Perelman’s entropy on ancient Ricci flows

Zilu Ma and Yongjia Zhang

July 9, 2021

Abstract

In [ZY2], the second author proved Perelman’s assertion, namely, for an ancient Ricci flow with bounded and nonnegative curvature operator, bounded entropy is equivalent to noncollapsing on all scales. In this paper, we continue this discussion. It turns out that the curvature operator nonnegativity is not a necessary condition, and we need only to assume a consequence of Hamilton’s trace Harnack. Furthermore, we show that this condition holds for steady Ricci solitons with nonnegative Ricci curvature.

1 Introduction

The entropy formula for the Ricci flow was introduced by Perelman [Per02], with the help of which he proved the no local collapsing theorem and some other famous theorems, such as the pseudolocality theorem. Along with the entropy formula, Perelman also invented the reduced geometry for the Ricci flow. Both of these two monotonicity formulas have ever since been central techniques in this field. At the first glance, it appears that Perelman’s entropy is neatly formulated, whereas the reduced geometry deals with a subsolution—rather than a solution—to the conjugate heat equation. Nevertheless, the analysis of the reduced geometry needs only the local geometric information, whereas the analysis of Perelman’s entropy requires one to handle the heat equation, whose solution is sensitive to the global geometry. In view of these facts, it is not surprising that the reduced geometry is more tractable in the localization. Indeed, in the construction of the Ricci flow with surgeries, Perelman chiefly applied the reduced geometry.

On the other hand, it turns out that many theorems proved by the reduced geometry method can also be proved by Perelman’s entropy method. To begin with, Zhang [ZQ10] showed that in the proof of the Poincaré conjecture, the reduced geometry can be replaced with Perelman’s entropy. Furthermore, under the condition of either Type I curvature bound or bounded and nonnegative curvature operator, a noncollapsed Ricci flow always has an asymptotic shrinking soliton (c.f. [Per02] and [N10]). The original proofs of these asymptotic soliton theorems applied the reduced geometry method, yet they can both be proved by implementing Perelman’s entropy method (c.f. [CZ11] and [ZY2]). Hereby we would like to point out Bamler’s recent groundbreaking works [Bam20a]—[Bam20c], in which he largely refined the analysis of the Nash entropy (the time average of Perelman’s entropy) and proved a nice structure theorem for singularity models of the Ricci flow.

In this paper, we continue the discussion initiated in [ZY2], where the second author proved Perelman’s assertion (section 11 in [Per02]).
We impose one more requirement on the solutions; namely, we fix some \( \kappa > 0 \) and require that \( g_{ij}(t) \) to be \( \kappa \)-noncollapsed on all scales. It is not hard to show that this requirement is equivalent to a uniform bound on the entropy \( S \), defined as in 5.1 using an arbitrary fundamental solution to the conjugate heat equation.

In this assertion Perelman assumes bounded and nonnegative curvature operator. It is still interesting to ask whether the nonnegative curvature operator condition is necessary. In the present paper, we will show that this condition can indeed be relaxed.

To present the statements of our main results, let us recall the several definitions. Given a complete Ricci flow \((M, g(t))_{t \in I}\), for any \( x, y \in M, s, t \in I, s < t \), we denote by \( K(x, t | y, s) \) the minimal heat kernel coupled with the flow, namely,

\[
(\partial_t - \Delta_{g(t)})K(\cdot, t | y, s) = 0, \quad \lim_{t \downarrow s} K(x, t | y, s) = \delta_y(x),
\]

\[
(-\partial_s - \Delta_{g(s)} + R(\cdot, s)K(x, t | \cdot, s) = 0, \quad \lim_{s \uparrow t} K(x, t | y, s) = \delta_x(y).
\]

See, for example, [CCGG+10, Theorem 24.40] for the existence of such heat kernels.

Following [Per02], we denote by \( \Box := \partial_t - \Delta_{g(t)} \) the heat operator coupled with the Ricci flow \((M, g(t))_{t \in I}\), and by \( \Box^* = -\partial_t - \Delta_{g(t)} + R_{g(t)} \) the conjugate heat operator. It follows from the Stokes theorem that for any \( u, v \in C^2_c(M \times I) \), i.e., \( C^2 \) functions over some interval \( I \subset [0, T] \) with compact supports, we have

\[
\frac{d}{dt} \int_M uv \, dg_t = \int_M (\Box u)v - u(\Box^* v) \, dg_t,
\]

where we denote by \( dg_t \) the volume form induced by the metric \( g(t) \).

**Definition 1.1.** Let \((M, g(t))_{t \in [0, T]}\) be a complete Ricci flow. Let

\[
u(x, t) = K(x_0, t_0 | x, t) = (4\pi\tau)^{-\frac{n}{2}} e^{-f(x, t)}
\]

be the conjugate heat kernel based at \((x_0, t_0) \in M \times (0, T]\), where \( \tau = t_0 - t \in (0, t_0] \) is the backward time. Then Perelman’s entropy and the Nash entropy based at \((x_0, t_0)\) are respectively defined as

\[
W_{(x_0, t_0)}(\tau) = \int_M \left( \tau(|\nabla f|^2 + R) + f - \frac{n}{2} \right) u \, dg_t,
\]

\[
N_{(x_0, t_0)}(\tau) = \int_M fu \, dg_t - \frac{n}{2},
\]

for all \( \tau \in (0, t_0] \).

Perelman’s well-known monotonicity formula indicates that, unless \((M, g(t))\) is the static Euclidean space, in which case \( W_{(x_0, t_0)}(\tau) \equiv 0 \), \( W_{(x_0, t_0)}(\tau) \) is always negative and monotonically decreasing in \( \tau \). Furthermore, Perelman’s entropy always converges to zero at its base time, namely,

\[
\lim_{\tau \to 0} W_{(x_0, t_0)}(\tau) = 0.
\]
The Nash entropy bears the same monotonicity property and satisfies \( (1.5) \), since it is known to be the time average of Perelman’s entropy

\[
\mathcal{N}(x_0, t_0)(\tau) = \frac{1}{\tau} \int_0^\tau \mathcal{W}(x_0, t_0)(\eta) \, d\eta.
\]  

(1.5)

In Perelman’s assertion, he mentioned the notions of \textit{bounded entropy} and \textit{noncollapsing}. They precisely mean the following.

\textbf{Definition 1.6.} Let \((M, g(t))_{t \in (-\infty, 0]}\) be an ancient solution. Then \(g(t)\) is said to have \textit{bounded entropy}, if

\[
W := \inf_{(x, t) \in M \times (-\infty, 0]; \eta > 0} \mathcal{W}(x, t)(\eta) > -\infty,
\]

where the quantity \(W\) is called the \textit{entropy bound}.

\textbf{Definition 1.8.} Let \((M^n, g(t))\) be a complete Ricci flow. Then, \(g(t)\) is called \textit{weakly \(\kappa\)-noncollapsed}, if, for any \(r > 0\), it holds that \(\text{Vol}_{g(t)}(B_{g(t)}(x, r)) \geq \kappa r^n\) whenever \(|\text{Rm}(y, s)| \leq r^{-2}\) for any \((y, s) \in B_{g(t)}(x, r) \times [t - r^2, t]\). \(g(t)\) is called \textit{strongly \(\kappa\)-noncollapsed}, if, for any \(r > 0\), it holds that \(\text{Vol}_{g(t)}(B_{g(t)}(x, r)) \geq \kappa r^n\) whenever \(\text{R}_{g(t)} \leq r^{-2}\) on \(B_{g(t)}(x, r)\). Here \(B_{g(t)}(x, r)\) stands for the \(g(t)\)-geodesic ball centered at \(x\) with radius \(r\), and \(\text{Vol}_{g(t)}\) stands for the Riemannian volume with respect to \(g(t)\).

Next, we review several definitions involved with Perelman’s reduced geometry. Suppose that \((M, g(t))_{t \in [-T, 0]}\) is a complete Ricci flow. Fix a base point \((p_0, t_0) \in M \times (-T, 0]\) For any piecewisely smooth curves \(\gamma : [0, \tau] \to M\) with \(\gamma(0) = p\) and \(t_0 - \tau \geq -T\), we define

\[
\mathcal{L}(\gamma) := \int_0^\tau \sqrt{s} \left( R_{g(t_0-s)} + |\gamma'|_{g(t_0-s)}^2 \right) (\gamma(s)) \, ds.
\]

Then, let

\[
L(x, \tau) := L_{(p_0, t_0)}(x, \tau) := \inf_{\gamma} \mathcal{L}(\gamma),
\]

where the infimum is taken over all piecewisely smooth curve \(\gamma : [0, \tau] \to M\) with \(\gamma(0) = p\) and \(\gamma(\tau) = x\), and

\[
\ell(x, \tau) := \ell_{(p_0, t_0)}(x, \tau) := \frac{1}{2\sqrt{\tau}} L_{(p_0, t_0)}(x, \tau)
\]

is called the \textit{reduced distance} based at \((p_0, t_0)\). Perelman’s \textit{reduced volume} based at \((p_0, t_0)\) is defined as

\[
V_{(p_0, t_0)}(\tau) := (4\pi\tau)^{-n/2} \int_M \exp \left( -\ell_{(p_0, t_0)}(\cdot, \tau) \right) \, dg_{t_0 - \tau},
\]

for all \(\tau \in (0, T + t_0]\). \(V_{(p_0, t_0)}(\tau)\) is known to be monotonically decreasing in \(\tau\).

The main theorem of this paper is the following.

\textbf{Theorem 1.9.} Let \((M^n, g(t))_{t \in (-\infty, 0]}\) be an ancient solution to the Ricci flow with bounded curvature within each compact time interval. Let \(\ell\) be the reduced distance based at \((p, 0)\), where \(p \in M\) is a fixed point. Assume that there exists \(C > 0\), such that the following hold.

1. \(|\text{Rm}_{g(t)}| \leq CR_{g(t)}\) for all \(t \in (-\infty, 0]\).
(2) $|\nabla \ell|^2 + R \leq \frac{C\ell}{\tau}$ for all $\tau \in (0, \infty)$, where $\tau = -t$ is the backward time.

Then $g(t)$ is $\kappa$-noncollapsed if and only if it has bounded entropy; the quantity $\kappa$ and the entropy bound are mutually dependent on.

Remarks:

1. Because of the assumptions above, the notions of strong noncollapsing and weak noncollapsing are equivalent.

2. Assumption (2) is implied by Hamilton’s trace Harnack estimate [Ha1]:

$$\frac{\partial R}{\partial t} - 2\langle X, \nabla R \rangle + 2 \text{Ric}(X, X) \geq 0$$

for any smooth vector field $X$ on $M$. See the argument in Section 7.2 of [Per02].

In [ZY2], the second author used the nonnegativity of the curvature operator for two reasons.

Let $(M, g(\tau)), \tau \in [0, \infty)$ be an ancient Ricci flow with bounded and nonnegative curvature operator, where $\tau$ is the backward time, then

(1) Hamilton’s trace Harnack implies assumption (2) in the statement of Theorem 1.9. This inequality implies that the ancient solution is locally Type I wherever the reduced distance is bounded. This fact, along with the noncollapsing assumption, implies the existence of an asymptotic shrinker.

(2) If, in addition to the bounded and nonnegative curvature operator condition, the ancient solution is $\kappa$-noncollapsed, then by Perelman’s bounded curvature at bounded distance theorem (c.f. Section 11 of [Per02]), the curvature scale $R(x) := R(x)^{-\frac{1}{2}}$ is comparable with the curvature radius $r_B(x) := \sup\{s : |Rm| \leq s^{-2} \text{ on } B(x, s)\}$, and hence the geometry of the parabolic cube $B_{g(\tau)}(x, r) \times [\tau - r^2, \tau + r^2]$, where $r := R(x, \tau)^{-\frac{1}{2}}$, is bounded in terms of $r$. This fact is crucial to the point-wise estimate of the conjugate heat kernel.

As we will find out later in this paper, among the above two points, the latter is not as essential as the former. Indeed, the main technique of proving Theorem 1.9 is to replace the curvature scale $R^{-\frac{1}{2}}$ by $r(x) = \sqrt{\tau} -\frac{1}{2}$. The latter scale, though not directly related to the curvature, serves well for the purpose mentioned in point (2) above. Once this is established, we can use a similar argument as in [ZY2] to estimate the conjugate heat kernel. In fact, this scale is also implemented in the proof of the quadratic lower bound of $\ell$; see Lemma 3.2 in [Y].

The proof of Theorem 1.9 together with some of Bamler’s [Bam20a] results on the Nash entropy also implies the following interesting corollary, which is an extension of a result of Xu [Xu17].

**Corollary 1.11** (Entropy uniqueness of the asymptotic shrinker). Let $(M, g(t))_{t \in (-\infty, 0]}$ be a $\kappa$-noncollapsed ancient solution with bounded curvature within each compact time interval. Assume either

(1) $g(t)$ has Type I curvature bound, i.e., there is a constant $C_1 < \infty$ such that

$$\sup_M |Rm|_{g(t)} \leq \frac{C_1}{1 + |t|},$$

for all $t \leq 0$, or
(2) $|Rm|_{g(t)} \leq CRg(t)$ for some positive constant $C$ and for all $t \in (-\infty, 0]$, and Hamilton’s trace Harnack (1.10) holds on $(M, g(t))$.

Then, all asymptotic shrinkers of $(M, g(t))$ based at any point in $M \times (-\infty, 0]$ have the same entropy, which is also equal to the logarithmic of the Gaussian density as defined in [CHI07]. Furthermore, for any $(x_1, t_1)$ and $(x_2, t_2) \in M \times (-\infty, 0]$, the following holds:

$$
\lim_{\tau \to \infty} W_{(x_1, t_1)}(\tau) = \lim_{\tau \to \infty} W_{(x_2, t_2)}(\tau) = \lim_{\tau \to \infty} N_{(x_1, t_1)}(\tau) = \lim_{\tau \to \infty} N_{(x_2, t_2)}(\tau) = \lim_{\tau \to \infty} \log V_{(x_1, t_1)}(\tau) = \lim_{\tau \to \infty} \log V_{(x_2, t_2)}(\tau).
$$

One may naturally ask, why assumption (2) in the statement of Theorem 1.9 is meaningful at all, since Hamilton’s trace Harnack is not known to hold without any strong curvature positivity assumption; see [Ha1] and [Br09]. In the present paper, we also show that this condition holds on steady Ricci solitons assuming only nonnegative Ricci curvature. Recall that a Riemannian manifold $(M^n, g)$ is said to have a Ricci soliton structure, if there is a smooth vector field $X$ such that

$$
2 \text{Ric} + \mathcal{L}_X g = \lambda g,
$$

for some real number $\lambda$. The Ricci soliton is called shrinking, steady, or expanding if $\lambda$ is positive, zero, or negative, respectively. The Ricci soliton is gradient if $X = \nabla f$ for some smooth function $f$, which is usually called the potential function. Let $(M^n, g, f)$ be a complete gradient Ricci soliton with potential function $f$, that is,

$$
\text{Ric} + \nabla^2 f = \frac{\lambda}{2} g,
$$

for some real number $\lambda$. Set $\tau(t) = 1 - \lambda t$, and define $\Phi_t$ to be the 1-parameter family of diffeomorphisms generated by $\frac{\tau'(t)}{\tau(t)} \nabla f$ with $\Phi_0 = \text{id}$. Then $g(t) := \tau(t) \Phi_t^* g$ is a Ricci flow and is called the canonical form of the gradient Ricci soliton.

For the standard properties of Ricci solitons, see, for example, [CCG+07] and the references therein.

The following theorem is an application of our main theorem.

**Theorem 1.13.** Let $(M^n, g(\tau))_{\tau \in [0, \infty)}$ be the canonical form of a steady gradient Ricci soliton with nonnegative Ricci curvature, where $\tau$ is the backward time. Then Hamilton’s trace Harnack inequality (1.10) holds. Consequently, if $|Rm| \leq CR < \infty$ for some constant $C$, then $g(\tau)$ is $\kappa$-noncollapsed if and only if it has bounded entropy, where $\kappa$ and the entropy bound are mutually dependent on. In the case where $g(\tau)$ is non-flat and $\kappa$-noncollapsed, it has a non-flat asymptotic shrinker.

This paper is organized as follows. In section 2 we show that a noncollapsed ancient solution satisfying the assumptions in Theorem 1.9 has an asymptotic shrinker, and its entropy converges to that of the asymptotic shrinker. In section 3 we verify Bamler’s gradient estimates [Bam20a] on noncompact Ricci flows with bounded curvature. In section 4 we show that on an ancient solution, Perelman’s entropy and the Nash entropy based at varying points must all converge to the same number; Theorem 1.10 and Corollary 1.11 are proved in this section. In section 5 we apply our main theorem to steady solitons with nonnegative Ricci curvature.
2 Asymptotic shrinking gradient Ricci soliton

Let \((M, g(\tau)), \tau \in [0, \infty)\) be a \(\kappa\)-noncollapsed ancient Ricci flow with bounded curvature within each compact time interval, where \(\tau\) is the backward time, satisfying

\[ |Rm| \leq CR. \tag{2.1} \]

Here and henceforth in this section, \(C\) stands for a positive constant which may differ from line to line. Furthermore, we assume that \(p \in M\) is a fixed point such that item (2) in the statement of Theorem 1.9 holds for \(\ell\), the reduced distance based at \((p, 0)\). That is, we assume the following estimate for \(\ell\).

\[ |\nabla \ell|^2 + R \leq \frac{C\ell}{\tau}. \tag{2.2} \]

From Perelman’s proof of the existence of an asymptotic shrinker (c.f. [Y]), one can easily show that a noncollapsed ancient solution satisfying (2.1) and (2.2) also has an asymptotic shrinker. Indeed, the conditions above are sufficient to prove quadratic upper and lower bounds for \(\ell\) (c.f. Lemma 3.2 in [Y]):

\[ \ell(x, \tau) \sim \frac{1}{\tau} \text{dist}^2(x, y) + C, \tag{2.3} \]

where \(y \in M\) is a point at which \(\ell(\cdot, \tau)\) is bounded by \(C\). However, it is worth mentioning that (2.3) itself is not sufficient to show that the integral \((4\pi\tau)^{-\frac{1}{2}}e^{-\ell}\) of the reduced volume is uniformly negligible outside a large ball, yet this fact can be obtained by Hein-Naber’s Gaussian concentration theorem [HN14]; see the argument in the line above (2.34); all these facts are sufficient to show the existence of an asymptotic shrinker.

In this section, we will show that the entropy of the ancient solution converges to that of its asymptotic shrinker. The main result of this section is the following.

**Proposition 2.4.** Let \((M, g(\tau)), \tau \in [0, \infty)\) be a \(\kappa\)-noncollapsed ancient solution with bounded curvature within each compact time interval and satisfying (2.1). Let \(p \in M\) be a fixed point and \(\ell\) the reduced distance based at \((p, 0)\) such that (2.2) holds. Then, for any sequence of positive numbers \(\tau_i \nearrow \infty\), if the sequence of points \(\{x_i\}_{i=1}^\infty \subset M\) satisfies

\[ \limsup_{i \to \infty} \ell(x_i, \tau_i) < \infty, \]

then the following convergence happens after passing to a subsequence

\[ \left( M, g_i(\tau), (x_i, 1), \ell_i \right) \tau \in [1, 2] \to \left( M_\infty, g_\infty(\tau), (x_\infty, 1), \ell_\infty \right) \tau \in [1, 2], \]

where the limit is the canonical form of a Ricci shrinker, \(g_i(\tau) = \tau_i^{-1}g(\tau \tau_i)\), and \(\ell_i(\cdot, \tau) = \ell(\cdot, \tau \tau_i)\). Here the Ricci flows converge in the Cheeger-Gromov-Hamilton sense [Ha2], and the functions \(\ell_i\) converge in the weak \(\ast W^{1,2}_{loc}(M_\infty \times [1, 2])\) sense as well as in the \(C^{0,0}_{loc}(M_\infty \times [1, 2])\) sense, with arbitrarily fixed \(\alpha \in (0, 1)\).

Furthermore, we have

\[ \lim_{\eta \to 1} W_{(p, 0)}(\eta) = \lim_{\eta \to 1} \log V_{(p, 0)}(\eta) = \log \left( \int_{M_\infty} (4\pi\tau)^{-\frac{1}{2}}e^{-\ell} \, dg_\infty(\tau) \right), \tag{2.5} \]

where \(W_{(p, 0)}(\eta)\) is Perelman’s entropy and \(V_{(p, 0)}(\eta)\) is Perelman’s reduced volume, both based at \((p, 0)\).
Remark: As indicated above, due to (2.1), (2.2), and (2.3), the existence statement of an asymptotic shrinker in Proposition 2.4 is almost straightforward. The main focus of Proposition 2.4 is the equality (2.5). Xu [Xu17] first proved it for Type I noncollapsing ancient solutions, and the second author [ZY2] proved it in the bounded and nonnegative curvature operator case.

In the rest of this section,

\[ u := (4\pi \tau)^{-\frac{n}{2}} e^{-f} \]  

(2.6)
denotes the conjugate heat kernel based at \((p,0)\)—the same base point as that of \( \ell \). Similar to the idea in [ZY2], the major effort of proving (2.5) is to obtain the Gaussian upper and lower bounds for \( u \).

**Lemma 2.7.**

\[ (4\pi \tau)^{-\frac{n}{2}} e^{-\ell} \leq u, \]  

(2.8)

\[ \inf_M \ell(\cdot, \tau) \leq \frac{n}{2}. \]  

(2.9)

**Proof.** (2.8) follows from the fact that the left-hand-side is a subsolution to the conjugate heat equation, and when \( \tau \to 0^+ \), both sides converge to the Dirac delta measure concentrated at \( p \in M \). (2.9) simply follows from a maximum principle; see [Per02].

Henceforth we will use \( p_{\tau} \in M \) to denote a point satisfying \( \ell(p_{\tau}, \tau) \leq \frac{n}{2} \).

**Lemma 2.10.**

\[ \left| \frac{\partial \ell}{\partial \tau} \right| \leq C \ell \frac{\tau}{\ell}. \]  

(2.11)

**Proof.** This inequality follows from the assumption (2.2), in combination with the following well-known formula of Perelman (see Lemma 2.22 in [Y])

\[ 2 \frac{\partial \ell}{\partial \tau} + |\nabla \ell|^2 - R + \ell \frac{\tau}{\ell} = 0. \]

(2.22)

In view of inequalities (2.22) and (2.11), let us define the scale

\[ r(x, \tau) = \sqrt{\tau} \min \{ c_0, \ell(x, \tau)^{-\frac{1}{2}} \}, \]

where \( c_0 \in (0, \frac{1}{2}) \) is a small constant which we will determine in the course of the proof. Note that \( \ell \) is always positive on ancient solutions, hence \( r \) is well-defined. Let us then consider the parabolic neighborhood centered at \((x, \tau)\) with radius \( r(x, \tau) \).

**Lemma 2.12.** If \( c_0 \) is taken to be small enough, then the following holds. Let \((x_0, \tau_0) \in M \times (0, \infty)\) and \( r_0 := r(x_0, \tau_0) \). Then

\[ r(y, \tau_0) \geq \frac{4}{5} r_0, \quad \text{for all} \quad y \in B_{g(\tau_0)}(x_0, r_0). \]  

(2.13)
Proof. By (2.2) and the definition of $r$, if ever $r(\cdot, \tau) < c_0 \sqrt{\tau}$, then the following computation is valid.

$$|\nabla r| = \frac{\tau^\frac{1}{2}}{\sqrt{\ell}} |\nabla \ell^\frac{1}{2}| = \frac{\tau^\frac{1}{2}}{2\sqrt{\ell}} |\nabla \ell| \leq \frac{C}{2} \ell^{-1} = \frac{C}{2\tau} r^2,$$

namely,

$$|\nabla r^{-1}| \leq \frac{C}{2\tau} \text{ whenever } r^{-1} > \frac{1}{c_0 \sqrt{\tau}}.$$

Therefore, for any $y \in B_{g(\tau_0)}(x_0, r_0)$, integrating the above inequality along a $g(\tau_0)$-geodesic connecting $x_0$ and $y$, we have

$$\frac{1}{r(y, \tau_0)} \leq \frac{1}{r_0} + \frac{C}{2\tau_0} r_0 = \frac{1}{r_0} + \frac{C r_0^2}{2\tau_0} \frac{1}{r_0} \leq \left(1 + \frac{C c_0^2}{2}\right) \frac{1}{r_0} \leq \frac{5}{4r_0},$$

where the last inequality holds if $c_0$ is taken to be small enough. We have also used the fact that $r_0 \leq c_0 \sqrt{\tau_0}$ by definition. \qed

Lemma 2.14. If $c_0$ is taken to be small enough, then the following holds. Let $(x_0, \tau_0) \in M \times (0, \infty)$ and $r_0 := r(x_0, \tau_0)$. Then

$$r(x_0, \tau) \geq \frac{4}{5} r_0 \text{ for all } \tau \in [\tau_0 - r_0^2, \tau_0 + r_0^2].$$

Proof. By (2.11) and the definition of $r$, so long as $r(\cdot, \tau) < c_0 \sqrt{\tau}$, the following computation is valid

$$\left| \frac{\partial}{\partial \tau} \right| \leq \frac{C}{\tau^\frac{1}{2} \tau^\frac{1}{2}} = \frac{C}{\tau} r,$$

namely,

$$\left| \frac{\partial}{\partial \tau} \log \left( \frac{r}{\sqrt{\tau}} \right) \right| \leq \frac{C}{\tau} \text{ whenever } \log \left( \frac{r}{\sqrt{\tau}} \right) < \log c_0 \ll 0.$$

Integrating the above inequality from $\tau_0$ to $\tau \in [\tau_0 - r_0^2, \tau_0 + r_0^2]$, we have

$$\log \left( \frac{r(x_0, \tau)}{\sqrt{\tau}} \right) \geq \log \left( \frac{r(x_0, \tau_0)}{\sqrt{\tau_0}} \right) - C \int_{\tau_0}^{\tau_0} \frac{1}{\eta} d\eta \geq \log \left( \frac{r(x_0, \tau_0)}{\sqrt{\tau_0}} \right) - C \log \left( \frac{1 + c_0^2}{1 - c_0^2} \right),$$

where the last inequality is because of the fact $\tau \in [\tau_0 - r_0^2, \tau_0 + r_0^2] \subset [(1 - c_0^2)r_0, (1 + c_0^2)r_0].$ Therefore, taking $c_0$ to be small enough, the lemma follows. \qed
Summarizing the previous two lemmas, we obtain the following result. This shows that the scale \( r \) which we defined above serves effectively as a local curvature scale. This fact can make up for the lack of bounded curvature at bounded distance.

**Proposition 2.15.** Let \( c_0 \) be a small constant such that Lemma 2.12 and Lemma 2.14 both hold. Let \((x_0, \tau_0) \in M \times (0, \infty)\) and \( r_0 = r(x_0, \tau_0) \). Then the following are true.

1. \( r \geq \frac{1}{2} r_0 \) on \( B_{g(\tau_0)}(x_0, r_0) \times [\tau_0 - r_0^2, \tau_0 + r_0^2] \),
2. \( |Rm| \leq C r_0^{-2} \) on \( B_{g(\tau_0)}(x_0, r_0) \times [\tau_0 - r_0^2, \tau_0 + r_0^2] \),
3. \( |\partial^p \nabla^q Rm| \leq C_{p,q} r_0^{2-2p-q} \) on \( B_{g(\tau_0)}(x, \frac{1}{2} r_0) \times [\tau - \frac{1}{4} r_0^2, \tau + \frac{1}{4} r_0^2] \), where \( C_{p,q} \) is a constant depending on \( p \) and \( q \).
4. \( \text{Vol}(B_{g(\tau_0)}(x_0, r_0)) \geq c r_0^n \), where \( c \) is a constant depending on \( \kappa \).

**Proof.** (1) is a combination of Lemma 2.12 and Lemma 2.14. (2) follows from the definition of \( r \) and the assumptions (2.1) and (2.2). (3) follows from the localized Shi’s estimates \[Sh\]. (4) follows from (2) and the noncollapsing assumption.

The following Gaussian concentration theorem proved by Hein and Naber in [HN14] is the fundamental technique in the proof of the integral upper bound for the conjugate heat kernel.

**Proposition 2.16 (Hein-Naber’s Gaussian Concentration).** We have

\[
\nu_\tau(A) \nu_\tau(B) \leq \exp\left(-\frac{1}{8\tau} \text{dist}_\tau^2(A, B)\right),
\]
for any subsets \( A \) and \( B \subset M \) and for all \( \tau > 0 \), where

\[
\nu_\tau(A) = \int_A u \text{dg}(\tau)
\]
is a probability measure on \( M \).

We are ready to prove the following Gaussian upper bound for the conjugate heat kernel.

**Proposition 2.17.** Let \( u \) be the conjugate heat kernel based at \((p, 0)\) and \( p_\tau \in M \) a time-dependent point satisfying \( \ell(p_\tau, \tau) \leq \frac{n}{2\tau} \). Then \( u \) has a time-invariant Gaussian upper bound centered at \( p_\tau \), namely,

\[
u_\tau(A) \nu_\tau(B) \leq \exp\left(-\frac{1}{8\tau} \text{dist}_\tau^2(A, B)\right),
\]

for all \((x, \tau) \in M \times (0, \infty)\), where \( C \) is a constant independent of time.

**Proof.** Let \( \tau \in (0, \infty) \) be arbitrarily fixed. Then, by (2.2) and the noncollapsing assumption, the following hold on \( B_{g(\tau)}(p_\tau, \sqrt{\tau}) \).

\[
l \leq C,
\]
\[
|Rm| \leq C R \leq \frac{C}{\tau},
\]
\[
\text{Vol}(B_{g(\tau)}(p_\tau, \sqrt{\tau})) \geq c \tau^{\frac{n}{2}}.
\]
By (2.18), if we take \( A = B_g(\tau)(p_\tau, \sqrt{\tau}) \), then we have
\[
\nu_\tau(A) \geq \int_{B_g(\tau)(p_\tau, \sqrt{\tau})} (4\pi\tau)^{-\frac{n}{2}} e^{-\xi} dg_\tau \geq C.
\]

Fix an arbitrary \( x \in M \) and let \( r_0 = r(x, \tau) \leq c_0 \sqrt{\tau} \). Then, by applying Proposition 2.16 with \( A = B_g(\tau)(p_\tau, \sqrt{\tau}) \) and \( B = B_g(\tau)(x, r_0) \), we have
\[
\int_{B_g(\tau)(x,r_0)} u \ dg(\tau) = \nu_\tau(B) \leq \nu_\tau(A)^{-1} \exp \left( -\frac{1}{8\tau} \text{dist}_\tau^2(A, B) \right) \leq C \exp \left( -\frac{1}{8\tau} (\text{dist}_\tau(p_\tau, x) - \sqrt{\tau} - r_0)^2 \right) \leq C \exp \left( -\frac{1}{16\tau} \text{dist}_\tau^2(p_\tau, x) \right).
\]

Next, we extend the above integral bound to space-time. (2.11) implies that
\[
\ell(p_\tau, \tau') \leq C \quad \text{for all} \quad \tau' \in [\tau - r_0^2, \tau] \subset [(1 - c_0^2)\tau, \tau].
\]

Then, implementing the same argument as before at \( \tau' \in [\tau - r_0^2, \tau] \subset [(1 - c_0^2)\tau, \tau] \), with \( A = B_g(\tau')(p_\tau, \sqrt{\tau'}) \) and \( B = B_g(\tau')(x, \tau') \), where \( \tau' = r(x, \tau') \), we obtain
\[
\int_{B_g(\tau')(x,r')} u \ dg(\tau') \leq C \exp \left( -\frac{1}{16\tau} \text{dist}_{\tau'}^2(p_\tau, x) \right),
\]
for all \( \tau \in [\tau - r_0^2, \tau] \). Because of Proposition 2.15 (1)(2), we may find a small constant \( c_1 \in (0, \frac{1}{4}) \), such that
\[
B_g(\tau)(x, cr_0) \subset B_g(\tau')(x, r') \quad \text{for all} \quad \tau \in [\tau - r_0^2, \tau].
\]

Hence, (2.20) can be rewritten as
\[
\int_{B_g(\tau')(x, cr_0)} u \ dg(\tau') \leq C \exp \left( -\frac{1}{16\tau} \text{dist}_{\tau'}^2(p_\tau, x) \right).
\]

Note that the distance on the right-hand-side is in terms of \( g(\tau') \). This can be dealt with by Perelman’s distance distortion estimate. Indeed, the definition of \( r_0 \) and (2.19) imply that\( r(p_\tau, \tau') \geq c \sqrt{\tau} \geq cr_0 \) for all \( \tau' \in [\tau - r_0^2, \tau] \). In combination with the fact that \( r(x, \tau') \geq \frac{\tau}{2} r_0 \) for all \( \tau' \in [\tau - r_0^2, \tau] \) as well as Proposition 2.15(2), we have
\[
\text{Ric}_{g(\tau')} \leq C r_0^{-2} \quad \text{on} \quad B_g(\tau')(x, cr_0) \cup B_g(\tau')(p_\tau, cr_0)
\]
for all \( \tau' \in [\tau - r_0^2, \tau] \). Hence, by using Lemma 8.3 in [Per02], we obtain
\[
\frac{d}{ds} \text{dist}_s(p_\tau, x) \leq C r_0^{-2} \quad \text{for all} \quad s \in [\tau - r_0^2, \tau].
\]

Integrating the above inequality from \( \tau' \in [\tau - r_0^2, \tau] \) to \( \tau \), we have
\[
\text{dist}_\tau(p_\tau, x) \geq \text{dist}_\tau(p_\tau, x) - C.
\]
Combining (2.21) and (2.22), we obtain
\[ \int_{r_0^2}^{\tau} \int_{B_{g(r)}(x, c_1 r_0)} u \, dg, \, dr' \leq C r_0^2 \exp \left( -\frac{1}{16 \tau} \text{dist}^2_r(p_r, x) \right). \quad (2.23) \]

Since Proposition 2.15 provides good geometry bounds on the parabolic ball \( B_{g(r)}(x, c_1 r_0) \times [\tau - r_0^2, \tau] \), we then obtain the pointwise bound by using the standard parabolic mean value inequality (c.f. Lemma 3.1 in [CTY]).

\[
\begin{align*}
 u(x, \tau) & \leq \frac{C}{r_0^{n+2}} \int_{\tau - r_0^2}^{\tau} \int_{B_{g(r)}(x, c_1 r_0)} u \, dg, \, dr' \quad (2.24) \\
 & \leq C r_0^{-n} \exp \left( -\frac{1}{16 \tau} \text{dist}^2_r(p_r, x) \right) \\
 & \leq C (4\pi \tau)^{-\frac{n}{2}} \exp \left( -\frac{1}{16 \tau} \text{dist}^2_r(p_r, x) - C \log \frac{r_0}{\sqrt{\tau}} \right). 
\end{align*}
\]

Finally, to deal with the last term in the above formula, recall the definition of \( r_0 := r(x, \tau) = \sqrt{\tau \min \{c_0, \ell(x, \tau)^{-\frac{1}{n}} \}} \), we then compute
\[
\begin{align*}
-C \log \left( \frac{r_0}{\sqrt{\tau}} \right) & \leq C \log \left( \max \left\{ c_0^{-1}, \ell(x, \tau)^{\frac{1}{n}} \right\} \right) \quad (2.25) \\
& \leq C \log \left( c_0^{-1} + C + C \frac{\text{dist}^2_r(p_r, x)}{\sqrt{\tau}} \right) \\
& \leq \frac{1}{32} \frac{\text{dist}^2_r(p_r, x)}{\tau} + C,
\end{align*}
\]
where we have used the trivial observation that the logarithmic function grows slower than the quadratic function. Precisely, for any \( \epsilon > 0 \), there exists \( C(\epsilon) > 0 \), such that
\[
\log x \leq \epsilon x^2 + C(\epsilon) \quad \text{for all} \quad x > 0; \]

note that the quadratic upper bound of \( \ell \) is merely a consequence of (2.2). Combining (2.24) and (2.25), the conclusion of the proposition follows.

\[ \square \]

Proposition 2.17 has the following immediate implication.

**Corollary 2.26.** For every \( \tau \in (0, \infty) \), let \( p_\tau \in M \) be a point such that \( \ell(p_\tau, \tau) \leq \frac{\tau}{2} \). Then u has the following Gaussian upper and lower bounds
\[ \frac{1}{C(4\pi \tau)^{\frac{n}{2}}} \exp \left( -\frac{C}{\tau} \text{dist}^2_r(p_r, x) \right) \leq u(x, \tau) \leq \frac{C}{(4\pi \tau)^{\frac{n}{2}}} \exp \left( -\frac{1}{C \tau} \text{dist}^2_r(p_r, x) \right), \quad (2.27) \]
for all \((x, \tau) \in M \times (0, \infty)\), where \( C \) is a constant independent of \( \tau \). Furthermore, the reduced distance \( \ell \) satisfies the following quadratic estimates
\[ \frac{1}{C \tau} \text{dist}^2_r(p_r, x) - C \leq \ell(x, \tau) \leq \frac{C}{\tau} \text{dist}^2_r(p_r, x) + C, \quad (2.28) \]
for all \((x, \tau) \in M \times (0, \infty)\), where \( C \) is a constant independent of \( \tau \).
Proof. The second inequality of (2.28) follows from the assumption (2.2); this fact, in combination with (2.8), implies the first inequality of (2.27). The first inequality of (2.28) follows from Proposition 2.17 and (2.8).

Proof of Proposition 2.4. We only prove the equality (2.5). As to the existence of the asymptotic shrinker, one may refer to either the comments at the beginning of this section, or the arguments of Proposition 4.2 in [ZY2]. First of all, let us fix the sequences of points \(\{x_i\}_{i=1}^{\infty} \subset M\) and positive numbers \(\tau_i \nearrow \infty\) as in the statement of the proposition. Then, we define

\[
f_i(\cdot, \tau) = f(\cdot, \tau \tau_i), \quad u_i = (4\pi \tau)^{-\frac{n}{2}} e^{-f_i},
\]

where, as before,

\[
u := (4\pi \tau)^{-\frac{n}{2}} e^{-f}
\]
is the conjugate heat kernel based at \((p_0, 0)\). We will prove that \(\{f_i\}_{i=1}^{\infty}\) also smoothly converges to the potential function on the asymptotic shrinker.

First, we verify that the base points \(x_i\) are equivalent to \(p_{\tau_i}\). By (2.11), we have

\[
\ell(x_i, \tau) \leq \ell(x_i, \tau_i) \left(\frac{\tau}{\tau_i}\right)^C \leq \frac{n}{2} \cdot 2^C \leq C \text{ for all } \tau \in \left[\frac{1}{2} \tau_i, 2 \tau_i\right].
\]

(2.29)

It then follows from (2.28) that

\[
\text{dist}_{\tau}(x_i, p_{\tau}) \leq C\sqrt{\tau_i} \quad \text{for all } \tau \in \left[\frac{1}{2} \tau_i, 2 \tau_i\right].
\]

Hence, after the parabolic scaling, the Gaussian upper and lower bounds (2.27) become

\[
\frac{1}{C} \cdot \text{dist}^2_{g_\tau}(x_i, x) - C \leq f_i(x, \tau) \leq C \cdot \text{dist}^2_{g_\tau}(x_i, x) + C,
\]

(2.30)

for all \(\tau \in \left[\frac{1}{2}, 2\right]\).

Arguing as Proposition 3.3 in [ZY2], one may immediately obtain uniform growth estimates for the derivatives of \(f_i\); indeed, all gradients of \(f_i\) have uniformly polynomial growth bounds. Hence we have

\[
f_i \to f_{\infty} \quad \text{and} \quad u_i \to u_{\infty} = (4\pi \tau)^{-\frac{n}{2}} e^{-f_{\infty}}
\]

locally smoothly on \(M \times [1, 2]\). For the convergence of the entropy and of the integral of \(u_i\), we need to show that the integral of \(u_i\) outside a large ball is negligible. Note that unlike the case of [ZY2], we do not have a lower curvature bound to apply volume comparison theorem. Nevertheless, the Gaussian concentration theorem is sufficient. Fixing a large number \(\rho \gg 1\), we apply Proposition 2.16 to \(A = B_{g_\tau}(x_i, \sqrt{\tau})\) and \(B = M \setminus B_{g_\tau}(x_i, \rho \sqrt{\tau_i})\) for all \(\tau \in [\tau_i, 2 \tau_i]\). Since, by (2.29), \(\ell(x_i, \tau)\) is bounded from above by a constant, we may argue as the beginning of the proof of Proposition 2.17 to obtain that \(\nu_{\tau}(A) \geq c > 0\) for some \(c > 0\). Hence

\[
\nu_{\tau}(A) \leq \exp\left(-\frac{1}{16 \tau} \rho^2\right).
\]
3 Bamler’s gradient estimate on noncompact manifolds

In this section, we vindicate Bamler’s sharp gradient estimates [Bam20a] on noncompact manifolds. Assuming bounded curvature, these results should be readily established by using his method. We include this proof for the convenience of the readers.

For a function $u$ defined on $M \times [a, b]$, we write $u_t = u(\cdot, t)$. Following [Bam20a], we consider the function $\Phi(x) := \frac{1}{\sqrt{4\pi \tau}} \int_x^\infty \exp(-y^2/4) \, dy$. Then $\Phi_t(x) := \Phi(t^{-1/2}x)$ is a solution to the 1-dimensional heat equation $\partial_t \Phi_t = \Phi_t''$ with initial condition $\chi_{[0, \infty)}$. Using Bamler’s method, we have the following.

**Theorem 3.1** (Theorem 4.1 in [Bam20a]). Let $(M, g(t))_{t \in [t_0, t_1]}$ be a complete Ricci flow with $|\text{Rm}| \leq \Lambda$ on $M \times [t_0, t_1]$ for some constant $\Lambda < \infty$. Consider a solution $u$ to the heat equation coupled with $g(t)$ and assume that $u$ takes values in $(0, 1)$. Let $T \geq 0$ and suppose $|\nabla \Phi_T^{-1}(u_{t_0})|_{g(t_0)} \leq 1$ if $T > 0$.

Then $|\nabla \Phi_{T+t-t_0}^{-1}(u_t)|_{g(t)} \leq 1$ for any $t \in (t_0, t_1)$.
Following Bamler, we wish to apply a version of maximum principle for noncompact Ricci flows [CCGG+08, Theorem 12.14] to $|\nabla \Phi_T^{-1}(u_t)|^2$. To do so an a priori bound is necessary. The following lemma, whose detailed proof is left to the readers, is a consequence of the Bernstein-Bando-Shi technique (c.f. [Ko07]) as well as the well know Bochner formula for a solution $u$ to the heat equation coupled with the Ricci flow

$$\Box |\nabla u|^2 = -2|\nabla^2 u|^2 \leq 0,$$

where $\Box = \partial_t - \Delta_{g(t)}$ is the heat operator coupled with the Ricci flow.

**Lemma 3.2.** Let $(M^n, g(t))_{t \in [0,1]}$ be a complete Ricci flow with bounded curvature. Suppose that $u$ is a solution to the heat equation coupled with $g(t)$ satisfying $|u| \leq 1$ on $M \times [0,1]$. If $|\nabla u_0|^2 \leq A$, where $A$ is a positive number, then $|\nabla u_t|^2 \leq A$ for all $t \in [0,1]$. If there is no initial gradient bound, then we have $t|\nabla u_t|^2 \leq 25$ on $M \times [0,1]$.

**Proof of Theorem 3.1** We first consider the case where $T > 0$. By parabolic rescaling and time shifting, we may assume $0 < t_0 = T < 1$ and $t_1 \geq 1$. It suffices to show $|\nabla \Phi_T^{-1}(u_t)| \leq 1$ under the assumption $|\nabla \Phi_T^{-1}(u_t)|_{g(T)} \leq 1$. In the following, we shall omit the subindices and the reader should keep in mind that the norms of the gradients are computed using the evolving metric.

We write $h_t = \Phi_T^{-1}(u_t)$. For any small $\epsilon > 0$, let

$$u_t^{(\epsilon)} := \epsilon + (1 - 2\epsilon)u_t \in (\epsilon, 1 - \epsilon), \quad h_t^{(\epsilon)} := \Phi_T^{-1}(u_t^{(\epsilon)}).$$

Define $A_\epsilon = \Phi^{-1}(1 - \epsilon)$, then $|h_t^{(\epsilon)}| \leq A_\epsilon \sqrt{t}$ for all $t \in [T,1]$. Note that

$$u^{(\epsilon)} - \frac{1}{2} = (1 - 2\epsilon)\left(u - \frac{1}{2}\right),$$

which means that $u^{(\epsilon)}$ is closer to $1/2$ than $u$. It follows that $|h_t^{(\epsilon)}| = |\Phi_T^{-1}(u_t^{(\epsilon)})| \leq |\Phi^{-1}_T(u_t)| = |h_t|$ and

$$|\nabla h_t^{(\epsilon)}| = (1 - 2\epsilon) \frac{|\nabla u_T|_{\Phi_T'(h_T^{(\epsilon)})}}{|\Phi_T'(h_T^{(\epsilon)})|} \leq \frac{|\nabla u_T|}{|\Phi_T'(h_T)|} = |\nabla h_T| \leq 1,$$

where the first inequality above is due to the definition of $\Phi$, and the last inequality is simply the assumption of the theorem. Hence

$$|\nabla u_T^{(\epsilon)}| \leq \Phi_T'(h_T^{(\epsilon)}) \leq (4\pi T)^{-1/2}.$$

By Lemma 3.2, we have

$$|\nabla u_t^{(\epsilon)}| \leq C \quad \text{for all} \quad t \in [T,1],$$

where $C = (4\pi T)^{-1/2}$. It follows that

$$|\nabla h_t^{(\epsilon)}| \leq (4\pi t)^{1/2} C \exp(A_\epsilon^2/4),$$

where we have used the fact that $h_t^{(\epsilon)} \leq A_\epsilon \sqrt{T}$. Hence, $\left|\nabla \frac{(h_t^{(\epsilon)})^2}{2t}\right|$ is bounded on $M \times [T,1]$. By the computations in Theorem 4.1 of [Bam20a], we have

$$\left(\partial_t - \Delta_t\right) \left|\nabla h_t^{(\epsilon)}\right|^2 + \nabla h_t^{(\epsilon)} \cdot \nabla \left|\nabla h_t^{(\epsilon)}\right|^2 = \frac{1 - \left|\nabla h_t^{(\epsilon)}\right|^2}{t} \left|\nabla h_t^{(\epsilon)}\right|^2.$$
We can apply the maximum principle [CCGG+08, Theorem 12.14] to $|\nabla h^{(\epsilon)}|^2 - 1$, and thereby conclude that $|\nabla h^{(\epsilon)}_1| \leq 1$. Taking $\epsilon \to 0$, we have that $|\nabla h_1| \leq 1$ and we have finished the proof in this case.

We now consider the case where $T = 0$. This can be proved by a limiting argument as observed in [Bam20b, Lemma 3.8]. By parabolic rescaling and time shifting, we may assume that $T = t_0 = 0$ and $t_1 \geq 1$. It suffices to prove $|\nabla h_1| = |\nabla \Phi^{-1}_1(u_1)| \leq 1$, where we use the same notations as above.

By Lemma 3.2 we have $|\nabla u_s|^2 \leq \frac{25}{4}$ on $M \times (0,1]$. Let $\epsilon \in (0,1/2)$ be arbitrarily fixed. Then, for any $s \in (0,1/2)$ we will define $T = T(\epsilon,s) \in (0,s)$, so that the result in the first case can be applied. To this end, let us compute as follows.

$$
|\nabla \Phi^{-1}_T\left(u^{(\epsilon)}_s\right)| = (4\pi T)^{1/2} \exp\left(\left(\Phi^{-1}_T(u^{(\epsilon)}_s)\right)^2 / 4T\right) (1 - 2\epsilon)|\nabla u_s| \leq C(T/s)^{1/2} \exp(A^2_t/4).
$$

We take $T = T(\epsilon,s)$ small enough so that $|\nabla \Phi^{-1}_T\left(u^{(\epsilon)}_s\right)| \leq 1$. By the result in the first case, we have

$$
|\nabla \Phi^{-1}_{T+1-s}\left(u^{(\epsilon)}_1\right)| \leq 1.
$$

Letting $s \to 0$ and then $\epsilon \to 0$, we have $|\nabla \Phi^{-1}_1(u_1)| \leq 1$. $
\square$

We state the following standard result.

**Lemma 3.3.** Suppose that $\mu$ is a probability measure on some measure space $(X, \mathcal{B})$. Let $q : X \to \mathbb{R}$ be a measurable function and let $\phi : \mathbb{R} \to \mathbb{R}$ be absolutely continuous on each compact interval. For any real number $\alpha$, if $\phi|_{[\alpha, \infty)}$ is monotone, then

$$
\int_{\{q > \alpha\}} \phi(q) \, d\mu = \phi(\alpha)\mu(q > \alpha) + \int_{\alpha}^{\infty} \phi'(t)\mu(q > t) \, dt.
$$

**Proof.** This is a slightly different statement of [Rui, Theorem 8.16]. The proof is the same. One can consider $\phi'(t)\chi_E(x,t)$ with $E = \{(x,t) \in X \times [\alpha, \infty) : q(x) > t \geq \alpha\}$ and apply Tonelli’s theorem to its integral. $
\square$

**Proposition 3.4** (Proposition 4.2 in [Bam20a]). Suppose that the same condition as in Theorem 3.7 holds for a complete Ricci flow $(M, g(\cdot))$ defined on $[s,t]$. Write $d\nu = K(x,t|\cdot, s) \, dg_s$. Then for any $x \in M$, $1 \leq p < \infty$, and any measurable subset $\Omega \subseteq M$, we have

$$
(t-s)^{p/2} \int_{\Omega} \left(\frac{\nabla_x K(x,t|\cdot, s)}{K(x,t|\cdot, s)}\right)^p \, d\nu \leq C(n,p)\nu(\Omega) \left(-\log(\nu(\Omega)/2)\right)^{p/2}. \tag{3.5}
$$

Moreover, for any $x \in M$ and $v \in T_x M$ with $|v|_{g(t)} = 1$, it holds that

$$
(t-s) \int_M \left|\frac{\partial_v K(x,t|\cdot, s)}{K(x,t|\cdot, s)}\right|^2 \, d\nu \leq 1/2. \tag{3.6}
$$

**Proof.** By parabolic rescaling and time shifting, we may assume that $[s,t] = [0,1]$. Let $q = \frac{\partial_v K(x,1|\cdot,0)}{K(x,1|\cdot,0)}$. The proof of [Bam20a, Proposition 4.2] applies here almost without many
variations. For completeness, we will sketch the proof and will refer to Bamler’s original proof wherever his argument comes through. Let 
\[ \lambda(\alpha) = \nu(q \geq \alpha), \quad h(t) = \sup\{\alpha : \lambda(\alpha) \geq t\}, \]
where we have written \( \nu(q \geq \alpha) = \nu(\{q \geq \alpha\}) \), and the notations such as \( \nu(q > \alpha), \nu(q = \alpha) \), etc., shall also be defined accordingly. Obviously, \( \lambda \) and \( h \) are non-increasing and left continuous.

For any measurable subset \( \Omega \subset M \), we have, by Proposition 3.1,
\[
\left| \int_{\Omega} q \, d\nu \right| = \left| \partial_v \int_{M} \chi_{\Omega} K(x, 1|x, 0) \, dg_0 \right| \leq F(\nu(\Omega)), \tag{3.7}
\]
where \( F(s) = \Phi'(\Phi^{-1}(s)) \). The integration and differentiation are interchangeable by similar (and indeed, much simpler) arguments as in the appendix.

By the definition of \( h \), Bamler proved that
\[
\nu(q \geq \alpha) = |h \geq \alpha|, \quad \nu(q > \alpha) = |h > \alpha|, \quad \nu(q = \alpha) = |h = \alpha|.
\]
It follows directly from Lemma 3.3 that
\[
\int_{\{h > \alpha\}} \phi(h) = \int_{\{q > \alpha\}} \phi(q) \, d\nu,
\]
whenever \( \phi|_{[\alpha, \infty)} \) is monotone and absolutely continuous on any compact intervals.

Claim: Assume \( f : \mathbb{R} \to \mathbb{R} \) is absolutely continuous on any compact interval and is monotone on both \([A, \infty)\) and \((\infty, -A]\) for some \( A \in \mathbb{R}_+ \). Note that we do not require \( f \) to have the same monotonicity on \([A, \infty)\) and \((\infty, -A]\). For any \( \alpha \in \mathbb{R} \cup \{\pm \infty\} \) and any measurable subset \( \Omega \subset M \) with
\[
\{q > \alpha\} \subset \Omega \subset \{q \geq \alpha\},
\]
it holds that
\[
\int_{\Omega} f(q) \, d\nu = \int_{0}^{\nu(\Omega)} f(h). \tag{3.8}
\]

Proof of the claim. The case where \( \alpha = +\infty \) is obvious. We only need to consider the case where \( \alpha \in \mathbb{R} \). This is because the case where \( \alpha = -\infty \) follows from the case where \( \alpha \in \mathbb{R} \) and the following observation: when \( \alpha = -\infty \), we have \( \Omega = M \) and
\[
\int_{M} f(q) \, d\nu = \lim_{\alpha \to -\infty} \int_{\{q > \alpha\}} f(q) \, d\nu = \lim_{\alpha \to -\infty} \int_{\{h > \alpha\}} f(h) = \int_{0}^{1} f(h),
\]
where the first and the third equalities can be justified by applying the monotone convergence theorem to \( f(q) \) and \( f(h) \) using the monotonicity of \( f|_{(-\infty, -A]} \).

Fix an arbitrary \( \epsilon > 0 \). Without loss of generality, we may assume \( A > \alpha \), for otherwise we may always enlarge \( A \). Then \( \{q > A\} \subset \{q > \alpha\} \subset \Omega \), and there is a partition \( \alpha = b_0 < b_1 < \cdots <
$b_m = A$ such that $\text{osc}_{[b_{i-1}, b_i]} f \leq \epsilon/2$ for all $i = 1, \ldots, m$. As in [Bam20a], we have

$$
\left| \int_{\Omega} f(q) \, d\nu - \int_{\nu(\Omega)} f(h) \right| \leq \epsilon
$$

(3.9)

where we have applied Lemma 3.3 to conclude that the term on the third line of the above formula is indeed zero.

By the same arguments as in [Bam20a], the claim above implies that for $a \in (0, 1)$,

$$ah(a) \leq \int_0^a h \leq F(a),$$

$$(1-a)h(a) \geq \int_a^1 h \geq -F(a).$$

When $a \in (0, 1/4)$, we also have

$$h(a) \leq \frac{F(a)}{a} \leq C(-\log a)^{1/2}, \quad h(1-a) \geq -C(-\log a)^{1/2}.$$

Then, applying the claim to $f(t) = |t|^p$ for any $1 \leq p < \infty$, (3.5) can be proved using the same arguments as in [Bam20a, Proposition 4.2].

To prove (3.6), let us define

$$H(a) := \int_0^a h \leq F(a),$$

and clearly $H(0) = 0$. Applying the claim with $f(t) = t$ and $\alpha = -\infty$, we have $H(1) = \int_0^1 h = \int_M q \, dv = 0$. From the definition of $H$, we have $H'' \leq 0$ in the weak sense. Furthermore, we may argue as [Bam20a, Proposition 4.2] and obtain

$$\int_0^1 h(a)^2 \, da \leq \frac{1}{2}.$$

(3.10)

Note that the boundary terms produced by integration by parts in [Bam20a, Proposition 4.2] vanish because of the fact that

$$H'(a) = h(a) \leq \frac{F(a)}{a} \leq C(-\log a)^{1/2},$$

when $a \in (0, 1/4)$. Applying the claim above with $f(t) = t^2$ and $\alpha = -\infty$, (3.6) follows from (3.10).
4 The Nash entropy based at varying points

After having established Bamler’s gradient estimates on noncompact manifolds, we subsequently prove several results for the Nash entropy. Following [Bam20a], we use the notation

\[ N^*_s(x, t) := N(x, t)(t - s) \] (4.1)

to manifest the dependence of the Nash entropy on its base point. Here \( N(x, t) \) is as defined in (1.3). Henceforth we will also define the time-dependent probability measure

\[ \nu_{x,t}(s)(A) = \int_A K(x, t | \cdot, s) \, dg(s) \]

for all measurable subset \( A \subset M \). The \( L^1 \) Wasserstein distance is also a very important tool in [Bam20a]:

**Definition 4.2.** Let \((X, d)\) be a complete metric space and \(\mu, \nu\) probability measures on \(X\). Then the \(L^1\) Wasserstein distance between \(\mu\) and \(\nu\) is defined as

\[ d_{W_1}(\mu, \nu) = \sup_f \left( \int_X f \, d\mu - \int_X f \, d\nu \right), \]

where the supremum is taken over all bounded 1-Lipschitz functions.

The following observation made by Bamler is merely a consequence of Lemma 3.2.

**Proposition 4.3** (Lemma 2.7 in [Bam20a]). Let \((M, g(t))_{t \in [0, T]}, \) where \(0 < T < \infty\), be a Ricci flow with bounded curvature. Assume that \(s, t, t_1, \) and \(t_2 \in [0, T]\) satisfy \(s < t \leq t_1, t_2\). Let \(x_1, x_2 \in M\). Then the following holds

\[ d_{W_1}^{g(s)}(\nu_{x_1, t_1}(s), \nu_{x_2, t_2}(s)) \leq d_{W_1}^{g(t)}(\nu_{x_1, t_1}(t), \nu_{x_2, t_2}(t)). \]

Our generalization of Bamler’s gradient estimate (Proposition 3.1) also leads to the generalization of the following theorem and corollary in [Bam20a]. These results indicate that, fixing the evaluating time, the Nash entropy satisfies good properties with respect to its base point.

**Theorem 4.4** (Theorem 5.9 in [Bam20a]). Let \((M, g(t))_{t \in [0, T]}, \) where \(0 < T < \infty\), be a complete Ricci flow with bounded curvature. Let \(s \in [0, T]\) and assume \(R(\cdot, s) \geq R_{\min}, \) where \(R_{\min}\) is a real number. Then, on \(M \times (s, T]\), it holds that

\[ |\nabla N^*_s| \leq \left( \frac{n}{2(t - s)} - R_{\min} \right)^{\frac{1}{2}}, \quad -\frac{n}{2(t - s)} \leq \left( \frac{\partial}{\partial t} - \Delta_t \right) N^*_s \leq 0. \]

**Proof.** The proof of Theorem 4.1 shows that the integration and the differentiation are always interchangeable in the computations of both \(\nabla N^*_s(\cdot, t)\) and \(\square_{x,t} N^*_s(x, t)\). Hence, the proof follows after [Bam20a] line-by-line, using (3.9) above and Proposition 2.1 in [CMZ21]. \(\square\)

As a direct consequence, we have the following Corollary.
Corollary 4.5 (Corollary 5.11 in [Bam20a]). In the same settings as the previous theorem, if \( R(\cdot, t^*) \geq R_{\min} \) and \( s < t^* \leq t_1, t_2 \), where \( s, t^*, t_1, t_2 \in [0, T] \), then for \( x_1 \) and \( x_2 \in M \), we have

\[
N_s^*(x_1, t_1) - N_s^*(x_2, t_2) \leq \left( \frac{n}{2(t^* - s)} - R_{\min} \right)^{\frac{1}{2}} d_{W_1}^{(t^*)}(\nu_{x_1, t_1}(t^*), \nu_{x_2, t_2}(t^*)) + \frac{n}{2} \log \left( \frac{t_2 - s}{t^* - s} \right).
\]

We now apply Corollary 4.5 to an ancient solution.

Proposition 4.6. Let \((M, g(t))_{t \in (-\infty, 0]}\) be an ancient solution with bounded curvature within each compact time interval. Then for any \((x_1, t_1)\) and \((x_2, t_2) \in M \times (-\infty, 0]\), it holds that

\[
\lim_{\tau \to \infty} W_{(x_1, t_1)}(\tau) = \lim_{\tau \to \infty} N_{x_1, t_1}(\tau) = \lim_{\tau \to \infty} N_{x_2, t_2}(\tau) = \lim_{\tau \to \infty} W_{(x_2, t_2)}(\tau). \tag{4.7}
\]

Proof. We first prove the second equality of (4.7). Observe that this is equivalent to

\[
\lim_{s \to -\infty} \left( N_s^*(x_1, t_1) - N_s^*(x_2, t_2) \right) = 0. \tag{4.8}
\]

Without loss of generality, we may assume \( t_1 \leq t_2 \). Let \( \varepsilon > 0 \) be arbitrarily fixed, we then apply Corollary 4.5 with \( s \ll -1 \) and \( t^* = \varepsilon s \leq t_1 \). Since we can take \( R_{\min} = 0 \) by [CB], we have

\[
N_s^*(x_1, t_1) - N_s^*(x_2, t_2) \leq \left( \frac{n}{2(1 - \varepsilon)|s|} \right)^{\frac{1}{2}} d_{W_1}^{(t_1)}(\delta_{x_1}, \nu_{x_2, t_2}(t_1)) + \frac{n}{2} \log \left( \frac{t_2 - s}{t^* - s} \right).
\]

where we have used Proposition 4.3. By the definition of the \( L^1 \) Wasserstein distance, we have that

\[
d_{W_1}^{(t_1)}(\delta_{x_1}, \nu_{x_2, t_2}(t_1)) \leq \int_M \text{dist}_{t_1}(x, \cdot) d\nu_{x_2, t_2}(t_1) \leq C \sum_{j=1}^{\infty} j \exp(-c_0j^2) < \infty,
\]

where the second inequality is due to the Gaussian concentration theorem (Proposition 2.10 see also Proposition 2.2 in [CMZ21] under the current assumption). Taking \( s \to -\infty \) on both sides of (4.8), we obtain

\[
\lim_{s \to -\infty} \left( N_s^*(x_1, t_1) - N_s^*(x_2, t_2) \right) \leq \frac{n}{2} \log \left( \frac{1}{1 - \varepsilon} \right).
\]

Since \( \varepsilon > 0 \) is arbitrary, we then have

\[
\lim_{s \to -\infty} \left( N_s^*(x_1, t_1) - N_s^*(x_2, t_2) \right) \leq 0.
\]

Reversing the order of \((x_1, t_1)\) and \((x_2, t_2)\), we obtain (4.8).
is proved in [ZY1, Corollary 4.5]; we include its proof here. Fix an arbitrary $\varepsilon > 0$, then (1.5) implies
\[
W(x,t; \tau) \leq N(x,t; \tau) = 1 = \int_0^\tau W(x,t; \eta) d\eta \leq (1 - \varepsilon) W(x,t; \varepsilon \tau).
\]
By first taking $\tau \to \infty$ and then $\varepsilon \to 0$, we obtain
\[
\lim_{\tau \to \infty} W(x,t; \tau) = \lim_{\tau \to \infty} N(x,t; \tau).
\]

Finally, we are ready to prove Theorem 1.9.

**Proof of Theorem 1.9.** Let $(M, g(t))_{t \in (-\infty, 0]}$ be an ancient solution satisfying the conditions of this theorem. The sufficiency direction follows immediately from Proposition 3.3 in [ZY1], namely, if $(M, g(t))_{t \in (-\infty, 0]}$ has a finite entropy bound, then it is $\kappa$-noncollapsed on all scales, where $\kappa > 0$ depends on the entropy bound.

On the other hand, if $(M, g(t))_{t \in (-\infty, 0]}$ is $\kappa$-noncollapsed on all scales, where $\kappa > 0$, then, by Proposition (2.4), $(M, g(t))_{t \in (-\infty, 0]}$ has an asymptotic shrinker. In combination with Proposition 4.6, we have that Perelman’s entropy defined on $(M, g(t))_{t \in (-\infty, 0]}$ with arbitrary base point must have the same lower bound—the entropy on the asymptotic shrinker. This finishes the proof.

**Proof of Corollary 1.11.** We consider only case (2). Let $(x_0, t_0)$ be an arbitrary point in $M \times (-\infty, 0]$ and $N$ the reduced distance based at $(x_0, t_0)$. Assumption (2) in Theorem 1.9 is implied by Hamilton’s trace Harnack as mentioned in Remarks below Theorem 1.9. Then, Corollary 1.11 is a combination of Proposition 2.4 and Proposition 4.6.

5 Applications to steady solitons

In this section, we prove Theorem 1.13 together with some other results concerning steady solitons with nonnegative Ricci curvature. Let $(M^n, g, f)$ be a complete steady gradient Ricci soliton with nonnegative Ricci curvature, normalized so that
\[
\text{Ric} = \nabla^2 f, \quad R + |\nabla f|^2 = 1.
\]
Throughout this section, we will consider its canonical form $g(\tau)$; see section 1 for the definition. Note that $\tau$ stands for the backward time.

We first observe that the Ricci nonnegativity condition implies Hamilton’s trace Harnack inequality, which in turn implies item (2) in Theorem 1.9 according to the remarks thereof. After establishing the following lemma, Theorem 1.13 follows immediately from Theorem 1.9.

**Lemma 5.1.** Let $(M^n, g, f)$ be a complete steady gradient Ricci soliton with nonnegative Ricci curvature and let $(M, g(\tau))_{\tau \in [0, \infty)}$ be its canonical form. Then Hamilton’s trace Harnack inequality (1.10) holds.
5 APPLICATIONS TO STEADY SOLITONS

Proof. Since \( g(\tau) = \Phi^*_\tau g \) moves only by diffeomorphisms, it suffices to verify for the metric \( g = g(0) \). Recall that \( \nabla R = -2\text{Ric}(\nabla f, \cdot) \) on steady gradient Ricci solitons; see, for example, (1.27) in [CCGG+07]. (Note that we use a different sign convention on \( f \) here.) Let \( X \) be an arbitrary vector field on \( M \). We can then simplify the expression of the trace Harnack quantity as:

\[
- \frac{\partial R}{\partial \tau} \bigg|_{\tau=0} = -2\langle X, \nabla R \rangle + 2\text{Ric}(X, X)
\]

\[
= 2\text{Ric}(\nabla f, \nabla f) + 4\text{Ric}(X, \nabla f) + 2\text{Ric}(X, X)
\]

\[
= 2\text{Ric}(X + \nabla f, X + \nabla f) \geq 0,
\]

where we have applied the assumption \( \text{Ric} \geq 0 \).

Under the assumption of this section, the AVR (asymptotic volume ratio) is well defined, and hence the existence of a non-flat asymptotic shrinker implies zero AVR as observed in [Ni05], which in turn implies infinite ASCR (asymptotic scalar curvature ratio) as observed in [DZ18]. Recall that for a Riemannian manifold \((M, g)\) with nonnegative Ricci curvature, we have

\[
\text{AVR}(g) := \lim_{r \to \infty} \frac{\text{Vol}(B_r(p))}{r^n}, \quad \text{ASCR}(g) := \limsup_{x \to \infty} R(x)\text{dist}^2(x, p).
\]

One can easily show that the two definitions do not depend on the choice of the base point \( p \in M \).

Corollary 5.2. Let \((M, g, f)\) be a complete and \( \kappa \)-noncollasped steady gradient Ricci soliton with nonnegative Ricci curvature. Assume \( |R_m| \leq CR \) for some constant \( C \). If \( g \) is not flat, then \( \text{AVR}(g) = 0 \) and \( \text{ASCR}(g) = \infty \).

Proof. Suppose to the contrary that \( \text{AVR}(g) = c > 0 \). Let \( g(\tau) = \Phi^*_\tau g \) be the canonical form of \((M, g, f)\), where \( \Phi^*_\tau \) is the 1-parameter family of diffeomorphisms generated by \( \nabla f \). Then, by the Bishop-Gromov comparison theorem and the self-similarity of \( g(\tau) \), we have

\[
\text{Vol} B_{g(\tau)}(x, r) \geq cr^n
\]

for all \((x, \tau) \in M \times [0, \infty)\) and for all \( r > 0 \). Let \( \ell \) be the reduced distance based at some \((p, 0)\).

Let \( \tau_i \nearrow \infty \) and \( x_i \in M \) be such that \( \ell(x_i, \tau_i) \leq n/2 \). By Proposition 2.4 after passing to a subsequence, we have that

\[
(M, g_i(\tau_i), (x_i, 1), \ell_i) \to (M_\infty, g_\infty(\tau_i), (x_\infty, 1), \ell_\infty),
\]

in the sense as indicated in that proposition, where \( g_i(\tau_i) = \tau_i^{-1} g(\tau_i) \) and \( \ell_i(\cdot, \tau_i) = \ell(\cdot, \tau_i) \). Here \((M_\infty, g_\infty, \ell_\infty)\) is the canonical form of a non-flat Ricci shrinker. Obviously, (5.3) holds for every \( g_i(\tau_i) \), and consequently it also holds for the asymptotic shrinker. It follows that \( \text{AVR}(g_\infty) \geq c > 0 \), which is a contradiction to the fact that non-flat shrinking gradient Ricci solitons with nonnegative Ricci curvature must have zero AVR (see [CaN09 Corollary 1.1]). Hence \( \text{AVR}(g) = 0 \).

To prove \( \text{ASCR}(g) = \infty \), we can argue as [DZ18 Proposition 2.4]. Suppose that \( \text{ASCR}(g) < \infty \). Write \( \rho(x) := \text{dist}(p, x) \) for some base point \( p \in M \). Then there are constants \( A > 1 \) and \( r_0 > 0 \) such that

\[
R(x) \leq \frac{A}{\rho^2(x)} \quad \text{whenever} \quad \rho(x) \geq r_0.
\]
Without loss of generality, we assume that the constant $C$ in (2.1) is no less than 1. For large $r > r_0$, pick $y \in \partial B(p, 2\sqrt{ACr})$. Then we have

$$B(y, r) \subset B(p, 3\sqrt{ACr}) \setminus B(p, \sqrt{ACr}).$$

For any $x \in B(y, r)$, we have

$$|Rm|(x) \leq CR(x) \leq \frac{AC}{\rho^2(x)} \leq \frac{AC}{ACr^2} = r^{-2}.$$

Since $g$ is $\kappa$-noncollapsed, we have

$$\frac{Vol[B(p, 3\sqrt{ACr})]}{(3\sqrt{ACr})^n} \geq \frac{Vol[B(y, r)]}{(3\sqrt{ACr})^n} \geq \frac{\kappa}{(3\sqrt{AC})^n}.$$  

Taking $r \to \infty$, we obtain $AVR(g) > 0$, which is a contradiction to what we just proved. Hence $ASCR(g) = \infty$.

The above result extends [CLN06, Theorem 9.44], while the latter proved $ASCR = \infty$ assuming $\sec \geq 0, \text{Ric} > 0$, and that $R$ attains its maximum somewhere. Here, apart from $|Rm| \leq CR$, we assume $\kappa$-noncollapsing, a condition which holds for singularity models. We also generalized the previous results in [CDM20 Theorem 1.10] to higher dimensions assuming nonnegative Ricci curvature from a different approach.

Finally, we remark that P.-Y. Chan [Cha20] recently proved that for any 4-dimensional steady gradient Ricci soliton which is a singularity model, we must have that $|Rm| \leq CR$ for some constant $C$ (without assuming a curvature decaying condition as in [Cha19]).

Appendices

A Interchangeability of integration and differentiation

In the proofs of Proposition 3.4 and Theorem 4.4 it must be applied that the integration and the differentiation are interchangeable. In Bamler’s [Bam20a] original proof, this is obviously valid since the Ricci flow in question is on a closed manifold. However, under our assumption, we cannot make the same conclusion so easily. Nonetheless, given the curvature boundedness of the Ricci flow, the estimates for the heat kernel have made it possible for us to prove the interchangeability of the integration and the differentiation. We shall only prove the following theorem, and all other similar arguments required in the proofs of Proposition 3.4 and Theorem 4.4 can be proved with similar (and indeed, much easier) method.

**Theorem A.1.** Let $(M^n, g(t))_{t \in I}$ be a complete Ricci flow with bounded curvature within each time interval compact in $I$. Then, for all $s, t \in I$ with $s < t$, we have

$$\Box_{x,t} \mathcal{N}^*_s(x,t) = - \int_M \Box_{x,t} \left( K(x,t \mid s, 0) \log K(x,t \mid s, s) \right) dg_s - \frac{n}{2t},$$

where, as before, we have defined $\mathcal{N}^*_s(x,t) := \mathcal{N}_{(x,t)}(t-s)$. 

By a parabolic scaling, we shall assume that $s = 0$ and $t = 1$. We shall henceforth fix a point $x \in M$, a small convex open neighborhood $x \in U \subset B_{g_0}(x, 1)$, and a small positive number $\varepsilon \ll 1$, where $g_0 := g(0)$ will be used as the reference metric throughout the proof. Indeed, we need only to show that

$$\begin{align*}
y \to \sup_{z \in U} \left| \nabla_z \left( K(z, 1 \mid y, 0) \log K(z, 1 \mid y, 0) \right) \right|,
y \to \sup_{z \in U} \left| \nabla_z^2 \left( K(z, 1 \mid y, 0) \log K(z, 1 \mid y, 0) \right) \right|,
y \to \sup_{t \in (1-\varepsilon,1+\varepsilon)} \left| \partial_t \left( K(x, t \mid y, 0) \log K(x, t \mid y, 0) \right) \right|,
\end{align*}$$

are all dominated by integrable functions, and the conclusion of the theorem follows from Lebesgue’s dominated convergence theorem.

**Lemma A.2.** There are constants $C$, depending on the curvature bound on $M \times [0, 1]$ and $\text{Vol}_{g_0} B_0(x, 1)$, such that

$$\sup_{z \in U} \left| \nabla_z \left( K(z, 1 \mid y, 0) \log K(z, 1 \mid y, 0) \right) \right| \leq C \exp \left( -C^{-1} \text{dist}_0^2(x, y) \right),$$

for all $y \in M$.

**Proof.** Let us first of all fix an arbitrary $y \in M$. By Lemma 26.17 in [CCGG+10], we have

$$K(\cdot, \cdot \mid y, 0) \leq \frac{C}{\text{Vol}_{g_0} B_0 \left( y, \frac{\sqrt{t}}{2} \right)} \leq \frac{C}{t^{\frac{n}{2}} \text{Vol}_{g_0} B_0(y, 1)} \leq \frac{C \exp(\text{dist}_0(x, y))}{t^{\frac{n}{2}}}$$

on $M \times (0, 1]$, where we have applied the Bishop-Gromov comparison theorem. Then, applying [BCP10, ZQ06] (c.f. Lemma 2.4 in [ZY1]) on $M \times [\frac{1}{2}, 1]$, we have

$$\left| \nabla_z K(z, t \mid y, 0) \right| \leq \frac{1}{\sqrt{t - \frac{1}{4}}} \cdot K(z, t \mid y, 0) \cdot \sqrt{\log \left( \frac{C \exp(\text{dist}_0(x, y))}{K(z, t \mid y, 0)} \right)} \left( A.3 \right)$$

for all $(z, t) \in M \times [\frac{1}{2}, 1]$. If we restrict $(z, t) \in B_0(x, 2) \times [\frac{1}{2}, 1]$, then, by Theorem 26.31 in [CCGG+10], we have

$$K(z, t \mid y, 0) \geq C^{-1} e^{-C \text{dist}_0^2(y, z)} \geq C^{-1} e^{-C \text{dist}_0^2(x, y)}, \left( A.4 \right)$$

and (A.3) becomes

$$\left| \nabla_z K(z, t \mid y, 0) \right| \leq C K(z, t \mid y, 0) \left( C + C \text{dist}_0^2(x, y) \right) \left( A.5 \right)$$

for all $(z, t) \in B_0(x, 2) \times [\frac{1}{2}, 1]$. This further implies that, if $z \in U \subset B_0(x, 1)$, then we have

$$\left| \nabla_z \left( K(z, 1 \mid y, 0) \log K(z, 1 \mid y, 0) \right) \right| \left( A.6 \right)$$

$$\leq \left| \nabla_z K(z, 1 \mid y, 0) \right| (1 + \left| \log K(z, 1 \mid y, 0) \right|)$$

$$\leq C \left( C + C \text{dist}_0^2(x, y) \right) \cdot \left( K(z, 1 \mid y, 0) + K^2(z, 1 \mid y, 0) + K^2(z, 1 \mid y, 0) \right),$$

and the conclusion of the theorem follows from Lebesgue’s dominated convergence theorem.
where we have applied the fact the fact that $|u \log u| \leq u^{\frac{1}{2}} + u^2$ for all $u > 0$. Finally, by Corollary 2.26 in [CCGG+10], for all $(z, t) \in B_0(x, 2) \times [\frac{1}{2}, 1]$, it holds that

$$K(z, t \mid y, 0) \leq Ce^{-C^{-1}\text{dist}_{0}^2(z, y)},$$

(A.7)

where $C$ depends on the curvature bound and $\inf_{z \in B_0(x, 2)} \operatorname{Vol}_{g_0} B_0(z, 1)$, which in turn depends only on the curvature bound and $\operatorname{Vol}_{g_0} B_0(x, 1)$ by the Bishop-Gromov theorem. The lemma then follows from combining (A.6) and (A.7).

\begin{lemma}
There are constants $C$, depending on the curvature bound on $M \times [0, 1]$ and $\operatorname{Vol}_{g_0} B_0(x, 1)$, such that

$$\sup_{y \in \mathbb{U}} |\nabla_{z} K(z, 1 \mid y, 0) \log K(z, t \mid y, 0)| \leq C \exp \left(-C^{-1}\text{dist}_{0}^2(x, y)\right),$$

for all $y \in M$.

\end{lemma}

\begin{proof}
Let us fix an arbitrary $y \in M$. Combining (A.3) and (A.7), we have

$$|\nabla_{z} K(z, t \mid y, 0)| \leq C \exp \left(-C^{-1}\text{dist}_{0}^2(x, y)\right) \quad \text{for all} \quad (z, t) \in B_0(x, 2) \times [\frac{1}{2}, 1].$$

(A.9)

We shall then consider the function $u(z, t) := K(z, t \mid y, 0)$, which is a solution to the heat equation. Indeed, since

$$\left(\frac{\partial}{\partial t} - \Delta_L\right) \nabla^2 u = 0,$$

where $\Delta_L$ is the Lichnerowicz Laplacian operator, we have

$$\Box |\nabla^2 u|^2 = -2|\nabla^3 u| + 4 \text{Rm}(\nabla^2 u, \nabla^2 u)$$

$$\leq -2|\nabla^3 u|^2 + C|\nabla^2 u|^2,$$

and $C$ depends on the curvature bound. On the other hand, we have

$$\Box |\nabla u|^2 = -2|\nabla^2 u|^2.$$

We may then apply Shi’s gradient estimate (c.f. Theorem 14.10 in [CCGG+10], using the cut-off function constructed by Lemma 14.3 therein) on $B_0(x, 2) \times [\frac{1}{2}, 1]$ and obtain that

$$|\nabla_{z}^2 K(z, 1 \mid y, 0)| \leq C \exp \left(-C^{-1}\text{dist}_{0}^2(x, y)\right) \quad \text{for all} \quad z \in B_0(x, 1).$$

(A.10)

Since, for all $z \in B_0(x, 1)$, we have

$$|\nabla_{z}^2 (K(z, 1 \mid y, 0) \log K(z, 1 \mid y, 0))| \leq |\nabla^2 K(z, 1 \mid y, 0)|||\log K(z, 1 \mid y, 0)| + 1 + \frac{|\nabla_{z} K(z, 1 \mid y, 0)|^2}{K(z, 1 \mid y, 0)}$$

$$\leq |\nabla^2 K(z, 1 \mid y, 0)|||\log K(z, 1 \mid y, 0)| + 1 + C \exp \left(-C^{-1}\text{dist}_{0}^2(x, y)\right),$$

(A.11)
where in the last inequality, we have applied a consequence of \( \text{(A.4)} \) and \( \text{(A.7)} \):

\[
\frac{|\nabla_z K(z, 1 | y, 0)|^2}{K(z, 1 | y, 0)} = \left( \frac{|\nabla_z K(z, 1 | y, 0)|}{K(z, 1 | y, 0)} \right)^2 \cdot K(z, 1 | y, 0) \\
\leq C(1 + \text{dist}_0^2(x, y))^2 \exp \left(-C^{-1} \text{dist}_0^2(x, y) \right) \\
\leq C \exp \left(-C^{-1} \text{dist}_0^2(x, y) \right),
\]

for all \( z \in B_0(x, 1) \), the lemma then follows from \( \text{(A.11)} \) together with \( \text{(A.4)}, \text{(A.7)}, \text{and (A.10)}. \)

**Lemma A.12.** There are constants \( C \), depending on the curvature bound on \( M \times [0, 1] \) and \( \text{Vol}_g B_0(x, 1) \), such that

\[
\sup_{t \in (1-x, 1+x)} | \partial_t (K(x, t | y, 0) \log K(x, t | y, 0)) | \leq C \exp \left(-C^{-1} \text{dist}_0^2(x, y) \right),
\]

for all \( y \in M \), where \( \varepsilon \) is a small positive constant, which could be taken to be, say \( 100^{-1} \).

**Proof.** Since

\[
\partial_t (K(x, t | y, 0) \log K(x, t | y, 0)) = (\log K(x, t | y, 0) + 1) \cdot \Delta_x K(x, t | y, 0),
\]

the proof is not essentially different from the above lemma.

**Proof of Theorem A.1.** Let \( v \in T_x M \) be a unit vector with respect to \( g(1) \), and let \( \gamma(s) \) be a unit speed \( g(1) \)-geodesic emanating from \( x \) with \( \gamma'(0) = v \). Letting \( u_z(y) := K(z, 1 | y, 0) \) \( \log K(z, 1 | y, 0) \), we have

\[
\left| \frac{1}{s} (u_{\gamma(s)}(y) - u_x(y)) \right| \leq \frac{1}{s} \int_0^s \sup_{z \in B_0(x, 1)} |\nabla_z u_z(y)| \, ds \leq C \exp \left(-C^{-1} \text{dist}_0^2(x, y) \right),
\]

for all \( s \) small enough, where we have applied \( \text{(A.2)} \), and the right-hand-side is obviously integrable in \( y \) because of the curvature boundedness assumption. Hence, by Lebesgue’s dominated convergence theorem, we have

\[
\nabla_v \mathcal{N}_0^*(x, 1) = \lim_{\varepsilon \to 0^+} \frac{1}{s} \left( \mathcal{N}_0^*(\gamma(s), 1) - \mathcal{N}_0^*(x, 1) \right) \\
= \lim_{\varepsilon \to 0^+} \int_M \frac{1}{s} (u_{\gamma(s)}(y) - u_x(y)) \, dg_0(y) \\
= \int_M \lim_{\varepsilon \to 0^+} \frac{1}{s} (u_{\gamma(s)}(y) - u_x(y)) \, dg_0(y) \\
= \int_M \nabla_v (K(x, 1 | y, 0) \log K(x, t | y, 0)) \, dg_0(y).
\]

We have proved that the first spatial derivative is interchangeable with the integration. In like manner, with the help of Lemma A.8 and Lemma A.12, one may also verify that both the second spatial derivative and the first time derivative are interchangeable with the integration. Once this is done, the theorem follows from Bamler’s original proof of Theorem 5.9 in \( \text{[Bam20a]} \).
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).

[BCP10] Mihai Bailesteanu, Xiaodong Cao, and Artem Pulemotov. "Gradient estimates for the heat equation under the Ricci flow." Journal of Functional Analysis 258.10 (2010): 3517-3542.

[Br09] S. Brendle. *A generalization of Hamilton’s differential Harnack inequality for the Ricci flow*. J. Differential Geom. 2009, 82(1):207-227.

[CZ11] Xiaodong Cao and Qi S Zhang, *The conjugate heat equation and ancient solutions of the Ricci flow*. Advances in Mathematics, 2011, 228(5): 2891-2919.

[CHI04] Huai-Dong Cao, Richard S. Hamilton, and Tom Ilmanen. *Gaussian densities and stability for some Ricci solitons*, arXiv preprint math/0404165, 2004.

[CaN09] José Carrillo and Lei Ni, *Sharp logarithmic Sobolev inequalities on gradient solitons and applications*, Comm. Anal. Geom. 17 (2009), 721–753.

[Cha19] Pak-Yeung Chan, *Curvature estimates for steady gradient Ricci solitons*, Trans. Amer. Math. Soc. 372 (2019), no. 12, 8985–9008.

[Cha20] ________, *in preparation*.

[CMZ21] Pak-Yeung Chan, Zilu Ma, and Yongjia Zhang, *Ancient Ricci flows with asymptotic solitons*. arXiv: 2106.06904.

[CTY] Albert Chau, Luen-Fai Tam, and Chengjie Yu, *Pseudolocality for the Ricci flow and applications*, arXiv preprint math/0701153, 2007.

[CBl] Bing-Long Chen. *Strong uniqueness of the Ricci flow*, Journal of Differential Geometry, 82(2):363-382, 2009.

[CDM20] Bennett Chow, Yuxing Deng, and Zilu Ma. *On four-dimensional steady gradient Ricci solitons that dimension reduce*, https://arxiv.org/abs/arXiv:2009.11456 (2020).

[CCGG+07] 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 I. Geometric Aspects*, Mathematical Surveys and Monographs, vol. 135, AMS, Providence, RI, 2008.

[CCGG+08] ________, *The Ricci flow: techniques and applications. Part II. Analytic Aspects*, Mathematical Surveys and Monographs, vol. 144, AMS, Providence, RI, 2008.

[CCGG+10] ________, *The Ricci flow: techniques and applications. Part III: Geometric-Analytic Aspects*. Mathematical Surveys and Monographs, 163, AMS, Providence, RI, 2010.
REFERENCES

[CLN06] Bennett Chow, Peng Lu, and Lei Ni, Hamilton’s Ricci flow, Lectures in Contemporary Mathematics, 3. Science Press and Graduate Studies in Mathematics, 77, American Mathematical Society (co-publication), 2006.

[DZ18] Yuxing Deng and Xiaohua Zhu, Asymptotic behavior of positively curved steady Ricci solitons, Trans. Amer. Math. Soc. 370 (2018), 2855-2877.

[Ha1] Richard S. Hamilton, Harnack estimate for the mean curvature flow, J. Differential Geom. Volume 41, Number 1 (1995), 215-226.

[Ha2] ________, A Compactness Property for Solutions of the Ricci Flow, American Journal of Mathematics, 1995, 117(3): 545-572

[HN14] Hans-Joachim Hein and Aaron Naber. New Logarithmic Sobolev Inequalities and an $\varepsilon$-Regularity Theorem for the Ricci Flow, Communications on Pure and Applied Mathematics 67, no. 9 (2014), 1543-1561.

[KP82] Leon Karp and Peter Li, The heat equation on complete riemannian manifolds. Unpublished notes, 1982.

[KL08] Bruce Kleiner and John Lott, Notes on Perelman’s papers, Geom. Topol.12 (2008), no. 5, 2587–2855.

[Ko07] Brett Kotschwar, Hamilton’s gradient estimate for the heat kernel on complete manifolds, Proceedings of the American Mathematical Society 135, no. 9 (2007), 3013-3019.

[N10] Aaron Naber, Noncompact shrinking four solitons with nonnegative curvature, Journal fur die reine und angewandte Mathematik (Crelles Journal), 2010, 645: 125–153.

[Ni05] Lei Ni, Ancient solutions to Kähler-Ricci flow, Math. Res. Lett. 12 (2005), 633-654.

[Per02] Grisha Perelman, The entropy formula for the Ricci flow and its geometric applications, arXiv:math.DG/0211159

[Per03] ________, Ricci flow with surgery on three-manifolds, arXiv:math/0303109

[Ru] Walter Rudin, Real and complex analysis. Tata McGraw-hill education, 2006.

[Sh] Wan-Xiong Shi, Deforming the metric on complete Riemannian manifolds. J. Differential Geom. 1989, 30(1):223-301.

[Xu17] Guoyi Xu, An equation linking $W$-entropy with reduced volume , Journal fur die reine und angewandte Mathematik, Volume 2017: Issue 727.

[Y] Rugang Ye, On the $l$-Function and the Reduced Volume of Perelman I, Trans. Amer. Math. Soc. 360 (2008), 507–531

[ZQ06] Qi S Zhang, Some gradient estimates for the heat equation on domains and for an equation by Perelman. International Mathematics Research Notices 2006 (2006), 1-39.

[ZQ10] ________, Sobolev inequalities, heat kernels under Ricci flow, and the Poincaré conjecture. CRC Press, 2010.
[ZY1] Yongjia Zhang, *Entropy, noncollapsing, and a gap theorem for ancient solutions to the Ricci flow*. Communications in Analysis and Geometry, to appear.

[ZY2] , *On the equivalence between noncollapsing and bounded entropy for ancient solutions to the Ricci flow*. Journal fur die reine und angewandte Mathematik, Volume 2020, Issue 762, Pages 35.

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

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