a_3)\langle\nabla f_t,\nabla V\rangle+\frac{\dot a_2}{a_2}F_t+\left(\dot a_3-\frac{\dot a_2a_3}{a_2}\right)V+\dot a_4-\frac{\dot a_2a_4}{a_2}\cdot$$

Let $r:[0,\infty)\to[0,1]$ be a function such that $r(x)=1$ if $x\le R$, $r(x)=0$ if $x\ge 2R$, $r'\le 0$, $\frac{r'(x)^2}{r(x)}\le\frac{C}{R^2}$, and $|r''|\le\frac{C}{R^2}$, where $C>0$ is a constant. Let us fix a point x_0 and let us denote the ball of radius R centered at x_0 by B_R. Let $\eta=r(d(x_0,x))$, where $d(x_0,x)$ is the distance from x_0 to x. It follows that

$$\frac{|\nabla\eta|^2}{\eta}\le\frac{C}{R^2}\cdot$$

Since the Ricci curvature is non-negative, we have

$$\Delta\eta\ge-\frac{C}{R^2}$$

by the Laplacian comparison theorem.

Then

$$-\frac{3CF_t}{\eta R^2}-\frac{2\sqrt{C}F_t^{3/2}}{\sqrt{\eta}R}\frac{|\nabla f_t|}{\sqrt{F_t}}\le-\frac{2a_2}{n}\left(\frac{1}{\alpha a_2}F_t+\left(1-\frac{1}{\alpha}\right)|\nabla f_t|^2-\frac{a_4}{\alpha a_2}\right)^2$$
$$+\frac{\alpha nk^2}{2}a_2+a_2(\alpha-1)|\nabla f_t||\nabla V|+\frac{\dot a_2}{a_2}F_t+\dot a_4-\frac{\dot a_2a_4}{a_2}\cdot$$

Let $H_t=\frac{|\nabla f_t|^2}{F_t}$ and assume that $|\nabla V|\le C$, then

$$-\frac{3CF_t}{\eta R^2}-\frac{2\sqrt{C}\sqrt{H_t}F_t^{3/2}}{\sqrt{\eta}R}\le-\frac{2a_2F_t^2}{n}\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H\right)^2+\frac{4a_4F_t}{n\alpha}\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)$$
$$+a_2C(\alpha-1)\sqrt{F_tH_t}+\frac{\dot a_2}{a_2}F_t+\dot a_4-\frac{\dot a_2a_4}{a_2}-\frac{2a_4^2}{\alpha^2na_2}+\frac{\alpha nk^2}{2}a_2.$$

Let us choose a_4 such that $\frac{a_4}{a_2}=-\frac{n\alpha^{3/2}k}{2}\coth\left(\frac{kt}{\sqrt{\alpha}}\right)$. Note that $\frac{a_4}{a_2}$ satisfies the following Riccati equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{a_4}{a_2}\right)-\frac{2}{\alpha^2n}\left(\frac{a_4}{a_2}\right)^2+\frac{\alpha nk^2}{2}=0.$$

Then

$$0\le-\frac{2a_2G_t^2}{n}\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)^2+\frac{4a_4\eta G_t}{n\alpha}\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)$$
$$+a_2C(\alpha-1)\sqrt{G_tH_t}+\frac{\dot a_2}{a_2}\eta G_t+\frac{3CG_t}{R^2}+\frac{2\sqrt{C}\sqrt{H_t}G_t^{3/2}}{R}\cdot$$

Since $a_4\le 0$, it follows that

$$0\le-\frac{2a_2G_t^2}{n}\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)^2+\left(\frac{\dot a_2}{a_2}+\frac{4a_4}{n\alpha^2a_2}\right)\eta G_t$$
$$+a_2C(\alpha-1)\sqrt{G_tH_t}+\frac{3CG_t}{R^2}+\frac{2\sqrt{C}\sqrt{H_t}G_t^{3/2}}{R}\cdot$$

Since $\alpha \geq 1$, a computation shows that

$$\frac{\dot{a}_2}{a_2} + \frac{4a_4}{n\alpha^2 a_2} = 2k\coth(kt) - \frac{2k}{\sqrt{\alpha}}\coth\left(\frac{kt}{\sqrt{\alpha}}\right) \geq 0.$$

Therefore,

$$0 \leq -\frac{2a_2G_t^2}{n}\left(\frac{1}{\alpha a_2} + \left(1-\frac{1}{\alpha}\right)H_t\right)^2 + \left(\frac{\dot{a}_2}{a_2} + \frac{4a_4}{n\alpha^2 a_2}\right)G_t$$
$$+ a_2C(\alpha-1)\sqrt{G_tH_t} + \frac{3CG_t}{R^2} + \frac{2\sqrt{C}\sqrt{H_t}G_t^{3/2}}{R}\cdot$$

It follows that G_t and hence the restriction of F_t to the ball B_R are less than or equal to the largest zero of the function

$$x \mapsto -\frac{2a_2x^2}{n} + \left(\frac{\dot{a}_2}{a_2} + \frac{4a_4}{n\alpha^2 a_2}\right)Bx$$
$$+ a_2AC(\alpha-1)\sqrt{x} + \frac{3BCx}{R^2} + \frac{2A\sqrt{C}x^{3/2}}{R},$$

where $A = \frac{\sqrt{H_t}}{\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)^2}$ and $B = \frac{1}{\left(\frac{1}{\alpha a_2}+\left(1-\frac{1}{\alpha}\right)H_t\right)^2}$

Since $B \leq \alpha^2 a_2^2$ and $A \leq C$ are bounded independent of R, we can let $R \to \infty$. Therefore, F_t is less than or equal to the largest zero of the function

$$x \mapsto -\frac{2a_2x^2}{n} + \left(\alpha^2 a_2\dot{a}_2 + \frac{4a_2a_4}{n}\right)x + a_2AC(\alpha-1)\sqrt{x}. \tag{9.1}$$

Now let $\alpha \to 1$ in (9.1). Then we have $F_t \leq 0$. The result follows. □

References

[1] A.A. Agrachev and Y.L. Sachkov, Control theory from the geometric viewpoint. Vol. 87 of *Encyclopaedia of Mathematical Sciences*, *Control Theory and Optimization, II*. Springer-Verlag, Berlin (2004).

[2] B. Andrews, Harnack inequalities for evolving hypersurfaces. *Math. Z.* **217** (1994) 179–197.

[3] D.G. Aronson and P. Bénilan, Régularité des solutions de l'q́uation des milieux poreux dans RN. *C. R. Acad. Sci. Paris Sér. A-B* **288** (1979) A103–A105.

[4] D. Bakry and M. Ledoux, A logarithmic Sobolev form of the Li–Yau parabolic inequality. *Rev. Mat. Iberoam.* **22** (2006) 683–702.

[5] A.A. Borisenko and V. Miquel, Gaussian mean curvature flow. *J. Evol. Eqn.* **10** (2010) 413–423.

[6] S. Brendle, A generalization of Hamilton's differential Harnack inequality for the Ricci flow. *J. Differ. Geom.* **82** (2009) 207–227.

[7] P. Cannarsa and C. Sinestrari, Semiconcave functions, Hamilton–Jacobi equations, and optimal control. Vol. 58 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Boston, Inc., Boston, MA (2004).

[8] H.-D. Cao, On Harnack's inequalities for the Kähler-Ricci flow. *Invent. Math.* **109** (1992) 247–263.

[9] H.-D. Cao and L. Ni, Matrix Li–Yau-Hamilton estimates for heat equation on Kähler manifolds. *Math. Ann.* **331** (2005) 795–807.

[10] H.-D. Cao and S.-T. Yau, Gradient estimates, Harnack inequalities and estimates for heat kernels of the sum of squares of vector fields. *Math. Z.* **211** (1992) 485–504.

[11] J. Cheeger and S.-T. Yau, A lower bound for the heat kernel. *Commun. Pure Appl. Math.* **34** (1981) 465–480.

[14] B. Chow, On Harnack's inequality and entropy for the Gaussian curvature flow. *Commun. Pure Appl. Math.* **44** (1991) 469–483.

[13] B. Chow, The Yamabe flow on locally conformally flat manifolds with positive Ricci curvature. *Commun. Pure Appl. Math.* **45** (1992) 1003–1014.

[14] B. Chow, On Harnack's Inequality and Entropy for the Gaussian curvature flow. *Commun. Pure Appl. Math.* **44** (1991) 469–483.

[15] L.C. Evans, Partial differential equations. Vol. 19 of *Grad. Stud. Math.*, 2nd edition. American Mathematical Society, Providence, RI (2010).

[16] P.B. Gilkey, The spectral geometry of a Riemannian manifold. *J. Differ. Geom.* **10** (1975) 601–618.

[17] M. Gromov, Isoperimetry of waists and concentration of maps. *Geom. Funct. Anal.* **13** (2003) 178–215.

[18] R. Hamilton, A matrix Harnack estimate for the heat equation. *Commun. Anal. Geom.* **1** (1993) 113–126.

[19] R. Hamilton, The Harnack estimate for the Ricci flow. *J. Differ. Geom.* **37** (1993) 225–243.

[20] R. Hamilton, Monotonicity formulas for parabolic flows on manifolds. *Commun. Anal. Geom.* **1** (1993) 127–137.

[21] R. Hamilton, Harnack estimate for the mean curvature flow. *J. Differ. Geom.* **41** (1995) 215–226.

[22] G. Huisken, Asymptotic behavior for singularities of the mean curvature flow. *J. Differ. Geom.* **31** (1990) 285–299.

[23] P.W.Y. Lee, Differential Harnack inequality for a family of sub-elliptic diffusion equations on Sasakian manifolds. Preprint arXiv:1302.3315 (2013).

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

[25] P. Lu, L. Ni, J-L Vázquez and C. Villani, Local Aronson–Bénilan estimates and entropy formulae for porous medium and fast diffusion equations on manifolds. *J. Math. Pures Appl.* **91** (2009) 1–19.

[26] J. Moser, A Harnack Inequality for Parabolic Differential Equations. *Commun. Pure Appl. Math.* **17** (1964) 101–134.

[27] L. Ni, A matrix Li−Yau-Hamilton inequality for Kähler-Ricci flow. *J. Differ. Geom.* **75** (2007) 303–358.

[28] L. Ni, Monotonicity and Li−Yau-Hamilton inequalities. Geometric flows. Vol. 12 of *Surv. Differ. Geom.* Int. Press, Somerville, MA (2008) 251–301.

[29] G. Perelmann, The entropy formula for the Ricci flow and its geometric applications. Preprint arXiv:math/0211159 (2002).

[30] O.C. Schnürer and K. Smoczyk, Evolution of hypersurfaces in central force fields. *J. Reine Angew. Math.* **550** (2002) 77–95.

[31] R.T. Seeley, Complex powers of an elliptic operator. Singular Integrals Proc. of Sympos. Pure Math., Chicago, Ill., 1966. Amer. Math. Soc., Providence, R.I. (1967) 288–307.

[32] C. Villani, Optimal transport. Old and new. Vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (2009).

[33] S.T. Yau, Harmonic functions on complete Riemannian manifolds. *Commun. Pure Appl. Math.* **28** (1975) 201–228.