

# A local Sobolev inequality on Ricci flow and its applications

Pak-Yeung Chan, Zilu Ma, Yongjia Zhang

December 7, 2021

## Abstract

In this article, we prove a local Sobolev inequality for complete Ricci flows. Our main result is that the local  $\nu$ -functional of a disk on a Ricci flow depends only on the Nash entropy based at the center of the disk, and consequently depends only on the volume of the disk. Furthermore, we introduce some applications of this local Sobolev inequality. These applications reveal the way in which the local geometry evolves along Ricci flow. In particular, we show that several classical theorems related to Perelman’s monotonicity formula can be derived from our results.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Preliminaries</b>	<b>10</b>
<b>3</b>	<b>Local Sobolev inequalities on Ricci flow</b>	<b>12</b>
<b>4</b>	<b>The “reaction-diffusion” property of the Ricci flow</b>	<b>17</b>
<b>5</b>	<b>Local monotonicity and Nash entropy</b>	<b>18</b>
<b>6</b>	<b>Perelman’s noncollapsing improving and pseudolocality theorems</b>	<b>23</b>
<b>7</b>	<b>Application to ancient Ricci flows</b>	<b>28</b>

## 1 Introduction

### 1.1 Local Sobolev inequalities on Ricci flow

As interpreted by Q. Zhang [Zhq07, Zhq10], Perelman’s entropy functional [Per02] is a family of logarithmic Sobolev inequalities in disguise. By applying Perelman’s monotonicity formula, Zhang [Zhq07] proved a uniform Sobolev inequality for compact Ricci flows. The main idea of his proof, intuitively speaking, is that Perelman’s  $\nu$ -functional (which is also a Sobolev constant; see the definition in Section 2) is monotonically increasing along a Ricci flow. Hence, for any compact Ricci flow  $(M^n, g_t)_{t \in [0, T]}$ , we have

$$\nu(g_t) \geq \nu(g_0) \quad \text{for all } t \in [0, T],$$

and consequently the Sobolev constant on each time-slice is determined by the geometry of the initial metric  $g_0$ .

This idea has also been adapted by the authors to prove a uniform Sobolev inequality for ancient Ricci flows. Recall that an ancient Ricci flow “starts” from  $t = -\infty$ , and hence there is no initial metric to begin with. We proved that, however, by improving Bamler’s entropy estimates [Bam20a] on ancient Ricci flows, Perelman’s asymptotic shrinking gradient Ricci soliton [Per02, Proposition 11.2] works equally well as an “initial manifold”, with which we may apply Q. Zhang’s argument; this leads to an initial version of the uniform Sobolev inequality for ancient Ricci flows [CMZ21a], where we assumed that the ancient Ricci flow in question has an asymptotic shrinking gradient Ricci soliton in the sense of Perelman [Per02, Proposition 11.2].

Later, we improved this uniform Sobolev inequality in [CMZ21b]. We found that one does not need to assume the existence of Perelman’s asymptotic shrinking gradient Ricci soliton, since, as is proved by Bamler [Bam20c], every ancient Ricci flow with bounded Nash entropy has a tangent flow at infinity, which is a metric soliton regular almost everywhere. It turns out that this metric soliton can also serve as an “initial manifold” for the purpose of applying Q. Zhang’s argument. Hence, the main result in [CMZ21b] says that every ancient Ricci flow with bounded Nash entropy has uniformly bounded  $\nu$ -functional, and therefore on it holds a uniform Sobolev inequality. With a different approach, Li-Wang [LW20] also proved a uniform Sobolev inequality on shrinking gradient Ricci solitons, which are special ancient Ricci flows; their result is nice in that it does not assume any curvature condition on the shrinking gradient Ricci soliton.

All the results mentioned above are global Sobolev inequalities on the whole manifold, and their validity requires some explicit or implicit assumptions on the geometry of the Ricci flow. Q. Zhang’s theorem assumes the manifold to be compact. Li-Wang’s result holds only on shrinking gradient Ricci solitons. Our Sobolev inequality requires the ancient solution to have time-wise bounded curvature and uniformly bounded Nash entropy; these two assumptions also imply that the ancient solution in question has a uniform lower bound for unit balls on each time-slice. However, in the most general case, the existence of a Ricci flow is guaranteed by the boundedness of curvature of the initial manifold alone (c.f. [Sh89]), and the curvature remains bounded until the first singular time (c.f. [Ham93b]). Therefore, it makes much sense to consider a complete Ricci flow with bounded curvature on compact time intervals, without making any further geometric assumptions, such as a lower bound for the volume of unit balls. The first main theorem in this article provides some analytic tools for such a Ricci flow—we show that the local Sobolev constant depends only on the bound of the Nash entropy or on the local volume ratio.

Most of the definitions in the statements of results can be found in Section 2 below. Throughout this paper, we shall always consider a complete Ricci flow  $(M^n, g_t)_{t \in I}$ , where  $I$  is an interval, with *bounded curvature on each compact time interval*; this assumption is made purely for the technical convenience. We denote by  $B_t(x, r)$ , where  $t \in I$ , the ball centered at  $x \in M$ , of radius  $r > 0$ , and with respect to the metric  $g_t$ , and by  $|\cdot|_t$  the volume of a set measured with  $g_t$ .

**Theorem 1.1.** *Assume that  $[-r^2, 0] \subseteq I$ . Then for any point  $x_0 \in M$  and any  $\tau, A > 0$ , we have*

$$\nu(B_0(x_0, Ar), g_0, \tau r^2) \geq N_{x_0, 0}(r^2) - \sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2} \log(1 + \tau). \quad (1.1)$$

Applying a standard argument (c.f. [Zhq10, Ye15]), we may rewrite the above local bound of Perelman’s  $\nu$ -functional as the following local Sobolev inequality.

**Corollary 1.2.** *Under the same assumptions as Theorem 1.1, we have that for any  $n \geq 3$ , there is a dimensional constant  $c = c(n)$ , such that for any  $u \in W_0^{1,2}(B_0(x_0, Ar))$ , it holds that*

$$\left( \int_{B_0(x_0, Ar)} |u|^{\frac{2n}{n-2}} dg_0 \right)^{\frac{n-2}{n}} \leq ce^{-\frac{2\nu_0}{n} - c\tau r^2 R_{\min}} \int_{B_0(x_0, Ar)} \left( 4|\nabla u|^2 + \left( R_{g_0} - R_{\min} + \frac{c}{\tau r^2} \right) u^2 \right) dg_0, \quad (1.2)$$

where  $\nu_0 = \mathcal{N}_{x_0,0}(r^2) - \sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2}\log(1 + \tau)$  and  $R_{\min} = \min \{ \inf_{B_0(x_0, Ar)} R_{g_0}, 0 \} \geq -\frac{n}{2r^2}$ .

Previously, Q. Zhang [Zhq10, Theorem 6.3.2] also proved a local Sobolev inequality assuming local curvature and volume bounds. We shall show that Zhang's theorem is actually a special case of our local Sobolev inequality above.

**Corollary 1.3** (Q. Zhang's local Sobolev inequality [Zhq10, Theorem 6.3.2]). *Assume that  $[0, r^2] \subset I$ . Let  $x_0 \in M$  be a point and  $A > 0$  be a positive number satisfying*

$$\begin{aligned} |B_0(x_0, r)|_0 &\geq A^{-1}r^n, \\ |\text{Ric}| &\leq \frac{1}{(nr)^2} \quad \text{on } B_0(x_0, r) \times [0, r^2]. \end{aligned} \quad (1.3)$$

Then we have

$$\nu(B_{r^2}(x_0, Ar), g_{r^2}, A^2r^2) \geq -C(n, A).$$

Bamler [Bam20a, Theorem 8.1] proved that the Nash entropy based at a point can be estimated by the local volume ratio. Therefore, we can rewrite Theorem 1.1 and Corollary 1.2 as follows.

**Corollary 1.4.** *Assume that  $[-r^2, 0] \subseteq I$ . Then for any positive numbers  $\alpha, \tau$ , and  $A$ , the following holds. For any  $x_0 \in M$  satisfying*

$$|B_0(x_0, r)|_0 \geq \alpha r^n,$$

we have

$$\nu(B_0(x_0, Ar), g_0, \tau r^2) \geq \log \alpha - 4n(A + \tau) - C(n).$$

Furthermore, (1.2) also holds with  $\nu_0 = \log \alpha - 4n(A + \tau) - C(n)$  instead. As a standard consequence, the ball  $B_0(x_0, Ar)$  is strongly noncollapsed at scale  $Ar$ , i.e., there is a constant  $\kappa = \kappa(n, \alpha, A) > 0$ , such that for any  $\rho \in (0, Ar]$  and  $y_0 \in B_0(x_0, Ar)$ , it holds that

$$\sup_{B_0(y_0, \rho)} R_{g_0} \leq \rho^{-2} \implies |B_0(y_0, \rho)|_0 \geq \kappa \rho^n.$$

## 1.2 The ‘‘reaction-diffusion’’ property of the Ricci flow

The Ricci flow is a weakly parabolic quasilinear partial differential equation, which is sometimes known as a reaction-diffusion system in other fields of study. In a pure diffusion process, Fick's first law of diffusion says that the flux of a chemical substance always goes from regions of high concentration to regions of low concentration, and eventually, the substance tends to be distributed evenly everywhere. Such phenomenon can also be observed in the Ricci flow, especially when the initial metric is good enough. In this respect, see, for instance, [BW08, BS09, Ham82, Ham86].

In general, when the Ricci flow is noncompact without good initial condition, it is less clear how the reaction-diffusion mechanism operates. One may simply regard the Ricci flow as a background

and consider the heat equation or the conjugate heat equation. For instance, the (conjugate) heat kernel coupled with the Ricci flow has some Gaussian upper and lower estimates [CCGGIKLLN10, Theorem 26.25, Theorem 26.31]. However, these estimates are rather coarse, and they depend not so much on the parabolicity of the Ricci flow itself than on that of the (conjugate) heat equation. What, then, should we expect from the Ricci flow as a reaction-diffusion system? Whither is the “geometric information” diffused?

A hint of the answer can already be seen from the works of Perelman [Per02, Per03]. For instance, on a noncollapsed ancient solution with nonnegative curvature operator, one can always observe a Ricci shrinker arising along the path of  $\ell$ -centers, namely, the points where Perelman’s  $\ell$ -function does not exceed  $\frac{n}{2}$ ; see [CMZ21a, Definition 2.6]. Furthermore, Perelman proved the noncollapsing property at a point  $(x, t)$  after surgery time by estimating the geometry near its  $\ell$ -center  $(z, s)$  before surgery time. If the geometry near  $(z, s)$  can be controlled, then the geometry near  $(x, t)$  can be controlled accordingly. These facts indicate that the “geometric information” is diffused along the path of  $\ell$ -centers.

Recently, by a delicate estimate on the conjugate heat kernel, Bamler [Bam20a] discovered a property of the Ricci flow called  $H_n$ -concentration—a conjugate heat kernel on a Ricci flow is centered around some points, called  $H_n$ -centers ([Bam20a, Definition 3.10]), in the same way an Euclidean Gaussian heat kernel is centered around its base point. However, in contrast to the case of heat equation on the Euclidean space, an  $H_n$ -center is not a fixed point in space. In a pure diffusion process, if the initial distribution of a chemical substance is a Dirac delta function at the origin, then the substance is diffused forward in time and outward in space, or in other words, loosely speaking, along the family of parabolic neighborhoods of the form  $B(O, r) \times [0, r^2]$ ,  $r > 0$ . Analogous to the diffusion process in the Euclidean space, the Ricci flow bears a similar property along parabolic neighborhoods constructed around  $H_n$ -centers, called  $P^*$ -parabolic neighborhoods. This can be seen from Bamler’s Harnack inequality of the Nash entropy (c.f. Theorem 2.1); this result indicates that the difference of Nash entropy based at two different points is determined by the “ $P^*$ -distance” between the base points.

As supported by some previous results ([Bam20a, Theorem 6.1, Theorem 6.2, Theorem 8.1], [CMZ21a, Theorem 1.8]), the Nash entropy is closely related to the local geometry. We shall come back to this point later, and it suffices us now to view it in combination with the observation in the previous paragraph: as we shall see in the following corollary, the geometric information around one point in space-time determines the geometric information in  $P^*$ -parabolic neighborhoods centered at this point.

**Corollary 1.5.** *Assume that  $[-r^2, 0] \subseteq I$ . Let  $x_0 \in M$  and  $A, \tau > 0$ . Then for any  $(y, s) \in P^*(x_0, 0 \mid Ar, 0, A^2r^2)$ , we have*

$$\nu(B_s(y, Ar), g_s, \tau r^2) \geq \mathcal{N}_{x_0, 0}(r^2) - 2\sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2}\log(1 + \tau) - \frac{n}{2}\log(1 + A^2).$$

*As a consequence, there is a constant  $\kappa = \kappa(n, A) > 0$  such that for any ball  $B := B_s(y, \rho)$  with  $(y, s) \in P^*(x_0, 0 \mid Ar, 0, A^2r^2)$  and  $\rho \in (0, Ar]$ , we have*

$$\sup_B R_{g_s} \leq \rho^{-2} \implies |B|_s \geq \kappa \rho^n.$$

In spite of the formal nice expression of the above result, however, without any further geometric assumption, the shape of a  $P^*$ -parabolic neighborhood could be very complicated. In fact, given an arbitrary space-time point  $(y, s) \in M \times I$  and an arbitrary time  $t \in (s, \infty) \cap I$ , one may not be

able to find a point  $x \in M$ , such that  $(y, s)$  is (or is at least close to) an  $H_n$ -center of  $(x, t)$ . This is due to the property of parabolic systems (i.e. one may not solve a parabolic equation backwards, etc.). This observation is supported by Lai's recent examples [Lai20]. Taking a higher-dimensional noncollapsed Lai's steady soliton, a point  $(y, s)$  on the edge, where  $s \ll -1$ , can never be close to an  $H_n$ -center of  $(x, 0)$ , where  $x$  is any point. This is because along the  $H_n$ -centers of a fixed point, one always sees an asymptotic shrinking gradient Ricci soliton forming (c.f. [CMZ21a]), which can never arise from the edges of such examples. In fact, Perelman's reduced distance has a similar problem, namely, given a point in space-time, this point may not be an  $\ell$ -center of any point on a later slice.

Another perspective from which we may understand the "reaction-diffusion" property of the Ricci flow along  $P^*$ -parabolic neighborhoods is the following Harnack inequality of the heat equation. Recall that a Harnack inequality of a parabolic equation provides some backward estimate of a heat-type equation. Some results in this respects are [Ca08, Ham93a, LY86], [Per02, §9], [W18, §6]. Interestingly, the differential Harnack inequalities of Cao [Ca08] and Perelman [Per02, §9] do not depend on any local geometric condition. Nonetheless, if one wants to apply these differential Harnack inequalities to obtain some backward control, namely, to estimate the earlier slices of a solution using its later slices, then some local geometric assumption is indispensable.

Wang's Harnack inequality [W18, Proposition 6.6] shows that, for a positive supersolution to the heat equation coupled with Ricci flow, the local bound at a earlier time-slice can be controlled by that of a later slice. This Harnack inequality depends only on a very weak assumption of the Ricci curvature. In fact, inspired by Wang's result, since, as we shall see again later, the geometric information is diffused along the path of  $H_n$ -centers, and since the local geometric assumption is nothing but that which controls the position of an  $H_n$ -center, it is reasonable to hope for a Harnack inequality of heat equation along Ricci flow in terms of  $P^*$ -parabolic neighborhoods. We leave it to the reader to check that, from our theorem below, one can derive B. Wang's Harnack inequality, or at least a qualitatively similar one; the method is not different from the proof of Corollary 1.9, which we do present in the sequel.

**Theorem 1.6** (A Harnack inequality). *Let  $[-r^2, 0] \subset I$  and let  $H : M \times [-r^2, 0]$  be a positive supersolution to the heat equation coupled with Ricci flow, namely,  $\square_t H \geq 0$ . Let  $x_0 \in M$  be any point and  $(z, -r^2)$  be an  $H_n$ -center of  $(x_0, 0)$ . Then we have*

$$\Phi_0(\Lambda, A) \cdot \min_{\bar{B}_{-r^2}(z, Ar)} H(\cdot, -r^2) \leq \min_{\bar{B}_0(x_0, \Lambda r)} H(\cdot, 0),$$

for all  $A > \sqrt{2H_n} + 4\sqrt{2\pi}$  and  $\Lambda > 0$ , where

$$\Phi_0(\Lambda, A) := \Phi\left(\frac{1}{\sqrt{2}}A - \Lambda - \sqrt{H_n} - 4\sqrt{\pi}\right),$$

and  $\Phi$  is the function defined in [Bam20a, §4] satisfying

$$\Phi'(t) = (4\pi)^{-\frac{1}{2}} e^{-\frac{t^2}{4}}, \quad \lim_{t \rightarrow -\infty} \Phi(t) = 0, \quad \lim_{t \rightarrow +\infty} \Phi(t) = 1.$$

Here (and throughout the paper)  $H_n := \frac{(n-1)\pi^2}{2} + 4$ .

### 1.3 Local monotonicity and Nash entropy

Because of the interdependence of the Nash entropy and the geometry near the base point ([Bam20a, Theorem 6.1, Theorem 8.1]), and because of the interdependence of the Nash entropy and the geometry near the  $H_n$ -center ([Bam20a, Theorem 6.2], [CMZ21a, Theorem 1.8]), we know that the geometry near a point in space-time depends on the geometry near its  $H_n$ -center, and the Nash entropy is that which relates the geometry near these two points; this, indeed, is the philosophy underlying Theorem 1.1. However, given this fact, we can directly estimate some geometric constants near a point in terms of what we know about its  $H_n$ -center—instead of going through the Nash entropy. This leads to a local monotonicity theorem.

Some examples in this respect are [W18, §5] and [TZ21, §2.2]. These results all show how local geometry on later time-slices is dependent on that of earlier time-slices. However, because the property of  $H_n$ -concentration was not well understood by the time of their publication, these results all make local geometric assumptions in one way or another. In retrospective, these local geometric assumptions can be regarded as means of controlling the possible position of an  $H_n$ -center. One can see this point more clearly when one views the proof of Theorem 1.12 below (see also [J21]). For other local monotonicity formulas along geometric flows, see, for example, [E01, EKNT08].

Our local monotonicity formula, on the contrary, does not depend on any local geometric assumptions, since we are making use of the properties of  $H_n$ -centers observed above. In particular, we show that the local  $\mu$ -functional around a space-time point is solely determined by the local  $\mu$ -functional around its  $H_n$ -center. This result is in the spirit of, though slightly stronger than, [W18, Theorem 5.4].

**Theorem 1.7.** *Assume that  $[-r^2, 0] \subseteq I$ . Then for any  $x_0 \in M$ , any  $H_n$ -center  $(z, -r^2)$  of  $(x_0, 0)$ , any  $\tau > 0$ , and any  $A \geq 16$ , we have*

$$\mu\left(B_{-r^2}\left(z, 2A\sqrt{H_n r}\right), g_{-r^2}, (1+\tau)r^2\right) \leq \mu\left(B_0\left(x_0, A\sqrt{H_n r}\right), g_0, \tau r^2\right) + \frac{C_n}{A^2} e^{-\frac{A^2}{20}},$$

where  $C_n$  is a positive dimensional constant.

The sharpness of the above result can be seen from the fact that, taking  $A \rightarrow \infty$ , it reduces to the classical monotonicity of Perelman's  $\mu$ -functional [Per02].

**Corollary 1.8** (Monotonicity of Perelman's  $\mu$ -functional). *Assume that  $[-r^2, 0] \subset I$ . Then, for all  $\tau > 0$ , we have*

$$\mu(g_{-r^2}, (1+\tau)r^2) \leq \mu(g_0, \tau r^2).$$

B. Wang's local monotonicity formula can also be derived from Theorem 1.7. It is worth noting that, in the corollary below, our error term  $\frac{C_n}{A^2} e^{-c_n A^2}$  is better than Wang's, and neither is there any restriction on the positive scale  $\tau$ .

**Corollary 1.9** (Local monotonicity formula of Wang [W18, Theorem 5.4]). *Assume that  $[0, T] \subset I$ . Let  $A \geq 1000n$  be a large constant and let  $x_0 \in M$  be a fixed point satisfying*

$$\text{Ric}_{g_t}(x) \leq A/t \quad \text{for all } t \in (0, T] \text{ and } x \in B_t(x_0, \sqrt{t}).$$

Then, for any  $\tau > 0$ , we have

$$\mu\left(g_T, B_T\left(x_0, 8A\sqrt{T}\right), \tau\right) - \mu\left(g_0, B_0\left(x_0, 20A\sqrt{T}\right), \tau + T\right) \geq -\frac{C_n}{A^2} e^{-c_n A^2},$$

where  $C_n < \infty$  and  $c_n > 0$  are both dimensional constants.

Next, we improve our previous result [CMZ21a, Theorem 1.8], which says that the Nash entropy depends on the local volume lower bound at an  $H_n$ -center. Seeing that it determines not only the Nash entropy (c.f. [Bam20a, Theorem 8.1]), but also the local  $\nu$ -functional (c.f. Corollary 1.4), the local volume lower bound is a relatively stronger assumption. We shall weaken the assumption by showing that the Nash entropy depends only on the local  $\mu$ -functional near an  $H_n$ -center of the base point.

**Theorem 1.10.** *Assume that  $[-r^2, 0] \subseteq I$ . Furthermore, assume that  $R_{g_{-r^2}} \geq R_{\min}$ . Then, for any  $x_0 \in M$ , any  $H_n$ -center  $(z, -r^2)$  of  $(x_0, 0)$ , and any  $A \geq 8$ , we have*

$$\mu \left( B_{-r^2} \left( z, 2A\sqrt{H_n}r \right), g_{-r^2}, r^2 \right) \leq \mathcal{N}_{x_0,0}(r^2) + C(n, R_{\min}r^2, A), \quad (1.4)$$

where

$$C(n, R_{\min}r^2, A) = \frac{C_n}{A^2} e^{-\frac{A^2}{20}} + 8 \left( e^{-\frac{A^2}{20}} \cdot (n - 2R_{\min}r^2) + e^{-\frac{A^2}{40}} \cdot (n - 2R_{\min}r^2)^{\frac{1}{2}} \right),$$

and  $C_n$  is a dimensional constant.

The sharpness of the above theorem can be seen from the fact that, if we take  $A \rightarrow \infty$ , then (1.4) reduces to a consequence of (2.4)

$$\mathcal{N}_{x_0,0}(r^2) \geq \mathcal{W}_{x_0,0}(r^2) \geq \mu(g_{-r^2}, r^2).$$

It can be easily observed that [CMZ21a, Theorem 1.8] becomes a corollary of Theorem 1.10 and Corollary 1.4; this shows that Theorem 1.10 is indeed stronger than [CMZ21a, Theorem 1.8]. Furthermore, later we shall see that Perelman's pseudolocality theorem [Per02, Theorem 10.1] is also a corollary of Theorem 1.7 and Theorem 1.10.

**Corollary 1.11** ([CMZ21a, Theorem 1.8]). *Assume that  $[-2r^2, 0] \subseteq I$ . Let  $x_0$  be a point on  $M$  and  $(z, -r^2)$  be an  $H_n$ -center of  $(x_0, 0)$ . Furthermore, assume that*

$$|B_{-r^2}(z, r)|_{-r^2} \geq \alpha r^n,$$

where  $\alpha$  is a positive constant. Then we have

$$\mathcal{N}_{x_0,0}(r^2) \geq -\beta(n, \alpha),$$

where  $\beta$  is a positive constant depending only on  $n$  and  $\alpha$ .

*Proof of Corollary 1.11 assuming Theorem 1.10 and Corollary 1.4.* By applying the maximum principle to the evolution equation of  $R$ , we have

$$r^2 R_{g_{-r^2}} \geq -\frac{n}{2}.$$

By Corollary 1.4, we have

$$\mu \left( B_{-r^2} \left( z, 8\sqrt{H_n}r \right), g_{-r^2}, r^2 \right) \geq -C(n, \alpha).$$

The conclusion then follows from Theorem 1.10.  $\square$

## 1.4 Perelman's noncollapsing improving and pseudolocality theorems

As we have mentioned above, previous local monotonicity theorems can be understood as special examples of ours, in which the local geometric assumptions control the position of an  $H_n$ -center. Perelman's noncollapsing improving theorem [Per02, Theorem 8.2] is one of them. Recall that, by a second-variation argument, Perelman discovered that the distance function along Ricci flow is a supersolution to a linear parabolic equation, and subsequently proved that the geometric information in a space-time neighborhood of a point determines the geometric information of a larger space-time neighborhood. Bamler [Bam20a, Theorem 3.5] has refined the second variation argument, and the consequence is a much better property called  $H_n$ -concentration which we already mentioned above. As we have seen in Subsection 1.2, it is along the  $P^*$ -parabolic neighborhoods, not the classical parabolic neighborhoods, that the geometric information is diffused, and the local geometric assumptions near a point is nothing but that which controls the shape of a  $P^*$ -parabolic neighborhood.

Therefore, it is no surprising that Perelman's noncollapsing improving theorem is a consequence of our local monotonicity formula. In fact, the point emphasized in the above paragraph is inspired by the work of Jian [J21], who provided an improved version of [Per02, Theorem 8.2] and [W18, Theorem 1.1]. Hereby we also prove a stronger version of Jian's result, which is simply a combination of Jian [J21] and Corollary 1.5 above.

**Theorem 1.12** (Noncollapsing improving). *Assume that  $[-2r^2, 0] \subseteq I$ . Let  $x_0 \in M$  be a fixed point. If*

$$\int_{-r^2}^0 \sqrt{|t|} R(x_0, t) dt \leq Ar,$$

$$|B_{-r^2}(x_0, r)|_{-r^2} \geq A^{-1}r^n,$$

then

$$\nu(B_0(x_0, Ar), g_0, A^2r^2) \geq -C(n, A).$$

As a consequence, there is a constant  $\kappa = \kappa(n, A) > 0$ , such that for any ball  $B := B_0(y, \rho)$  satisfying  $y \in B_0(x_0, Ar)$  and  $\rho \in (0, Ar]$ , we have

$$\sup_B R_{g_0} \leq \rho^{-2} \implies |B|_0 \geq \kappa \rho^n.$$

Another application of our local monotonicity formula is to simplify the proof of Perelman's pseudolocality theorem [Per02, Theorem 10.1]. By Theorem 1.10, if the geometry at an  $H_n$ -center is regular enough, namely, almost Euclidean, then the Nash entropy is almost equal to zero. Furthermore, Bamler's  $\varepsilon$ -regularity theorem [Bam20a, Theorem 10.3] shows that the smallness of Nash entropy implies the regularity at the base point. Combining these two facts, Perelman's pseudolocality theorem follows naturally. Here we present B. Wang's improved version of Perelman's pseudolocality theorem [W20, Theorem 1.2].

**Theorem 1.13** (Pseudolocality theorem [Per02, W20]). *For any  $\alpha \in (0, \frac{1}{100n})$ , there is a  $\delta = \delta(n, \alpha) > 0$ , such that the following holds. Assume that  $[0, T] \subset I$ . Let  $x_0 \in M$  be a point satisfying*

$$\inf_{0 < t \leq T} \mu \left( B_0 \left( x_0, \delta^{-1} \sqrt{t} \right), g_0, t \right) \geq -\delta^2. \quad (1.5)$$



Then, for any  $t \in (0, T]$  and any  $x \in B_t(x_0, \alpha^{-1}\sqrt{t})$ , we have

$$|\text{Rm}_{g_t}|(x) \leq \alpha t^{-1} \quad (1.6)$$

$$\inf \left\{ \rho^{-n} |B_t(x, \rho)|_t : \rho \in (0, \alpha^{-1}\sqrt{t}) \right\} \geq (1 - \alpha)\omega_n, \quad (1.7)$$

$$\text{inj}_{g_t}(x) \geq \alpha^{-1}\sqrt{t}, \quad (1.8)$$

where  $\omega_n$  is the volume of the  $n$ -dimensional unit disk in the Euclidean space.

*Remark.* Another pseudolocality theorem of Perelman [Per02, Theorem 10.3] (c.f. [Lu10]) can also be simplified by our method. One may combine the method in the proof of Theorem 1.13 with the argument in [Lu10] to obtain a proof slightly simpler than Lu's—at least, like the proof of Theorem 1.13, the point picking argument can be avoided, and the division into multiple cases is not necessary. The details are left to the readers.

## 1.5 Application to ancient Ricci flows

Finally, we show that our uniform Sobolev inequality for ancient Ricci flows [CMZ21b] is merely a consequence of our local  $\nu$ -functional estimate in Theorem 1.1.

**Theorem 1.14.** ([CMZ21b, Theorem 1.1]) *Let  $I = (-\infty, 0]$  and assume that  $(M^n, g_t)_{t \in I}$  has uniformly bounded Nash entropy, namely, there is a point  $(x_0, t_0) \in M \times (-\infty, 0]$  such that*

$$\mu_\infty := \lim_{\tau \rightarrow \infty} \mathcal{N}_{x_0, t_0}(\tau) > -\infty.$$

Then we have

$$\inf_{t \leq 0} \nu(g_t) = \mu_\infty > -\infty.$$

The following corollary is then a standard consequence by applying the computations in [Zhq10, Ye15].

**Corollary 1.15.** ([CMZ21b, Corollary 1.2]) *Under the same condition as the above theorem, we have*

(1) *Logarithmic Sobolev inequality: for any compactly supported locally Lipschitz function  $u$  on  $(M, g_t)$ , where  $t \leq 0$ , and positive scale  $\tau > 0$ , we have*

$$\begin{aligned} \int_M u^2 \log u^2 dg_t - \log \left( \int_M u^2 dg_t \right) \int_M u^2 dg_t + \left( \mu_\infty + n + \frac{n}{2} \log(4\pi\tau) \right) \int_M u^2 dg_t \\ \leq \tau \int_M (4|\nabla u|^2 + Ru^2) dg_t. \end{aligned}$$

(2) *Sobolev inequality: for any compactly supported locally Lipschitz function  $u$  on  $(M, g_t)$ , where  $t \leq 0$ , we have*

$$\left( \int_M |u|^{\frac{2n}{n-2}} dg_t \right)^{\frac{n-2}{n}} \leq C(n) e^{-\frac{2\mu_\infty}{n}} \int_M (4|\nabla u|^2 + Ru^2) dg_t.$$

Recall that in [CMZ21b], we have applied Bamler’s tangent flow at infinity as an “initial manifold” to apply Q. Zhang’s argument [Zhq10]. However, this requires that Bamler’s theory of singular spaces [Bam20c] can be extended to noncompact manifolds. This, though almost certainly true, is not yet fully technically verified in every detail. Indeed, Bamler [Bam21] has sketched how his theory can be generalized to noncompact manifolds. The proof of Theorem 1.14 which we shall present in this article does not rely on tangent flow at infinity, and hence is much more technically transparent and reader-friendly, for it is nothing but taking  $r$  to infinity in formula (1.1).

To help the reader navigate our paper, the titles of Sections 3–7 are made identical to the titles of the subsections of Section 1. For a result presented in a subsection of Section 1, one may expect its proof to appear in the section bearing the same title.

**Acknowledgement.** The third-named author would like to thank Professor Jiaping Wang for many helpful discussions, and for suggesting this problem. The authors would like to thank Professor Richard Bamler for suggesting a statement as presented in Theorem 1.10 after the first draft of this paper. The authors would like to thank Professor Richard Bamler and Professor Bennett Chow for their constant and helpful communicating of ideas.

## 2 Preliminaries

### 2.1 Perelman’s entropy and Nash entropy

Let  $(M^n, g)$  be a Riemannian manifold,  $\tau > 0$ , and  $u$  be a nonnegative function with unit integral such that  $\sqrt{u}$  is smooth, then we define the  $\overline{\mathcal{W}}$  functional as

$$\overline{\mathcal{W}}(g, u, \tau) := \int_M (\tau (|\nabla \log u|^2 + R) - \log u) u \, dg - \frac{n}{2} \log(4\pi\tau) - n.$$

In fact, writing  $u := (4\pi\tau)^{-\frac{n}{2}} e^{-f}$ , it is easy to observe that  $\overline{\mathcal{W}}$  is the same as Perelman’s  $\mathcal{W}$ -functional

$$\mathcal{W}(g, f, \tau) := \int_M (\tau (|\nabla f|^2 + R) + f - n) (4\pi\tau)^{-\frac{n}{2}} e^{-f} \, dg.$$

For any region  $\Omega \subseteq M$ , following [W18], we define

$$\begin{aligned} \mu(\Omega, g, \tau) &:= \inf \left\{ \overline{\mathcal{W}}(g, u, \tau) : u \geq 0, \sqrt{u} \in C_0^\infty(\Omega), \int_M u \, dg = 1 \right\}, \\ \nu(\Omega, g, \tau) &:= \inf_{0 < s \leq \tau} \mu(\Omega, g, s). \end{aligned}$$

Indeed,  $\nu(\Omega, g, \tau)$  is a local Sobolev constant of the region  $\Omega$ . If  $\Omega = M$  and  $\tau = \infty$ , then we shall denote by  $\nu(g) := \nu(M, g, \infty)$  the  $\nu$ -functional of Perelman [Per02].

By Perelman’s monotonicity formula, if  $u := (4\pi\tau_t)^{-\frac{n}{2}} e^{-f}$  is a positive solution to the conjugate heat equation

$$\square_t^* u := -\partial_t u - \Delta_{g_t} u + R_{g_t} u = 0$$

with unit integral, where  $\frac{d}{dt}\tau_t = -1$ , then  $\overline{\mathcal{W}}(g_t, u(\cdot, t), \tau_t)$  is, in general, non-decreasing in  $t$ . More precisely, assuming the validity of integration by parts at infinity, we have

$$\frac{d}{dt} \overline{\mathcal{W}}(g_t, u(\cdot, t), \tau_t) = \int_M 2\tau_t \left| \text{Ric}_{g_t} + \nabla^2 f - \frac{1}{2\tau_t} g_t \right|^2 u_t \, dg_t \geq 0. \quad (2.1)$$

Here the conjugate heat operator  $\square_t^*$  is the conjugate of the heat operator

$$\square_t := \partial_t - \Delta_{g_t},$$

since, provided that integration by parts is valid, we have

$$\frac{d}{dt} \int_M uv \, dg_t = \int_M (v \square_t u - u \square_t^* v) \, dg_t.$$

Next, we introduce the definition of Perelman's entropy and the Nash entropy. Let  $(x_0, t_0) \in M \times I$  be a fixed point in space-time. Then, letting  $u_t = (4\pi(t_0 - t))^{-\frac{n}{2}} e^{-f_t} := K(x_0, t_0 | \cdot, t)$  be the conjugate heat kernel based at  $(x_0, t_0)$ , Perelman's entropy and the Nash entropy based at  $(x_0, t_0)$  are respectively defined as

$$\mathcal{W}_{x_0, t_0}(\tau) = \overline{\mathcal{W}}(g_{t_0 - \tau}, u_{t_0 - \tau}, \tau), \quad (2.2)$$

$$\mathcal{N}_{x_0, t_0}(\tau) = \int_M f_{t_0 - \tau} \, d\nu_{x_0, t_0 | t_0 - \tau} - \frac{n}{2}, \quad (2.3)$$

where  $\tau > 0$  and  $t_0 - \tau \in I$ , and the evolving probability measure

$$\nu_{x_0, t_0 | t}(\Omega) := \int_{\Omega} u_t \, dg_t, \quad \Omega \subseteq M \text{ is measurable, and } t \in (-\infty, t_0] \cap I,$$

is also called the *conjugate heat kernel* based at  $(x_0, t_0)$  when there is no ambiguity. It is well-known that both  $\mathcal{W}_{x_0, t_0}(\tau)$  and  $\mathcal{N}_{x_0, t_0}(\tau)$  are monotonically decreasing in  $\tau$ , and that

$$0 \geq \mathcal{N}_{x_0, t_0}(\tau) \geq \mathcal{W}_{x_0, t_0}(\tau) \quad \text{for all } \tau > 0 \quad \text{and} \quad t_0 - \tau \in I. \quad (2.4)$$

## 2.2 Bamler's Harnack inequality for the Nash entropy

Bamler's Harnack inequality [Bam20a, Corollary 5.11] enables us to compare the Nash entropy based not only at different points on the same time-slice, but also at points on different time-slices. Let us recall this result. Note that the second and the third authors [MZ21, Corollary 4.5] have already generalized this theorem to the noncompact setting, assuming only bounded curvature on each compact time interval.

**Theorem 2.1.** ([Bam20a, Corollary 5.11]) *If  $R_{g_{t^*}} \geq R_{\min}$ ,  $s < t^* \leq \min\{t_1, t_2\}$ , and  $s, t^*, t_1, t_2 \in I$ , then for any  $x_1, x_2 \in M$ , we have*

$$\mathcal{N}_{x_1, t_1}(t_1 - s) - \mathcal{N}_{x_2, t_2}(t_2 - s) \leq \left( \frac{n}{2(t^* - s)} - R_{\min} \right)^{\frac{1}{2}} \text{dist}_{W_1}^{g_{t^*}}(\nu_{x_1, t_1 | t^*}, \nu_{x_2, t_2 | t^*}) + \frac{n}{2} \log \left( \frac{t_2 - s}{t^* - s} \right).$$

In the statement of the above theorem,  $\text{dist}_{W_1}$  is the  $W_1$ -Wasserstein distance between two probability measures on a metric space. Recall that the  $P^*$ -parabolic neighborhood of Bamler [Bam20a, Definition 9.2] is defined in terms of the  $W_1$ -Wasserstein distance. Let  $(x_0, t_0) \in M \times I$ . Then, for all  $A > 0$  and  $T^-, T^+ \geq 0$ ,  $P^*(x_0, t_0 | A, -T^-, T^+)$  is the set of all points  $(x, t) \in M \times (I \cap [t_0 - T^-, t_0 + T^+])$ , satisfying

$$\text{dist}_{W_1}^{g_{t_0 - T^-}}(\nu_{x_0, t_0 | t_0 - T^-}, \nu_{x, t | t_0 - T^-}) < A.$$

Therefore, Theorem 2.1 can be stated in the way that the difference of the Nash entropy is determined by the “ $P^*$ -distance” between the base points.

The notion of  $H_n$ -center, where  $H_n := \frac{(n-1)\pi^2}{2} + 4$ , is exactly that which characterizes “shortest  $P^*$ -distance” ([Bam20a, Definition 3.10]). Let us recall that  $(z, s) \in M \times I$  is an  $H_n$ -center of  $(x, t) \in M \times I$ , where  $s < t$ , if

$$\text{Var}(\delta_z, \nu_{x,t}|_s) \leq H_n(t-s),$$

where  $\text{Var}$  is the variance of two probability measures ([Bam20a, Definition 3.1]). By the definitions of variance and  $W_1$ -Wasserstein distance, we also have

$$\text{dist}_{W_1}^{g_s}(\delta_z, \nu_{x,t}|_s) \leq \sqrt{H_n(t-s)}. \quad (2.5)$$

Therefore, a path of  $H_n$ -centers, heuristically speaking, can be understood as a “ $P^*$ -geodesic”. Given an arbitrary point  $(x, t) \in M \times I$  and an arbitrary time  $s \in I \cap (-\infty, t)$ , there always exists a point  $z$ , such that  $(z, s)$  is an  $H_n$ -center of  $(x, t)$ .

### 2.3 Reduced distance as a “sub”-conjugate heat kernel

Perelman’s reduced distance, though a very interesting and important object of study in its own right, is also a central tool in the analysis of the conjugate heat equation. Let  $(x_0, t_0) \in M \times I$  be a fixed point in space-time, then the reduced distance function based at  $(x_0, t_0)$  is defined to be

$$\ell_{x_0, t_0}(x, \tau) := \frac{1}{2\sqrt{\tau}} \inf_{\gamma} \int_0^{\tau} \sqrt{\eta} \left( R_{g_{t_0-\eta}}(\gamma(\eta)) + |\gamma'(\eta)|_{g_{t_0-\eta}}^2 \right) d\eta,$$

where  $x \in M$ ,  $\tau > 0$ ,  $t_0 - \tau \in I$ , and the infimum is taken over all piecewise smooth curves  $\gamma : [0, \tau] \rightarrow M$  satisfying  $\gamma(0) = x_0$  and  $\gamma(\tau) = x$ .

The following Theorem is a restatement of [Per02, Corollary 9.5].

**Theorem 2.2.** *For any  $(x, t)$  and  $(x_0, t_0) \in M \times I$  with  $t < t_0$ , we have*

$$(4\pi(t_0 - t))^{-\frac{n}{2}} e^{\ell_{x_0, t_0}(x, t)} \leq K(x_0, t_0 | x, t),$$

where  $K$  is the fundamental solution to the conjugate heat equation.

## 3 Local Sobolev inequalities on Ricci flow

In this section, we present the proof of Theorem 1.1, Corollary 1.3, and Corollary 1.4, where the last two results follow immediately from the first. Our main technique is Bamler’s Harnack inequality of the Nash entropy [Bam20a, Corollary 5.11] (see Theorem 2.1 above), which shows how the Nash entropy is dependent on the base point. Given a bound of the Nash entropy based at a fixed point  $(x_0, t_0)$  in space-time, we may, by applying [Bam20a, Corollary 5.11], show that the Nash entropy based everywhere nearby is similarly bounded. Since the Nash entropy can also be regarded as a concave functional of the conjugate heat kernel, we may show, by applying Jensen’s inequality, that any test function whose support is in a disk center at  $(x_0, t_0)$  has bounded “Nash functional”, which also implies a bound of the local  $\nu$ -functional up to a certain scale.

By parabolic rescaling, we assume that  $r = 1$ , and we let  $\tau > 0, A < \infty$  be arbitrarily fixed constants. We shall pick an arbitrary test function and verify that its  $\overline{W}$  functional is bounded

from below by the Nash entropy in the way as stated in the theorem. To this end, let  $u_0$  be an arbitrary nonnegative test function such that  $\sqrt{u_0} \in C_0^\infty(B_0(x_0, A))$  and  $\int_M u_0 dg_0 = 1$ . We solve the backward conjugate heat equation

$$\square_t^* u_t := (-\partial_t - \Delta_{g_t} + R_{g_t})u_t = 0 \quad \text{on} \quad M \times [-1, 0],$$

with the initial data being  $u_0$  at time  $t = 0$ . Then for all  $t \in [-1, 0]$ ,  $u_t$  can be written as

$$u_t(x) = \int_M K(y, 0 | x, t) u_0(y) dg_0(y), \quad (3.1)$$

where  $K$  is the fundamental solution to the conjugate heat equation. Then, the evolving probability measure

$$\mu_t(\Omega) := \int_\Omega u_t dg_t, \quad \Omega \subseteq M \text{ is measurable and } t \in [-1, 0]$$

is what Bamler [Bam20b] calls a *conjugate heat flow*. For  $t \in [-1, 0)$ , we define

$$\begin{aligned} \tau_t &:= \tau - t, \quad u_t :=: (4\pi\tau_t)^{-n/2} e^{-f_t}, \\ \overline{\mathcal{N}}(t) &:= \int_M f_t d\mu_t - \frac{n}{2}, \\ \overline{\mathcal{W}}(t) &:= \int_M (\tau_t(|\nabla f_t|^2 + R_{g_t}) + f_t - n) d\mu_t. \end{aligned}$$

**Lemma 3.1.**

$$\overline{\mathcal{W}}(0) := \lim_{t \rightarrow 0^-} \overline{\mathcal{W}}(t) = \overline{\mathcal{W}}(g_0, u_0, \tau), \quad (3.2)$$

$$-\frac{d}{dt} (\tau_t \overline{\mathcal{N}}(t)) = \overline{\mathcal{W}}(t), \quad (3.3)$$

$$\frac{d}{dt} \overline{\mathcal{W}}(t) \geq 0. \quad (3.4)$$

*Proof.* The proof of the lemma, especially of (3.2), is not essentially different from [W18, §4], and a detailed treatment can be found in [CMZ21a, §9]. Indeed, by [CCGGIHKLLN10, Theorem 26.25 and Theorem 26.31], there is a constant  $C$  depending only on the local geometry bound in  $B_0(x_0, 2A) \times [-1, 0]$ , such that

$$\frac{1}{C(t-s)^{\frac{n}{2}}} \exp\left(-\frac{C \text{dist}_t^2(x, y)}{(t-s)}\right) \leq K(x, t | y, s) \leq \frac{C}{(t-s)^{\frac{n}{2}}} \exp\left(-\frac{\text{dist}_t^2(x, y)}{C(t-s)}\right), \quad (3.5)$$

whenever  $-1 \leq s < t \leq 0$  and either  $x$  or  $y$  is in  $B_0(x_0, A)$ . As a consequence, we have

$$u(x, t) \leq \frac{C}{(-t)^{\frac{n}{2}}} \exp\left(-\frac{\text{dist}_0^2(x, B_0(x_0, A))}{C(-t)}\right), \quad \text{for all } (x, t) \in M \times [-1, 0). \quad (3.6)$$

One may then follow the argument in [CMZ21a, §9] line by line to obtain (3.2).

For (3.4), one needs only to observe that the bounds in (3.6) together with some standard gradient estimates (c.f. [EKNT08, Theorem 10] or [CMZ21a, Lemma 9.11]) enable integration by parts at infinity. Hence, (3.4) is a consequence of Perelman's monotonicity formula (2.1).

Finally, given the validity of integration by parts at infinity, (3.3) follows from standard computation; one may refer to [CMZ21b] for more details.  $\square$

**Proof of Theorem 1.1.** Our goal is to estimate the lower bound of  $\overline{\mathcal{W}}(0)$ . By (3.2)—(3.4), we compute

$$\begin{aligned} & \tau_{-1}\overline{\mathcal{N}}(-1) - \tau_0\overline{\mathcal{N}}(0) \\ &= - \int_{-1}^0 \frac{d}{dt} (\tau_t\overline{\mathcal{N}}(t)) dt = \int_{-1}^0 \overline{\mathcal{W}}(t) dt \\ &\leq \overline{\mathcal{W}}(0) = \overline{\mathcal{W}}(g_0, u_0, \tau), \end{aligned}$$

where we have defined

$$\overline{\mathcal{N}}(0) := \lim_{t \rightarrow 0^-} \overline{\mathcal{N}}(t).$$

Therefore, it remains to estimate  $\overline{\mathcal{N}}(-1)$  and  $\overline{\mathcal{N}}(0)$ .

First, we estimate  $\overline{\mathcal{N}}(0)$ . By the maximum principle, for  $s \in (-1, 0]$ , we have

$$R_{g_s} \geq -\frac{n}{2(s+1)} \quad \text{on } M. \quad (3.7)$$

For each  $s \in (-1, 0)$  close to 0, we have

$$\begin{aligned} \overline{\mathcal{N}}(s) &= \overline{\mathcal{W}}(s) + \frac{n}{2} - \int \tau_s (|\nabla f_s|^2 + R_{g_s}) d\mu_s \\ &\leq \overline{\mathcal{W}}(s) + \frac{n}{2} + \frac{n\tau_s}{2(s+1)}. \end{aligned}$$

By taking  $s \rightarrow 0^-$ , we have

$$\overline{\mathcal{N}}(0) \leq \overline{\mathcal{W}}(0) + \frac{n}{2}(1 + \tau).$$

Then

$$\begin{aligned} \overline{\mathcal{W}}(0) &\geq \tau_{-1}\overline{\mathcal{N}}(-1) - \tau_0\overline{\mathcal{N}}(0) \\ &\geq (1 + \tau)\overline{\mathcal{N}}(-1) - \tau\overline{\mathcal{W}}(0) - \frac{n}{2}\tau(1 + \tau), \end{aligned}$$

and hence

$$\overline{\mathcal{W}}(0) \geq \overline{\mathcal{N}}(-1) - \frac{n}{2}\tau. \quad (3.8)$$

It remains to estimate  $\overline{\mathcal{N}}(-1)$ . To this end, we recall the definition of  $\overline{\mathcal{N}}$ :

$$\begin{aligned} \overline{\mathcal{N}}(-1) &= \int f_{-1}u_{-1} dg_{-1} - \frac{n}{2} \\ &= - \int u_{-1} \log u_{-1} dg_{-1} - \frac{n}{2} - \frac{n}{2} \log(4\pi(1 + \tau)). \end{aligned} \quad (3.9)$$

By (3.1), we have

$$u_{-1}(x) = \int K(y, 0 | x, -1) u_0(y) dg_0(y) = \int K(y, 0 | x, -1) d\mu_0(y),$$

where  $\mu_0$  is a probability measure. Hence, by Jensen's inequality, we have

$$u_{-1} \log u_{-1}(x) \leq \int K(y, 0 | x, -1) \log K(y, 0 | x, -1) d\mu_0(y).$$

It follows from integrating both sides with the measure  $dg_{-1}$  that

$$\begin{aligned} & \int u_{-1} \log u_{-1}(x) dg_{-1}(x) \\ & \leq \int \int K(y, 0 | x, -1) \log K(y, 0 | x, -1) d\mu_0(y) dg_{-1}(x) \\ & = \int \int K(y, 0 | x, -1) \log K(y, 0 | x, -1) dg_{-1}(x) d\mu_0(y) \\ & = - \int \mathcal{N}_{y,0}(1) d\mu_0(y) - \frac{n}{2} - \frac{n}{2} \log(4\pi). \end{aligned} \tag{3.10}$$

Here, it is easy to verify that the change of order of the integration is valid, since  $\mu_0$  is supported in  $B_0(x_0, A)$  and since, by (3.5),  $K(y, 0 | \cdot, -1)$  has rapid decay at infinity.

Combining (3.9) and (3.10), we have

$$\overline{\mathcal{N}}(-1) \geq \int \mathcal{N}_{y,0}(1) d\mu_0(y) - \frac{n}{2} \log(1 + \tau). \tag{3.11}$$

We shall estimate the integral on the right-hand-side. Let us fix an arbitrary  $y \in \text{spt}(\mu_0) \subseteq B_0(x_0, A)$ . By the definition of  $W_1$ -distance, we have

$$\text{dist}_{W_1}^{g_0}(\nu_{y,0|0}, \nu_{x_0,0|0}) = \text{dist}_{W_1}^{g_0}(\delta_y, \delta_{x_0}) = \text{dist}_{g_0}(x_0, y) < A.$$

We may then apply Theorem 2.1 with  $x_0 \rightarrow x_1$ ,  $y \rightarrow x_2$ ,  $0 \rightarrow t_1 = t_2 = t^*$ ,  $-1 \rightarrow s$ , and obtain

$$\begin{aligned} \mathcal{N}_{x_0,0}(1) & \leq \mathcal{N}_{y,0}(1) + \left(\frac{n}{2} - \inf R_{g_0}\right)^{\frac{1}{2}} \text{dist}_{W_1}^{g_0}(\nu_{y,0|0}, \nu_{x_0,0|0}) + \frac{n}{2} \log \frac{1}{1} \\ & \leq \mathcal{N}_{y,0}(1) + \sqrt{n}A, \end{aligned} \tag{3.12}$$

where we have also applied (3.7). It follows from (3.8), (3.11), and (3.12) that

$$\begin{aligned} \overline{\mathcal{W}}(0) & \geq \overline{\mathcal{N}}(-1) - \frac{n}{2}\tau \\ & \geq \int \mathcal{N}_{y,0}(1) d\mu_0(y) - \frac{n}{2}\tau - \frac{n}{2} \log(1 + \tau) \\ & \geq \mathcal{N}_{x_0,0}(1) - \sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2} \log(1 + \tau). \end{aligned}$$

Since the test function  $u_0$  is arbitrarily fixed, we have

$$\mu(B_0(x_0, A), g_0, \tau) \geq \mathcal{N}_{x_0,0}(1) - \sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2} \log(1 + \tau).$$

To see that the local  $\nu$ -functional is also bounded, one needs only to observe that, for any  $s \in (0, \tau]$ , we have

$$\begin{aligned} \mu(B_0(x_0, A), g_0, s) & \geq \mathcal{N}_{x_0,0}(1) - \sqrt{n}A - \frac{n}{2}s - \frac{n}{2} \log(1 + s) \\ & \geq \mathcal{N}_{x_0,0}(1) - \sqrt{n}A - \frac{n}{2}\tau - \frac{n}{2} \log(1 + \tau). \end{aligned}$$

This finishes the proof of the theorem.  $\square$

**Proof of Corollary 1.3.** We still let  $r = 1$ . The idea of the proof is to use the local geometric conditions (1.3) to estimate the distance between  $(x_0, 1/2)$  and an  $H_n$ -center of  $(x_0, 1)$ . First of all, applying (1.3) and the standard distortion estimates of distance and volume, we may easily obtain

$$\begin{aligned} |B_{1/2}(x_0, c_n)|_{1/2} &\geq c(n, A), \\ |\text{Ric}| &\leq \frac{1}{n^2} \quad \text{on} \quad B_{1/2}(x_0, c_n) \times \left[\frac{1}{2}, 1\right], \end{aligned} \quad (3.13)$$

where  $c_n$  is a positive constant depending only on  $n$ . We shift the base time from 0 to  $1/2$  for the purpose of applying [CMZ21a, Theorem 1.8].

Let us fix any point  $x \in B_{1/2}(x_0, c_n)$ , and let  $\gamma : [0, 1/2] \rightarrow M$  be a minimizing  $g_{1/2}$ -geodesic with constant speed connecting  $x_0$  and  $x$ . Then, by (3.13), we estimate

$$|\gamma'(s)|_{g_{1-s}}^2 \leq e^{\frac{1}{n^2}} |\gamma'(s)|_{g_{1/2}}^2 \leq 4c_n^2 e^{\frac{1}{n^2}} \quad \text{for all} \quad s \in \left[0, \frac{1}{2}\right],$$

and consequently

$$\ell_{x_0, 1} \left(x, \frac{1}{2}\right) \leq \frac{1}{2} \int_0^{1/2} \sqrt{\eta} \left( R_{g_{1-\eta}}(\gamma(\eta)) + |\gamma'(s)|_{g_{1-\eta}}^2 \right) d\eta \leq C(n).$$

Hence, Theorem 2.2 implies that

$$K \left(x_0, 1 \mid x, \frac{1}{2}\right) \geq (4\pi)^{-\frac{n}{2}} e^{-\ell_{x_0, 1}(x, 1/2)} \geq c(n) \quad \text{for all} \quad x \in B_{1/2}(x_0, c_n),$$

and by (3.13), we have

$$\nu_{x_0, 1} |_{1/2} (B_{1/2}(x_0, c_n)) \geq c(n) |B_{1/2}(x_0, c_n)|_{1/2} \geq c_0(n, A).$$

Letting  $(z, 1/2)$  be an  $H_n$ -center of  $(x_0, 1)$ , we have, by [Bam20a, Proposition 3.13],

$$\text{dist}_{1/2}(x_0, z) \leq c_n + \sqrt{\frac{1}{2c_0(n, A)} H_n} \leq C(n, A).$$

By (3.13) again, we have

$$|B_{1/2}(z, C(n, A) + c_n)|_{1/2} \geq |B_{1/2}(x_0, c_n)|_{1/2} \geq c(n, A).$$

It then follows from [CMZ21a, Theorem 1.8] and the remark thereof that

$$\mathcal{N}_{x_0, 1}(1/2) \geq -C(n, A).$$

The corollary follows immediately from Theorem 1.1. □

**Proof of Corollary 1.4.** We still assume  $r = 1$  by applying a parabolic scaling. We shall only prove the bound of the local  $\nu$ -functional, and the noncollapsing result is a straightforward consequence. By (3.7) again,  $R_{g_s} \geq -n$  for  $s \in [-1/2, 0]$ . Hence, [Bam20a, Theorem 8.1] implies that

$$\alpha \leq |B_0(x_0, 1)|_0 \leq C(n) \exp(\mathcal{N}_{x_0, 0}(1/2)). \quad (3.14)$$

Applying Theorem 1.1, we have

$$\begin{aligned} \nu(B_0(x_0, A), g_0, \tau) &\geq \mathcal{N}_{x_0, 0}(1/2) - 2\sqrt{n}A - 4n\tau \\ &\geq \log \alpha - 4n(A + \tau) - C(n). \end{aligned}$$

This finishes the proof. □



## 4 The “reaction-diffusion” property of the Ricci flow

In this section, we prove Corollary 1.5 and Theorem 1.6. The proof of Corollary 1.5 is no more than an application of Theorem 1.1 and Theorem 2.1. In fact, given the assumption in Corollary 1.5, Bamler’s Harnack inequality shows that the Nash entropy based at any point in  $P^*(x_0, 0 | Ar, 0, A^2r^2)$  is bounded, and this corollary follows from Theorem 1.1. The proof of Theorem 1.6 uses the fact that the conjugate heat kernel concentrates around an  $H_n$ -center, and that a solution to the heat equation can be viewed as a convolution with a conjugate heat kernel.

**Proof of Corollary 1.5.** The proof is no more than an application of the definition of  $P^*$ -parabolic neighborhood. As usual, we assume, by parabolic rescaling, that  $r = 1$ . Let  $(y, s)$  be an arbitrary point in  $P^*(x_0, 0 | A, 0, A^2)$ , and we shall estimate the Nash entropy based at  $(y, s)$ . By the definition of  $P^*$ -parabolic neighborhood, we have

$$\text{dist}_{W_1}^{g_0}(\delta_{x_0}, \nu_{y,s|0}) \leq A.$$

Since (3.7) implies

$$R_{g_0} \geq -\frac{n}{2} \quad \text{on } M,$$

we may apply Theorem 2.1 with  $(x_0, 0) \rightarrow (x_1, t_1)$ ,  $(y, s) \rightarrow (x_2, t_2)$ ,  $0 \rightarrow t^*$ ,  $-1 \rightarrow s$ ,  $-\frac{n}{2} \rightarrow R_{\min}$ , and this yields

$$\begin{aligned} \mathcal{N}_{x_0,0}(1) &\leq \mathcal{N}_{y,s}(1+s) + \left( \frac{n}{2(0-(-1))} - (-\frac{n}{2}) \right)^{\frac{1}{2}} \text{dist}_{W_1}^{g_0}(\nu_{y,s|0}, \delta_{x_0}) + \frac{n}{2} \log \left( \frac{s-(-1)}{0-(-1)} \right) \\ &\leq \mathcal{N}_{y,s}(1+s) + \sqrt{n}A + \frac{n}{2} \log(1+A^2), \end{aligned}$$

where we have used the fact that  $s \in [0, A^2]$ . Therefore, we have

$$\mathcal{N}_{y,s}(1) \geq \mathcal{N}_{y,s}(1+s) \geq \mathcal{N}_{x_0,0}(1) - \sqrt{n}A - \frac{n}{2} \log(1+A^2),$$

and the conclusion follows from Theorem 1.1.  $\square$

**Proof of Theorem 1.6.** By parabolic scaling, we assume  $r = 1$ . Let us define the function  $\tilde{H} : M \times [-1, 0] \rightarrow \mathbb{R}$  as

$$\tilde{H}(x, t) := \int_{\bar{B}_{-1}(z, A)} H(\cdot, -1) d\nu_{x,t|-1} \quad \text{for } (x, t) \in M \times [-1, 0].$$

Obviously, since  $\tilde{H}$  solves the heat equation, and since  $H(\cdot, -1) \geq \tilde{H}(\cdot, -1)$  everywhere on  $M$ , we have that  $H \geq \tilde{H}$  everywhere on  $M \times (0, 1]$  by the maximum principle. Hence, we need only to estimate  $\tilde{H}$  on  $\bar{B}_0(x_0, \Lambda)$ .

By [Bam20a, Proposition 3.14], we have

$$\nu_{x_0,0|-1}(\bar{B}_{-1}(z, A)) \geq 1 - 2 \exp \left( -\frac{1}{8} \left( A - \sqrt{2H_n} \right)^2 \right).$$

Therefore, regarding

$$(x, t) \mapsto \nu_{x,t|-1}(\bar{B}_{-1}(z, A))$$

as a solution to the heat equation, and applying [Bam20a, Theorem 4.1], we have

$$\begin{aligned} \nu_{x,0|-1}(\bar{B}_{-1}(z,A)) &\geq \Phi\left(\Phi^{-1}\left(\nu_{x_0,0|-1}(\bar{B}_{-1}(z,A))\right) - \Lambda\right) \\ &\geq \Phi\left(\Phi^{-1}\left(1 - 2\exp\left(-\frac{1}{8}\left(A - \sqrt{2H_n}\right)^2\right)\right) - \Lambda\right), \end{aligned} \quad (4.1)$$

for all  $x \in B_0(x_0, \Lambda)$ . We will now estimate the right-hand-side. Let

$$\eta = \Phi^{-1}\left(1 - 2\exp\left(-\frac{1}{8}\left(A - \sqrt{2H_n}\right)^2\right)\right).$$

Then, by the definition of  $\Phi$ , we have

$$\begin{aligned} 2\exp\left(-\frac{1}{8}\left(A - \sqrt{2H_n}\right)^2\right) &= \int_{\eta}^{\infty} (4\pi)^{-\frac{1}{2}} e^{-\frac{t^2}{4}} dt \\ &\geq \int_{\eta}^{\eta+4\sqrt{\pi}} (4\pi)^{-\frac{1}{2}} e^{-\frac{t^2}{4}} dt \\ &\geq 2\exp\left(-\frac{1}{4}\left(\eta + 4\sqrt{\pi}\right)^2\right), \end{aligned}$$

and consequently

$$\eta \geq \frac{1}{\sqrt{2}}A - \sqrt{H_n} - 4\sqrt{\pi}. \quad (4.2)$$

Combining (4.1) and (4.2), we have

$$\nu_{x,0|-1}(\bar{B}_{-1}(z,A)) \geq \Phi\left(\frac{1}{\sqrt{2}}A - \Lambda - \sqrt{H_n} - 4\sqrt{\pi}\right) := \Phi_0 \quad \text{for all } x \in B_0(x_0, \Lambda). \quad (4.3)$$

Finally, letting  $x' \in \bar{B}_0(x_0, \Lambda)$  be the point where  $\tilde{H}(\cdot, 0)$  attains its minimum on  $\bar{B}_0(x_0, \Lambda)$ , we have

$$\begin{aligned} \min_{\bar{B}_0(x_0, \Lambda)} H(\cdot, 0) &\geq \min_{\bar{B}_0(x_0, \Lambda)} \tilde{H}(\cdot, 0) = \tilde{H}(x', 0) \\ &= \int_{\bar{B}_{-1}(z,A)} H(\cdot, -1) d\nu_{x',0|-1} \\ &\geq \nu_{x',0|-1}(\bar{B}_{-1}(z,A)) \cdot \min_{\bar{B}_{-1}(z,A)} H(\cdot, -1) \\ &\geq \Phi_0 \cdot \min_{\bar{B}_{-1}(z,A)} H(\cdot, -1). \end{aligned}$$

This finishes the proof of the Theorem.  $\square$

## 5 Local monotonicity and Nash entropy

In this section, we prove Theorem 1.7, Theorem 1.10, and Corollary 1.9. The proof of Theorem 1.7 is inspired by [W18] and [TZ21]. The idea is to evolve the minimizer of the local  $\mu$ -functional at  $t = 0$

using conjugate heat flow, and we may then compare the local  $\mu$ -functional at different time-slices using the differential Harnack inequality in [W18]. The error terms come from the concentration estimate of the heat kernel measure near an  $H_n$ -center [Bam20a, Proposition 3.14]. The proof of Theorem 1.10 is similar to that of Theorem 1.7. Given Theorem 1.7, the proof of Corollary 1.9 is reduced to estimating the distance between  $(x_0, 0)$  and an  $H_n$ -center of  $(x_0, T)$ .

**Proof of Theorem 1.7.** By parabolic rescaling, we may assume that  $r = 1$ . For simplicity, we write

$$B = B_{-1} \left( z, \frac{7}{4}A\sqrt{H_n} \right), \quad B' = B_{-1} \left( z, 2A\sqrt{H_n} \right),$$

where  $(z, -1)$  is an  $H_n$ -center of  $(x_0, 0)$ . Let  $u_0$  be the minimizer of  $\mu := \mu(B_0(x_0, A\sqrt{H_n}), g_0, \tau)$ . Following [W18], we let  $u_t$  be the solution to the conjugate heat equation  $\square_t^* u_t = 0$  on  $M \times [-1, 0]$  with initial data being  $u_0$  at  $t = 0$ . Defining  $\tau_t := \tau - t$ , by [W18, Theorem 4.2], we have

$$\left\{ \tau_t \left( -2 \frac{\Delta u_t}{u_t} + |\nabla \log u_t|^2 + R \right) - \log u - n - \mu - \frac{n}{2} \log(4\pi\tau_t) \right\} u_t \leq 0 \quad \text{on} \quad M \times [-1, 0].$$

As before,  $d\mu_t := u_t dg_t$  is a probability measure for  $t \in [-1, 0]$ .

**Claim:** If  $A \geq 16$ , then we have

$$\mu_{-1}(M \setminus B) \leq 2e^{-\frac{A^2}{20}} < \frac{1}{2}.$$

*Proof of the Claim.* For any  $x \in B_0(x_0, A\sqrt{H_n})$ , let  $(z_x, -1)$  be an  $H_n$ -center of  $(x, 0)$ . Then

$$\begin{aligned} \text{dist}_{-1}(z_x, z) &= \text{dist}_{W_1}^{g_{-1}}(\delta_{z_x}, \delta_z) \\ &\leq \text{dist}_{W_1}^{g_{-1}}(\delta_{z_x}, \nu_{x,0|-1}) + \text{dist}_{W_1}^{g_{-1}}(\nu_{x,0|-1}, \nu_{x_0,0|-1}) + \text{dist}_{W_1}^{g_{-1}}(\delta_z, \nu_{x_0,0|-1}) \\ &\leq 2\sqrt{H_n} + \text{dist}_{W_1}^{g_0}(\nu_{x,0|0}, \nu_{x_0,0|0}) \\ &\leq 2\sqrt{H_n} + A\sqrt{H_n} \\ &\leq \frac{5}{4}A\sqrt{H_n}, \end{aligned}$$

where we used the monotonicity of the  $W_1$ -Wassernstein distance (c.f. [Bam20a, Lemma 2.7]). Then  $B_{-1}(z_x, \frac{1}{2}A\sqrt{H_n}) \subseteq B$  for any  $x \in B_0(x_0, A\sqrt{H_n})$ . By [Bam20a, Proposition 3.14], for any  $x \in B_0(x_0, A\sqrt{H_n})$ , if  $A \geq 16$ , then we have

$$\nu_{x,0|-1}(M \setminus B) \leq \nu_{x,0|-1} \left( (M \setminus B_{-1} \left( z_x, \frac{1}{2}A\sqrt{H_n} \right)) \right) \leq 2e^{-\frac{A^2}{20}}.$$

By the standard semi-group property, we have

$$\mu_{-1}(M \setminus B) = \int_{B_0(x_0, A\sqrt{H_n})} \nu_{x,0|-1}(M \setminus B) \cdot u_0(x) dg_0(x) \leq 2e^{-\frac{A^2}{20}}.$$

□

Let  $\eta$  be a smooth cutoff function supported in  $B'$  such that  $\eta|_B = 1$ ,  $0 \leq \eta \leq 1$ , and  $|\nabla \eta|^2 \leq \frac{160}{A^2 H_n} \eta$ . Define

$$\alpha := \int_M \eta^2 u_{-1} dg_{-1} = \int_M \eta^2 d\mu_{-1}.$$

By the Claim, we have

$$\frac{1}{2} \leq 1 - 2e^{-\frac{A^2}{20}} \leq \alpha \leq 1.$$

Let  $\tilde{u} = \eta^2 u_{-1} / \alpha$ . Then  $\tilde{u} \in C_0^\infty(B')$  and  $\int \tilde{u} dg_{-1} = 1$ . In the following, we omit the measure  $dg_{-1}$  and write  $u = u_{-1}$  when there is no ambiguity. Integrating by parts, we have

$$\begin{aligned} \int |\nabla \log \tilde{u}|^2 \tilde{u} &= \alpha^{-1} \int |\nabla \log u|^2 \eta^2 u + 2\langle \nabla \eta^2, \nabla u \rangle + 4|\nabla \eta|^2 u \\ &= \alpha^{-1} \int (|\nabla \log u|^2 u - 2\Delta u) \eta^2 + 4|\nabla \eta|^2 u. \end{aligned}$$

Then

$$\begin{aligned} \mu(B', g_{-1}, 1 + \tau) &\leq \int (\tau_{-1}(|\nabla \log \tilde{u}|^2 + R)\tilde{u} - \log \tilde{u} \cdot \tilde{u}) dg_{-1} - n - \frac{n}{2} \log(4\pi\tau_{-1}) \quad (5.1) \\ &= \alpha^{-1} \int \left\{ \tau_{-1}(|\nabla \log u|^2 - 2\frac{\Delta u}{u} + R) - \log u - n - \frac{n}{2} \log(4\pi\tau_{-1}) \right\} \eta^2 u \\ &\quad + \alpha^{-1} \int 4|\nabla \eta|^2 u - \int \log \frac{\eta^2}{\alpha} \cdot \frac{\eta^2}{\alpha} u \\ &\leq \mu + \frac{C_n}{\alpha A^2} \mu_{-1}(M \setminus B) \leq \mu + \frac{C_n}{A^2} e^{-\frac{A^2}{20}}, \end{aligned}$$

where we have implemented the following consequence of Jensen's inequality applied to the convex function  $t \mapsto t \log t$  and the probability measure  $u_{-1} dg_{-1}$

$$- \int \log \frac{\eta^2}{\alpha} \cdot \frac{\eta^2}{\alpha} u \leq - \int \frac{\eta^2}{\alpha} u \cdot \log \int \frac{\eta^2}{\alpha} u = 0.$$

□

**Proof of Theorem 1.10.** The proof of Theorem 1.10 is similar to the above proof. We shall again assume  $r = 1$  by parabolic rescaling. We define

$$\begin{aligned} B &= B_{-1} \left( z, A\sqrt{H_n} \right), \quad B' = B_{-1} \left( z, 2A\sqrt{H_n} \right), \\ u_t &= (-4\pi t)^{-\frac{n}{2}} e^{-f_t} =: K(x_0, 0 | \cdot, t) \quad \text{for } t \in [-1, 0), \end{aligned}$$

where  $(z, -1)$  is an  $H_n$ -center of  $(x_0, 0)$ . As before, let  $\eta$  be the cut-off function defined on  $(M, g_{-1})$ , satisfying  $\eta|_B = 1$ ,  $\eta|_{M \setminus B'} = 0$ ,  $0 \leq \eta \leq 1$ , and  $|\nabla \eta|^2 \leq \frac{10}{A^2 H_n} \eta$ , and let

$$\alpha := \int_M u_{-1} \eta^2 dg_{-1}.$$

We shall then use

$$\tilde{u} := \alpha^{-1} u_{-1} \eta^2$$

as a test function to estimate  $\mu(g_{-1}, B', 1)$ .

First of all, [Bam20a, Proposition 3.14] implies that

$$\frac{1}{2} \leq 1 - 2e^{-\frac{A^2}{20}} \leq \nu_{x_0, 0 | -1}(B) \leq \alpha \leq \nu_{x_0, 0 | -1}(M) = 1. \quad (5.2)$$

Then, we may follow the same computation as in (5.1). Since this computation is on the fixed time-slice at  $t = -1$ , we shall omit the subindex  $-1$  and the measure notation  $dg_{-1}$  when there is no ambiguity.

$$\begin{aligned}
\mu(B', g_{-1}, 1) &\leq \int (|\nabla \log \tilde{u}|^2 + R)\tilde{u} - \log \tilde{u} \cdot \tilde{u} \, dg_{-1} - n - \frac{n}{2} \log(4\pi) \\
&= \alpha^{-1} \int \left\{ (|\nabla \log u|^2 - 2\frac{\Delta u}{u} + R) - \log u - n - \frac{n}{2} \log(4\pi) \right\} \eta^2 u \\
&\quad + \alpha^{-1} \int 4|\nabla \eta|^2 u - \int \log \frac{v^2}{\alpha} \cdot \frac{v^2}{\alpha} u \\
&= \alpha^{-1} \int (2\Delta f - |\nabla f|^2 + R + f - n) \eta^2 u + \alpha^{-1} \int 4|\nabla \eta|^2 u - \int \log \frac{v^2}{\alpha} \cdot \frac{v^2}{\alpha} u \\
&= \alpha^{-1} \int (|\nabla f|^2 + R + f - n) \eta^2 u - \alpha^{-1} \int 4\langle \nabla f, \nabla \eta \rangle \eta u \\
&\quad + \alpha^{-1} \int 4|\nabla \eta|^2 u - \int \log \frac{v^2}{\alpha} \cdot \frac{v^2}{\alpha} u.
\end{aligned} \tag{5.3}$$

The last two terms are easily dealt with using the same argument as in the proof of the previous theorem, we shall consider the first two terms.

First of all, we observe that the first term can be split into two terms.

$$\begin{aligned}
&\alpha^{-1} \int (|\nabla f|^2 + R + f - n) \eta^2 u \\
&= \alpha^{-1} \int (|\nabla f|^2 + R) \eta^2 u + \alpha^{-1} \int (f - \frac{n}{2}) \eta^2 u - \frac{n}{2} \\
&=: \text{I} + \text{II} - \frac{n}{2}.
\end{aligned} \tag{5.4}$$

Applying [Bam20a, Proposition 5.13] (see also [CMZ21a, Proposition 3.3] for the proof under our current condition), we may estimate the term I as follows.

$$\begin{aligned}
\text{I} &= \alpha^{-1} \int (|\nabla f|^2 + R) \eta^2 u \\
&= \alpha^{-1} \int (|\nabla f|^2 + R - R_{\min}) \eta^2 u + R_{\min} \\
&\leq \alpha^{-1} \int (|\nabla f|^2 + R - R_{\min}) u + R_{\min} \\
&= \alpha^{-1} \int (|\nabla f|^2 + R) u - (\alpha^{-1} - 1) R_{\min} \\
&\leq \alpha^{-1} \cdot \frac{n}{2} - (\alpha^{-1} - 1) R_{\min}.
\end{aligned} \tag{5.5}$$

We may apply [Bam20a, Proposition 5.13] again as well as the definition of the Nash entropy (2.3) to

estimate the term II as follows. For the sake of notational simplicity, we have defined  $\mathcal{N} := \mathcal{N}_{x_0,0}(1)$ .

$$\begin{aligned}
\Pi &= \alpha^{-1} \int (f - \frac{n}{2} - \mathcal{N}) \eta^2 u + \mathcal{N} \\
&= \alpha^{-1} \int (f - \frac{n}{2} - \mathcal{N}) (\eta^2 - 1) u + \mathcal{N} \\
&\leq \alpha^{-1} \left( \int (f - \frac{n}{2} - \mathcal{N})^2 u \right)^{\frac{1}{2}} \left( \int (\eta^2 - 1)^2 u \right)^{\frac{1}{2}} + \mathcal{N} \\
&\leq \alpha^{-1} (\nu_{-1}(M \setminus B))^{\frac{1}{2}} (n - 2R_{\min})^{\frac{1}{2}} + \mathcal{N} \\
&\leq \sqrt{2} \alpha^{-1} \cdot (n - 2R_{\min})^{\frac{1}{2}} \cdot e^{-\frac{A^2}{40}} + \mathcal{N},
\end{aligned} \tag{5.6}$$

where we have applied (5.2). Combining (5.4), (5.5), and (5.6), we have

$$\begin{aligned}
&\alpha^{-1} \int (|\nabla f|^2 + R + f - n) \eta^2 u \\
&\leq \mathcal{N} + \frac{1}{2} (\alpha^{-1} - 1) \cdot (n - 2R_{\min}) + \sqrt{2} \alpha^{-1} e^{-\frac{A^2}{40}} \cdot (n - 2R_{\min})^{\frac{1}{2}} \\
&\leq \mathcal{N} + 4 \left( e^{-\frac{A^2}{20}} \cdot (n - 2R_{\min}) + e^{-\frac{A^2}{40}} \cdot (n - 2R_{\min})^{\frac{1}{2}} \right),
\end{aligned} \tag{5.7}$$

where we have applied (5.2).

Next, we apply the Cauchy-Schwarz inequality to estimate the second term on the right-hand-side of (5.3).

$$\begin{aligned}
-\alpha^{-1} \int 4 \langle \nabla f, \nabla \eta \rangle \eta u &\leq 4\alpha^{-1} \left( \int |\nabla f|^2 u \right)^{\frac{1}{2}} \left( \int |\nabla \eta|^2 \eta^2 u \right)^{\frac{1}{2}} \\
&\leq 4\alpha^{-1} \left( \int (|\nabla f|^2 + R) u - R_{\min} \right)^{\frac{1}{2}} (\nu_{-1}(M \setminus B))^{\frac{1}{2}} \\
&\leq 4\alpha^{-1} e^{-\frac{A^2}{40}} \cdot (n - 2R_{\min})^{\frac{1}{2}}.
\end{aligned} \tag{5.8}$$

Finally, arguing as in the proof of Theorem 1.7, we have

$$\alpha^{-1} \int 4 |\nabla \eta|^2 u \leq \alpha^{-1} \frac{C_n}{A^2} e^{-\frac{A^2}{20}}, \quad - \int \log \frac{\eta^2}{\alpha} \cdot \frac{\eta^2}{\alpha} u \leq 0. \tag{5.9}$$

Combining (5.3), (5.7), (5.8), and (5.9), the conclusion follows.  $\square$

**Proof of Corollary 1.9.** Let us assume  $T = 1$ . By Theorem 1.7, the proof is reduced to estimating the distance between  $(x_0, 0)$  and an  $H_n$ -center of  $(x_0, 1)$ . Let  $\psi : \mathbb{R} \rightarrow [0, 1]$  be the cut-off function applied by [W18], satisfying  $\psi|_{(-\infty, 1)} \equiv 1$ ,  $\psi|_{(2, \infty)} \equiv 0$ , and

$$\psi'' \geq -10\psi \quad \text{and} \quad (\psi')^2 \leq 10\psi. \tag{5.10}$$

Then, arguing as the proof of [W18, Theorem 5.4] (which is merely an application of [Per02, Lemma 8.3(a)]), we have that the function

$$\psi(x, t) := \psi \left( \frac{\text{dist}_t(x_0, x) + A\sqrt{t}}{A + 1} \right), \quad \text{where} \quad (x, t) \in M \times [0, 1],$$

satisfies the inequality

$$\square\psi \leq \frac{10}{(A+1)^2}\psi.$$

Therefore, we may compute

$$\frac{d}{dt} \int_M \psi(\cdot, t) d\nu_{x_0,1|t} = \int_M \square_t \psi(\cdot, t) d\nu_{x_0,1|t} \leq \frac{10}{(A+1)^2} \int_M \psi(\cdot, t) d\nu_{x_0,1|t},$$

and consequently

$$\nu_{x_0,1|0}(B_0(x_0, 3A)) \geq \int_M \psi(\cdot, 0) d\nu_{x_0,1|0} \geq e^{-\frac{10}{(A+1)^2}} \int_M \psi(\cdot, 1) d\nu_{x_0,1|1} = e^{-\frac{10}{(A+1)^2}}. \quad (5.11)$$

Letting  $(z, 0)$  be an  $H_n$ -center of  $(x_0, 1)$  and  $r := \text{dist}_0(x_0, z)$ . Then, by [Bam20a, Proposition 3.14], whenever  $r > 3A$ , we have

$$\begin{aligned} \nu_{x_0,1|0}(B_0(x_0, 3A)) &\leq \nu_{x_0,1|0}(M \setminus B_0(z, r - 3A)) \\ &\leq 2 \exp\left(-\frac{(r - 3A - \sqrt{2H_n})_+^2}{8}\right). \end{aligned} \quad (5.12)$$

Combining (5.11) and (5.12), we have

$$r \leq 3A + \sqrt{2H_n} + \left(\frac{80}{(A+1)^2} + 8 \log 2\right)^{\frac{1}{2}} \leq 4A,$$

where we have applied the fact that  $A \geq 1000n$ .

Finally, applying Theorem 1.7, we have

$$\begin{aligned} \mu(B_1(x_0, 8A), g_1, \tau) &\geq \mu(B_0(z, 16A), g_0, 1 + \tau) - \frac{C_n}{A^2} e^{-c_n A^2} \\ &\geq \mu(B_0(x_0, 20A), g_0, 1 + \tau) - \frac{C_n}{A^2} e^{-c_n A^2}. \end{aligned}$$

The corollary then follows.  $\square$

## 6 Perelman's noncollapsing improving and pseudolocality theorems

### 6.1 Noncollapsing improving theorem

In this subsection, we prove Theorem 1.12. The proof is almost identical to [J21], but we sketch it for the convenience of the reader. The main idea is that the integral bound on the scalar curvature guarantees that  $x_0$  is not far away from any  $H_n$ -center at time  $-r^2$ . The theorem follows from the Harnack property of the Nash entropy and Theorem 1.1.

**Proof of Theorem 1.12.** As usual, we assume  $r = 1$  by parabolic rescaling. Let  $(z, -1)$  be an  $H_n$ -center of  $(x_0, 0)$ , then Bamler's conjugate heat kernel estimate (c.f. [Bam20a, Theorem 7.2];

see [CMZ21a, Theorem 3.2] for the proof under the bounded curvature assumption) together with Theorem 2.2 implies that

$$\frac{1}{(4\pi)^{\frac{n}{2}}} e^{-\ell_{x_0,0}(x_0,1)} \leq K(x_0,0 | x_0, -1) \leq \frac{C}{(4\pi)^{\frac{n}{2}}} \exp\left(-\mathcal{N}_{x_0,0}(1) - \frac{\text{dist}_{-1}^2(x_0,z)}{C}\right), \quad (6.1)$$

where  $C$  is a numerical constant depending only on the dimension, and  $\ell_{x_0,0}$  is Perelman's reduced distance based at  $(x_0,0)$  which can be estimated by

$$\ell_{x_0,0}(x_0,1) \leq \frac{1}{2\sqrt{1}} \int_0^1 \sqrt{\tau} R_{g_{-\tau}}(x_0) d\tau \leq \frac{A}{2}, \quad (6.2)$$

where we have used the static curve at  $x_0$  as a test curve. Combining (6.1) and (6.2), we have

$$\begin{aligned} \text{dist}_{-1}(x_0,z) &\leq C(n,A) (-\mathcal{N}_{x_0,0}(1) + 1)^{\frac{1}{2}} \\ &\leq C(n,A) (-\mathcal{N}_{x_0,0}(2) + 1)^{\frac{1}{2}}. \end{aligned} \quad (6.3)$$

On the other hand, since  $(z,-1)$  is an  $H_n$ -center of  $(x_0,0)$ , by (2.5), we have

$$\begin{aligned} \text{dist}_{W_1}^{g^{-1}}(\nu_{x_0,0|-1}, \delta_{x_0}) &\leq \text{dist}_{W_1}^{g^{-1}}(\nu_{x_0,0|-1}, \delta_z) + \text{dist}_{W_1}^{g^{-1}}(\delta_{x_0}, \delta_z) \\ &\leq \sqrt{H_n} + \text{dist}_{-1}(x_0,z) \\ &\leq C(n) + C(n,A) (-\mathcal{N}_{x_0,0}(2) + 1)^{\frac{1}{2}}, \end{aligned} \quad (6.4)$$

where we have also applied (6.3). Applying Theorem 2.1 with  $(x_0, -1) \rightarrow (x_1, t_1)$ ,  $(x_0, 0) \rightarrow (x_2, t_2)$ ,  $-1 \rightarrow t^*$ ,  $-2 \rightarrow s$ , and taking (6.4) and the fact that  $R_{g_{-1}} \geq -\frac{n}{2}$  (obtained in the same ways as (3.7)) in to account, we have

$$\begin{aligned} \mathcal{N}_{x_0,-1}(1) &\leq \mathcal{N}_{x_0,0}(2) + \left(\frac{n}{2(-1-(-2))} - (-\frac{n}{2})\right)^{\frac{1}{2}} \text{dist}_{W_1}^{g^{-1}}(\nu_{x_0,0|-1}, \delta_{x_0}) + \frac{n}{2} \log\left(\frac{0-(-2)}{-1-(-2)}\right) \\ &\leq \mathcal{N}_{x_0,0}(2) + C(n) + C(n,A) (-\mathcal{N}_{x_0,0}(2) + 1)^{\frac{1}{2}} \\ &\leq \mathcal{N}_{x_0,0}(2) + C(n) + \frac{1}{2} (-\mathcal{N}_{x_0,0}(2) + 1) + C(n,A) \\ &\leq \frac{1}{2} \mathcal{N}_{x_0,0}(2) + C(n,A). \end{aligned}$$

Hence, applying [Bam20a, Theorem 8.1] in the same way as in (3.14), we have

$$\mathcal{N}_{x_0,0}(2) \geq 2\mathcal{N}_{x_0,-1}(1) - C(n,A) \geq -C(n,A),$$

and the rest of the proof is but a straightforward application of Theorem 1.1.  $\square$

## 6.2 Pseudolocality theorem

In this subsection, we present the proof of Theorem 1.13 using our local monotonicity theorem. The proof goes by contradiction. We show that under the assumption of Theorem 1.13, a point at which (1.6) fails has very small Nash entropy, and this is a contradiction. To this end, we need the following preparatory result due to Bamler.



**Lemma 6.1** ([Bam20a, Theorem 10.3]). *For any  $\varepsilon > 0$  there is  $\delta = \delta(n, \varepsilon) > 0$ , such that the following holds. Let  $(x, t) \in M \times I$  and  $r > 0$  satisfy  $[t - r^2, t] \subset I$  and  $\mathcal{N}_{x,t}(r^2) \geq -\delta$ . Then we have*

$$|\text{Rm}| \leq \varepsilon r^{-2} \quad \text{on} \quad B_t(x, \varepsilon^{-1}r) \times ([t - (1 - \varepsilon)r^2, \varepsilon^{-1}r^2] \cap I), \quad (6.5)$$

$$\inf_{\rho \in (0, \varepsilon^{-1}r)} \rho^{-n} |B_t(x, \rho)|_t \geq (1 - \varepsilon)\omega_n, \quad (6.6)$$

where  $\omega_n$  is the volume of the  $n$ -dimensional unit disk in the Euclidean space.

*Proof.* (6.5) is the same as the statement of [Bam20a, Theorem 10.3], and we show that (6.6) is a straightforward consequence of (6.5).

Without loss of generality, we assume  $t = 0$  and  $r = 1$ . Let us consider a sequence of counter examples  $(M_i^n, g_{i,t}, x_i)_{t \in [-1, 0]}$  with  $\mathcal{N}_{x_i, 0(1)} = -\delta_i \nearrow 0$ , but

$$\inf_{\rho \in (0, \varepsilon^{-1})} \rho^{-n} |B_{g_{i,0}}(x_i, \rho)|_0 \leq (1 - \varepsilon)\omega_n \quad (6.7)$$

for some fixed  $\varepsilon > 0$ .

By (6.5), we can find a sequence  $\varepsilon_i \searrow 0$ , such that

$$|\text{Rm}_{g_i}| \leq \varepsilon_i \quad \text{on} \quad B_{g_{i,0}}(x_i, \varepsilon_i^{-1}) \times [-1 + \varepsilon_i, 0]$$

and hence we may extract a subsequence from  $(M_i, g_{i,t}, x_i)$  and obtain a limit Ricci flow. This limit must be the Euclidean space, and this contradicts (6.7).  $\square$

Since, to apply Theorem 1.10, we need a global scalar curvature lower bound on the initial slice, which is not available in the assumption of Theorem 1.13, we combine Theorem 1.10 and Theorem 1.7 to obtain the following result for the convenience of application.

**Proposition 6.2.** *Assume that  $[-r^2, 0] \subseteq I$ . Then, for any  $x_0 \in M$ , any  $H_n$ -center  $(z, -r^2)$  of  $(x_0, 0)$ , and any  $A \geq \underline{A}(n)$ , we have*

$$\mu(B_{-r^2}(z, Ar), g_{-r^2, r^2}) \leq \mathcal{N}_{x_0, 0}(\frac{1}{2}r^2) + C_n e^{-c_n A^2},$$

where  $\underline{A}(n)$ ,  $c_n$ , and  $C_n$  are dimensional constants.

To prove the above proposition we need the following observation, namely, that an  $H_n$ -center of an  $H_n$ -center is almost an  $H_n$ -center; this result is interesting in its own right, since the conclusion is independent of  $t_2 \in (t_1, t_3)$ .

**Lemma 6.3.** *Let  $x_1, x_2, x_3, z \in M$  and  $t_1, t_2, t_3 \in I$  such that  $t_1 < t_2 < t_3$ . Furthermore, assume that  $(x_1, t_1)$  is an  $H_n$ -center of  $(x_2, t_2)$ , that  $(x_2, t_2)$  is an  $H_n$ -center of  $(x_3, t_3)$ , and that  $(z, t_1)$  is an  $H_n$ -center of  $(x_3, t_3)$ . Then we have*

$$\text{dist}_{t_1}(x_1, z) \leq C_n \sqrt{t_3 - t_1},$$

where  $C_n$  is a dimensional constant.

*Proof.* Letting  $A > \sqrt{2H_n} + 4\sqrt{2\pi}$  and  $\Lambda > 0$ , we may argue as the proof of (4.3) and obtain

$$\nu_{x,t_2|t_1} \left( B_{t_1} \left( x_1, A\sqrt{t_2 - t_1} \right) \right) \geq \Phi \left( \frac{1}{\sqrt{2}}A - \Lambda - \sqrt{H_n} - 4\sqrt{\pi} \right)$$

for all  $x \in B_{t_2} \left( x_2, \Lambda\sqrt{t_2 - t_1} \right)$ , where  $\Phi$  is the function in the statement of Theorem 1.6.

Taking  $\Lambda := \sqrt{2H_n \frac{t_3 - t_2}{t_2 - t_1}}$ , and applying the fact (c.f. [Bam20a, Proposition 3.13]) that

$$\nu_{x_3, t_3 | t_2} \left( B_{t_2} \left( x_2, \sqrt{2H_n(t_3 - t_2)} \right) \right) \geq \frac{1}{2},$$

we have

$$\begin{aligned} \nu_{x_3, t_3 | t_1} \left( B_{t_1} \left( x_1, A\sqrt{t_2 - t_1} \right) \right) &\geq \int_{B_{t_2} \left( x_2, \sqrt{2H_n(t_3 - t_2)} \right)} \nu_{\cdot, t_2 | t_1} \left( B_{t_1} \left( x_1, A\sqrt{t_2 - t_1} \right) \right) d\nu_{x_3, t_3 | t_2} \\ &\geq \frac{1}{2} \Phi \left( \frac{1}{\sqrt{2}}A - \sqrt{2H_n \frac{t_3 - t_2}{t_2 - t_1}} - \sqrt{H_n} - 4\sqrt{\pi} \right). \end{aligned}$$

Finally, taking  $A := \sqrt{4H_n \frac{t_3 - t_1}{t_2 - t_1}}$ , we have

$$\begin{aligned} \nu_{x_3, t_3 | t_1} \left( B_{t_1} \left( x_1, \sqrt{4H_n(t_3 - t_1)} \right) \right) &\geq \frac{1}{2} \Phi \left( \sqrt{2H_n \frac{t_3 - t_1}{t_2 - t_1}} - \sqrt{2H_n \frac{t_3 - t_2}{t_2 - t_1}} - \sqrt{H_n} - 4\sqrt{\pi} \right) \\ &= \Phi \left( \frac{2H_n}{\sqrt{2H_n \frac{t_3 - t_1}{t_2 - t_1}} + \sqrt{2H_n \frac{t_3 - t_2}{t_2 - t_1}}} - \sqrt{H_n} - 4\sqrt{\pi} \right) \\ &\geq \Phi \left( -\sqrt{H_n} - 4\sqrt{\pi} \right) =: c_n. \end{aligned}$$

It then follows from [Bam20a, Proposition 3.13] that

$$\text{dist}_{t_1}(x_1, z) \leq (c_n^{-\frac{1}{2}} + 2)\sqrt{H_n(t_3 - t_1)}.$$

□

*Proof of Proposition 6.2.* By parabolic scaling, let us assume  $r = 1$ . Let  $(y, -1/2)$  be an  $H_n$ -center of  $(x_0, 0)$ , and  $(z', -1)$  be an  $H_n$ -center of  $(y, -1/2)$ . Then, by Lemma 6.3, we have

$$\text{dist}_{-1}(z', z) \leq C_n, \tag{6.8}$$

where  $(z, -1)$  is an  $H_n$ -center of  $(x_0, 0)$ .

By the maximum principle, we have  $R_{g_{-1/2}} \geq -n$ . Hence, we may apply Theorem 1.10 to obtain

$$\mathcal{N}_{x_0, 0}(1/2) \geq \mu \left( B_{-1/2} \left( y, \frac{1}{4}A \right), g_{-1/2}, 1/2 \right) - C_n e^{-c_n A^2}, \tag{6.9}$$

if  $A \geq \underline{A}(n)$ , where  $c_n$  and  $C_n$  are both dimensional constants. Applying Theorem 1.7 and (6.8), we have

$$\begin{aligned} \mu \left( B_{-1/2} \left( y, \frac{1}{4}A \right), g_{-1/2}, 1/2 \right) &\geq \mu \left( B_{-1} \left( z', \frac{1}{2}A \right), g_{-1}, 1 \right) - C_n e^{-c_n A^2} \\ &\geq \mu \left( B_{-1} \left( z, A \right), g_{-1}, 1 \right) - C_n e^{-c_n A^2}, \end{aligned} \tag{6.10}$$

if  $A \geq \underline{A}(n)$ . The proposition then follows from (6.8) and (6.10). □

**Proof of Theorem 1.13.** Let us first prove (1.6) by contradiction, since (1.7) and (1.8) follow from it. Fixing an  $\alpha \in (0, \frac{1}{100n})$ , we shall show that under the assumption of (1.5), if (1.6) fails, then there is a contradiction arising when  $\delta$  is small enough.

Let  $\bar{t} \in (0, T]$  be the first time such that the following statement fails

$$|\text{Rm}_{g_t}|(x) < \alpha t^{-1} \quad \text{for all } t \in (0, \bar{t}] \text{ and } x \in \bar{B}_t(x_0, \alpha^{-1}\sqrt{t}).$$

Since the Ricci flow is smooth, we must have  $\bar{t} > 0$ . By the contradictory assumption, we also have  $\bar{t} < T$ . Henceforth we shall assume  $\bar{t} = 1$  by a parabolic scaling. Then, by the definition of  $\bar{t}$ , we have

$$|\text{Rm}_{g_t}|(x) \leq \frac{\alpha}{t} \quad \text{for all } t \in (0, 1] \text{ and } x \in B_t(x_0, \alpha^{-1}\sqrt{t}),$$

and there exists  $y \in \bar{B}_1(x_0, \alpha^{-1})$ , such that

$$|\text{Rm}_{g_1}|(y) = \alpha. \tag{6.11}$$

Let  $z$  be an  $H_n$ -center of  $y$ . Then, we may apply the same argument as in the proof of Corollary 1.9 by using the cutoff function of the form

$$\psi(x, t) := \psi\left(\frac{\text{dist}_t(x_0, x) + 2\alpha^{-1}\sqrt{t}}{10\alpha^{-1}}\right),$$

where  $\psi$  is the function defined in (5.10). This leads to

$$\text{dist}_0(z, x_0) \leq C(n, \alpha).$$

Now, if  $\delta$  is taken to be small enough, such that  $\delta^{-1} \gg C(n, \alpha)$ , then we may apply Proposition 6.2 to obtain

$$\begin{aligned} \mathcal{N}_{y,1}(1/2) &\geq \mu(B_0(z, \frac{1}{2}\delta^{-1}), g_0, 1) - C_n e^{-c_n \delta^{-2}} \\ &\geq \mu(B_0(x_0, \delta^{-1}), g_0, 1) - C_n e^{-c_n \delta^{-2}} \\ &\geq -\delta^2 - C_n e^{-c_n \delta^{-2}}, \end{aligned} \tag{6.12}$$

where we have applied the assumption (1.5). Obviously, by Lemma 6.1, (6.12) contradicts (6.11) when  $\delta$  is small enough; this proves (1.6).

Next, we prove (1.7) assuming (1.6), and (1.8) is a consequence of (1.6) and (1.7) as a well-known result proved by Cheeger-Gromov-Taylor [CGT82]. In fact, fixing any  $(y, t) \in M \times (0, T]$  such that  $\text{dist}_t(x_0, y) \leq \alpha^{-1}\sqrt{t}$ . Let  $(z, 0)$  be an  $H_n$ -center of  $(y, t)$ . Given the curvature bound (1.6), we may apply the same argument as above to obtain

$$\text{dist}_0(x_0, z) \leq C(n, \alpha)\sqrt{t}.$$

If  $\delta$  is taken to be small enough such that  $\delta^{-1} \gg C(n, \alpha)$ , then we argue in the same way as (6.12) to obtain

$$\begin{aligned} \mathcal{N}_{y,t}(\frac{1}{2}t) &\geq \mu(B_0(z, \frac{1}{2}\delta^{-1}\sqrt{t}), g_0, t) - C_n e^{-c_n \delta^{-2}} \\ &\geq \mu(B_0(x_0, \delta^{-1}\sqrt{t}), g_0, t) - C_n e^{-c_n \delta^{-2}} \\ &\geq -\delta^2 - C_n e^{-c_n \delta^{-2}}, \end{aligned}$$

where we have also applied (1.5). (1.7) then follows from Lemma 6.1.  $\square$

## 7 Application to ancient Ricci flows

In this section, we provide a short proof of Theorem 1.14 using Theorem 1.1.

**Proof of Theorem 1.14.** Let us first of all show that

$$\inf_{t \leq 0} \nu(g_t) \geq \mu_\infty. \quad (7.1)$$

Fixing an arbitrary  $t \in (-\infty, 0]$  and  $x \in M$ , we have, by [MZ21, Proposition 4.6]

$$\lim_{\tau \rightarrow \infty} \mathcal{N}_{x,t}(\tau) = \lim_{\tau \rightarrow \infty} \mathcal{N}_{x_0,t_0}(\tau) = \mu_\infty.$$

Hence, by the monotonicity of the Nash entropy, we have

$$\mathcal{N}_{x,t}(r^2) \geq \mu_\infty \quad \text{for all } r > 0.$$

Applying Theorem 1.1 at  $(x, t)$  with  $A = \tau = \epsilon > 0$ , we have

$$\begin{aligned} \nu(B_t(x, \epsilon r), g_t, \epsilon r^2) &\geq \mathcal{N}_{x,t}(r^2) - \sqrt{n}\epsilon - \frac{n}{2}\epsilon - \frac{n}{2}\log(1 + \epsilon) \\ &\geq \mu_\infty - \sqrt{n}\epsilon - \frac{n}{2}\epsilon - \frac{n}{2}\log(1 + \epsilon) \quad \text{for all } r > 0. \end{aligned}$$

By first taking  $r \rightarrow \infty$  and then taking  $\epsilon \rightarrow 0$ , we have

$$\nu(g_t) \geq \mu_\infty.$$

Since  $t \in (-\infty, 0]$  is arbitrary, we have proved (7.1).

Next, we prove

$$\inf_{t \leq 0} \nu(g_t) \leq \mu_\infty. \quad (7.2)$$

By the assumption of the theorem, for any  $\epsilon > 0$ , we can find  $\tau < \infty$ , such that

$$\mathcal{N}_{x_0,t_0}(\tau) \leq \mu_\infty + \epsilon.$$

Consequently we have

$$\mu_\infty + \epsilon \geq \mathcal{N}_{x_0,t_0}(\tau) \geq \mathcal{W}_{x_0,t_0}(\tau) \geq \nu(g_{t_0-\tau}) \geq \inf_{t \leq 0} \nu(g_t).$$

Since  $\epsilon$  is arbitrary, (7.2) follows immediately.  $\square$

## References

- [Bam20a] Richard H. Bamler, *Entropy and heat kernel bounds on a Ricci flow background*, <https://arxiv.org/abs/2008.07093> (2020).
- [Bam20b] ———, *Compactness theory of the space of super Ricci flows*, <https://arxiv.org/abs/2008.09298> (2020).

- [Bam20c] ———, *Structure theory of non-collapsed limits of Ricci flows*, <https://arxiv.org/abs/2009.03243> (2020).
- [Bam21] ———, *On the fundamental group of non-collapsed ancient Ricci flows*, <https://arxiv.org/abs/2110.02254> (2021).
- [BW08] Böhm, Christoph, and Burkhard Wilking. *Manifolds with positive curvature operators are space forms*. *Annals of mathematics* (2008): 1079-1097.
- [BS09] Brendle, Simon, and Richard Schoen. *Manifolds with  $1/4$ -pinched curvature are space forms*. *Journal of the American Mathematical Society* 22.1 (2009): 287-307.
- [Ca08] Cao, Xiaodong. *Differential Harnack estimates for backward heat equations with potentials under the Ricci flow*. *Journal of Functional Analysis* 255.4 (2008): 1024-1038.
- [CMZ21a] Chan, Pak-Yeung, Zilu Ma, and Yongjia Zhang. *Ancient Ricci flows with asymptotic solitons*. arXiv preprint arXiv:2106.06904 (2021).
- [CMZ21b] Chan, Pak-Yeung, Zilu Ma, and Yongjia Zhang. *A uniform Sobolev inequality for ancient Ricci flows with bounded Nash entropy*. arXiv preprint arXiv:2107.01419 (2021).
- [CGT82] Cheeger, Jeff, Mikhail Gromov, and Michael Taylor. *Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds*. *Journal of Differential Geometry* 17.1 (1982): 15-53.
- [CCGGIIKLLN10] Chow, B.; Chu, S.; Glickenstein, D.; Guenther, C.; Isenberg, J.; Ivey, T.; Knopf, D.; Lu, P.; Luo, F.; Ni, L. *The Ricci flow: techniques and applications. Part III. Geometric-Analytic Aspects*, *Mathematical Surveys and Monographs*, vol. **163**, AMS, Providence, RI, 2010.
- [E01] Ecker, Klaus. *A local monotonicity formula for mean curvature flow*. *Annals of mathematics* 154.2 (2001): 503-525.
- [EKNT08] Ecker, Klaus; Knopf, Dan; Ni, Lei; Topping, Peter, *Local monotonicity and mean value formulas for evolving Riemannian manifolds*. *J. Reine Angew. Math.* 616 (2008), 89–130.
- [J21] Jian, Wangjian. *On the improved no-local-collapsing theorem of Ricci flow*, arXiv:2107.12515.
- [Lai20] Lai, Yi. *A family of 3d steady gradient solitons that are flying wings*. arXiv preprint arXiv:2010.07272 (2020).
- [LW20] Li, Yu; Wang, Bing. *Heat kernel on Ricci shrinkers*. *Calc. Var. Partial Differential Equations* 59 (2020), no. 6, Paper No. 194, 84 pp.
- [Lu10] Lu, Peng. *A local curvature bound in Ricci flow*. *Geometry & Topology* 14.2 (2010): 1095-1110.
- [Ham82] Hamilton, Richard S. *Three-manifolds with positive Ricci curvature*. *Journal of Differential geometry* 17.2 (1982): 255-306.

- [Ham86] Hamilton, Richard S. *Four-manifolds with positive curvature operator*. Journal of Differential Geometry 24.2 (1986): 153-179.
- [Ham93a] Richard S. Hamilton, *The Harnack estimate for the Ricci flow*, J. Differential Geom., 1993, 37(1): 225-243.
- [Ham93b] Richard S. Hamilton, *The formation of singularities in the Ricci flow*. Surveys in differential geometry, Vol. II (Cambridge, MA, 1993), 7–136, Internat. Press, Cambridge, MA, 1995.
- [LY86] Li, Peter, and Shing Tung Yau. *On the parabolic kernel of the Schrödinger operator*. Acta Mathematica 156 (1986): 153-201.
- [MZ21] Ma, Zilu; Zhang, Yongjia, *Perelman’s entropy on ancient Ricci flows*. J. Funct. Anal. 281 (2021), no. 9, Paper No. 109195, 31 pp
- [Per02] Perelman, Grisha, *The entropy formula for the Ricci flow and its geometric applications*, arXiv:math.DG/0211159.
- [Per03] Perelman, Grisha, *Ricci flow with surgery on three-manifolds*, arXiv:math.DG/0303109.
- [Sh89] Shi, Wan-Xiong. *Deforming the metric on complete Riemannian manifolds*. J. Diff. Geom. 1989, 30(1):223-301.
- [TZ21] Tian, Gang, and Zhenlei Zhang. *Relative volume comparison of Ricci flow*. Sci. China Math. 64, 1937–1950 (2021). <https://doi.org/10.1007/s11425-021-1869-5>
- [W18] Wang, Bing, *The local entropy along Ricci flow Part A: the no-local-collapsing theorems*. Camb. J. Math. 6 (2018), no. 3, 267–346.
- [W20] Wang, Bing. *The local entropy along Ricci flow—Part B: the pseudo-locality theorems*. arXiv preprint arXiv:2010.09981 (2020).
- [Ye15] Ye, Rugang. *The logarithmic Sobolev and Sobolev inequalities along the Ricci flow*. Commun. Math. Stat. 3 (2015), no. 1, 1–36.
- [Zhq07] Zhang, Qi S. *A uniform Sobolev inequality under Ricci flow*. International Mathematics Research Notices 2007 (2007).
- [Zhq10] Zhang, Qi S. *Sobolev inequalities, heat kernels under Ricci flow, and the Poincaré conjecture*. CRC Press, 2010.

Department of Mathematics, University of California, San Diego, CA 92093, USA  
E-mail address: pachan@ucsd.edu

Department of Mathematics, University of California, San Diego, CA 92093, USA  
E-mail address: zim022@ucsd.edu

School of Mathematics, University of Minnesota, Twin Cities, MN 55414, USA  
E-mail address: zhan7298@umn.edu