



HEAT KERNEL ON RICCI SHRINKERS (II)*

Yu LI (李宇)[†] Bing WANG (王兵)*Institute of Geometry and Physics, University of Science and Technology of China,
Hefei 230026, China;**Hefei National Laboratory, Hefei 230088, China**E-mail: yuli21@ustc.edu.cn; topspin@ustc.edu.cn*

Abstract This paper is the sequel to our study of heat kernel on Ricci shrinkers [29]. In this paper, we improve many estimates in [29] and extend the recent progress of Bamler [2]. In particular, we drop the compactness and curvature boundedness assumptions and show that the theory of \mathbb{F} -convergence holds naturally on any Ricci flows induced by Ricci shrinkers.

Key words Ricci flow; Ricci shrinker; heat kernel

2020 MR Subject Classification 53C44; 35K08

1 Introduction

A Ricci shrinker (M^n, g, f) is a complete Riemannian manifold (M^n, g) coupled with a smooth function f satisfying

$$\text{Rc} + \text{Hess } f = \frac{1}{2}g, \quad (1.1)$$

where the potential function f is normalized so that

$$R + |\nabla f|^2 = f. \quad (1.2)$$

The study of shrinkers is an essential component of analyzing the singularity formation of solutions to the Ricci flow. For a Ricci flow with type-I curvature bound, it was proven by Enders-Müller-Topping [20] that any proper blow-up sequence converges smoothly to a nontrivial Ricci shrinker. For general compact Ricci flows, it was proven by Bamler [4] that the finite-time singularities are modeled on Ricci shrinkers containing a singular set by using the theory of \mathbb{F} -convergence developed in [2–4].

In dimension 2 or 3, all Ricci shrinkers are completely classified (cf. [8, 22, 32, 34], etc.). We know that $\mathbb{R}^2, S^2, \mathbb{R}^3, S^3, S^2 \times \mathbb{R}$ and their quotients form the complete list. In particular, all low-dimensional Ricci shrinkers have bounded and nonnegative sectional curvature.

In higher dimensions, the complete classification of Ricci shrinkers seems out of reach. Subject to an additional curvature assumption, some partial classifications are also known (cf.

*Received February 7, 2023; revised May 29, 2024. Li's research was supported by the YSBR-001, the NSFC (12201597) and research funds from USTC (University of Science and Technology of China) and CAS (Chinese Academy of Sciences). Wang's research was supported by the YSBR-001, the NSFC (11971452, 12026251) and a research fund from USTC.

[†]Corresponding author

[18, 26, 27, 30, 32, 33]). In general, it is still unclear if there exists any Ricci shrinker with unbounded sectional curvature.

On the one hand, Ricci shrinkers can be regarded as critical metrics which generalize the classical positive Einstein manifolds. On the other hand, for any Ricci shrinker, there exists an associated self-similar solution to the Ricci flow (cf. Section 2). As a special class of Ricci flows, Ricci shrinkers have many known important properties of compact Ricci flows. In [29], many fundamental analytic tools, including the maximum principle, optimal log-Sobolev constant estimate, the no-local-collapsing theorems, etc., were established for Ricci flows associated with Ricci shrinkers. Many heat kernel estimates including the differential Harnack inequality and the pseudolocality theorem were also shown in [29].

In this paper, we continue to focus on Ricci flows associated with Ricci shrinkers without any curvature assumption. Based on the techniques and results in [29] and [2], we obtain further results, including a Gaussian bound on the heat kernel, no-local-collapsing and non-expanding estimates, an ϵ -regularity theorem, etc.. All of these results are stronger than their counterparts in [29]. It is important to notice that we have no assumption of curvature at all. If we assume bounded curvature on non-compact manifolds, then many results are already known (cf. [5, 10]).

The pointed Nash entropy (cf. Definition 3.18) plays an important role in [2], which first appeared in [35, Section 5] and was systematically studied in [24]. In [29], we use Perelman's entropy μ (see (2.1)) to characterize the optimal log-Sobolev constant and the local non-collapsing. The pointed Nash entropy, which is always bounded below by μ , has the advantage of being local in the spacetime of Ricci flows. In [24], it was proven that the Nash entropy is Lipschitz. Moreover, the oscillation of the Nash entropy in the spacetime was established in [2]. We generalize the Nash entropy and its fundamental estimates to the Ricci flows associated with Ricci shrinkers; see Theorem 3.23 and Corollary 4.19.

Theorem 1.1 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Then for any $s < t < 1$, the Nash entropy $\mathcal{N}_s^*(x, t) := \mathcal{N}_{(x,t)}(t-s)$ is smooth and satisfies the following estimates on $M \times (s, 1)$:

$$|\nabla \mathcal{N}_s^*| \leq \sqrt{\frac{n}{2(t-s)}} \quad \text{and} \quad -\frac{n}{2(t-s)} \leq \square \mathcal{N}_s^* \leq 0. \quad (1.3)$$

The proof of (1.3) is based on an integral estimate of the heat kernel (cf. Theorem 3.16), which was initially obtained in [2] for compact Ricci flows. A key application of Theorem 1.1 is to estimate the local oscillation of the Nash entropy (cf. Corollary 3.25). Using the Nash entropy properties and the heat kernel estimates, we obtain the improved no-local-collapsing and non-expanding result (cf. Theorem 4.2 and Theorem 4.7).

Theorem 1.2 (No-local-collapsing and non-expanding) Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. For any $x \in M$ and $t < 1$,

$$|B_t(x, r)|_t \leq C(n) \exp(\mathcal{N}_{(x,t)}(r^2)) r^n,$$

and if $R \leq r^{-2}$ on $B_t(x, r)$, then

$$|B_t(x, r)|_t \geq c(n) \exp(\mathcal{N}_{(x,t)}(r^2)) r^n.$$

Note that $\mathcal{N}_{(x,t)}(r^2) \leq 0$ (cf. Corollary 3.22), it is clear that Theorem 1.2 provides a uniform volume ratio upper bound, independent of base point and radius. This clearly improves the

known volume upper bounds (cf. [9, 23, 25]). On the other hand, as $\mu \leq \mathcal{N}_{(x,t)}(r^2)$, the non-collapsing estimate in Theorem 1.2 also improves the one in [29].

An important concept introduced in [2] is the H -center (cf. Definition 3.11). Roughly speaking, an H -center is a point around which the conjugate heat kernel is concentrated (cf. Proposition 3.13). In addition, for any two conjugate heat kernels, the W_1 -Wasserstein distance between them can be roughly measured by the distance between two H -centers. We prove the existence of an H_n -center, where $H_n = (n - 1)\pi^2/2 + 4$, for any conjugate heat kernel, by generalizing the monotonicity of the variance obtained in [2] to our setting (cf. Proposition 3.10, Proposition 3.12). By using these concepts and related techniques, we have the following heat kernel estimates (cf. Theorem 4.9, Theorem 4.15, Theorem 4.16).

Theorem 1.3 (Heat kernel estimates) Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker satisfying $\mu \geq -A$. Then the following properties hold:

(i) There exists a constant $C = C(n, A, \delta) > 1$ such that

$$\frac{C^{-1}}{(t - s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(x, y)}{C^{-1}(t - s)}\right) \leq H(x, t, y, s) \leq \frac{C}{(t - s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(x, y)}{C(t - s)}\right) \tag{1.4}$$

for any $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_t(p, x) \leq \delta^{-1}$. Here, the point p is a minimum point of f , regarded as Ricci shrinker’s base point.

(ii) For any $\epsilon > 0$, there exists a constant $C = C(n, \epsilon) > 0$ such that

$$H(x, t, y, s) \leq \frac{C \exp(-\mathcal{N}_{(x,t)}(t - s))}{(t - s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(z, y)}{(4 + \epsilon)(t - s)}\right) \tag{1.5}$$

for any $s < t < 1$ and any H_n -center (z, s) of (x, t) .

The Gaussian estimate (1.5) was previously proven in [2] for compact Ricci flows, with $4 + \epsilon$ replaced by $8 + \epsilon$. Our proof uses an iteration argument by showing that if (1.5) fails, one can find a new spacetime point (x', t') with an H_n -center (z', s) such that $H(x', t', y, s)$ has a worse bound than (1.5). Eventually, we will arrive at a contradiction if t' is sufficiently close to s . The proof in our case is more involved since we do not have a global heat kernel bound as (1.5) when t is close to s , which is always available for compact Ricci flows. Therefore, in the iteration process, we must carefully choose the sequence of spacetime points so that they all fall into a compact set. Then the contradiction comes from the local heat kernel estimate (cf. Corollary 4.12), since the scalar curvature is locally bounded.

Once we have the estimate (1.5), the upper bound in (1.4) follows since the distance between (x, s) and (z, s) can be well-controlled. Moreover, the lower bound in (1.4) is already contained in [29] in a different guise. We also obtain the gradient estimate of the heat kernel; see Theorem 4.6.

By the monotonicity of the W_1 -Wasserstein distance between two conjugate heat kernels (cf. Proposition 3.7), it is natural to consider new P^* -parabolic neighborhoods in the spacetime of the Ricci flow, as pointed out in [2] (cf. Definition 5.1, (5.1), (5.2)). Comparing the P^* -parabolic neighborhoods with the conventional ones, we have the following result (cf. Proposition 5.7, Proposition 5.9, Proposition 5.10, Proposition 5.13).

Theorem 1.4 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker satisfying $\mu \geq -A$. Then the following properties hold:

(i) Given $\delta \in (0, 1)$, $t_0 \in (-\infty, 1)$, $T^\pm \geq 0$ and $S \geq 0$, there exists a constant $C = C(n, A, \delta) > 1$ such that

$$P^*(p, t_0; S, -T^-, T^+) \subset Q(p, t_0; \sqrt{2}S + C, -T^-, T^+) \subset P^*(p, t_0; \sqrt{2}S + 2C, -T^-, T^+),$$

provided that $t_0 - T^- \geq -\delta^{-1}$.

(ii) There exists a constant $\rho = \rho(n, A) \in (0, 1)$ satisfying the following property: given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $R \leq r^{-2}$ on $P(x_0, t_0; r, -(\rho r)^2, (\rho r)^2)$. Then

$$P(x_0, t_0; \rho r) \subset P^*(x_0, t_0; r, -(\rho r)^2, (\rho r)^2) \quad \text{and} \quad P^*(x_0, t_0; \rho r) \subset P(x_0, t_0; r, -(\rho r)^2, (\rho r)^2).$$

The proof of Theorem 1.4 involves the distance distortion estimates globally with respect to p and locally under the scalar curvature control. Moreover, one needs to locate the H_n -center of (p, t_0) or (x_0, t_0) . Notice that, if $t_0 + T^+ < 1$, Theorem 1.4 implies that any $P^*(p, t_0; S, -T^-, T^+)$ is precompact, i.e., its closure is compact. By using the estimates of the Nash entropy and P^* -neighborhoods, one has the following ϵ -regularity theorem (cf. Theorem 5.15), which was proven in [2] for compact Ricci flows. Here, r_{Rm} is the spacetime curvature radius, whose definition can be found in Definition 5.14.

Theorem 1.5 (ϵ -regularity theorem) There exists a small constant $\epsilon = \epsilon(n) > 0$ satisfying the following property.

Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Given $(x, t) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $\mathcal{N}_{(x,t)}(r^2) \geq -\epsilon$, then $r_{\text{Rm}}(x, t) \geq \epsilon r$.

Based on the results and techniques generalized (or slightly improved) from [2], we can generalize the theory about metric flows and \mathbb{F} -convergence in [3] and [4] from compact Ricci flows to the setting of Ricci flows associated with or induced by Ricci shrinkers (cf. Definition 2.2). In particular, a pointed Ricci flow induced by a Ricci shrinker can be regarded as a metric flow pair in the sense of [3, Definition 5.1]. Therefore, any sequence of pointed Ricci shrinkers induced by Ricci shrinkers with $\mu \geq -A$, by taking a subsequence, will \mathbb{F} -converge to a limit metric flow admitting concrete structure theorems (cf. Theorem 6.10, Theorem 6.12). As an application of the theory of \mathbb{F} -convergence, we have the following two-sided pseudolocality theorem. Notice that the forward pseudolocality theorem was proven in [29, Theorem 24]. Thus, to obtain a two-sided pseudolocality, it suffices to obtain a backward pseudolocality, which is proven in Theorem 6.21.

Theorem 1.6 (Two-sided pseudolocality theorem) For any $\alpha > 0$, there is an $\epsilon(n, \alpha) > 0$ such that the following holds.

Let $(M^n, g(t))_{t < 1}$ be a Ricci flow associated with a Ricci shrinker. Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, if

$$|B_{t_0}(x_0, r)| \geq \alpha r^n, \quad |\text{Rm}| \leq (\alpha r)^{-2} \quad \text{on} \quad B_{t_0}(x_0, r),$$

then

$$|\text{Rm}| \leq (\epsilon r)^{-2} \quad \text{on} \quad P(x_0, t_0; (1 - \alpha)r, -(\epsilon r)^2, (\epsilon r)^2).$$

Another application of the \mathbb{F} -converge is the following integral estimate of curvature, which originates from the estimate of Cheeger-Naber [12]. For more details, see Theorem 6.23 and Corollary 6.24.

Theorem 1.7 Let (M^n, g, f, p) be a Ricci shrinker in $\mathcal{M}(A)$. Then

$$\int_{d(p, \cdot) \leq r} |\text{Rm}|^{2-\epsilon} dV \leq \int_{d(p, \cdot) \leq r} r_{\text{Rm}}^{-4+2\epsilon} dV \leq Cr^{n+2\epsilon-2},$$

$$\int_{d(p, \cdot) \geq 1} \frac{|\text{Rm}|^{2-\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq \int_{d(p, \cdot) \geq 1} \frac{r_{\text{Rm}}^{-4+2\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq C$$

for any $\epsilon > 0$ and $r \geq 1$, where $r_{\text{Rm}}(\cdot) = r_{\text{Rm}}(\cdot, 0)$ and $C = C(n, A, \epsilon)$.

This rest of the paper is organized as follows. Section 2 discusses some properties of Ricci flows associated with Ricci shrinkers, including the existence of cutoff functions and maximum principles. In Section 3, we prove some estimates and properties regarding the variance, H -centers and the Nash entropy. Section 4 focuses on various estimates of the heat kernel. In Section 5, we prove the theorems about the parabolic neighborhoods and the ϵ -regularity theorem. In the last section, we generalize the theory of \mathbb{F} -convergence in our setting and prove some applications in Ricci shrinkers.

2 Preliminaries

For any Ricci shrinker (M^n, g, f) , the scalar curvature $R \geq 0$ from [14, Corollary 2.5] and $R > 0$ unless (M^n, g) is isometric to the Gaussian soliton (\mathbb{R}^n, g_E) , by the strong maximum principle.

With the normalization (1.2), the entropy is defined as

$$\mu = \mu(g) := \log \int \frac{e^{-f}}{(4\pi)^{n/2}} dV. \tag{2.1}$$

Notice that e^μ is uniformly comparable to the volume of the unit ball $B(p, 1)$ (cf. [25, Lemma 2.5]). It was proven in [29, Theorem 1] that μ is the optimal log-Sobolev constant for all scales. Following [28], we have

Definition 2.1 Let $\mathcal{M}(A)$ be the family of Ricci shrinkers (M^n, g, f) satisfying

$$\mu(g) \geq -A. \tag{2.2}$$

Recall that any Ricci shrinker (M^n, g, f) can be considered a self-similar solution to the Ricci flow. Let $\psi^t : M \rightarrow M$ be a family of diffeomorphisms generated by $\frac{1}{1-t}\nabla f$ and $\psi^0 = \text{id}$. In other words, we have

$$\frac{\partial}{\partial t} \psi^t(x) = \frac{1}{1-t} \nabla f(\psi^t(x)). \tag{2.3}$$

It is well known that the rescaled pull-back metric $g(t) := (1-t)(\psi^t)^*g$ satisfies the Ricci flow equation for any $-\infty < t < 1$,

$$\partial_t g = -2\text{Rc}_{g(t)} \quad \text{and} \quad g(0) = g. \tag{2.4}$$

Sometimes we encounter Ricci flow obtained from the above Ricci flow through time-shifting and rescaling. We emphasize the possible extra time-shifting and rescaling by the following definition.

Definition 2.2 For any Ricci shrinker, the Ricci flow defined in (2.4) is called the associated Ricci flow. Any Ricci flow obtained from the associated Ricci flow via time-shifting and rescaling is called the Ricci flow induced by a Ricci shrinker.

Clearly, a Ricci flow associated with a Ricci shrinker must be a Ricci flow induced by a Ricci shrinker, but the reverse is generally not true. In this article, if not mentioned explicitly, the associated Ricci flow is the default.

Next, we recall the function $F(x, t) := \bar{\tau}f(x, t)$, where $\bar{\tau} := 1 - t$ and $f(x, t) := (\psi^t)^*f$, satisfies the following identities (see [29, Section 2] for proofs):

$$\partial_t f = |\nabla f|^2, \tag{2.5}$$

$$\partial_t F = -\bar{\tau}R, \tag{2.6}$$

$$\bar{\tau}R + \Delta F = \frac{n}{2}, \tag{2.7}$$

$$\bar{\tau}^2 R + |\nabla F|^2 = F, \tag{2.8}$$

$$\square F = -\frac{n}{2}. \tag{2.9}$$

Here, we define $\square := \partial_t - \Delta_t$ and drop the subscript $g(t)$ or t if there is no confusion. Based on these identities, we have the following estimates of F .

Lemma 2.3 (Lemma 1 of [29]) There exists a point $p \in M$ where F attains its infimum and F satisfies the quadratic growth estimate

$$\frac{1}{4} (d_t(x, p) - 5n\bar{\tau} - 4)_+^2 \leq F(x, t) \leq \frac{1}{4} (d_t(x, p) + \sqrt{2n\bar{\tau}})^2 \tag{2.10}$$

for all $x \in M$ and $t < 1$, where $a_+ := \max\{0, a\}$.

Thanks to (2.10), $F(x, t)$ grows like $d_t^2(x, p)/4$, and hence one can obtain a family of cutoff functions by composing F with a cutoff function on \mathbb{R} . More precisely, we fix a function $\eta \in C^\infty([0, \infty))$ such that $0 \leq \eta \leq 1$, $\eta = 1$ on $[0, 1]$ and $\eta = 0$ on $[2, \infty)$. Furthermore, $-C \leq \eta'/\eta^{\frac{1}{2}} \leq 0$ and $|\eta''| \leq C$ for a universal constant $C > 0$. For each $r \geq 1$, we define

$$\phi^r := \eta\left(\frac{F}{r}\right). \tag{2.11}$$

Then ϕ^r is a smooth function on $M \times (-\infty, 1)$. The following estimates of ϕ^r were proven in [29, Lemma 3]:

$$(\phi^r)^{-1}|\nabla\phi^r|^2 \leq Cr^{-1}, \tag{2.12}$$

$$|\phi^r_t| \leq C\bar{\tau}^{-1}, \tag{2.13}$$

$$|\Delta\phi^r| \leq C(\bar{\tau}^{-1} + r^{-1}), \tag{2.14}$$

$$|\square\phi^r| \leq Cr^{-1}, \tag{2.15}$$

where the constant C depends only on the dimension n .

For later applications, we recall the following volume estimate proven in [29, Lemma 2].

Lemma 2.4 There exists a constant $C = C(n) > 0$ such that for any Ricci shrinker (M^n, g, f) with $p \in M$ a minimum point of f ,

$$|B_t(p, r)|_t \leq Cr^n.$$

Next, we recall the following version of the maximum principle on Ricci shrinkers, which was proven in [29, Theorem 6] and will be frequently used.

Theorem 2.5 (Maximum principle on Ricci shrinkers I) Let $(M, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Given any closed interval $[a, b] \subset (-\infty, 1)$ and a function

u which satisfies $\square u \leq 0$ on $M \times [a, b]$, suppose that

$$\int_a^b \int_M u_+^2(x, t)e^{-2f(x, t)} dV_t(x)dt < \infty. \tag{2.16}$$

If $u(\cdot, a) \leq c$, then $u(\cdot, b) \leq c$.

We also need the following version of the maximum principle, which was proven in [19, Theorem 12.14] for Ricci flows with bounded curvature. Notice that if $X \equiv 0$, Theorem 2.6 follows from Theorem 2.5.

Theorem 2.6 (Maximum principle on Ricci shrinkers II) Let $(M, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Given any closed interval $[a, b] \subset (-\infty, 1)$ and a function u which satisfies

$$Lu := \square u - \langle \nabla u, X(t) \rangle \leq 0$$

on $M \times [a, b]$, suppose that $X(t)$ is a bounded vector field on $M \times [a, b]$ and

$$u(x, t) \leq Ke^{kf(x, t)} \tag{2.17}$$

on $M \times [a, b]$ for some constants $K > 0$ and $k < 1$. If $u(\cdot, a) \leq c$, then $u(\cdot, b) \leq c$.

Proof We first construct a barrier function

$$\phi(x, t) := Ke^{B(t-a)+(1-\epsilon)f(x, t)},$$

where $1 - \epsilon > k$ and B is a constant determined later.

Claim There exists a constant $B > 0$ such that

$$L\phi \geq \phi. \tag{2.18}$$

Proof of Claim By direct computations, we have

$$\begin{aligned} L\phi &= \phi \left(B + (1 - \epsilon)f_t - (1 - \epsilon)^2|\nabla f|^2 - (1 - \epsilon)\Delta f - (1 - \epsilon)\langle \nabla f, X \rangle \right) \\ &= \phi \left(B + \epsilon(1 - \epsilon)|\nabla f|^2 - \frac{n(1 - \epsilon)}{2\bar{r}} + (1 - \epsilon)R - (1 - \epsilon)\langle \nabla f, X \rangle \right) \\ &\geq \phi \left(B + \epsilon(1 - \epsilon)|\nabla f|^2 - C_1|\nabla f| - \frac{n(1 - \epsilon)}{2(1 - b)} \right), \end{aligned}$$

where we have used (2.6), (2.7) and the assumption that $|X| \leq C_1$. Therefore, (2.18) holds if we choose

$$B = \frac{C_1^2}{4\epsilon(1 - \epsilon)} + \frac{n(1 - \epsilon)}{2(1 - b)} + 1.$$

We assume $c = 0$ by considering $u - c$ instead of u . To complete the proof, we only need to verify that for any $\delta > 0$, $u \leq \delta\phi$ on $M \times [a, b]$. Otherwise, there exists $(x', t') \in M \times [a, b]$ such that $(u - \delta\phi)(x', t') > 0$. Due to the estimate (2.17) and our definition of ϕ , we know that $(u - \delta\phi)(x, t) \rightarrow -\infty$ as $d_t(x, p) \rightarrow +\infty$ uniformly in t , i.e., $u - \delta\phi < 0$ for $d_t(x, p)$ large enough independent of t . Moreover, $(u - \delta\phi)(x, a) < 0$ for all $x \in M$. Consequently, there exists $(x'', t'') \in M \times (a, t')$ such that $(u - \delta\phi)(x, t) \leq 0$ for all $(x, t) \in M \times [a, t'']$ and $(u - \delta\phi)(x'', t'') = 0$. At (x'', t'') , we compute

$$0 \leq L(u - \delta\phi) \leq -\delta\phi < 0,$$

which is a contradiction. In sum, our proof is complete. □

3 Variance, H -center and Nash Entropy

Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. It was proven in [29, Theorem 7] that there exists a positive heat kernel function $H(x, t, y, s)$ for $x, y \in M$ and $s < t < 1$. More precisely,

$$\square H(\cdot, \cdot, y, s) = 0, \quad \lim_{t \searrow s} H(\cdot, t, y, s) = \delta_y$$

and

$$\square^* H(x, t, \cdot, \cdot) = 0, \quad \lim_{s \nearrow t} H(x, t, \cdot, s) = \delta_x,$$

where $\square := \partial_t - \Delta$ and $\square^* := -\partial_t - \Delta + R$. Furthermore, the heat kernel H satisfies the semigroup property

$$H(x, t, y, s) = \int_M H(x, t, z, \rho) H(z, \rho, y, s) dV_\rho(z), \quad \forall x, y \in M, \rho \in (s, t) \subset (-\infty, 1), \quad (3.1)$$

and the following integral relationships:

$$\int_M H(x, t, y, s) dV_t(x) \leq 1, \quad (3.2)$$

$$\int_M H(x, t, y, s) dV_s(y) = 1. \quad (3.3)$$

For any $(x, t) \in M \times (-\infty, 1)$, we define the conjugate heat kernel measure $v_{x,t;s}$ by $dv_{x,t;s}(y) = H(x, t, y, s) dV_s(y)$. It follows immediately from (3.3) that $v_{x,t;s}$ is a probability measure on M . In particular, $v_{x,t;t} = \delta_x$.

With the help of the heat kernel, one can solve the (conjugate) heat solution from the given initial condition. More precisely, it follows from [29, Lemma 5, Lemma 6] that

Theorem 3.1 Suppose that $[a, b] \subset (-\infty, 1)$ and u_a is a bounded function on the time slice $(M, g(a))$. Then

$$u(x, t) := \int_M H(x, t, y, a) u_a(y) dV_a(y), \quad \forall t \in [a, b] \quad (3.4)$$

is the unique bounded heat solution with the initial value u_a . Similarly, suppose that w_b is an integrable function on the time slice $(M, g(b))$. Then

$$w(y, s) := \int_M H(x, b, y, s) w_b(x) dV_b(x) \quad (3.5)$$

is the unique conjugate heat solution with initial value w_b such that

$$\sup_{s \in [a, b]} \int |w| dV_s < \infty. \quad (3.6)$$

Next, we recall the following gradient estimate, which slightly strengthens [29, Corollary 1].

Lemma 3.2 Let u be a bounded heat solution on $M \times [a, b]$ such that $\sup_M |\nabla u(\cdot, a)| < \infty$. Then,

(i) we have

$$\sup_M |\nabla u(\cdot, b)| \leq \sup_M |\nabla u(\cdot, a)|. \quad (3.7)$$

(ii) assuming that w is a nonnegative conjugate heat solution on $M \times [a, b]$ such that

$$\sup_{t \in [a, b]} \int_M w \, dV_t < \infty, \tag{3.8}$$

then we have

$$2 \int_a^b \int_M |\text{Hess } u|^2 w \, dV_t dt = \int_M |\nabla u|^2 w \, dV \Big|_b^a < \infty. \tag{3.9}$$

Proof (i) From $\square u = 0$ and direct computation, we have

$$\square |\nabla u|^2 = -2|\text{Hess } u|^2 \leq 0. \tag{3.10}$$

Therefore, (3.7) follows from Theorem 2.5, provided that

$$\int_a^b \int_M |\nabla u|^2 e^{-2f} \, dV_t dt < \infty. \tag{3.11}$$

Now, we fix $r \gg 1$ and multiply both sides of $\square u = 0$ by $u(\phi^r)^2 e^{-2f}$. By integrating on $M \times [a, b]$, we obtain

$$\begin{aligned} & \frac{1}{2} \int_M u^2 (\phi^r)^2 e^{-2f} \, dV \Big|_a^b \\ & - \int_a^b \int_M u^2 \phi^r \phi_t^r e^{-2f} \, dV_t dt + \int_a^b \int_M u^2 (\phi^r)^2 f_t e^{-2f} \, dV_t dt + \frac{1}{2} \int_a^b \int_M u^2 (\phi^r)^2 R e^{-2f} \, dV_t dt \\ & = \int_a^b \int_M \{ -|\nabla(u\phi^r)|^2 + |\nabla\phi^r|^2 u^2 + \langle \nabla u^2, \nabla f \rangle (\phi^r)^2 \} e^{-2f} \, dV_t dt \\ & = \int_a^b \int_M \{ -|\nabla(u\phi^r)|^2 + |\nabla\phi^r|^2 u^2 + (2|\nabla f|^2 - \Delta f) u^2 (\phi^r)^2 - 2u^2 \phi^r \langle \nabla\phi^r, \nabla f \rangle \} e^{-2f} \, dV_t dt. \end{aligned}$$

Since $R \geq 0$ and $f_t = |\nabla f|^2$ by (2.5), we have

$$\begin{aligned} & \int_a^b \int_M |\nabla(u\phi^r)|^2 e^{-2f} \, dV_t dt + \frac{1}{2} \int_M u^2 (\phi^r)^2 e^{-2f} \, dV \Big|_a^b \\ & \leq \int_a^b \int_M \{ |\nabla\phi^r|^2 u^2 + (|\nabla f|^2 - \Delta f) u^2 (\phi^r)^2 + u^2 \phi^r (\phi_t^r - 2\langle \nabla\phi^r, \nabla f \rangle) \} e^{-2f} \, dV_t dt \\ & = \int_a^b \int_M \left\{ |\nabla\phi^r|^2 u^2 + \frac{1}{1-t} \left(f - \frac{n}{2} \right) u^2 (\phi^r)^2 + u^2 \phi^r (\phi_t^r - 2\langle \nabla\phi^r, \nabla f \rangle) \right\} e^{-2f} \, dV_t dt, \end{aligned}$$

where we have used the identity $\Delta f - |\nabla f|^2 = \bar{\tau}^{-1}(f - n/2)$ from (2.7) and (2.8).

Since u is bounded on $M \times [a, b]$ and $|\nabla f|^2 \leq f/(1 - b)$, it follows from (2.12), (2.13), Lemma 2.3 and Lemma 2.4 that, by letting $r \rightarrow +\infty$,

$$\int_a^b \int_M |\nabla u|^2 e^{-2f} \, dV_t dt \leq \frac{1}{2} \int_M u^2 e^{-2f} \, dV \Big|_b^a + \int_a^b \int_M \frac{1}{1-t} \left(f - \frac{n}{2} \right) u^2 e^{-2f} \, dV_t dt < \infty,$$

and hence (3.11) holds.

(ii) Fix $r \gg 1$ and $\epsilon \ll 1$. We calculate

$$\begin{aligned} \partial_t \int_M |\nabla u|^2 \phi^r w \, dV & = \int_M \{ \square(|\nabla u|^2 \phi^r) w - (|\nabla u|^2 \phi^r) \square^* w \} \, dV \\ & = \int_M \{ |\nabla u|^2 \square \phi^r + \phi^r \square |\nabla u|^2 - 2\langle \nabla |\nabla u|^2, \nabla \phi^r \rangle \} w \, dV \end{aligned}$$

$$\begin{aligned}
&\leq \int_M \{|\nabla u|^2 \square \phi^r - 2|\text{Hess } u|^2 \phi^r + 4|\text{Hess } u||\nabla u||\nabla \phi^r|\} w \, dV \\
&\leq \int_M \{|\nabla u|^2 |\square \phi^r| - (2 - \epsilon^2)|\text{Hess } u|^2 \phi^r + 4\epsilon^{-2}|\nabla u|^2 |\nabla \phi^r|^2 (\phi^r)^{-1}\} w \, dV.
\end{aligned} \tag{3.12}$$

Since $|\nabla u|$ is uniformly bounded by (3.7), it follows from (3.12), (2.12) and (2.15) that

$$2 \int_a^b \int_M |\text{Hess } u|^2 w \, dV_t dt \leq \int_M |\nabla u|^2 w \, dV \Big|_b^a$$

if we let $r \rightarrow +\infty$ and $\epsilon \rightarrow 0$. The other inequality can be proven similarly, and hence (3.9) holds. \square

Next, we prove

Proposition 3.3 For any $[a, b] \subset (-\infty, 1)$, suppose that u and w are two smooth functions on $M \times [a, b]$ satisfying $\square u = \square^* w = 0$. Then, the identity

$$\int_M uw \, dV_a = \int_M uw \, dV_b \tag{3.13}$$

holds under one of the following additional assumptions:

- (i) $\sup_{t \in [a, b]} \int_M |wu| \, dV_t + \int_a^b \int_M |w||\nabla u| \, dV_t dt < \infty$.
- (ii) $\sup_{t \in [a, b]} \int_M |wu| \, dV_t + \int_a^b \int_M |u||\nabla w| \, dV_t dt < \infty$.

Proof (i) We take $r \gg 1$ and calculate

$$\begin{aligned}
\partial_t \int_M wu\phi^r \, dV &= \int_M \{w\square(u\phi^r) - (u\phi^r)\square^* w\} \, dV \\
&= \int_M w \{u\square\phi^r + \phi^r\square u - 2\langle \nabla u, \nabla \phi^r \rangle\} \, dV \\
&= \int_M w \{u\square\phi^r - 2\langle \nabla u, \nabla \phi^r \rangle\} \, dV.
\end{aligned}$$

By using (2.12) and (2.15), we conclude

$$\left| \int_M wu\phi^r \, dV \Big|_a^b \right| \leq C(r^{-1} + r^{-\frac{1}{2}}) \int_a^b \int_M |w|(|u| + |\nabla u|) \, dV_t dt.$$

By taking $r \rightarrow \infty$, we arrive at (3.13).

(ii) Similarly, we have

$$\begin{aligned}
\partial_t \int_M uw\phi^r \, dV &= \int_M \{(\square u)w\phi^r - u\square^*(w\phi^r)\} \, dV \\
&= \int_M u \{-\square^* w\phi^r + w(\Delta\phi^r + \phi_t^r) + 2\langle \nabla w, \nabla \phi^r \rangle\} \, dV \\
&= \int_M u \{w(\Delta\phi^r + \phi_t^r) + 2\langle \nabla w, \nabla \phi^r \rangle\} \, dV.
\end{aligned}$$

Therefore, by (2.12), (2.13) and (2.14), we have

$$\left| \int_M uw\phi^r \, dV \Big|_a^b \right| \leq C(r^{-1} + r^{-\frac{1}{2}} + (1-b)^{-1}) \iint_{K_r} |u|(|w| + |\nabla w|) \, dV_t dt,$$

where $K_r := \{r \leq F(x, r) \leq 2r, a \leq t \leq b\}$. Consequently, by our assumption, (3.13) holds if $r \rightarrow \infty$. \square

Remark 3.4 Suppose $\square u = \square^* w = 0$.

(i) If $\sup_M |\nabla u(\cdot, a)| + \sup_{M \times [a, b]} |u| + \sup_{t \in [a, b]} \int_M |w| dV_t < \infty$, then assumption (i) holds by (3.7).

If $\sup_{M \times [a, b]} |u| + \sup_{t \in [a, b]} \int_M |w| dV_t < \infty$ and u is positive, then $|\nabla u| \leq C/\sqrt{t-a}$ by [29, Lemma 18]. Therefore, (3.13) also holds by taking the limit for $t \searrow a$.

(ii) If $\sup_{M \times [a, b]} |u| + \sup_{t \in [a, b]} \int_M |w| dV_t < \infty$ and $w(\cdot, b)$ is a nonnegative function with compact support, then assumption (ii) holds. Indeed, it follows from [29, Lemma 9] that

$$\int_a^b \int_M \frac{|\nabla w|^2}{w} dV_t dt < \infty,$$

and hence

$$\int_a^b \int_M |u| |\nabla w| dV_t dt \leq C \left(\sup_{M \times [a, b]} |u| \right) \left(\int_a^b \int_M \frac{|\nabla w|^2}{w} dV_t dt \right)^{\frac{1}{2}} \left(\int_a^b \int_M w dV_t dt \right)^{\frac{1}{2}} < \infty.$$

For later applications, we prove the following estimate of the heat kernel.

Lemma 3.5 For any $y \in M$ and $s < t < 1$, we set $u(x, t) := H(x, t, y, s)$ and $\bar{w}(x, t) := (4\pi\tau)^{-\frac{n}{2}} e^{-f(x, t)}$. Then

$$\int_M u(x, t) \bar{w}(x, t) dV_t(x) = \bar{w}(y, s). \tag{3.14}$$

Proof It is clear from the definition of \bar{w} that $\square^* \bar{w} = 0$; see [29, Equation (28)]. Moreover, for any $[a, b] \subset (s, t]$, the assumption (ii) of Proposition 3.3 holds since

$$\begin{aligned} \int_a^b \int_M u |\nabla \bar{w}| dV_t dt &\leq \int_a^b \int_M u \bar{w} (1 + |\nabla f|) dV_t dt \\ &\leq (4\pi(1-b))^{-\frac{n}{2}} \int_a^b \int_M u (1 + (1-b)^{-\frac{1}{2}} f^{\frac{1}{2}}) e^{-f} dV_t dt \\ &\leq C \int_a^b \int u dV_t dt \leq C(b-a), \end{aligned}$$

where we have used (2.8) and (3.2) and the constant C depends only on n and b .

By choosing $b = t$ and letting $a \searrow s$, the proof of Proposition 3.3 yields

$$\int_M u \bar{w} \phi^r dV_t - \bar{w}(y, s) \phi^r(y, s) = o(r),$$

where $o(r) \rightarrow 0$ as $r \rightarrow \infty$. Therefore, we immediately obtain (3.14) by letting $r \rightarrow \infty$. □

Next, we recall the definition of W_1 -Wasserstein distance.

Definition 3.6 Let (X, d) be a complete metric space and μ_1, μ_2 two probability measures on X . Then the W_1 -Wasserstein distance between μ_1 and μ_2 is defined by

$$d_{W_1}(\mu_1, \mu_2) := \sup_f \left(\int f d\mu_1 - \int f d\mu_2 \right),$$

where the supremum is taken for all bounded 1-Lipschitz functions f . We also use $d_{W_1}^t$ to denote the W_1 -distance with respect to $g(t)$.

We prove the following monotonicity of the Wasserstein distance as [2, Lemma 2.7].

Proposition 3.7 Let $(M^n, g(t))_{t < 1}$ be a Ricci flow associated with a Ricci shrinker. For $[a, b] \subset (-\infty, 1)$, let $w_1, w_2 \in C^\infty(M \times [a, b])$ be two nonnegative conjugate heat solutions such

that $\int_M w_i dV_t = 1$ for any $t \in [a, b]$ and $i = 1, 2$. We define the probability measures with $d\mu_{i,t} = w_i(\cdot, t) dV_t$, $i = 1, 2$. Then

$$d_{W_1}^t(\mu_{1,t}, \mu_{2,t})$$

is increasing for $t \in [a, b]$. In particular, if $t_1 \leq t_2 < 1$, then for any $x_1, x_2 \in M$ and $t \leq t_1$,

$$d_{W_1}^t(v_{x_1,t_1;t}, v_{x_2,t_2;t})$$

is increasing and

$$d_{W_1}^t(v_{x_1,t_1;t}, v_{x_2,t_2;t}) \leq d_{t_1}(x_1, x_2).$$

Proof Let $t_1 \leq t_2$, $t_1, t_2 \in [a, b]$ and consider a bounded function $u_1 \in C^\infty(M)$ with $\sup_M |\nabla u_1(\cdot, t_1)| \leq 1$. Suppose that u is the unique bounded heat solution on $M \times [t_1, t_2]$ starting from u_1 . Then it follows from Lemma 3.2 (i) that

$$\sup_M |\nabla u(\cdot, t)| \leq 1$$

for any $t \in [t_1, t_2]$. Clearly, we have

$$\begin{aligned} \int_M u d\mu_{1,t_1} - \int_M u d\mu_{2,t_1} &= \int_M u(x, t_1)w_1(x, t_1) dV_{t_1}(x) - \int_M u(x, t_1)w_2(x, t_1) dV_{t_1}(x) \\ &= \int_M u(x, t_2)w_1(x, t_2) dV_{t_2}(x) - \int_M u(x, t_2)w_2(x, t_2) dV_{t_2}(x) \\ &= \int_M u d\mu_{1,t_2} - \int_M u d\mu_{2,t_2} \leq d_{W_1}^{t_2}(\mu_{1,t_2}, \mu_{2,t_2}). \end{aligned}$$

Here, we have used [29, Proposition 1] for the second equality. By taking the supremum over all such u_1 , one obtains

$$d_{W_1}^{t_1}(\mu_{1,t_1}, \mu_{2,t_1}) \leq d_{W_1}^{t_2}(\mu_{1,t_2}, \mu_{2,t_2}).$$

□

Next, we recall the following definition from [2, Definition 3.1].

Definition 3.8 (Variance) The variance between two probability measures μ_1, μ_2 on a Riemannian manifold (M, g) is defined as

$$\text{Var}(\mu_1, \mu_2) := \int_M \int_M d^2(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2).$$

In the case $\mu_1 = \mu_2 = \mu$, we write

$$\text{Var}(\mu) = \text{Var}(\mu, \mu) = \int_M \int_M d^2(x_1, x_2) d\mu(x_1) d\mu(x_2).$$

We also define Var_t as the variance with respect to the metric $g(t)$.

For some basic properties of the variance, we refer the readers to [2, Lemma 3.2]. Next, we prove some results which originate from [2, Corollary 3.7, Corollary 3.8]. Before that, we first prove the following maximum principle on the product manifold (cf. [1, 7] for the related survey).

Theorem 3.9 (Maximum principle on the product) Let $(M^n, g(t))_{t < 1}$ be a Ricci flow associated with a Ricci shrinker. Given any closed interval $[a, b] \subset (-\infty, 1)$ and a function u on $M \times M \times [a, b]$ such that

$$(\partial_t - \Delta_x - \Delta_y)u(x, y, t) \leq 0. \tag{3.15}$$

Suppose that

$$\int_a^b \int_{M \times M} u_+^2(x, y, t) e^{-2f(x,t)-2f(y,t)} dV_t(x) dV_t(y) dt < \infty. \tag{3.16}$$

If $u(\cdot, a) \leq c$, then $u(\cdot, b) \leq c$.

Proof The proof follows almost verbatim from [29, Theorem 6], except that we multiply (3.15) by $u_+(x, y, t)(\phi^r(x)\phi^r(y))^2 e^{-2f(x,t)-2f(y,t)}$ and do the integration. Since no other new ingredient is needed, we omit the details here. \square

Proposition 3.10 Under the same assumptions as in Proposition 3.7, if we further assume $w_1(\cdot, b)$ and $w_2(\cdot, b)$ have compact supports, then

$$\text{Var}_t(\mu_{1,t}, \mu_{2,t}) + H_n t$$

is increasing for $t \in [a, b]$, where $H_n := (n - 1)\pi^2/2 + 4$. Moreover, for any $x_1, x_2 \in M$,

$$\text{Var}_t(v_{x_1,b;t}, v_{x_2,b;t}) + H_n t$$

is increasing for $t \leq b$. In particular,

$$\text{Var}_t(v_{x_1,b;t}, v_{x_2,b;t}) \leq d_b^2(x_1, x_2) + H_n(b - t) \quad \text{and} \quad \text{Var}_t(v_{x,b;t}) \leq H_n(b - t).$$

Proof For any $[c, d] \subset [a, b]$, we set that $u \in C^0(M \times M \times [c, b]) \cap C^\infty(M \times M \times (c, b])$ is the solution to the heat equation

$$(\partial_t - \Delta_x - \Delta_y)u = -H_n, \quad u(\cdot, c) = d_c^2.$$

Indeed, by the existence of the heat kernel, one may define

$$u(x, y, t) := \int_M \int_M H(x, t, z, c) H(y, t, w, c) d_c^2(z, w) dV_c(z) dV_c(w) - H_n(t - c). \tag{3.17}$$

We first show that (3.17) is well-defined. In fact, it is clear that

$$\begin{aligned} & \int_M \int_M H(x, t, z, c) H(y, t, w, c) d_c^2(z, w) dV_c(z) dV_c(w) \\ & \leq 2 \left(\int_M H(x, t, z, c) d_c^2(z, p) dV_c(z) + \int_M H(y, t, w, c) d_c^2(w, p) dV_c(w) \right), \end{aligned} \tag{3.18}$$

and the convergence of the last two integrals follows from [29, Corollary 5].

On the other hand, it follows from [2, Theorem 3.5] that

$$(\partial_t - \Delta_x - \Delta_y) d_t^2(x, y) \geq -H_n. \tag{3.19}$$

Combining (3.17) and (3.19), we claim that $u(x, y, t) \leq d_t^2(x, y)$ for any $t \in [c, b]$. Indeed, this follows from the maximum principle Theorem 3.9 as long as the condition (3.16) is satisfied.

First, notice that

$$\begin{aligned} & \int_c^b \int_{M \times M} d_t^4(x, y) e^{-2f(x,t)-2f(y,t)} dV_t(x) dV_t(y) dt \\ & \leq 8 \int_c^b \int_{M \times M} (d_t^4(x, p) + d_t^4(y, p)) e^{-2f(x,t)-2f(y,t)} dV_t(x) dV_t(y) dt. \end{aligned} \tag{3.20}$$

From Lemmas 2.3 and 2.4, it is clear that (3.20) is bounded. In addition, it follows from (3.18) that

$$\int_c^b \int_{M \times M} (u(x, y, t) + H_n(t - c))^2 e^{-2f(x,t)-2f(y,t)} dV_t(x) dV_t(y) dt$$

$$\begin{aligned}
&\leq 8 \int_c^b \int_{M \times M} \left(\int_M H(x, t, z, c) d_c^2(z, p) dV_c(z) \right)^2 e^{-2f(x, t) - 2f(y, t)} dV_t(x) dV_t(y) dt \\
&\quad + 8 \int_c^b \int_{M \times M} \left(\int_M H(y, t, w, c) d_c^2(w, p) dV_c(w) \right)^2 e^{-2f(x, t) - 2f(y, t)} dV_t(x) dV_t(y) dt \\
&\leq 8 \int_c^b \int_{M \times M} \int_M H(x, t, z, c) d_c^4(z, p) e^{-2f(x, t) - 2f(y, t)} dV_c(z) dV_t(x) dV_t(y) dt \\
&\quad + 8 \int_c^b \int_{M \times M} \int_M H(y, t, w, c) d_c^4(w, p) e^{-2f(x, t) - 2f(y, t)} dV_c(w) dV_t(x) dV_t(y) dt, \quad (3.21)
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality for the last inequality. From Lemma 3.5, we obtain

$$\begin{aligned}
&\int_c^b \int_{M \times M} \int_M H(x, t, z, c) d_c^4(z, p) e^{-2f(x, t) - 2f(y, t)} dV_c(z) dV_t(x) dV_t(y) dt \\
&\leq \int_c^b \int_M \int_M \int_M H(x, t, z, c) d_c^4(z, p) e^{-f(x, t) - 2f(y, t)} dV_t(x) dV_c(z) dV_t(y) dt \\
&\leq \int_c^b \int_M \int_M \left(\frac{1-t}{1-c} \right)^{\frac{3}{2}} d_c^4(z, p) e^{-f(z, c) - 2f(y, t)} dV_c(z) dV_t(y) dt \\
&\leq \int_c^b \int_M \int_M d_c^4(z, p) e^{-f(z, c) - 2f(y, t)} dV_c(z) dV_t(y) dt < \infty,
\end{aligned}$$

by Lemmas 2.3 and 2.4. Similarly, the second term in (3.21) is also bounded. Therefore, we have proven that $u(x, y, t) \leq d_t^2(x, y)$ for any $t \in [c, b]$.

By our assumption, $w_1(\cdot, b)$ and $w_2(\cdot, b)$ have compact supports. Then it follows from [29, Lemma 8, Lemma 9] that

$$w_i(x, t) \leq C \bar{w}(x, t) \quad (3.22)$$

for any $c \leq t \leq b$ and

$$\int_c^b \int_M \frac{|\nabla w_i|^2}{w_i} dV_t dt \leq C \quad (3.23)$$

for some constant $C > 0$.

Next, we set $w_1 = w_1(x, t)$, $w_2 = w_2(y, t)$, $\phi_x^r = \phi^r(x)$ and $\phi_y^r = \phi^r(y)$. Then we compute

$$\begin{aligned}
&\partial_t \int_M \int_M u w_1 w_2 \phi_x^r \phi_y^r dV_t(x) dV_t(y) \\
&= \int_M \int_M (\partial_t - \Delta_x - \Delta_y) (u \phi_x^r \phi_y^r) w_1 w_2 dV_t(x) dV_t(y) \\
&= \int_M \int_M (-H_n \phi_x^r \phi_y^r - u (\Delta_x \phi_x^r \phi_y^r + \Delta_y \phi_y^r \phi_x^r)) w_1 w_2 dV_t(x) dV_t(y) \\
&\quad + 2 \int_M \int_M (\Delta_x \phi_x^r + \langle \nabla \phi_x^r, \nabla w_1 \rangle) u \phi_y^r w_2 + (\Delta_y \phi_y^r + \langle \nabla \phi_y^r, \nabla w_2 \rangle) u \phi_x^r w_1 dV_t(x) dV_t(y). \quad (3.24)
\end{aligned}$$

From (3.23), we have

$$\left(\int_c^b \int_M \int_M |\nabla \phi_x^r| |\nabla w_1| |u| \phi_y^r w_2 dV_t(x) dV_t(y) dt \right)^2$$

$$\begin{aligned} &\leq \left(\int_c^b \int_M \int_M |\nabla \phi_x^r|^2 u^2 (\phi_y^r)^2 w_1 w_2 \, dV_t(x) dV_t(y) dt \right) \left(\int_c^b \int_M \frac{|\nabla w_1|^2}{w_1} \, dV_t dt \right) \\ &\leq C \int_c^b \int_M \int_M |\nabla \phi_x^r|^2 u^2 w_1 w_2 \, dV_t(x) dV_t(y) dt. \end{aligned} \tag{3.25}$$

Similarly, we have

$$\begin{aligned} &\left(\int_c^b \int_M \int_M |\nabla \phi_y^r| |\nabla w_2| |u| \phi_x^r w_1 \, dV_t(x) dV_t(y) dt \right)^2 \\ &\leq C \int_c^b \int_M \int_M |\nabla \phi_y^r|^2 u^2 w_1 w_2 \, dV_t(x) dV_t(y) dt. \end{aligned} \tag{3.26}$$

Combining (3.22), (3.24), (3.25), (3.26) and the fact that $-H_n(t - c) \leq u \leq d_t^2(x, y)$, we conclude, by letting $r \rightarrow \infty$, that

$$\int_M \int_M u w_1 w_2 \, dV_d(x) dV_d(y) - \int_M \int_M u w_1 w_2 \, dV_c(x) dV_c(y) = -H_n(d - c). \tag{3.27}$$

Since $u \leq d_t^2(x, y)$, it follows from (3.27) and the definition of the variance that

$$\text{Var}_d(\mu_{1,d}, \mu_{2,d}) + H_n d \geq \text{Var}_c(\mu_{1,c}, \mu_{2,c}) + H_n c.$$

Now, we assume $w_i = H(x_i, b, \cdot, \cdot)$ for $i = 1, 2$. Then it follows from [29, Lemma 23] that

$$\int_a^{b-\epsilon} \int_M \frac{|\nabla w_i|^2}{w_i} \, dV_t dt \leq C \log \epsilon^{-1}. \tag{3.28}$$

Therefore, one can use the same arguments as above, thanks to (3.28) and [29, Corollary 5], to conclude that (3.27) still holds if $[c, d] \subset [a, b - \epsilon]$. Since ϵ is arbitrary, we immediately show that

$$\text{Var}_t(v_{x_1, b; t}, v_{x_2, b; t}) + H_n t$$

is increasing for any $t \leq b$. □

Next, we recall the definition of H -center, where the conjugate heat kernel measure is concentrated.

Definition 3.11 (H -center) Given a constant $H > 0$, a point $(z, t) \in M \times (-\infty, 1)$ is called an H -center of $(x_0, t_0) \in M \times (-\infty, 1)$ if $t \leq t_0$ and

$$\text{Var}_t(\delta_z, v_{x_0, t_0; t}) \leq H(t_0 - t).$$

In particular, we have

$$d_{W_1}^t(\delta_z, v_{x_0, t_0; t}) \leq \sqrt{H(t_0 - t)}. \tag{3.29}$$

From Proposition 3.10, the following result is immediate; see [2, Proposition 3.12].

Proposition 3.12 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $t \leq t_0$, there is at least one point $z \in M$ such that (z, t) is an H_n -center of (x_0, t_0) and for any two such points $z_1, z_2 \in M$, we have $d_t(z_1, z_2) \leq 2\sqrt{H_n(t_0 - t)}$.

The following result ensures that the conjugate heat kernel measure is concentrated around an H -center; see [2, Proposition 3.13].

Proposition 3.13 If (z, t) is an H -center of (x_0, t_0) , then for any $L > 0$,

$$v_{x_0, t_0; t} \left(B_t(z, \sqrt{LH(t_0 - t)}) \right) \geq 1 - \frac{1}{L}. \tag{3.30}$$

Combining the above Proposition with [29, Theorem 14], we obtain the following integral bound for the conjugate heat kernel; see also [2, Theorem 3.14].

Proposition 3.14 If (z, t) is an H_n -center of (x_0, t_0) , then for all $r \geq 0$ and $\epsilon > 0$ we have

$$v_{x_0, t_0; t}(M \setminus B_t(z, r)) \leq C(n, \epsilon) \exp \left(- \frac{r^2}{(4 + \epsilon)(t_0 - t)} \right).$$

Proof We apply [29, Theorem 14] for $A = M \setminus B_t(z, r)$, $B = B_t(z, \sqrt{2H_n(t_0 - t)})$ and $\sigma = \epsilon/8$ to obtain

$$\begin{aligned} v_{x_0, t_0; t}(M \setminus B_t(z, r)) &\leq v_{x_0, t_0; t}^{\frac{\epsilon}{8}}(B_t(z, \sqrt{2H_n(t_0 - t)})) \exp \left(- \frac{(r - \sqrt{2H_n(t_0 - t)})_+^2}{(4 + \epsilon/2)(t_0 - t)} \right) \\ &\leq C(n, \epsilon) \exp \left(- \frac{r^2}{(4 + \epsilon)(t_0 - t)} \right), \end{aligned}$$

where we have used (3.30) for $L = 2$ and $H = H_n$. □

In order to obtain the estimates on the Nash entropy, we first generalize the improved gradient estimate [2, Theorem 4.1] to our setting. We define the following antiderivative of the 1-dimensional heat kernel:

$$\Phi(x) = \int_{-\infty}^x (4\pi)^{-1/2} e^{-t^2/4} dt. \tag{3.31}$$

Notice that $\Phi_t(x) := \Phi(t^{-1/2}x)$ is a solution to the 1-dimensional heat equation with initial condition $\chi_{[0, \infty)}$.

Theorem 3.15 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Given $[a, b] \subset (-\infty, 1)$ and a solution $u \in C^\infty(M \times [a, b])$ to the heat equation $\square u = 0$ and a constant $T \geq 0$, suppose that u only takes values in $(0, 1)$ and $|\nabla(\Phi_T^{-1}(u(\cdot, a)))| \leq 1$ if $T > 0$. Then $|\nabla(\Phi_{T+t-a}^{-1}(u(\cdot, t)))| \leq 1$ for all $t \in [a, b]$.

Proof We may assume that u takes values in $(\epsilon, 1 - \epsilon)$. Indeed, we can consider $(1 - 2\epsilon)u + \epsilon$ instead and let $\epsilon \searrow 0$. With the extra assumption, it follows from [29, Lemma 18] that

$$|\nabla u| \leq \frac{C_1}{\sqrt{t - a}} \tag{3.32}$$

on $M \times (a, b]$. It is clear from the definition of Φ_t that $\sup_M |\nabla(\Phi_T^{-1}(u(\cdot, a + \epsilon)))| \rightarrow 0$ if $T \searrow 0$. Therefore, we only need to prove the case for $T > 0$ and then let $T \searrow 0$ and $\epsilon \searrow 0$.

Now, we set $u(x, t) = \Phi_{T+t-a} \circ h(x, t)$. It follows from the definition of Φ_t that

$$|h| \leq C_2 \tag{3.33}$$

on $M \times [a, b]$. Moreover, since $|\nabla h(\cdot, a)| \leq 1$, it follows from (3.33) and Lemma 3.2(i) that

$$|\nabla h| \leq C_3 \tag{3.34}$$

on $M \times [a, b]$. By direct computation, see [2, Theorem 4.1] for details, we have

$$\square |\nabla h|^2 = -2|\text{Hess } h|^2 - \frac{1}{T + t - a} \langle \nabla h^2, \nabla |\nabla h|^2 \rangle + \frac{1}{2(T + t - a)} (1 - |\nabla h|^2) |\nabla h|^2. \tag{3.35}$$

Therefore, if we set $v = (|\nabla h|^2 - 1)_+$, then it follows from (3.35) that

$$\square v + \frac{1}{T + t - a} \langle \nabla h^2, \nabla v \rangle \leq 0.$$

Since $|\nabla h^2|$ and v are uniformly bounded on $M \times [a, b]$ by (3.33) and (3.34), it follows from Theorem 2.6 that $v \leq 0$ on $M \times [a, b]$. In other words, $|\nabla h| \leq 1$ on $M \times [a, b]$. Thus the proof is complete. \square

With the help of Theorem 3.15, one can follow verbatim as [2, Proposition 4.2] and [31, Proposition 3.4] to obtain the following estimate.

Theorem 3.16 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker and $[s, t] \subset (-\infty, 1)$. Then, for any $x \in M$, $1 \leq p < \infty$ and measurable subset $X \subset M$, we have

$$(t - s)^{\frac{p}{2}} \int_X \left(\frac{|\nabla_x H(x, t, \cdot, s)|}{H(x, t, \cdot, s)} \right)^p dv \leq C(n, p)v(X) \left(-\log \left(\frac{v(X)}{2} \right) \right)^{\frac{p}{2}},$$

where $dv = H(x, t, \cdot, s) dV_s$ is the conjugate heat kernel measure. Moreover, for any $x \in M$ and $w \in T_x M$ with $|w|_t = 1$, it holds that

$$(t - s) \int_M \left(\frac{\partial_w H(x, t, \cdot, s)}{H(x, t, \cdot, s)} \right)^2 dv \leq \frac{1}{2}. \tag{3.36}$$

In particular, we have

$$(t - s) \int_M \left| \frac{\nabla_x H(x, t, \cdot, s)}{H(x, t, \cdot, s)} \right|^2 dv \leq \frac{n}{2}. \tag{3.37}$$

Another application of Theorem 3.15 is the following L^p -Poincaré inequality; see [2, Theorem 11.1].

Theorem 3.17 (L^p -Poincaré inequality) Let $(M^n, g(t))_{t < 1}$ be a Ricci flow associated with a Ricci shrinker. Then, for $p \geq 1$ and any $[s, t] \subset (-\infty, 1)$, we have

$$\int_M u^p dv_s \leq C(p)(t - s)^{\frac{p}{2}} \int_M |\nabla u|^p dv_s,$$

for any $u \in W^{1,p}(M, dv_s)$ with $\int_M u dv_s = 0$. Here, $dv_s(y) = H(x, t, y, s) dV_s(y)$. One may choose $C(1) = \sqrt{\pi}$ and $C(2) = 2$.

Proof The proof for $p \neq 2$ follows verbatim from [2, Theorem 11.1]. Only the last statement for $p = 2$ needs to be proven. It follows from [29, Theorem 13] that the probability measure dv_s satisfies the log-Sobolev inequality with the constant $\frac{1}{2(t-s)}$. It is a standard fact that the log-Sobolev condition implies the Poincaré inequality; see [36, Theorem 22.17]. \square

Next, we recall the definitions of the Nash entropy and \mathcal{W} -entropy based at (x_0, t_0) .

Definition 3.18 Given a Ricci flow $(M^n, g(t))_{t < 1}$ associated with a Ricci shrinker and a point $(x_0, t_0) \in M \times (-\infty, 1)$, let

$$dv = dv_{x_0, t_0, t}(x) = (4\pi\tau)^{-\frac{n}{2}} e^{-b(x,t)} dV_t = H(x_0, t_0, x, t) dV_t,$$

where $\tau = t_0 - t$. Then Perelman’s \mathcal{W} -entropy and the Nash entropy based at (x_0, t_0) are, respectively, defined as

$$\mathcal{W}_{(x_0, t_0)}(\tau) = \int_M (\tau(2\Delta b - |\nabla b|^2 + R) + b - n) dv, \tag{3.38}$$

$$\mathcal{N}_{(x_0, t_0)}(\tau) = \int_M b dv - \frac{n}{2}. \tag{3.39}$$

Now we prove some basic properties of \mathcal{N} and \mathcal{W} .

Proposition 3.19 The following properties hold with Definition 3.18.

(a) $\mathcal{W}_{(x_0,t_0)}(0) = 0$, and for any $\tau_0 > 0$,

$$\mathcal{W}_{(x_0,t_0)}(\tau_0) = -2 \int_0^{\tau_0} \tau \int_M \left| \text{Rc} + \text{Hess } b - \frac{g}{2\tau} \right|^2 \text{d}v \text{d}\tau. \tag{3.40}$$

In particular, $\mathcal{W}_{(x_0,t_0)}(\tau)$ is nonpositive and decreasing.

(b) $\mathcal{N}_{(x_0,t_0)}(0) = 0$, and for any $\tau_0 > 0$,

$$\mathcal{N}_{(x_0,t_0)}(\tau_0) = \frac{1}{\tau_0} \int_0^{\tau_0} \mathcal{W}_{(x_0,t_0)}(\tau) \text{d}\tau \geq \mathcal{W}_{(x_0,t_0)}(\tau_0). \tag{3.41}$$

(c) For any $0 < \tau_1 \leq \tau_2$,

$$\mathcal{N}_{(x_0,t_0)}(\tau_1) - \frac{n}{2} \log \left(\frac{\tau_2}{\tau_1} \right) \leq \mathcal{N}_{(x_0,t_0)}(\tau_2) \leq \mathcal{N}_{(x_0,t_0)}(\tau_1). \tag{3.42}$$

Proof Given (x_0, t_0) and τ , we first prove that $\mathcal{N}_{(x_0,t_0)}(\tau)$ and $\mathcal{W}_{(x_0,t_0)}(\tau)$ are well-defined.

In what follows, all constants $C_i > 1$ depend on (x_0, t_0) , τ and the given Ricci shrinker.

It follows from [29, Theorem 19] that for any $r \geq 1$,

$$\int_{d_t(x_0,x) \geq r\sqrt{\tau}} \text{d}v_t(x) \leq C_1 e^{-\frac{r^2}{8}}. \tag{3.43}$$

Therefore, there exists $C_2 > 1$ such that

$$\int_{d_t(p,x) \geq r} \text{d}v_t(x) \leq C_2 e^{-\frac{r^2}{C_2}} \tag{3.44}$$

if $r \geq C_2$. In addition, it follows from [29, Theorem 15, Formula (203)] that

$$\mu \leq b(x, t) \leq -3\mu + \frac{d_{t_0}^2(x_0, x)}{3\tau} + \frac{4\tau}{3(1-t_0)^2} F(x, t_0). \tag{3.45}$$

From (3.45) and Lemma 2.3, there exists $C_3 > 1$ such that

$$-C_3 \leq b(x, t) \leq C_3(1 + F(x, t_0)). \tag{3.46}$$

Since F is decreasing with respect to t by (2.6), it follows from (3.46) and Lemma 2.3 that

$$b(x, t) \leq C_3(1 + F(x, t)) \leq C_4(1 + d_t^2(p, x))$$

for some $C_4 \geq C_3$. Consequently, we obtain

$$|b(x, t)| \leq C_4(1 + d_t^2(p, x)). \tag{3.47}$$

Combining (3.44) and (3.47), we can estimate

$$\begin{aligned} \int_M |b(x, t)| \text{d}v_t(x) &\leq C_4 + C_4 \int_M d_t^2(p, x) \text{d}v_t(x) \\ &= C_4 + C_4 \int_{d_t(p,x) \leq C_2} d_t^2(p, x) \text{d}v_t(x) \\ &\quad + C_4 \sum_{k=1}^{\infty} \int_{2^{k-1}C_2 \leq d_t(p,x) \leq 2^k C_2} d_t^2(p, x) \text{d}v_t(x) \\ &\leq C_4 + C_4 C_2^2 + C_4 \sum_{k=1}^{\infty} (2^k C_2)^2 C_2 e^{-2^{2k-2} C_2} < \infty. \end{aligned} \tag{3.48}$$

Therefore, it follows from the definition (3.39) that $\mathcal{N}_{(x_0,t_0)}(\tau)$ is finite. Now, the fact that $\mathcal{W}_{(x_0,t_0)}(\tau)$ is well-defined follows from Perelman’s differential Harnack inequality [29, Theorem 21].

(a) The identity (3.40) follows from [29, Remark 6]. Notice that the integral in (3.40) is always finite by [29, Lemma 30]. In particular, $\mathcal{W}_{(x_0,t_0)}(0) = \lim_{\tau \searrow 0} \mathcal{W}_{(x_0,t_0)}(\tau) = 0$.

(b) We fix $r \gg 1$ and compute

$$\begin{aligned} & \partial_\tau \left(\tau \int_M b\phi^r \, dv \right) - \frac{n}{2} \\ &= \int_M b\phi^r \, dv - \tau \int_M \square(b\phi^r) \, dv - \frac{n}{2} \\ &= \int_M \left(\tau(2\Delta b - |\nabla b|^2 + R)\phi^r + b\phi^r + \tau b \square \phi^r - 2\tau \langle \nabla b, \nabla \phi^r \rangle - \frac{n}{2}(1 + \phi^r) \right) \, dv, \end{aligned} \tag{3.49}$$

where we have used the fact that $\square b = -2\Delta b + |\nabla b|^2 - R + \frac{n}{2r}$. For $\tau_0 > 0$, we integrate (3.49) from 0 to τ_0 and obtain

$$\begin{aligned} & \tau_0 \left(\int_M b\phi^r \, dv - \frac{n}{2} \right) \\ &= \int_0^{\tau_0} \int_M \left(\tau(2\Delta b - |\nabla b|^2 + R)\phi^r + b\phi^r + \tau b \square \phi^r - 2\tau \langle \nabla b, \nabla \phi^r \rangle - \frac{n}{2}(1 + \phi^r) \right) \, dv \, d\tau. \end{aligned} \tag{3.50}$$

On the one hand, it follows from (2.15), (3.43) and (3.47) that

$$\lim_{r \rightarrow \infty} \int_0^{\tau_0} \int_M \tau |b| |\square \phi^r| \, dv \, d\tau = 0. \tag{3.51}$$

On the other hand, we estimate

$$\int_0^{\tau_0} \int_M \tau |\nabla b| |\nabla \phi^r| \, dv \, d\tau \leq Cr^{-\frac{1}{2}} \tau_0^{\frac{1}{2}} \left(\int_0^{\tau_0} \int_M \tau^2 |\nabla b|^2 \, dv \, d\tau \right)^2. \tag{3.52}$$

Since the last integral is finite by [29, Lemma 25], it follows from (3.52) that

$$\lim_{r \rightarrow \infty} \int_0^{\tau_0} \int_M \tau |\nabla b| |\nabla \phi^r| \, dv \, d\tau = 0. \tag{3.53}$$

Combining (3.50), (3.51) and (3.53), if we let $r \rightarrow \infty$, then

$$\tau_0 \left(\int_M b \, dv - \frac{n}{2} \right) = \int_0^{\tau_0} \int_M (\tau(2\Delta b - |\nabla b|^2 + R) + b - n) \, dv \, d\tau,$$

which is exactly (3.41). Notice that the last inequality in (3.41) follows from the fact that $\mathcal{W}_{(x_0,t_0)}(\tau)$ is decreasing. Moreover, it follows from (3.41) and $\mathcal{W}_{(x_0,t_0)}(0) = 0$ that $\mathcal{N}_{(x_0,t_0)}(0) = 0$.

(c) The inequality (3.42) follows in exactly the same way as in [2, Proposition 5.2 (5.7)], so we omit the proof. □

Corollary 3.20 Under the same assumptions, we have

$$\int_M (|\nabla b|^2 + R) \, dv \leq \frac{n}{2\tau}. \tag{3.54}$$

$$\int_M \left(b - \mathcal{N}_{(x_0,t_0)}(\tau) - \frac{n}{2} \right)^2 \, dv \leq n. \tag{3.55}$$

Proof From the fact that $\mathcal{N}_{(x_0,t_0)}(\tau) \geq \mathcal{W}_{(x_0,t_0)}(\tau)$, we conclude that

$$\lim_{r \rightarrow \infty} \int_M (2\Delta b - |\nabla b|^2)\phi^r + Rdv \leq \frac{n}{2\tau}, \tag{3.56}$$

where we have used the differential Harnack inequality [29, Theorem 21]. From integration by parts, we have

$$\int_M (2\Delta b - |\nabla b|^2)\phi^r dv = \int_M |\nabla b|^2\phi^r - 2\langle \nabla b, \nabla \phi^r \rangle dv. \tag{3.57}$$

In addition, we can estimate

$$2 \int_M |\nabla b| |\nabla \phi^r| dv \leq \int_M \epsilon |\nabla b|^2 \phi^r + \epsilon^{-1} \frac{|\nabla \phi^r|^2}{\phi^r} dv. \tag{3.58}$$

Therefore, it follows from (2.12), (3.56), (3.57) and (3.58) that

$$\int_M (1 - \epsilon) |\nabla b|^2 + Rdv \leq \frac{n}{2\tau}.$$

By letting $\epsilon \searrow 0$, we obtain (3.54).

Now, it follows from the Poincaré inequality Theorem 3.17 and (3.54) that

$$\int_M \left(b - \mathcal{N}_{(x_0,t_0)}(\tau) - \frac{n}{2} \right)^2 dv \leq 2\tau \int |\nabla b|^2 dv \leq n,$$

and (3.55) is proven. □

Remark 3.21 From the proof of (3.54), \mathcal{W} can be rewritten as

$$\mathcal{W}_{(x_0,t_0)}(\tau) = \int_M (\tau(|\nabla b|^2 + R) + b - n) dv,$$

which agrees with the original definition of Perelman [35, Formula (3.1)].

Corollary 3.22 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker $(M^n, g, f) \in \mathcal{M}(A)$, then

$$0 \geq \mathcal{N}_{(x_0,t_0)}(\tau) \geq \mathcal{W}_{(x_0,t_0)}(\tau) \geq \mu \geq -A \tag{3.59}$$

for any $(x_0, t_0) \in M \times (-\infty, 1)$ and $\tau > 0$. In particular, given a Ricci shrinker, the Nash entropy is always uniformly bounded.

Proof For fixed (x_0, t_0) and $\tau > 0$, it follows from [29, Theorem 20] that b increases quadratically. Therefore, it is easy to see that the function u , defined by $u^2 = (4\pi\tau)^{-\frac{n}{2}} e^{-b}$, belongs to $W_*^{1,2}(M)$ defined in [29, (92)]. From [29, Theorem 1] and the definition [29, (91)], we immediately conclude that

$$\mathcal{W}_{(x_0,t_0)}(\tau) \geq \mu(g(t_0 - \tau), \tau) \geq \mu \geq -A.$$

□

Following [2], we use the notation

$$\mathcal{N}_s^*(x, t) := \mathcal{N}_{(x,t)}(t - s).$$

Similarly to [2, Theorem 5.9], we have

Theorem 3.23 Let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker. Then, for any $s < t < 1$, the following properties hold.

- (i) \mathcal{N}_s^* is a Lipschitz function with Lipschitz constant $\sqrt{\frac{n}{2(t-s)}}$.
- (ii) In the distribution sense, we have

$$-\frac{n}{2(t-s)} \leq \square \mathcal{N}_s^* \leq 0. \tag{3.60}$$

Proof Without loss of generality, we assume $s = 0$ and consider $t \in (0, 1)$. We first define the following modified Nash entropy:

$$\mathcal{N}^r = \mathcal{N}^r(x, t) := \int b\phi^r dv - \frac{n}{2}, \tag{3.61}$$

where, as before, $b = b_{(x,t)}(y, 0) = -\frac{n}{2} \log(4\pi t) - \log H(x, t, y, 0)$ and $dv = H(x, t, y, 0) dV_0(y)$.

Claim \mathcal{N}^r converges to \mathcal{N}_0^* in C_{loc}^0 on $M \times (0, 1)$, as $r \rightarrow \infty$.

Proof of Claim Given a spacetime compact set $K \subset M \times (0, 1)$, all constants $C_i > 1$ below depend only K and the Ricci shrinker.

Similarly to (3.44), there exists $C_1 > 1$ such that

$$\int_{d_0(p,y) \geq r} dv(y) \leq C_1 e^{-\frac{r^2}{C_1}} \tag{3.62}$$

for any $r \geq C_1$. From the same argument leading to (3.47), we have

$$|b_{(x,t)}(y, 0)| \leq C_2(1 + d_0^2(p, y)). \tag{3.63}$$

Combining (3.62), (3.63) and the fact that $\text{supp}(\phi^r) \cap M \times \{0\} \subset \{C_3r \leq d_0^2(p, \cdot) \leq C_4r\}$, it is easy to show as (3.48) that

$$\lim_{r \rightarrow \infty} \int_M |b_{(x,t)}(y, 0)|(1 - \phi^r(y))dv(y) = 0 \tag{3.64}$$

uniformly for $(x, t) \in K$. From (3.64), the Claim is proven.

Next, for any vector $w \in T_x M$ with $|w|_t = 1$, we compute

$$\begin{aligned} \partial_w \mathcal{N}^r(x, t) &= \int_M \{(\partial_w b)H\phi^r + b(\partial_w H)\phi^r\} dV_0 \\ &= \int_M \{-\partial_w H\phi^r + b(\partial_w H)\phi^r\} dV_0 =: I + II, \end{aligned} \tag{3.65}$$

where $H = H(x, t, y, 0)$. Notice that

$$\int_M H\phi^r dV_0 = \int_M H(x, t, y, 0)\phi^r(y) dV_0(y)$$

is the heat solution starting from ϕ^r . Therefore, it follows from (3.7) and (2.12) that

$$|I| \leq \left| \nabla_x \int_M H(x, t, y, 0)\phi^r(y) dV_0(y) \right| \leq Cr^{-\frac{1}{2}}. \tag{3.66}$$

Next, we estimate

$$II = \int_M b \frac{\partial_w H}{H} \phi^r dv = \int_M \left(b - \mathcal{N}_0^* - \frac{n}{2}\right) \frac{\partial_w H}{H} \phi^r dv - \left(\mathcal{N}_0^* + \frac{n}{2}\right) I.$$

Therefore, we have

$$|II| \leq \left(\int_M \left(b - \mathcal{N}_0^* - \frac{n}{2}\right)^2 dv\right)^{\frac{1}{2}} \left(\int_M \left(\frac{\partial_w H}{H}\right)^2 dv\right)^{\frac{1}{2}} + Cr^{-\frac{1}{2}} \leq \sqrt{\frac{n}{2t}} + Cr^{-\frac{1}{2}}, \tag{3.67}$$

where we have used (3.36), (3.55) and (3.66). Combining (3.65), (3.66) and (3.67), we obtain

$$|\nabla_x \mathcal{N}^r(x, t)| \leq \sqrt{\frac{n}{2t}} + Cr^{-\frac{1}{2}}. \tag{3.68}$$

Since \mathcal{N}^r converges to \mathcal{N}_0^* locally uniformly by the Claim, we immediately conclude from (3.68) that \mathcal{N}_0^* is $\sqrt{\frac{n}{2t}}$ -Lipschitz.

Next, by direct computation, we have

$$\square \mathcal{N}^r(x, t) = \int_M \left| \frac{\nabla_x H}{H} \right|^2 \phi^r dv - \frac{n}{2t} \int_M \phi^r dv. \tag{3.69}$$

Combining (3.37), (3.69) and the Claim, it follows immediately that

$$-\frac{n}{2t} \leq \square \mathcal{N}_0^* \leq 0$$

in the distribution sense. □

Remark 3.24 Later, we will show that the conclusions in Theorem 3.23 hold in the classical sense once we know the decay of the conjugate heat kernel; see Corollary 4.19.

As an application of Theorem 3.23, we prove the following oscillation of the Nash entropy.

Corollary 3.25 For any $x_1, x_2 \in M$ and $s < t^* \leq t_1, t_2 < 1$, we have

$$\mathcal{N}_s^*(x_1, t_1) - \mathcal{N}_s^*(x_2, t_2) \leq \sqrt{\frac{n}{2(t^* - s)}} d_{W_1}^{t^*}(v_{x_1, t_1; t^*}, v_{x_2, t_2; t^*}) + \frac{n}{2} \log \left(\frac{t_2 - s}{t^* - s} \right). \tag{3.70}$$

In particular, if $s < t^* = t_2 \leq t_1 < 1$, then

$$\mathcal{N}_s^*(x_1, t_1) - \mathcal{N}_s^*(x_2, t_2) \leq \sqrt{\frac{n}{2(t_2 - s)}} d_{W_1}^{t_2}(v_{x_1, t_1; t_2}, \delta_{x_2}). \tag{3.71}$$

If we further assume (x_2, t_2) is an H_n -center of (x_1, t_1) , then

$$\mathcal{N}_s^*(x_1, t_1) - \mathcal{N}_s^*(x_2, t_2) \leq \sqrt{\frac{nH_n(t_1 - t_2)}{2(t_2 - s)}}. \tag{3.72}$$

Proof The proof follows verbatim from [2, Corollary 5.11]. The only difference is that we consider \mathcal{N}^r as defined in (3.61) instead and let $r \rightarrow \infty$. □

4 Heat Kernel Estimates

Throughout this section, we assume that $(M^n, g(t))_{t < 1}$ is the Ricci flow associated with a Ricci shrinker in $\mathcal{M}(A)$. First, we recall the following no-local-collapsing theorem proven in [29, Theorem 22].

Theorem 4.1 For any $x \in M$ and $t < 1$, if $R \leq r^{-2}$ on $B_t(x, r)$, then

$$|B_t(x, r)|_t \geq ce^{\mu} r^n \tag{4.1}$$

for some constant $c = c(n) > 0$.

One can improve (4.1) by using the Nash entropy. Based on the Lipschitz property of the Nash entropy, we can follow the same proof as that of [2, Theorem 6.1] to obtain the following result. Notice that by (3.59), (4.2) is stronger than (4.1)

Theorem 4.2 For any $x \in M$ and $t < 1$, if $R \leq r^{-2}$ on $B_t(x, r)$, then

$$|B_t(x, r)|_t \geq c \exp(\mathcal{N}_{(x,t)}(r^2)) r^n \tag{4.2}$$

for some constant $c = c(n) > 0$.

By using (3.55), we also have the following volume estimate around an H_n -center by following the same proof of [2, Theorem 6.2].

Theorem 4.3 For any $x \in M$ and $t < 1$, if $(z, t - r^2)$ is an H_n -center of (x, t) , then

$$|B_{t-r^2}(z, \sqrt{2H_n}r)|_{t-r^2} \geq c \exp(\mathcal{N}_{(x,t)}(r^2)) r^n \tag{4.3}$$

for some constant $c = c(n) > 0$ and any $r \geq 0$.

Next, we recall the following upper bound estimate of the heat kernel proven in [29, Theorem 15], which has already been used in the last section.

Theorem 4.4 For any $x, y \in M$ and $s < t < 1$, we have

$$H(x, t, y, s) \leq \frac{e^{-\mu}}{(4\pi(t-s))^{\frac{n}{2}}}. \tag{4.4}$$

Instead of using the entropy μ , one can include the Nash entropy and obtain the following result; see [2, Theorem 7.1].

Theorem 4.5 For any $x, y \in M$ and $s < t < 1$, we have

$$H(x, t, y, s) \leq \frac{C(n)}{(t-s)^{\frac{n}{2}}} \exp(-\mathcal{N}_{(x,t)}(t-s)). \tag{4.5}$$

Proof The proof follows in almost the same way as [2, Theorem 7.1]. The main idea is to improve the bound Z of the estimate

$$H(x, t, y, s) \leq \frac{Z}{(t-s)^{\frac{n}{2}}} \exp(-\mathcal{N}_{(x,t)}(t-s)).$$

Notice that such Z always exists by (4.4) and (3.59), which may depend on the Ricci shrinker. Thanks to (3.70) and (4.3), we can follow the same argument as in [2, Theorem 7.1] to improve Z to be $Z/2$, if Z is no smaller than a constant $\bar{Z}(n)$ depending only on n . □

With the help of Theorem 3.16, Corollary 3.25 and Theorem 4.5, we obtain the following gradient estimate of the heat kernel as [2, Theorem 7.5], which improves [29, Lemma 18].

Theorem 4.6 For any $x, y \in M$ and $s < t < 1$, then

$$\frac{|\nabla_x H|(x, t, y, s)}{H(x, t, y, s)} \leq \sqrt{\frac{C}{t-s}} \sqrt{\log \left(\frac{C \exp(-\mathcal{N}_{(x,t)}(t-s))}{(t-s)^{\frac{n}{2}} H(x, t, y, s)} \right)} \tag{4.6}$$

for some constant $C = C(n) > 0$.

With the gradient estimate (4.6), one obtains the following non-expanding estimate as [2, Theorem 8.1]. Notice that (4.7) generalizes the global volume estimate Lemma 2.4.

Theorem 4.7 For any $x \in M$, $t < 1$ and $r \geq 0$, we have

$$|B_t(x, r)|_t \leq C(n) \exp(\mathcal{N}_{(x,t)}(r^2)) r^n \leq C(n)r^n. \tag{4.7}$$

Before we prove more refined heat kernel estimates, we first prove a series of lemmas.

Lemma 4.8 (Distance comparison) For any $\delta \in (0, 1)$, there exists a constant $L_1 = L_1(n, \delta) > 1$ such that

$$d_t(x, p) \leq d_s(x, p) + L_1 \leq \frac{L_1}{\sqrt{1-t}}(d_t(x, p) + 1) + L_1^2 \tag{4.8}$$

for any $x \in M$ and $-\delta^{-1} \leq s \leq t < 1$.

Proof From (2.6) and (2.8), we have

$$-\frac{F}{1-t} \leq \partial_t F = -(1-t)R \leq 0.$$

Therefore, for any $x \in M$,

$$\frac{1-t}{1-s}F(x, s) \leq F(x, t) \leq F(x, s). \tag{4.9}$$

Consequently, (4.8) follows from the combination of Lemma 2.3 and (4.9). □

As an application of the distance comparison, we have the following lower bound of the heat kernel.

Theorem 4.9 For any $K > 1$, $\delta \in (0, 1)$ and $A > 0$, there exists a constant $C = C(n, K, \delta, A) > 1$ satisfying the following property.

Suppose $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_t(x, p) \leq K$. Then

$$H(x, t, y, s) \geq \frac{C^{-1}}{(t-s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(x, y)}{C^{-1}(t-s)}\right). \tag{4.10}$$

Proof In the proof, all constants $C_i > 1$ depend on n, K, δ and A .

It follows from [29, Formula (203)] that

$$H(x, t, y, s) \geq \frac{C_1^{-1}}{(t-s)^{\frac{n}{2}}} \exp\left(-\frac{d_t^2(x, y)}{3(t-s)} - \frac{4(t-s)}{3(1-t)^2}F(y, t)\right). \tag{4.11}$$

From (4.8), we have

$$d_t^2(x, y) \leq 2(d_t^2(x, p) + d_t^2(p, y)) \leq C_2(d_s^2(x, p) + d_s^2(p, y) + 1) \leq C_3(d_s^2(x, y) + 1), \tag{4.12}$$

where we have used $d_s(x, p) \leq C(d_t(x, p) + 1) \leq C(K + 1)$ by (4.8).

In addition, since F is decreasing with respect to t ,

$$F(y, t) \leq F(y, s) \leq C_4(d_s^2(p, y) + 1) \leq C_5(d_s^2(x, y) + 1) \tag{4.13}$$

by Lemma 2.3. Combining (4.11), (4.12) and (4.13), it is easy to see (4.10) holds for some C . □

Lemma 4.10 For any $K > 1$, $\delta \in (0, 1)$ and $A > 0$, there exist constants $L_2 = L_2(n, K, \delta, A) > 1$ and $L_3 = L_3(n, \delta, A) > 1$ satisfying the following property.

Suppose $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_t(p, x) \leq K$, then for any H_n -center (z, s) of (x, t) , we have

$$d_s(x, z) \leq L_2\sqrt{t-s} \tag{4.14}$$

and

$$d_s(z, p) \leq d_t(x, p) + L_3\sqrt{t-s}. \tag{4.15}$$

Proof Since $d_t(p, x) \leq K$, it follows from [29, Theorem 19] that

$$v_{x,t;s}(M \setminus B_s(x, r\sqrt{t-s})) \leq C_2 \exp\left(-\frac{r^2}{5}\right) \tag{4.16}$$

for any $r \geq 1$ and $C_2 = C_2(n, K, \delta, A) > 0$. On the other hand, by Proposition 3.14, we have

$$v_{x,t;s}(M \setminus B_s(z, r\sqrt{t-s})) \leq C(n) \exp\left(-\frac{r^2}{5}\right) \tag{4.17}$$

for any $r \geq 0$. Combining (4.16) and (4.17), (4.14) follows immediately.

If we assume (z', s) to be an H_n -center of (p, t) , then (4.14) indicates that

$$d_s(z', p) \leq C_3(n, \delta, A)\sqrt{t-s}. \tag{4.18}$$

Then it follows from Proposition 3.7 and (3.29) that

$$\begin{aligned} d_s(z, p) &\leq d_s(z, z') + d_s(z', p) \\ &\leq d_{W_1}^s(\delta_z, \delta_{z'}) + C_3\sqrt{t-s} \\ &\leq d_{W_1}^s(v_{x,t;s}, v_{p,t;s}) + d_{W_1}^s(\delta_z, v_{x,t;s}) + d_{W_1}^s(\delta_{z'}, v_{p,t;s}) + C_3\sqrt{t-s} \\ &\leq d_t(x, p) + 2\sqrt{H_n(t-s)} + C_3\sqrt{t-s}. \end{aligned}$$

Therefore, (4.15) holds for $L_3 = C_3 + 2\sqrt{H_n}$. □

Next, we prove the following rough heat kernel estimate.

Proposition 4.11 For any $K > 1$, $\delta \in (0, 1)$ and $A > 0$, there exists a constant $L_4 = L_4(n, K, \delta, A) > 1$ satisfying the following property.

Suppose $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_s(x, p) + d_s(y, p) \leq K$, then

$$H(x, t, y, s) \leq \frac{L_4}{(t-s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(x, y)}{L_4(t-s)}\right). \tag{4.19}$$

Proof Without loss of generality, we assume $s = 0$. In the proof, all constants C_i depend on n, K, δ and A .

Given $0 < t \leq 1 - \delta$ and $x, y \in M$ with $d_0(x, p) + d_0(y, p) \leq K$, we set $d := d_0(x, y)$. It follows from Lemma 4.8 that

$$d_l(x, p) + d_l(y, p) \leq C_1 \tag{4.20}$$

for any $l \in [0, t]$. Therefore, it follows from the local distance distortion estimate [29, Theorem 18] that there exists $C_2 > 1$ such that if $d \geq C_2\sqrt{t}$,

$$C_2^{-1}d \leq d_l(x, y) \leq C_2d \tag{4.21}$$

for any $l \in [0, t]$. Notice that if $d \leq C_2\sqrt{t}$, (4.19) follows immediately from (4.4). Consequently, we may assume $d \geq C_2\sqrt{t}$ and hence (4.21) holds.

For any $l \in [0, t/2]$, we apply [29, Theorem 14] for sets $B_l(x, \sqrt{t})$, $B_l(y, \sqrt{t})$ and parameter $\sigma = 1$ to obtain

$$v_{x,t;l}(B_l(x, \sqrt{t})) v_{x,t;l}(B_l(y, \sqrt{t})) \leq \exp\left(-\frac{(d_l(x, y) - 2\sqrt{t})_+^2}{16t}\right) \leq C_3 \exp\left(-\frac{d^2}{C_3t}\right) \tag{4.22}$$

for some $C_3 > 1$, where we have used (4.21). In addition, for any $l \in [0, t/2]$ and $d_l(x, z) \leq \sqrt{t}$, it follows from [29, Theorem 18] that $d_t(x, z) \leq C_4\sqrt{t}$. Therefore, it follows from [29, Theorem

17] that

$$H(x, t, z, l) \geq C_5^{-1} t^{-\frac{n}{2}},$$

and hence

$$v_{x,t;l} \left(B_l(x, \sqrt{t}) \right) \geq C_5^{-1} t^{-\frac{n}{2}} |B_l(x, \sqrt{t})| \geq C_6^{-1}, \quad (4.23)$$

where we have used the fact that R is bounded on $B_l(x, \sqrt{t})$ and the no-local-collapsing Theorem 4.1.

Combining (4.22) and (4.23), we obtain for any $l \in [0, t/2]$ that

$$\int_{B_l(y, \sqrt{t})} H(x, t, z, l) dV_l(z) \leq C_7 \exp \left(-\frac{d^2}{C_3 t} \right). \quad (4.24)$$

In light of (4.4), for any $l \in [0, t/2]$, the above inequality implies that

$$\int_{B_l(y, \sqrt{t})} H^2(x, t, z, l) dV_l(z) \leq \frac{C_8}{t^{\frac{n}{2}}} \exp \left(-\frac{d^2}{C_3 t} \right). \quad (4.25)$$

Integrating l from 0 to $t/2$, we have

$$\int_0^{\frac{t}{2}} \int_{B_l(y, \sqrt{t})} H^2(x, t, z, l) dV_l(z) ds \leq \frac{C_9}{t^{\frac{n}{2}-1}} \exp \left(-\frac{d^2}{C_3 t} \right). \quad (4.26)$$

Consequently, the desired heat kernel estimate (4.19) follows from (4.26) and a parabolic mean value inequality [6, Lemma 4.2]. Here, [6, Lemma 4.2] can be applied in our setting, since the key ingredient is the existence of a nice local cutoff function, which is constructed in [6, Theorem 1.3] (see also Proposition 5.12). Once the existence of the local cutoff function is guaranteed, one can follow verbatim the proof of [6, Lemma 4.2] to obtain the mean value inequality. \square

We immediately obtain the following result by combining Lemma 4.10 and Proposition 4.11.

Corollary 4.12 For any $K > 1$, $\delta \in (0, 1)$ and $A > 0$, there exists a constant $L_5 = L_5(n, K, \delta, A) > 1$ satisfying the following property.

Suppose $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_t(x, p) + d_s(y, p) \leq K$, then for any H_n -center (z, s) of (x, t) , we have

$$H(x, t, y, s) \leq \frac{L_5}{(t-s)^{\frac{n}{2}}} \exp \left(-\frac{d_s^2(z, y)}{L_5(t-s)} \right). \quad (4.27)$$

Proof From (4.8), we have $d_s(x, p) + d_s(y, p) \leq C$ for some $C = C(n, K, \delta) > 0$. Then (4.27) follows from (4.14) and (4.19). \square

Next, we prove the following technical result.

Lemma 4.13 There exists a positive constant $\bar{Q} = \bar{Q}(n) > 0$ satisfying the following property.

Suppose $x, y \in M$, $T \in (0, 1)$ and there exists an H_n -center $(z, 0)$ of (x, T) such that

$$H(x, T, y, 0) \geq Q \frac{\exp(-\mathcal{N}_0^*(x, T))}{T^{\frac{n}{2}}} \exp \left(-\frac{d_0^2(z, y)}{QT} \right) \quad (4.28)$$

for some $Q \geq \bar{Q}$. Then, for any H_n -center (z', T_1) of (x, T) , there exist a point $x_1 \in M$ and an H_n -center $(z_1, 0)$ of (x_1, T_1) such that

$$d_{T_1}(x_1, z') \leq \frac{10}{\sqrt{Q}} d_0(z, y) \quad (4.29)$$

and

$$H(x_1, T_1, y, 0) \geq Q_1 \frac{\exp(-\mathcal{N}_0^*(x_1, T_1))}{T_1^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z_1, y)}{Q_1 T_1}\right), \tag{4.30}$$

where $T_1 = T/8$ and $Q_1 = 2Q$.

Proof In the proof, all constants $C_i > 1$ depend only on n . We set

$$d := d_0(z, y), \quad a := H(x, T, y, 0), \quad v_t := v_{x, T; t}, \quad \text{and} \quad V := \left\{w \in M \mid H(w, T_1, y, 0) \geq \frac{a}{2}\right\}.$$

Notice that by (4.28) and (4.5), we have

$$C_1 \frac{\exp(-\mathcal{N}_0^*(x, T))}{T^{\frac{n}{2}}} \geq a \geq \sqrt{Q} \cdot \sqrt{Q} \frac{\exp(-\mathcal{N}_0^*(x, T))}{T^{\frac{n}{2}}} \exp\left(-\frac{d^2}{QT}\right).$$

Thus if \bar{Q} is sufficiently large, we have $Q > \bar{Q} > C_1^2$ and derive from the above inequality that

$$\frac{d^2}{T} \geq \frac{Q \log Q}{2}. \tag{4.31}$$

It follows from the semigroup property (3.1) that

$$\begin{aligned} a &= \int_M H(w, T_1, y, 0) dv_{T_1}(w) \\ &= \int_{M \setminus V} H(w, T_1, y, 0) dv_{T_1}(w) + \int_V H(w, T_1, y, 0) dv_{T_1}(w) \\ &\leq \frac{a}{2} v_{T_1}(M \setminus V) + \int_V H(w, T_1, y, 0) dv_{T_1}(w) \\ &\leq \frac{a}{2} + C_1 T_1^{-\frac{n}{2}} \int_V \exp(-\mathcal{N}_0^*(w, T_1)) dv_{T_1}(w), \end{aligned} \tag{4.32}$$

where we have used (4.5) for the last inequality. Moreover, it follows from (3.72) and the Lipschitz property of \mathcal{N}_0^* that

$$-\mathcal{N}_0^*(z', T_1) \leq -\mathcal{N}_0^*(x, T) + C_2 \sqrt{\frac{T - T_1}{T_1}} \leq -\mathcal{N}_0^*(x, T) + C_3 \tag{4.33}$$

and

$$-\mathcal{N}_0^*(w, T_1) \leq -\mathcal{N}_0^*(z', T_1) + \sqrt{\frac{n}{2T_1}} d_{T_1}(w, z') \leq -\mathcal{N}_0^*(x, T) + C_4 T^{-\frac{1}{2}} d_{T_1}(w, z') + C_4. \tag{4.34}$$

Now, we define $B := B_{T_1}(z', 10Q^{-\frac{1}{2}}d)$. Then it follows from (4.34) that

$$\begin{aligned} &\int_V \exp(-\mathcal{N}_0^*(w, T_1)) dv_{T_1}(w) \\ &\leq e^{C_4} \exp(-\mathcal{N}_0^*(x, T)) \int_V e^{C_4 T^{-\frac{1}{2}} d_{T_1}(w, z')} dv_{T_1}(w) \\ &\leq e^{C_4} \exp(-\mathcal{N}_0^*(x, T)) \left(e^{10C_4 T^{-\frac{1}{2}} Q^{-\frac{1}{2}} d} v_{T_1}(V \cap B) + \int_{M \setminus B} e^{C_4 T^{-\frac{1}{2}} d_{T_1}(w, z')} dv_{T_1}(w) \right). \end{aligned} \tag{4.35}$$

It follows from Proposition 3.14 that

$$\begin{aligned} &\int_{M \setminus B} e^{C_4 T^{-\frac{1}{2}} d_{T_1}(w, z')} dv_{T_1}(w) \\ &= \sum_{k=1}^{\infty} \int_{2^{k-1}(10Q^{-\frac{1}{2}}d) \leq d_{T_1}(w, z') \leq 2^k(10Q^{-\frac{1}{2}}d)} e^{C_4 T^{-\frac{1}{2}} d_{T_1}(w, z')} dv_{T_1}(w) \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{k=1}^{\infty} e^{C_4 2^k T^{-\frac{1}{2}} 10Q^{-\frac{1}{2}} d} \int_{d_{T_1}(w, z') \geq 2^{k-1}(10Q^{-\frac{1}{2}} d)} dv_{T_1}(w) \\
 &\leq C_5 \sum_{k=1}^{\infty} \exp\left(C_4 2^k T^{-\frac{1}{2}} 10Q^{-\frac{1}{2}} d - \frac{(2^{k-1} 10Q^{-\frac{1}{2}} d)^2}{5(T - T_1)}\right) \\
 &\leq C_5 \sum_{k=1}^{\infty} \exp\left(-\frac{(2^{k-1} 10Q^{-\frac{1}{2}} d)^2}{5T} + C_6\right) \leq C_7 \exp\left(-\frac{20d^2}{QT}\right), \tag{4.36}
 \end{aligned}$$

where we have used the fact that $\exp\left(-\frac{20d^2}{QT}\right) \leq Q^{-10} \ll 1$ by (4.31).

Combining (4.28), (4.32), (4.35) and (4.36), we have

$$QT^{-\frac{n}{2}} \exp\left(-\frac{d^2}{QT}\right) \leq C_8 T_1^{-\frac{n}{2}} \left(e^{10C_4 T^{-\frac{1}{2}} Q^{-\frac{1}{2}} d} v_{T_1}(V \cap B) + \exp\left(-\frac{20d^2}{QT}\right)\right). \tag{4.37}$$

Since Q is large, by (4.31) we have

$$Q \exp\left(\frac{19d^2}{QT}\right) \geq Q^{\frac{21}{2}} \geq 2C_8 8^{\frac{n}{2}}.$$

Then it is not hard to see from (4.37) that $v_{T_1}(V \cap B) > 0$. Thus there exists a point $x_1 \in V \cap B$ which satisfies (4.29). Then we take an H_n -center $(z_1, 0)$ of (x_1, T_1) . The point-selecting process is illustrated in Figure 1.

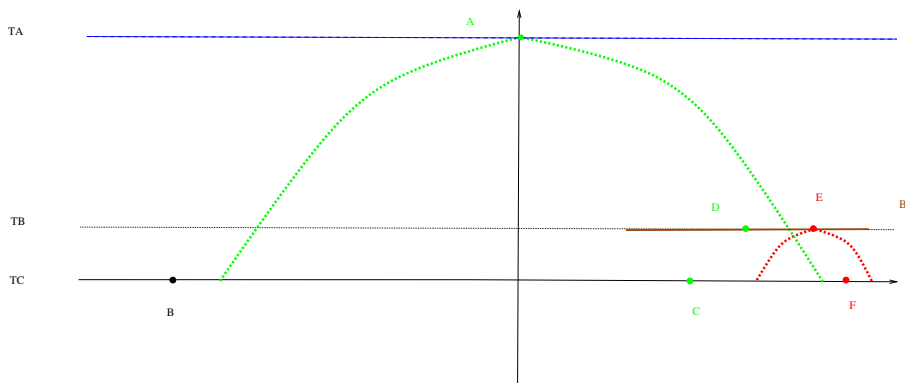


Figure 1

It follows from Proposition 3.7 and (3.29) that

$$\begin{aligned}
 d_0(z, z_1) &= d_{W_1}^0(\delta_z, \delta_{z_1}) \\
 &\leq d_{W_1}^0(v_{x, T; 0}, v_{x_1, T_1; 0}) + d_{W_1}^0(\delta_z, v_{x, T; 0}) + d_{W_1}^0(\delta_{z_1}, v_{x_1, T_1; 0}) \\
 &\leq d_{W_1}^{T_1}(v_{x, T; T_1}, \delta_{x_1}) + \sqrt{H_n T} + \sqrt{H_n T_1} \\
 &\leq d_{W_1}^{T_1}(v_{x, T; T_1}, \delta_{z'}) + d_{T_1}(z', x_1) + \sqrt{H_n T} + \sqrt{H_n T_1} \\
 &\leq \sqrt{H_n(T - T_1)} + \sqrt{H_n T} + \sqrt{H_n T_1} + 10Q^{-\frac{1}{2}} d \\
 &\leq 3\sqrt{H_n T} + 10Q^{-\frac{1}{2}} d. \tag{4.38}
 \end{aligned}$$

Therefore, we conclude

$$d_0(z_1, y) \geq d - d_0(z, z_1) \geq (1 - 10Q^{-\frac{1}{2}})d - 3\sqrt{H_n T}, \tag{4.39}$$

and hence

$$d_0^2(z_1, y) \geq \frac{(1 - 10Q^{-\frac{1}{2}})^2}{2} d^2 - 9H_n T. \tag{4.40}$$

Since $x_1 \in V$, from the definition of V and (4.28), we have

$$H(x_1, T_1, y, 0) \geq \frac{a}{2} \geq Q \frac{\exp(-\mathcal{N}_0^*(x, T))}{2T^{\frac{n}{2}}} \exp\left(-\frac{d^2}{QT}\right). \tag{4.41}$$

We claim

$$Q \frac{\exp(-\mathcal{N}_0^*(x, T))}{2T^{\frac{n}{2}}} \exp\left(-\frac{d^2}{QT}\right) \geq Q_1 \frac{\exp(-\mathcal{N}_0^*(x_1, T_1))}{T_1^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z_1, y)}{Q_1 T_1}\right). \tag{4.42}$$

Indeed, it follows from (4.34) that

$$-\mathcal{N}_0^*(x_1, T_1) \leq -\mathcal{N}_0^*(x, T) + C_9(1 + d(QT)^{-\frac{1}{2}}). \tag{4.43}$$

On the other hand, by (4.40), we have

$$\begin{aligned} & \exp\left(\frac{d_0^2(z_1, y)}{Q_1 T_1} - \frac{d^2}{QT} - C_9 \sqrt{\frac{d^2}{QT}} - C_9\right) \\ & \geq \exp\left(\frac{1}{QT} \left(2(1 - 10Q^{-\frac{1}{2}})^2 - 0.9\right) d^2 - \frac{36H_n}{Q} - C_{10}\right) \\ & \geq \exp\left(\frac{d^2}{QT} - \frac{36H_n}{Q} - C_{10}\right) \geq \sqrt{Q} \exp\left(-\frac{36H_n}{Q} - C_{10}\right) \geq 4 \cdot 8^{\frac{n}{2}}, \end{aligned} \tag{4.44}$$

where we have used (4.31) for the second last inequality. As \bar{Q} is sufficiently large, it is clear that (4.42) follows from the combination of (4.43) and (4.44). Consequently, we obtain (4.30) by (4.41). \square

Proposition 4.14 For any $x, y \in M$ and $t \in (0, 1)$,

$$H(x, t, y, 0) \leq \bar{Q} \frac{\exp(-\mathcal{N}_0^*(x, t))}{t^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z, y)}{\bar{Q}t}\right), \tag{4.45}$$

where $(z, 0)$ is any H_n -center of (x, t) and \bar{Q} is the same constant in Proposition 4.13.

Proof Supposing otherwise, there exist $x, y \in M$, $T \in (0, 1)$ and an H_n -center $(z, 0)$ of (x, T) such that

$$H(x, T, y, 0) \geq \bar{Q} \frac{\exp(-\mathcal{N}_0^*(x, T))}{T^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z, y)}{\bar{Q}T}\right). \tag{4.46}$$

Now, we define $Q_k = 2^k \bar{Q}$ and $T_k := 8^{-k} T$ for $k \in \mathbb{N}$. If we set $x_0 = x$ and $z_0 = z$, then we claim there are sequences x_k, z'_k and z_k satisfying

- (a) (z'_k, T_k) is an H_n -center of (x_{k-1}, T_{k-1}) .
- (b) $(z_k, 0)$ is an H_n -center of (x_k, T_k) .
- (c) $d_{T_k}(x_k, z'_k) \leq 10Q_{k-1}^{-\frac{1}{2}} d_0(z_{k-1}, y)$.
- (d) $d_0(z_k, z_{k-1}) \leq 3\sqrt{H_n T_{k-1}} + 10Q_{k-1}^{-\frac{1}{2}} d_0(z_{k-1}, y)$.
- (e) we have the heat kernel estimate

$$H(x_k, T_k, y, 0) \geq Q_k \frac{\exp(-\mathcal{N}_0^*(x_k, T_k))}{T_k^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z_k, y)}{Q_k T_k}\right). \tag{4.47}$$

The existence of x_k, z'_k and z_k satisfying (a)–(e) is obtained by Lemma 4.13 and an inductive argument. Notice that (d) is guaranteed by (4.38).

Claim $b_k := d_{T_k}(x_k, p)$ is uniformly bounded.

Proof of the Claim We set $d_k := d_0(z_k, y)$ for $k \in \mathbb{N}$. It follows from (d) that

$$d_k \leq d_{k-1} + d_0(z_k, z_{k-1}) \leq \left(1 + 10Q_{k-1}^{-\frac{1}{2}}\right) d_{k-1} + 3\sqrt{H_n T_{k-1}}. \tag{4.48}$$

Therefore, it is easy to derive from (4.48) and the definitions of Q_k and T_k that

$$d_k \leq K_1 < \infty \tag{4.49}$$

for some constant K_1 depending on $d_0(z, y), T, \bar{Q}$ and n . From (c) and (4.49), we have

$$d_{T_k}(x_k, z'_k) \leq 10Q_{k-1}^{-\frac{1}{2}} d_{k-1} \leq 10K_1 Q_{k-1}^{-\frac{1}{2}}. \tag{4.50}$$

Moreover, since (z'_k, T_k) is an H_n -center of (x_{k-1}, T_{k-1}) , it follows from (4.15) that

$$d_{T_k}(z'_k, p) \leq d_{T_{k-1}}(x_{k-1}, p) + L_3\sqrt{T_{k-1} - T_k} \leq b_{k-1} + L_3 T_{k-1}^{\frac{1}{2}}, \tag{4.51}$$

where $L_3 = L_3(n, \delta, A) > 0$ for some fixed constant $\delta \in (0, 1)$ with $T \leq 1 - \delta$. Combining (4.50) and (4.51), we obtain

$$b_k \leq b_{k-1} + 10K_1 Q_{k-1}^{-\frac{1}{2}} + L_3 T_{k-1}^{\frac{1}{2}}. \tag{4.52}$$

From (4.52), it is clear that b_k is uniformly bounded, and the Claim is proven.

Thanks to the Claim, we can apply Corollary 4.12 to obtain an upper bound of the heat kernel, which contradicts the lower bound (4.47) when k is sufficiently large. \square

Now, we state the main theorem of this section regarding the heat kernel upper bound, which generalizes and slightly improves [2, Theorem 7.2].

Theorem 4.15 $(M^n, g(t))_{t < 1}$ is the Ricci flow associated with a Ricci shrinker. For any $\epsilon > 0$, there exists a constant $C = C(n, \epsilon) > 0$ such that

$$H(x, t, y, s) \leq \frac{C \exp(-\mathcal{N}_{(x,t)}(t-s))}{(t-s)^{\frac{n}{2}}} \exp\left(-\frac{d_s^2(z, y)}{(4+\epsilon)(t-s)}\right), \tag{4.53}$$

for any $s < t < 1$ and any H_n -center (z, s) of (x, t) .

Proof Without loss of generality, we assume $s = 0$. The proof is a modification of the proof of Lemma 4.13 and all constants $C_i > 1$ depend on n and ϵ .

Supposing otherwise, there exist $x, y \in M, T \in (0, 1), \epsilon > 0$ and an H_n -center $(z, 0)$ of (x, T) such that

$$H(x, T, y, 0) \geq Q \frac{\exp(-\mathcal{N}_0^*(x, T))}{T^{\frac{n}{2}}} \exp\left(-\frac{d_0^2(z, y)}{(4+\epsilon)T}\right), \tag{4.54}$$

where Q is a large constant to be determined later. We also set $\theta \in (0, 1)$ as a small parameter and $\theta_1^3 = \theta$. Define

$$\begin{aligned} d &:= d_0(z, y), \quad a := H(x, T, y, 0), \quad v_t := v_{x, T; t}, \quad T_\theta := \theta T, \\ V &:= \left\{ w \in M \mid H(w, T_\theta, y, 0) \geq \frac{a}{2} \right\}. \end{aligned}$$

From (4.54) and (4.5), we have

$$\exp\left(\frac{d^2}{T}\right) \geq \{C_1^{-1}Q\}^{4+\epsilon}. \tag{4.55}$$

Now, we assume (z', T_θ) is an H_n -center of (x, T) and set $B := B_{T_\theta}(z', (1 - \theta_1)d)$. Similarly to (4.34), we have

$$-\mathcal{N}_0^*(w, T_\theta) \leq -\mathcal{N}_0^*(x, T) + C_2\theta^{-\frac{1}{2}}T^{-\frac{1}{2}}d_{T_\theta}(w, z') + C_2\theta^{-\frac{1}{2}}. \tag{4.56}$$

By the same argument as (4.36), we apply Proposition 3.14 for $\epsilon/4$ to obtain, if $\theta < \bar{\theta}(\epsilon)$,

$$\begin{aligned} & \int_{M \setminus B} e^{C_2\theta^{-\frac{1}{2}}T^{-\frac{1}{2}}d_{T_\theta}(w, z')} dv_{T_\theta}(w) \\ & \leq C_3 \sum_{k=1}^{\infty} \exp\left(C_2 2^k \theta^{-\frac{1}{2}} T^{-\frac{1}{2}} (1 - \theta_1) d - \frac{(2^{k-1}(1 - \theta_1)d)^2}{(4 + \epsilon/4)(1 - \theta)T}\right) \\ & \leq C_3 \sum_{k=1}^{\infty} \exp\left(-\frac{(2^{k-1}(1 - \theta_1)d)^2}{(4 + \epsilon/3)(1 - \theta)T} + C_4\theta^{-1}\right) \\ & \leq C_5 \exp\left(-\frac{((1 - \theta_1)d)^2}{(4 + \epsilon/3)(1 - \theta)T} + C_4\theta^{-1}\right). \end{aligned} \tag{4.57}$$

Similarly to (4.37), we obtain

$$\begin{aligned} & QT^{-\frac{n}{2}} \exp\left(-\frac{d^2}{(4 + \epsilon)T}\right) \\ & \leq C_6 T_\theta^{-\frac{n}{2}} \left(e^{C_2\theta^{-\frac{1}{2}}T^{-\frac{1}{2}}(1-\theta_1)d} v_{T_\theta}(V \cap B) + e^{C_6\theta^{-1}} \exp\left(-\frac{((1 - \theta_1)d)^2}{(4 + \epsilon/3)(1 - \theta)T}\right) \right). \end{aligned} \tag{4.58}$$

We claim that $v_{T_\theta}(V \cap B) > 0$. Indeed, it follows from (4.55) that

$$\begin{aligned} & Q \exp\left(\left(\frac{(1 - \theta_1)^2}{(4 + \epsilon/3)(1 - \theta)} - \frac{1}{(4 + \epsilon)}\right) \frac{d^2}{T}\right) \\ & \geq Q \exp\left(\frac{c(\epsilon)d^2}{T}\right) \geq Q^{1+c(\epsilon)} C_1^{-c(\epsilon)} \geq 2C_6\theta^{-\frac{n}{2}} e^{C_6\theta^{-1}}, \end{aligned}$$

where $c(\epsilon) > 0$ depends only on $\epsilon > 0$ and we choose $\theta < \bar{\theta}(\epsilon)$ and Q sufficiently large. Therefore, the claim follows from (4.58).

We choose a point $x_1 \in V \cap B$ and an H_n -center $(z_1, 0)$ of (x_1, T_θ) . Similarly to (4.38), we have

$$d_0(z, z_1) \leq 3\sqrt{H_n T} + (1 - \theta_1)d, \tag{4.59}$$

and hence

$$d_0(z_1, y) \geq \theta_1 d - 3\sqrt{H_n T}. \tag{4.60}$$

Moreover, as (4.43), we have by (4.56),

$$-\mathcal{N}_0^*(x_1, T_\theta) \leq -\mathcal{N}_0^*(x, T) + C_2\theta^{-\frac{1}{2}}(T^{-\frac{1}{2}}d + 1). \tag{4.61}$$

Now, by virtue of Proposition 4.14 and the definition of V , we have

$$\bar{Q} \exp\left(-\frac{d_0^2(z_1, y)}{Q T_\theta}\right) \geq H(x_1, T_\theta, y, 0) \cdot T^{\frac{n}{2}} \cdot \exp(\mathcal{N}_0^*(x_1, T_\theta)) \geq \frac{1}{2} Q \exp\left(-\frac{d^2}{(4 + \epsilon)T}\right). \tag{4.62}$$

Since $d_0^2(z_1, y) \geq \theta_1^2 d^2/2 - 9H_n T$ from (4.60), it follows from (4.61) and (4.62) that

$$Q \exp\left(\left(\frac{1}{2\theta_1 \bar{Q}} - \frac{1}{(4 + \epsilon)} - 1\right) \frac{d^2}{T} - C_7\theta^{-1}\right)$$

$$\leq Q \exp \left(\left(\frac{1}{2\theta_1 \bar{Q}} - \frac{1}{(4 + \epsilon)} \right) \frac{d^2}{T} - C_2 \theta^{-\frac{1}{2}} T^{-\frac{1}{2}} d \right) \leq 2\bar{Q} \theta^{-\frac{n}{2}} \exp \left(\frac{9H_n}{Q\theta} \right), \tag{4.63}$$

provided that $\theta \leq \bar{\theta}(\epsilon, \bar{Q})$. However, (4.63) is impossible by (4.55) if Q is sufficiently large.

In sum, we obtain a contradiction and (4.53) holds. □

Combining Lemma 4.10 and Theorem 4.15, we have the following estimate, which improves [29, Theorem 20].

Theorem 4.16 For any $K > 1$, $\delta \in (0, 1)$ and $A > 0$, there exists a constant $C = C(n, K, \delta, A) > 1$ satisfying the following property.

Suppose $-\delta^{-1} \leq s < t \leq 1 - \delta$ and $d_t(x, p) \leq K$, then

$$H(x, t, y, s) \leq \frac{C}{(t - s)^{\frac{n}{2}}} \exp \left(-\frac{d_s^2(x, y)}{C(t - s)} \right). \tag{4.64}$$

Remark 4.17 Given $(x_0, t_0) \in M \times (-\infty, 1)$, if we set

$$H(x_0, t_0, y, s) = (4\pi(t_0 - s))^{-\frac{n}{2}} e^{-b(y, s)},$$

then it follows from Theorem 4.9 and Theorem 4.16 that $b(y, s)$ increases quadratically.

Combining (4.64) and the standard regularity theory of the parabolic equation (cf. [21]), we have the following derivative estimate of higher orders.

Corollary 4.18 Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $s_0 < t_0$, there exists a small parabolic neighborhood $P = B_{t_0}(x_0, r) \times [t_0 - r^2, t_0 + r^2]$ such that for any $m_1, m_2 \in \mathbb{N}$

$$|\partial_t^{m_1} \nabla_x^{m_2} H(x, t, y, s_0)| \leq \frac{1}{r^{2m_1+m_2}} \cdot \frac{Q}{(t_0 - s_0)^{\frac{n}{2}}} \cdot \exp \left(-\frac{d_{s_0}^2(x_0, y)}{Q(t_0 - s_0)} \right) \tag{4.65}$$

for some constant $Q > 1$ and any $(x, t) \in P$ and $y \in M$.

Note that when (y, s_0) is fixed, $H(x, t, y, s_0)$ is a heat solution. The scale r in the above Corollary is a constant much smaller than the curvature radius at (x_0, t_0) . It indicates that one can take differentiation under the integral sign if the integrand involves the heat kernel in many cases. As an application, we can follow the same proof as in Theorem 3.23 to estimate $|\nabla \mathcal{N}_s^*|$ and $\square \mathcal{N}_s^*$ without using ϕ^r . Therefore, one obtains

Corollary 4.19 The Nash entropy $\mathcal{N}_s^*(x, t)$ is smooth on $M \times (s, 1)$ satisfying

$$|\nabla \mathcal{N}_s^*| \leq \sqrt{\frac{n}{2(t-s)}} \quad \text{and} \quad -\frac{n}{2(t-s)} \leq \square \mathcal{N}_s^* \leq 0$$

in the classical sense.

We end this section by proving the following hypercontractivity; see [2, Theorem 12.1].

Theorem 4.20 Suppose that $(x_0, t_0) \in M \times (-\infty, 1)$ and $0 < \tau_1 < \tau_2$. Let $u \in C^2(M \times [t_0 - \tau_2, t_0 - \tau_1])$ be a nonnegative function satisfying $\square u \leq 0$ and having at most polynomial spatial growth in the sense that

$$|u(x, t)| \leq m(d_t^m(p, x) + 1) \tag{4.66}$$

for some $m \in \mathbb{N}$. If $1 < q_0 \leq p_0 < \infty$ with

$$\frac{\tau_2}{\tau_1} \geq \frac{p_0 - 1}{q_0 - 1},$$

then for $dv_t := dv_{x_0, t_0; t}$,

$$\left(\int_M u^{p_0} dv_{t_0 - \tau_1} \right)^{1/p_0} \leq \left(\int_M u^{q_0} dv_{t_0 - \tau_2} \right)^{1/q_0}. \tag{4.67}$$

Proof Without loss of generality, we assume $t_0 = 0$. We set $p = p(t) = 1 + \tau_2(q_0 - 1)|t|^{-1}$ for $t < 0$. Notice that $p(-\tau_2) = q_0$ and $p(-\tau_1) \geq p_0$ by our assumption.

When $t < 0$, direct calculation shows that

$$\begin{aligned} & \partial_t \int_M u^p \phi^r dv_t - \int_M u^p \square \phi^r dv_t \\ &= \int_M (\dot{p}u^p \log u + pu^{p-1} \square u - p(p-1)|\nabla u|^2 u^{p-2}) \phi^r - 2 \langle \nabla u^p, \nabla \phi^r \rangle dv_t \\ &\leq \frac{\dot{p}}{p} \int_M \phi^r u^p \log u^p dv_t - \frac{p-1}{p} \int_M \frac{|\nabla u^p|^2}{u^p} \phi^r dv_t - 2 \int_M \langle \nabla u^p, \nabla \phi^r \rangle dv_t. \end{aligned} \tag{4.68}$$

Moreover, we have for any $\epsilon \ll 1$,

$$2 \int_M |\nabla u^p| |\nabla \phi^r| dv_t \leq \epsilon \int_M \frac{|\nabla u^p|^2}{u^p} \phi^r dv_t + \epsilon^{-1} \int_M u^p |\nabla \phi^r|^2 dv_t. \tag{4.69}$$

Combining (4.68) and (4.69), we have

$$\begin{aligned} & \partial_t \left(\int_M u^p \phi^r dv_t \right)^{\frac{1}{p}} \\ &= \frac{1}{p} \left(\int_M u^p \phi^r dv_t \right)^{\frac{1}{p}-1} \left(\partial_t \int_M u^p \phi^r dv_t - \frac{\dot{p}}{p} \left(\int_M u^p \phi^r dv_t \right) \log \left(\int_M u^p \phi^r dv_t \right) \right) \\ &\leq \frac{1}{p} \left(\int_M u^p \phi^r dv_t \right)^{\frac{1}{p}-1} \left(\frac{\dot{p}}{p} \int_M \phi^r u^p \log u^p dv_t - \frac{\dot{p}}{p} \left(\int_M u^p \phi^r dv_t \right) \log \left(\int_M u^p \phi^r dv_t \right) \right) \\ &\quad + \frac{1}{p} \left(\int_M u^p \phi^r dv_t \right)^{\frac{1}{p}-1} \left(\left(\epsilon - \frac{p-1}{p} \right) \int_M \frac{|\nabla u^p|^2}{u^p} \phi^r dv_t + \epsilon^{-1} \int_M u^p \{ |\nabla \phi^r|^2 + \square \phi^r \} dv_t \right). \end{aligned} \tag{4.70}$$

We integrate (4.70) from $-\tau_2$ to $-\tau_1$, let $r \rightarrow \infty$ and then let $\epsilon \rightarrow 0$. By Theorem 4.16, (4.66), (2.12) and (2.15), we obtain

$$\begin{aligned} & \left(\int_M u^p dv_t \right)^{\frac{1}{p}} \Big|_{-\tau_2}^{-\tau_1} \\ &\leq \int_{-\tau_2}^{-\tau_1} \frac{1}{p} \left(\int_M u^p dv_t \right)^{\frac{1}{p}-1} \left(\frac{\dot{p}}{p} \int_M u^p \log u^p dv_t - \frac{\dot{p}}{p} \left(\int_M u^p dv_t \right) \log \left(\int_M u^p dv_t \right) \right) dt \\ &\quad + \int_{-\tau_2}^{-\tau_1} \frac{1}{p} \left(\int_M u^p dv_t \right)^{\frac{1}{p}-1} \left(-\frac{p-1}{p} \int_M \frac{|\nabla u^p|^2}{u^p} dv_t \right) dt. \end{aligned} \tag{4.71}$$

Note that the log-Sobolev inequality [29, Theorem 13] implies that

$$\begin{aligned} & \frac{\dot{p}}{p} \int_M u^p \log u^p dv_t - \frac{\dot{p}}{p} \left(\int_M u^p dv_t \right) \log \left(\int_M u^p dv_t \right) \\ &\leq \frac{\dot{p}}{p} |t| \int_M \frac{|\nabla u^p|^2}{u^p} dv_t = \frac{p-1}{p} \int_M \frac{|\nabla u^p|^2}{u^p} dv_t. \end{aligned}$$

Therefore, it follows from (4.71) that

$$\left(\int_M u^{p_0} dv_{-\tau_1} \right)^{\frac{1}{p_0}} \leq \left(\int_M u^{p(-\tau_1)} dv_{-\tau_1} \right)^{\frac{1}{p(-\tau_1)}} \leq \left(\int_M u^{q_0} dv_{-\tau_2} \right)^{\frac{1}{q_0}},$$

and the proof is complete. □

Remark 4.21 If $u \in C^2(M \times [t_0 - \tau_2, t_0 - \tau_1])$ and satisfies $\square u = 0$ and (4.66), then

$$\left(\int_M |u|^{p_0} dv_{t_0 - \tau_1} \right)^{1/p_0} \leq \left(\int_M |u|^{q_0} dv_{t_0 - \tau_2} \right)^{1/q_0}.$$

Indeed, one can apply (4.67) to $\sqrt{u^2 + \epsilon}$, since $\square \sqrt{u^2 + \epsilon} \leq 0$, and let $\epsilon \rightarrow 0$.

5 Parabolic Neighborhoods and ϵ -regularity Theorem

In this section, we assume that $(M^n, g(t))_{t < 1}$ is the Ricci flow associated with a Ricci shrinker in $\mathcal{M}(A)$.

Given $(x_0, t_0) \in M \times (-\infty, 1)$, we first recall the conventional parabolic neighborhoods are defined by

$$P(x_0, t_0; S, -T^-, T^+) := B_{t_0}(x_0, S) \times ([t_0 - T^-, t_0 + T^+] \cap (-\infty, 1)) \tag{5.1}$$

$$Q(x_0, t_0; S, -T^-, T^+) := \{d_t(x, x_0) \leq S, t \in [t_0 - T^-, t_0 + T^+] \cap (-\infty, 1)\} \tag{5.2}$$

for any $S, T^\pm \geq 0$. Based on the monotonicity of W_1 -distance in Proposition 3.7, we follow [2] to define the following new parabolic neighborhoods.

Definition 5.1 (P^* -parabolic neighborhoods) Suppose that $(x_0, t_0) \in M \times (-\infty, 1)$ and $S, T^\pm \geq 0$. The P^* -parabolic neighborhood $P^*(x_0, t_0; S, -T^-, T^+) \subset M \times (-\infty, 1)$ is defined as the set of points $(x, t) \in M \times (-\infty, 1)$ with $t \in [t_0 - T^-, t_0 + T^+]$ and

$$d_{W_1}^{t_0 - T^-} (v_{x_0, t_0; t_0 - T^-}, v_{x, t; t_0 - T^-}) < S.$$

For any $r > 0$, we also define

$$P^*(x_0, t_0; r) := P^*(x_0, t_0; r, -r^2, r^2)$$

$$P^{*+}(x_0, t_0; r) := P^*(x_0, t_0; r, 0, r^2)$$

$$P^{*-}(x_0, t_0; r) := P^*(x_0, t_0; r, -r^2, 0).$$

Similar definitions are also made for P^\pm .

Some basic properties of P^* -parabolic neighborhoods can be found in [2, Proposition 9.4, Corollary 9.6]. We state the following containment result from [2, Proposition 9.4 (d)].

Lemma 5.2 If $A_1, A_2, T_1^\pm, T_2^\pm \geq 0$ and $(x_1, t_1) \in P^*(x_2, t_2; A_2, -T_2^-, T_2^+)$, then

$$P^*(x_1, t_1; A_1, -T_1^-, T_1^+) \subset P^*(x_2, t_2; A_1 + A_2, -(T_1^- + T_2^-), T_1^+ + T_2^+).$$

We immediately have the following result from the distance comparison Lemma 4.8.

Lemma 5.3 Given $\delta \in (0, 1)$, $t_0 \in (-\infty, 1)$, $T^\pm \geq 0$ and $S \geq 0$, there exists a constant $C = C(n, A, \delta) > 1$ such that

$$P(p, t_0; S, -T^-, T^+) \subset Q(p, t_0; C(S + 1), -T^-, T^+)$$

$$Q(p, t_0; S, -T^-, T^+) \subset P(p, t_0; C(S + 1), -T^-, T^+)$$

provided that $-\delta^{-1} \leq t_0 - T^- \leq t_0 + T^+ \leq 1 - \delta$.

In order to investigate the relation between P^* -parabolic neighborhoods and conventional ones, we first prove

Proposition 5.4 Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, suppose $R(x_0, t) \leq r^{-2}$ for any $t \in [t_0 - r^2, t_0]$. Then

$$d_{W_1}^{t_0-r^2}(v_{x_0, t_0; t_0-r^2}, \delta_{x_0}) \leq C(n, A)r. \tag{5.3}$$

Proof It follows from [29, Theorem 16] that

$$H(x_0, t_0, x_0, t_0 - r^2) \geq \frac{1}{(4\pi r^2)^{\frac{n}{2}}} \exp(-l_{(x_0, t_0)}(x_0, t_0 - r^2)), \tag{5.4}$$

where $l_{(x_0, t_0)}(x_0, t_0 - r^2)$ denotes the reduced distance (see [29, (187)]). From the definition we have

$$l_{(x_0, t_0)}(x_0, t_0 - r^2) \leq \frac{1}{2r} \int_{t_0-r^2}^{t_0} \sqrt{t_0 - s} R(x_0, s) ds \leq \frac{1}{3}. \tag{5.5}$$

Combining (4.53) for $\epsilon = 1$, (5.4) and (5.5), it is clear that

$$d_{t_0-r^2}^2(x_0, z) \leq C_1 r^2$$

for some constant $C_1 = C_1(n, A)$, where $(z, t_0 - r^2)$ is an H_n -center of (x_0, t_0) . Therefore,

$$d_{W_1}^{t_0-r^2}(v_{x_0, t_0; t_0-r^2}, \delta_{x_0}) \leq d_{W_1}^{t_0-r^2}(v_{x_0, t_0; t_0-r^2}, \delta_z) + d_{t_0-r^2}(x_0, z) \leq C_2 r,$$

where we have used (3.29) and $C_2 := \sqrt{H_n} + \sqrt{C_1}$. □

Remark 5.5 From the proof, we conclude that (5.3) also holds for a constant $C = C(n, A, \alpha)$ if we assume

$$R(x_0, t) \leq \frac{\alpha}{r^2(t_0 - t)}$$

for some $\alpha > 0$ and any $t \in [t_0 - r^2, t_0]$.

Corollary 5.6 For any $s_0 < t_0 < 1$, we have

$$d_{W_1}^{s_0}(v_{p, t_0; s_0}, \delta_p) \leq C(n, A)\sqrt{t_0 - s_0}. \tag{5.6}$$

Proof From the self-similarity of the flow, we know that

$$R(p, t) = \frac{R(p, 0)}{1 - t} \leq \frac{n}{2(1 - t)} \leq \frac{n}{2(t_0 - t)}$$

for any $t < t_0$. Therefore, the conclusion follows from Proposition 5.4 and Remark 5.5. □

Proposition 5.7 Given $\delta \in (0, 1)$, $t_0 \in (-\infty, 1)$, $T^\pm \geq 0$ and $S \geq 0$, there exists a constant $C = C(n, A, \delta) > 1$ such that

$$Q(p, t_0; S, -T^-, T^+) \subset P^*(p, t_0; S + C, -T^-, T^+)$$

provided that $t_0 - T^- \geq -\delta^{-1}$.

Proof For any $(x, t) \in Q(p, t_0; S, -T^-, T^+)$, we have

$$\begin{aligned} & d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, v_{x, t; t_0-T^-}) \\ & \leq d_{W_1}^{t_0-T^-}(v_{p, t; t_0-T^-}, v_{x, t; t_0-T^-}) + d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, v_{p, t; t_0-T^-}) \\ & \leq d_t(x, p) + d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, v_{p, t; t_0-T^-}) \leq S + d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, v_{p, t; t_0-T^-}), \end{aligned} \tag{5.7}$$

where we have used Proposition 3.7. In addition, it follows from Corollary 5.6 that

$$d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, v_{p, t; t_0-T^-}) \leq d_{W_1}^{t_0-T^-}(v_{p, t_0; t_0-T^-}, \delta_p) + d_{W_1}^{t_0-T^-}(v_{p, t; t_0-T^-}, \delta_p) \leq C(n, A, \delta). \tag{5.8}$$

Therefore, the conclusion follows from (5.7) and (5.8). \square

Next, we recall the following version of the local distance distortion estimate, which can be proven almost exactly as [29, Theorem 18]; see also [16, Section 4.3], [17, Theorem 3.1] and [6, Theorem 1.1].

Lemma 5.8 Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, suppose $R \leq r^{-2}$ on $P^-(x_0, t_0; r)$ (resp. $P(x_0, t_0; r)$). Then

$$\rho_1 d_t(x, x_0) \leq d_{t_0}(x, x_0) \leq \rho_1^{-1} d_t(x, x_0)$$

if $d_{t_0}(x, x_0) \leq \rho_1 r$ and $t \in [t_0 - (\rho_1 r)^2, t_0]$ (resp. $t \in [t_0 - (\rho_1 r)^2, t_0 + (\rho_1 r)^2] \cap (-\infty, 1)$), where $\rho_1 = \rho_1(n, A) \in (0, 1)$. In particular,

$$P^-(x_0, t_0; \rho_1^2 r) \subset Q^-(x_0, t_0; \rho_1 r) \subset P^-(x_0, t_0; r) \quad (5.9)$$

$$\text{(resp. } P(x_0, t_0; \rho_1^2 r) \subset Q(x_0, t_0; \rho_1 r) \subset P(x_0, t_0; r)\text{)} . \quad (5.10)$$

Thanks to Proposition 5.4 and Lemma 5.8, we have the following result.

Proposition 5.9 There exists a constant $\rho_2 = \rho_2(n, A) \in (0, 1)$ satisfying the following property. Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $R \leq r^{-2}$ on $P(x_0, t_0; r, -(\rho_2 r)^2, 0)$ (resp. $P(x_0, t_0; r, -(\rho_2 r)^2, (\rho_2 r)^2)$). Then

$$P^-(x_0, t_0; \rho_2 r) \subset P^*(x_0, t_0; r, -(\rho_2 r)^2, 0) \quad (5.11)$$

$$\text{(resp. } P(x_0, t_0; \rho_2 r) \subset P^*(x_0, t_0; r, -(\rho_2 r)^2, (\rho_2 r)^2)\text{)} . \quad (5.12)$$

Proof In the proof, all constants $C_i > 1$ depend on n and A , and ρ_1 is from Lemma 5.8. We only prove (5.11), and the proof of (5.12) is similar. Moreover, we set $0 < \tau \ll 1$ to be determined later.

For any $(y, s) \in P^-(x_0, t_0; \tau r)$, it follows from Lemma 5.8 that

$$d_t(y, x_0) \leq C_1 \tau r \quad (5.13)$$

for any $t \in [t_0 - (\tau r)^2, s]$. In particular, $R(y, t) \leq r^{-2}$ for any $t \in [t_0 - (\tau r)^2, s]$. Therefore, it follows from Proposition 5.4 that

$$d_{W_1}^{t_0 - (\tau r)^2}(v_{y, s; t_0 - (\tau r)^2}, \delta_y) \leq C_2 \tau r. \quad (5.14)$$

It follows from (5.13) and (5.14) that

$$\begin{aligned} & d_{W_1}^{t_0 - (\tau r)^2}(v_{y, s; t_0 - (\tau r)^2}, v_{x_0, t_0; t_0 - (\tau r)^2}) \\ & \leq d_{W_1}^{t_0 - (\tau r)^2}(v_{x_0, t_0; t_0 - (\tau r)^2}, \delta_{x_0}) + d_{W_1}^{t_0 - (\tau r)^2}(v_{y, s; t_0 - (\tau r)^2}, \delta_y) + d_{t_0 - (\tau r)^2}(y, x_0) \leq C_3 \tau r < r, \end{aligned}$$

if τ is sufficiently small. From this, it is immediate that (5.11) holds for small ρ_2 . \square

Now, we prove

Proposition 5.10 Given $\delta \in (0, 1)$, $t_0 \in (-\infty, 1)$, $T^\pm \geq 0$ and $S \geq 0$, there exists a constant $C = C(n, A, \delta) > 1$ such that

$$P^*(p, t_0; S, -T^-, T^+) \subset Q(p, t_0; \sqrt{2}S + C, -T^-, T^+) \quad (5.15)$$

provided that $t_0 - T^- \geq -\delta^{-1}$. In particular, it implies that $P^*(p, t_0; S, -T^-, T^+)$ is precompact in $M \times (-\infty, 1)$ if $t_0 + T^+ < 1$.

Proof In the proof, all constants $C_i > 1$ depend on n, A and δ . It follows from (5.6) that

$$d_{W_1}^{t_0-T^-}(\delta_p, v_{p,t_0;t_0-T^-}) \leq C_1. \tag{5.16}$$

For any $(x_1, t_1) \in P^*(p, t_0; S, -T^-, T^+)$, we assume $(z, t_0 - T^-)$ to be an H_n -center of (x_1, t_1) . By (5.16) and the definition of P^* neighborhood, we have

$$\begin{aligned} d_{t_0-T^-}(p, z) &\leq d_{W_1}^{t_0-T^-}(\delta_p, v_{p,t_0;t_0-T^-}) + d_{W_1}^{t_0-T^-}(v_{p,t_0;t_0-T^-}, \delta_z) \\ &\leq d_{W_1}^{t_0-T^-}(\delta_z, v_{x_1,t_1;t_0-T^-}) + d_{W_1}^{t_0-T^-}(v_{p,t_0;t_0-T^-}, v_{x_1,t_1;t_0-T^-}) + C_1 \\ &\leq S + C_2. \end{aligned} \tag{5.17}$$

Set $v_t = v_{x_1,t_1;t}$ and compute

$$\partial_t \int_M \phi^r dv_t = \int_M \square \phi^r dv_t.$$

By (2.15), we have

$$\phi^r(x_1, t_1) \geq \int_M \phi^r dv_{t_0-T^-} - C(n)r^{-1}(t_1 - t_0 + T^-) \geq \int_{F \leq r} 1 dv_{t_0-T^-} - C_3r^{-1}.$$

Note that $\phi^r = 1$ if $F \leq r$ and r is large. In light of (5.17) and Lemma 2.3, the set $\{F \leq r\}$ contains a large geodesic ball centered at z . Thus by Proposition 3.13, the above inequality implies that

$$\phi^r(x_1, t_1) \geq \int_{F \leq r} 1 dv_{t_0-T^-} - C_3r^{-1} \geq \frac{1}{2} \tag{5.18}$$

if $2\sqrt{r} = S + C_4$. Since ϕ^r is supported on $F \leq 2r$, we conclude from (5.18) that

$$F(x_1, t_1) \leq 2r = \frac{(S + C_4)^2}{2}.$$

From Lemma 2.3, we immediately conclude that

$$d_{t_1}(p, x_1) \leq \sqrt{2}S + C_5.$$

Now, the last conclusion follows from (5.15) and Lemma 5.3. □

Corollary 5.11 Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $S, T^\pm \geq 0$, $P^*(x_0, t_0; S, -T^-, T^+)$ is precompact in $M \times (-\infty, 1)$ if $t_0 + T^+ < 1$.

Proof It is clear that $(x_0, t_0) \in P^*(p, t_0; S', -1, 0)$ for some large $S' > 0$. Therefore, it follows from Lemma 5.2

$$P^*(x_0, t_0; S, -T^-, T^+) \subset P^*(p, t_0; S + S', -(1 + T^-), T^+).$$

Therefore, the conclusion follows from Proposition 5.10. □

Next, we recall the following existence of the local cutoff function from [6, Theorem 1.3].

Proposition 5.12 Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, there exists a constant $\rho_3 = \rho_3(n, A) \in (0, 1)$ satisfying the following property.

Suppose $R \leq r^{-2}$ on $P(x_0, t_0; r, -\tau, 0)$ with $0 < \tau \leq (\rho_3 r)^2$. Then there exists a function $\varphi \in C^\infty(M \times [t_0 - \tau, t_0])$ with the following properties:

- (a) $0 \leq \varphi < 1$ on $M \times [t_0 - \tau, t_0]$.
- (b) $\varphi > \rho_3$ on $P(x_0, t_0; \rho_3 r, -\tau, 0)$.
- (c) $\varphi = 0$ outside $P(x_0, t_0; r, -\tau, 0)$.

- (d) $|\nabla\varphi| \leq r^{-1}$ and $|\partial_t\varphi| + |\Delta\varphi| \leq r^{-2}$.
- (e) $\square\varphi \leq 0$ on $M \times [t_0 - \tau, t_0]$.

Proof We sketch the proof for the readers' convenience. In [6, Theorem 1.3], φ is constructed as the smoothing of ψ^2 , where

$$\psi(x, t) := c_1 \max\{K(x, t) - c, 0\}$$

for some constants $c, c_1 > 0$ on $U \times [t_0 - \tau, t_0]$ for some open set $U \subset B_{t_0}(x_0, r)$, where $\psi = 0$ on $\partial U \times [t_0 - \tau, t_0]$ and can be extended to be 0 outside $U \times [t_0 - \tau, t_0]$. Here, $K(x, t) = H(x, t, y, s)$ for some appropriate (y, s) such that $K(x_0, t_0) \geq (4\pi(t_0 - s))^{-\frac{n}{2}} e^{-n/2}$, and $t_0 - s$ is sufficiently small.

The estimates of (a)–(d) follow from [29, Lemma 20]. From the definitions of ψ and φ , it is clear that (e) also holds. □

Next, we prove

Proposition 5.13 There exists a constant $\rho_4 = \rho_4(n, A) \in (0, 1)$ satisfying the following property. Given $(x_0, t_0) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $R \leq r^{-2}$ on $P(x_0, t_0; r, -(\rho_4 r)^2, 0)$ (resp. $P(x_0, t_0; r, -(\rho_4 r)^2, (\rho_4 r)^2)$). Then

$$P^{*-}(x_0, t_0; \rho_4 r) \subset P(x_0, t_0; r, -(\rho_4 r)^2, 0) \tag{5.19}$$

$$\text{(resp. } P^*(x_0, t_0; \rho_4 r) \subset P(x_0, t_0; r, -(\rho_4 r)^2, (\rho_4 r)^2)\text{)}. \tag{5.20}$$

Proof In the proof, all positive constants $C_i > 1$ depend only on n and A . Moreover, we set $0 < \tau \ll 1$ to be determined later.

For any $(y, s) \in P^{*-}(x_0, t_0; \tau r)$, we assume $(z, t_0 - (\tau r)^2)$ to be its H_n -center. From Proposition 5.4, we have

$$\begin{aligned} d_{t_0 - (\tau r)^2}(z, x_0) &\leq d_{W_1}^{t_0 - (\tau r)^2}(\delta_{x_0}, v_{x_0, s; t_0 - (\tau r)^2}) + d_{W_1}^{t_0 - (\tau r)^2}(v_{y, s; t_0 - (\tau r)^2}, v_{x_0, s; t_0 - (\tau r)^2}) \\ &\quad + d_{W_1}^{t_0 - (\tau r)^2}(\delta_z, v_{y, s; t_0 - (\tau r)^2}) \\ &\leq C_1 \tau r. \end{aligned} \tag{5.21}$$

We assume $\tau < \rho_3$ and consider the cutoff function φ constructed in Proposition 5.12. If we set $v_t = v_{y, s; t}$, then by direct computation,

$$\partial_t \int_M \varphi dv_t = \int_M \square\varphi dv_t \geq -r^{-2},$$

where we have used Proposition 5.12(d). By integration, we have

$$\varphi(y, s) \geq \int_M \varphi dv_{t_0 - (\tau r)^2} - \tau. \tag{5.22}$$

Notice that $\varphi > \rho_3$ on $P(x_0, t_0; \rho_3 r, 0, (\tau r)^2)$. Combining this fact with (5.21) and Proposition 3.13, we conclude that if τ is sufficiently small,

$$\varphi(y, s) \geq \int_M \varphi dv_{t_0 - (\tau r)^2} - \tau \geq \frac{\rho_3}{2} > 0.$$

On the other hand, since $\varphi = 0$ outside $P(x_0, t_0; r, -(\tau r)^2)$, we conclude that

$$d_{t_0}(x_0, y) \leq r,$$

and hence (5.19) holds. □

Next, we recall the definition of the curvature radius.

Definition 5.14 (Curvature radius) For any $(x, t) \in M \times (-\infty, 1)$, the curvature radii at (x, t) are defined as

$$\begin{aligned} r_{\text{Rm}}(x, t) &:= \sup \{ r > 0 \mid |\text{Rm}| \leq r^{-2} \text{ on } P(x, t; r) \}, \\ r_{\text{Rm}}^-(x, t) &:= \sup \{ r > 0 \mid |\text{Rm}| \leq r^{-2} \text{ on } P^-(x, t; r) \}, \\ r_{\text{Rm}}^s(x, t) &:= \sup \{ r > 0 \mid |\text{Rm}| \leq r^{-2} \text{ on } B_t(x, r) \}. \end{aligned}$$

It is clear from the definition that $r_{\text{Rm}}(x, t) \leq r_{\text{Rm}}^-(x, t) \leq r_{\text{Rm}}^s(x, t)$. In addition, it follows from Theorem 4.1 and the pseudolocality theorem [29, Theorem 24] on Ricci shrinkers that there exists a constant $C = C(n, A) > 1$ such that

$$r_{\text{Rm}}^-(x, t) \leq Cr_{\text{Rm}}(x, t). \tag{5.23}$$

We are in a position to obtain the following ϵ -regularity theorem; see [2, Theorem 10.2].

Theorem 5.15 (ϵ -regularity) There exists a small constant $\epsilon = \epsilon(n) > 0$ satisfying the following property. Given $(x, t) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $\mathcal{N}_{(x,t)}(r^2) \geq -\epsilon$, then $r_{\text{Rm}}(x, t) \geq \epsilon r$.

Proof We only sketch the proof as the details can be found in [2, Theorem 10.2]. The key step is a point-picking argument in the spacetime with respect to the curvature radius r_{Rm} . More precisely, one needs to show that for any $A > 0$ with $10Ar_{\text{Rm}}(x, t) \leq 1/2$, there exists a point $(x', t') \in P^{*-}(x', t'; 10Ar_{\text{Rm}}(x, t))$ such that $r_{\text{Rm}}(x', t') \leq r_{\text{Rm}}(x, t)$ and $r_{\text{Rm}} \geq r_{\text{Rm}}(x', t')/10$ on $P^{*-}(x', t'; Ar_{\text{Rm}}(x', t'))$. Otherwise, one can iteratively pick a sequence of spacetime points (x_i, t_i) in a compact set of $M \times (-\infty, 1)$ satisfying $r_{\text{Rm}}(x_i, t_i) \rightarrow 0$. In light of Lemma 5.2, all (x_i, t_i) fall into a given P^{*-} -parabolic neighborhood, which is precompact by Corollary 5.11. Note that the curvature radius of (x_i, t_i) shrinks by a definite portion in each step, and the bounded geometry of a compact set implies that the process must terminate in finite steps, say $(x_k, t_k) = (x', t')$. Such a choice of (x', t') guarantees that it has an almost maximal curvature radius in the spacetime neighborhood. Notice that similar point-picking arguments can be found in [35, Theorem 10.1] and [16, Proposition 3.43].

If the ϵ -regularity theorem fails, we could obtain a sequence of pointed Ricci flows such that $r_{\text{Rm}} = 1$ at the base points after the point-picking and appropriate rescalings. Since nearby points have curvature radii uniformly bounded from below, the sequence converges smoothly to a limit Ricci flow which is the Euclidean spacetime by the assumption of the Nash entropy. Therefore, $r_{\text{Rm}} = 1$ must be violated and we obtain a contradiction. □

Using the ϵ -regularity theorem, one immediately has the following gap property, following the same proof of [29, Theorem 3].

Corollary 5.16 Suppose (M^n, g, f, p) is a non-flat Ricci shrinker. Then

$$\mathcal{N}_{(p,0)}(\epsilon^{-2}) < -\epsilon,$$

where ϵ is the same constant in Theorem 5.15.

Proof Suppose $\mathcal{N}_{(p,0)}(\epsilon^{-2}) \geq -\epsilon$ and (M, g) is a non-flat Ricci shrinker. Then it follows from Theorem 5.15 that $r_{\text{Rm}}(p, 0) \geq 1$. In particular, it implies that $|\text{Rm}(p, t)| \leq 1$ for any

$t \in [0, 1)$. By the self-similarity of the flow, we have $|\text{Rm}|(p, 0) = |\text{Rm}|(p, t)(1 - t)$, and hence $|\text{Rm}|(p, 0) = 0$, which contradicts the fact that $R > 0$ for non-flat Ricci shrinkers. \square

We conclude this section by stating the following two results, whose proofs are more or less standard; see [2, Theorem 10.3, Theorem 10.4].

Theorem 5.17 For any $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$ such that the following holds. Given $(x, t) \in M \times (-\infty, 1)$ and $r > 0$, if $\mathcal{N}_{(x,t)}(r^2) \geq -\delta$, then

$$|\text{Rm}| \leq \epsilon r^{-2} \quad \text{on} \quad P(x, t; \epsilon^{-1}r, -(1 - \epsilon)r^2, \epsilon^{-1}r^2).$$

Moreover, we have $\mathcal{N}_{t-r^2}^* \geq -\epsilon$ on $P(x, t; \epsilon^{-1}r, -(1 - \epsilon)r^2)$.

Theorem 5.18 For any $\epsilon > 0$ and $Y < \infty$ there is a $\delta = \delta(\epsilon, Y) > 0$ such that the following holds. Given $(x, t) \in M \times (-\infty, 1)$ and $r > 0$, suppose that $|\text{Rm}| \leq r^{-2}$ on $P^-(x, t; r)$ and $\mathcal{N}_{(x,t)}(r^2) \geq -Y$. Then $\mathcal{N}_{(x,t)}(\delta r^2) \geq -\epsilon$.

6 Metric Flows and \mathbb{F} -convergence

In previous sections, we have generalized (or slightly improved) the theorems and tools in [2]. Notice that these results also hold for Ricci flows induced by Ricci shrinkers (cf. Definition 2.2), since most of them are scaling-invariant. In a few cases, one needs to modify the assumptions correspondingly. For instance, the conditions in Theorem 4.9 and Theorem 4.16 need to be changed to $-\delta^{-1}\lambda \leq t < s \leq (1 - \delta)\lambda$ and $d_t(x, p) \leq K\lambda^{1/2}$, if the Ricci flow associated with a Ricci shrinker is parabolically rescaled by $\lambda > 0$.

Based on these results and techniques, one can generalize the theory of \mathbb{F} -convergence in [3] and [4] from compact Ricci flows to the setting of Ricci flows induced by Ricci shrinkers.

Notice that the results in [3] and [4] are already generalized by Bamler to Ricci flows with complete time-slices and bounded curvature on compact time-intervals (cf. [5]). In [5, Appendix A], some issues in the non-compact case are addressed and can be resolved similarly in the setting of Ricci shrinkers by the results and techniques developed in previous sections. For instance, by Theorem 4.9 and Theorem 4.16, it is known that the conjugate heat kernel decays exponentially and the function b induced by the conjugate heat kernel increases quadratically (cf. Remark 4.17). Therefore, the weak splitting maps (cf. [4, Definition 5.6]) constructed in [4, Section 10] have at most quadratic spatial growth. Moreover, it follows from [4, Proposition 12.1, Remark 12.3] that one can construct a bounded strong splitting map with the bounded gradient from a given weak splitting map.

At various places in [4], one also needs to consider the integral $\int u\phi^r$ instead of $\int u$, and take the limit for $r \rightarrow \infty$ after all the estimates (e.g., $u = \square|\omega_t|$ in [4, Lemma 17.37]). This technique has already appeared multiple times in previous sections. As a showcase, we generalize the integral estimates in [4, Section 6] to Ricci flows associated with Ricci shrinkers in Appendix 6. These estimates are frequently used in [4] and are of independent interest.

Now, we recall the following definition of the metric flow from [3, Definition 3.2].

Definition 6.1 (Metric flow) Let $I \subset \mathbb{R}$ be a subset. A metric flow over I is a tuple of the form

$$(\mathcal{X}, \mathfrak{t}, (d_t)_{t \in I}, (v_{x;s})_{x \in \mathcal{X}, s \in I, s \leq \mathfrak{t}(x)})$$

with the following properties:

- (1) \mathcal{X} is a set consisting of points.
- (2) $\mathfrak{t} : \mathcal{X} \rightarrow I$ is a map called time-function. Its level sets $\mathcal{X}_t := \mathfrak{t}^{-1}(t)$ are called time-slices and the preimages $\mathcal{X}_{I'} := \mathfrak{t}^{-1}(I')$, $I' \subset I$, are called time-slabs.
- (3) (\mathcal{X}_t, d_t) is a complete and separable metric space for all $t \in I$.
- (4) $v_{x;s}$ is a probability measure on \mathcal{X}_s for all $x \in \mathcal{X}$, $s \in I$, $s \leq \mathfrak{t}(x)$. For any $x \in \mathcal{X}$ the family $(v_{x;s})_{s \in I, s \leq \mathfrak{t}(x)}$ is called the conjugate heat kernel at x .
- (5) $v_{x;\mathfrak{t}(x)} = \delta_x$ for all $x \in \mathcal{X}$.
- (6) For all $s, t \in I$, $s < t$, $T \geq 0$ and any measurable function $u_s : \mathcal{X}_s \rightarrow [0, 1]$ with the property that if $T > 0$, then $u_s = \Phi \circ f_s$ for some $T^{-1/2}$ -Lipschitz function $f_s : \mathcal{X}_s \rightarrow \mathbb{R}$ (if $T = 0$, then there is no additional assumption on u_s), the following is true. The function

$$u_t : \mathcal{X}_t \longrightarrow \mathbb{R}, \quad x \longmapsto \int_{\mathcal{X}_s} u_s dv_{x;s}$$

is either constant or of the form $u_t = \Phi \circ f_t$, where $f_t : \mathcal{X}_t \rightarrow \mathbb{R}$ is $(t - s + T)^{-1/2}$ -Lipschitz. Here, Φ is given by (3.31).

- (7) For any $t_1, t_2, t_3 \in I$, $t_1 \leq t_2 \leq t_3$, $x \in \mathcal{X}_{t_3}$ we have the reproduction formula

$$v_{x;t_1} = \int_{\mathcal{X}_{t_2}} v_{\cdot;t_1} dv_{x;t_2},$$

meaning that for any Borel set $S \subset \mathcal{X}_{t_1}$

$$v_{x;t_1}(S) = \int_{\mathcal{X}_{t_2}} v_{y;t_1}(S) dv_{x;t_2}(y).$$

Given a metric flow \mathcal{X} over I , we recall the following definitions from [3, Definition 3.20, 3.30].

Definition 6.2 (Conjugate heat flow) A family of probability measures $(\mu_t \in \mathcal{P}(\mathcal{X}_t))_{t \in I'}$ over $I' \subset I$ is called a conjugate heat flow if for all $s, t \in I'$, $s \leq t$ we have

$$\mu_s = \int_{\mathcal{X}_t} v_{x;s} d\mu_t(x).$$

Definition 6.3 (H -Concentration) Given a constant $H > 0$, a metric flow \mathcal{X} is called H -concentrated if for any $s \leq t$, $s, t \in I$, $x_1, x_2 \in \mathcal{X}_t$

$$\text{Var}(v_{x_1;s}, v_{x_2;s}) \leq d_t^2(x_1, x_2) + H(t - s).$$

Next, we recall the definition of the metric flow pair from [3, Definition 5.1, 5.2]. Roughly speaking, two metric flow pairs are equivalent if they are the same in the metric measure sense almost everywhere.

Definition 6.4 (Metric flow pair) A pair $(\mathcal{X}, (\mu_t)_{t \in I'})$ is called a metric flow pair over $I \subset \mathbb{R}$ if

- 1. $I' \subset I$ with $|I \setminus I'| = 0$.
- 2. \mathcal{X} is a metric flow over I' .
- 3. $(\mu_t)_{t \in I'}$ is a conjugate heat flow on \mathcal{X} with $\text{supp } \mu_t = \mathcal{X}_t$ for all $t \in I'$.

If $J \subset I'$, then we say that $(\mathcal{X}, (\mu_t)_{t \in I'})$ is fully defined over J . We denote by \mathbb{F}_I^J the set of equivalence classes of metric flow pairs over I that are fully defined over J . Here, two metric flow pairs $(\mathcal{X}^i, (\mu_t^i)_{t \in I', i})$, $i = 1, 2$, that are fully defined over J are equivalent if there exists an

isometry $\phi : \mathcal{X}_I^1 \rightarrow \mathcal{X}_I^2$ (cf. [3, Definition 3.13]) such that $|I'^1 \setminus I'| = |I'^2 \setminus I'| = 0$, $(\phi_t)_* \mu_t^1 = \mu_t^2$ for all $t \in I'$ and $J \subset I'$.

We will only consider $I := (-\infty, 0]$, for simplicity. Then, for any pointed Ricci flow $(M^n, g(t), x_0)_{t \in I}$ induced by a Ricci shrinker, one can define $(\mathcal{X}, (\mu_t)_{t \in I})$ as follows.

$$(\mathcal{X} := M \times (I \setminus \{0\}) \sqcup x_0 \times \{0\}, \mathfrak{t} := \text{proj}_I, (d_t)_{t \in I}, (v_{x,t;s})_{(x,t) \in M \times I, s \in I, s \leq t}, \mu_t := v_{x_0,0;t}). \tag{6.1}$$

Then we have

Proposition 6.5 The pair $(\mathcal{X}, (\mu_t)_{t \in I})$ defined in (6.1) is an H_n -concentrated metric flow pair that is fully defined over I .

Proof The conditions (1)–(5) in the definition of the metric flow can be easily checked. Condition (6) follows from Theorem 3.15 and (7) from the semigroup property (3.1). The metric flow is H_n -concentrated by Proposition 3.10. \square

Next, we recall the definition of a correspondence between metric flows; see [3, Definition 5.4].

Definition 6.6 (Correspondence) Let $(\mathcal{X}^i, (\mu_t^i)_{t \in I',i})$ be metric flows over I , indexed by some $i \in \mathcal{I}$. A correspondence between these metric flows over I'' is a pair of the form

$$\mathfrak{C} := ((Z_t, d_t^Z)_{t \in I''}, (\varphi_t^i)_{t \in I'',i \in \mathcal{I}}),$$

where

1. (Z_t, d_t^Z) is a metric space for any $t \in I''$.
 2. $I'',i \subset I'' \cap I',i$ for any $i \in \mathcal{I}$.
 3. $\varphi_t^i : (\mathcal{X}_t^i, d_t^i) \rightarrow (Z_t, d_t^Z)$ is an isometric embedding for any $i \in \mathcal{I}$ and $t \in I'',i$.
- If $J \subset I'',i$ for all $i \in \mathcal{I}$, we say that \mathfrak{C} is fully defined over J .

Given a correspondence, one can define the \mathbb{F} -distance; see [3, Definition 5.6, 5.8].

Definition 6.7 (\mathbb{F} -distance within correspondence) We define the \mathbb{F} -distance between two metric flow pairs within \mathfrak{C} (uniform over J),

$$d_{\mathbb{F}}^{\mathfrak{C},J}((\mathcal{X}^1, (\mu_t^1)_{t \in I',1}), (\mathcal{X}^2, (\mu_t^2)_{t \in I',2})),$$

to be the infimum over all $r > 0$ with the property that there is a measurable subset $E \subset I''$ with

$$J \subset I'' \setminus E \subset I'',1 \cap I'',2$$

and a family of couplings $(q_t)_{t \in I'' \setminus E}$ between μ_t^1, μ_t^2 such that

- (1) $|E| \leq r^2$.
- (2) For all $s, t \in I'' \setminus E$, $s \leq t$, we have

$$\int_{\mathcal{X}_t^1 \times \mathcal{X}_t^2} d_{W_1}^{Z_s}((\varphi_s^1)_* \nu_{x^1;s}^1, (\varphi_s^2)_* \nu_{x^2;s}^2) dq_t(x^1, x^2) \leq r.$$

Notice that (2) above implies that for any $t \in I'' \setminus E$,

$$d_{GW_1}((\mathcal{X}_t^1, d_t^1, \mu_t^1), (\mathcal{X}_t^2, d_t^2, \mu_t^2)) \leq d_{W_1}^{Z_t}((\varphi_t^1)_* \mu_t^1, (\varphi_t^2)_* \mu_t^2) \leq r. \tag{6.2}$$

Here, d_{GW_1} denotes the Gromov- W_1 -Wasserstein distance; see [3, Definition 2.11] for the precise definition.

Definition 6.8 (\mathbb{F} -distance) The \mathbb{F} -distance between two metric flow pairs (uniform over J),

$$d_{\mathbb{F}}^J((\mathcal{X}^1, (\mu_t^1)_{t \in I^{',1}}), (\mathcal{X}^2, (\mu_t^2)_{t \in I^{',2}})),$$

is defined as the infimum of

$$d_{\mathbb{F}}^{\mathfrak{C},J}((\mathcal{X}^1, (\mu_t^1)_{t \in I^{',1}}), (\mathcal{X}^2, (\mu_t^2)_{t \in I^{',2}})),$$

over all correspondences \mathfrak{C} between $\mathcal{X}^1, \mathcal{X}^2$ over I'' that are fully defined over J .

With all those definitions, it can be proven (cf. [3, Theorem 5.13, 5.26]) that $(\mathbb{F}_I^J, d_{\mathbb{F}}^J)$ is a complete metric space, with possible infinite distances.

In addition, \mathbb{F} -convergence implies \mathbb{F} -convergence within a correspondence; see [3, Theorem 6.12]. More precisely,

Theorem 6.9 Let $(\mathcal{X}^i, (\mu_t^i)_{t \in I^{',i}}), i \in \mathbb{N} \cup \{\infty\}$ be metric flow pairs over I that are fully defined over some $J \subset I$. Suppose that for any compact subinterval $I_0 \subset I$

$$d_{\mathbb{F}}^{J \cap I_0}((\mathcal{X}^i, (\mu_t^i)_{t \in I_0 \cap I^{',i}}), (\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I_0 \cap I^{',\infty}})) \rightarrow 0.$$

Then there is a correspondence \mathfrak{C} between the metric flows $\mathcal{X}^i, i \in \mathbb{N} \cup \{\infty\}$ over I such that

$$(\mathcal{X}^i, (\mu_t^i)_{t \in I^{',i}}) \xrightarrow[i \rightarrow \infty]{\mathbb{F}, \mathfrak{C}, J} (\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I^{',\infty}})$$

on compact time intervals, in the sense that

$$d_{\mathbb{F}}^{\mathfrak{C}, J \cap I_0}((\mathcal{X}^i, (\mu_t^i)_{t \in I_0 \cap I^{',i}}), (\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I_0 \cap I^{',\infty}})) \rightarrow 0$$

for any compact subinterval $I_0 \subset I$.

For a sequence of Ricci flows $(M_i^n, g_i(t), x_i)_{t \in I}$ induced by Ricci shrinkers, one can use the \mathbb{F} -compactness theorem for metric flow pairs [3, Corollary 7.5, Theorem 7.6] to obtain the following result.

Theorem 6.10 (\mathbb{F} -compactness) Let $(M_i^n, g_i(t), x_i)_{t \in I}$ be a sequence of pointed Ricci flows induced by Ricci shrinkers with the corresponding metric flow pairs $(\mathcal{X}^i, (\mu_t^i)_{t \in I})$ as described in (6.1).

After passing to a subsequence, there exists an H_n -concentrated metric flow pair $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ for which \mathcal{X}^∞ is future continuous in the sense of [3, Definition 4.25] such that the following holds. There is a correspondence \mathfrak{C} between the metric flows $\mathcal{X}^i, i \in \mathbb{N} \cup \{\infty\}$, over I such that on compact time-intervals

$$(\mathcal{X}^i, (\mu_t^i)_{t \in I}) \xrightarrow[i \rightarrow \infty]{\mathbb{F}, \mathfrak{C}} (\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I}). \tag{6.3}$$

Moreover, the convergence (6.3) is uniform over any compact $J \subset I$ that only contains times at which \mathcal{X}^∞ is continuous; see [3, Definition 4.25]. Notice that \mathcal{X}^∞ is continuous everywhere except possibly at a countable set of times, by [3, Corollary 4.35].

We sketch the main ideas and steps of Theorem 6.10 modulo all technical details.

1. One needs a characterization of the compactness for a subset in (\mathbb{M}, d_{GW_1}) , the isometry classes of all metric measure space (X, d, μ) , where $\mu \in \mathcal{P}(X)$ with $\text{supp } \mu = X$ and d_{GW_1} denotes the Gromov- W_1 -Wasserstein distance (cf. [3, Definition 2.11]). Let $\mathbb{M}_r(V, b) \subset \mathbb{M}$ be the subset consisting of (X, d, μ) satisfying

$$\text{Var}(\mu) \leq Vr^2 \quad \text{and} \quad \mu(\{x \in X \mid \mu(D(x, \epsilon r)) < b(\epsilon)\}) \leq \epsilon, \quad \forall \epsilon \in (0, 1]. \tag{6.4}$$

Here, V, r are two positive constants and $b : (0, 1] \rightarrow (0, 1]$ is a function. Moreover, $D(x, \epsilon r)$ denotes a closed ball with center x and radius ϵr . It is proven by [3, Theorem 2.27] that $\mathbb{M}_r(V, b)$ is compact.

2. For any metric flow pair $(\mathcal{X}, (\mu_t)_{t \in I})$ defined in (6.1). It is clear by H_n -concentration that $\text{Var}(\mu_t) \leq H_n|t|$. It can be proven (cf. [3, Proposition 4.1] with $\tau = \frac{\epsilon^3}{8H_n}$) that for any $t < 0$,

$$(\mathcal{X}_t, d_t, \mu_t) \in \mathbb{M}_r(V, b), \quad (6.5)$$

where $V = 1/8$, $r = \sqrt{8H_n|t|}$, $b(\epsilon) = \Phi(\epsilon^{-2}\sqrt{8H_n})/2$ and Φ is given by (3.31). The proof of (6.5) uses Definition 6.1(6)(7) in an essential way. Therefore, for any $t \leq 0$, $(\mathcal{X}^i, d_t^i, \mu_t^i)$ subconverges in GW_1 to a limit metric measure space.

3. To compare different time-slices of $(\mathcal{X}, (\mu_t)_{t \in I})$, one considers the function

$$D(t) := \int_{\mathcal{X}_t} \int_{\mathcal{X}_t} d_t d\mu_t d\mu_t \quad (6.6)$$

for $t \in I$. It is not hard to prove (cf. [3, Lemma 4.7]) that for any $s \leq t \in I$,

$$-\sqrt{H_n(t-s)} \leq D(t) - D(s) \leq \sqrt{\text{Var}(\mu_t) - \text{Var}(\mu_s) + H_n(t-s)} + 2\sqrt{H_n(t-s)}. \quad (6.7)$$

It follows immediately from (6.7) that $D(t)$ is continuous on a complement of a countable subset of I . In addition, it is proven (cf. [3, Theorem 4.31]) that for any $t_0 \leq 0$, $D(t)$ is continuous at t_0 is equivalent to the statement that $(\mathcal{X}_t, d_t, \mu_t)$ is continuous at t_0 in the GW_1 sense. In this case, one can construct an isometric embedding of (\mathcal{X}_t, d_t) and $(\mathcal{X}_{t_0}, d_{t_0})$ into a metric space (Z_t, d_t^Z) with an explicit coupling q_t between μ_t and μ_{t_0} for t close to t_0 . Therefore, one concludes that the metric flow (\mathcal{X}_t, μ_t) is continuous on I except at a countable set of times.

4. For the sequence $(\mathcal{X}^i, (\mu_t^i)_{t \in I})$ in (6.3), we consider the limit

$$D^\infty(t) := \lim_{i \rightarrow \infty} D^i(t), \quad (6.8)$$

which exists for any $t \in I$ by taking a subsequence. Indeed, by (6.7), we may assume that $D^\infty(t)$ exists for $t \in I \cap \mathbb{Q}$ and $D^\infty(t) - D^\infty(s) \geq -\sqrt{H_n(t-s)}$ for any $s, t \in I \cap \mathbb{Q}$ with $s \leq t$, after taking a subsequence if necessary. Therefore, there exists a countable set $S \subset I$ such that D^∞ is continuous on $I \setminus S$, by extending the definition of D^∞ . Moreover, (6.8) holds for any $t \in I \setminus S$. Now, (6.8) also holds for $t \in S$, by further taking a subsequence.

The \mathbb{F} -convergence of $(\mathcal{X}^i, (\mu_t^i)_{t \in I})$ can be constructed as follows. We assume that $D^\infty(t)$ is continuous at $I \setminus S$ for a countable set S . For a large $k > 0$, we take a compact set $I_1 \subset [-k, 0] \setminus S$ so that $||[-k, 0] \setminus I_1|$ is small. Then I_1 is finitely covered by compact intervals I_{t_i} centered at $t_i \in I_1$ such that $|I_{t_i}|$ and the oscillation of all D^i and D^∞ on each I_{t_i} are sufficiently small. By steps 1 and 2 above, one can construct a correspondence \mathfrak{C}_0 that is fully defined on the finite set $I_0 := \{t_i\}$ between \mathcal{X}^i , so that

$$d_{\mathbb{F}}^{\mathfrak{C}_0, I_0}((\mathcal{X}^i, (\mu_t^i)_{t \in I_0}), (\mathcal{X}^j, (\mu_t^j)_{t \in I_0})) < \epsilon \quad (6.9)$$

for any $\epsilon > 0$, if i, j are sufficiently large. Then, by using the small oscillation of D^i on I_{t_i} , one can extend the correspondence \mathfrak{C}_0 to \mathfrak{C}_1 over I_1 so that $(\mathcal{X}^i, (\mu_t^i)_{t \in I_1})$ forms a Cauchy sequence over I_1 in the sense of (6.9) with respect to $d_{\mathbb{F}}^{\mathfrak{C}_1, I_1}$ (cf. [3, Lemma 7.24]). By letting $k \rightarrow \infty$ and taking a diagonal sequence, we obtain from the completeness of $(\mathbb{F}_I, d_{\mathbb{F}})$ a limit metric flow pair $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I \setminus S})$, which has an extended definition for all $t \in I$ by the future completion (cf.

[3, Section 4.4]) so that $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ is right continuous for $t \in I$. Notice that Definitions 6.1(1)–(7) for $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ are inherited from $(\mathcal{X}^i, (\mu_t^i)_{t \in I})$. In addition, one can construct a correspondence \mathfrak{C} so that

$$(\mathcal{X}^i, (\mu_t^i)_{t \in I}) \xrightarrow[i \rightarrow \infty]{\mathbb{F}, \mathfrak{C}} (\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$$

on compact time intervals and the convergence is uniform over the set on which $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ is continuous. Moreover, $(\mathcal{X}_t^\infty, d_t^\infty, \mu_t^\infty) \subset \mathbb{M}_r(V, b)$ as (6.5) and $\text{Var}(\mu_t^\infty) \leq H_n |t|$ for any $t \in I$. Notice that \mathcal{X}_0^∞ consists of a single point from which μ_t^∞ is the conjugate heat measure.

Remark 6.11 In [3, Theorem 7.4], a general compactness for a subset $\mathbb{F}_I^J(H, V, b, r) \subset \mathbb{F}_I^J$ is proven by the same method as described above.

It follows from [3, Theorem 8.2, 8.4] that the limit metric flow pair $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ obtained in (6.3) is a length space for any $t \in I$. In general, further geometric information contained in $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ is scarce. However, if $(M_i^n, g_i(t))$ are induced by Ricci shrinkers in $\mathcal{M}(A)$, then, in particular, their Nash entropies are uniformly bounded by Corollary 3.22. In this case, one obtains a much more concrete structure theorem regarding the limit metric flow obtained in (6.3); see [4, Theorem 2.3, 2.4, 2.5, 2.6, 2.46] and [3, Theorem 9.31].

Theorem 6.12 Let $(M_i^n, g_i(t), x_i)_{t \in I}$ be a sequence of pointed Ricci flows induced by Ricci shrinkers in $\mathcal{M}(A)$ and $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ the limit metric flow pair obtained in Theorem 6.10. Then the following properties hold.

- (1) There exists a decomposition

$$\mathcal{X}_0^\infty = \{x_\infty\}, \quad \mathcal{X}_{t < 0}^\infty = \mathcal{R} \sqcup \mathcal{S}, \tag{6.10}$$

such that \mathcal{R} is given by an n -dimensional Ricci flow spacetime $(\mathcal{R}, \mathfrak{t}, \partial_t^\infty, g^\infty)$, in the sense of [3, Definition 9.1] and $\dim_{\mathcal{M}^*}(\mathcal{S}) \leq n - 2$, where $\dim_{\mathcal{M}^*}$ denotes the $*$ -Minkowski dimension in [3, Definition 3.42]. Moreover, $\mu_t^\infty(\mathcal{S}_t) = 0$ for any $t < 0$.

- (2) Every tangent flow $(\mathcal{X}', (v_{x_\infty, t})_{t \leq 0})$ at every point $x \in \mathcal{X}^\infty$ is a metric soliton in the sense of [3, Definition 3.57]. Moreover, \mathcal{X}' is the Gaussian soliton iff $x \in \mathcal{R}$. If $x \in \mathcal{S}$, the singular set of $(\mathcal{X}', (v_{x_\infty, t})_{t \leq 0})$ on each $t < 0$ has Minkowski dimension at most $n - 4$. In particular, if $n = 3$, the metric soliton is a smooth Ricci flow associated with a 3-dimensional Ricci shrinker. If $n = 4$, each slice of the metric soliton is a smooth Ricci shrinker orbifold with isolated singularities.

- (3) $\mathcal{R}_t = \mathcal{R} \cap \mathcal{X}_t^\infty$ is open such that the restriction of d_t on \mathcal{R}_t agrees with the length metric of g_t .

- (4) The convergence (6.3) is smooth on \mathcal{R} , in the following sense. There exists an increasing sequence $U_1 \subset U_2 \subset \dots \subset \mathcal{R}$ of open subsets with $\bigcup_{i=1}^\infty U_i = \mathcal{R}$, open subsets $V_i \subset M_i \times I$, time-preserving diffeomorphisms $\phi_i : U_i \rightarrow V_i$ and a sequence $\epsilon_i \rightarrow 0$ such that the following holds:

- (a) We have

$$\begin{aligned} \|\phi_i^* g^i - g^\infty\|_{C^{[\epsilon_i^{-1}]}(U_i)} &\leq \epsilon_i, \\ \|\phi_i^* \partial_t^i - \partial_t^\infty\|_{C^{[\epsilon_i^{-1}]}(U_i)} &\leq \epsilon_i, \\ \|w^i \circ \phi_i - w^\infty\|_{C^{[\epsilon_i^{-1}]}(U_i)} &\leq \epsilon_i, \end{aligned}$$

where g^i is the spacetime metric induced by $g_i(t)$, and w^i is the conjugate heat kernel defined by $d\mu^i = w^i dg^i$, $i \in \mathbb{N} \cup \{\infty\}$.

(b) Let $y_\infty \in \mathcal{R}$ and $y_i \in M_i \times (-\infty, 0)$. Then y_i converges to y_∞ within \mathfrak{C} (cf. [3, Definition 6.18]) if and only if $y_i \in V_i$ for large i and $\phi_i^{-1}(y_i) \rightarrow y_\infty$ in \mathcal{R} .

(c) If the convergence (6.3) is uniform at some time $t \in I$, then for any compact subset $K \subset \mathcal{R}_t$ and for the same subsequence we have

$$\sup_{x \in K \cap U_i} d_t^Z(\varphi_t^i(\phi_i(x)), \varphi_t^\infty(x)) \rightarrow 0.$$

Theorem 6.12 is a flow version of the Cheeger-Colding theory (cf. [11–13]). Its proof shares a similar strategy as its elliptic counterparts. Many concepts also have counterparts. For example, the tangent flow corresponds to the tangent cone, and the metric soliton corresponds to the metric cone. We recall their definitions; see [3, Definition 6.55, 3.57].

Definition 6.13 (Tangent flow) Let \mathcal{X} be a metric flow over I and $x_0 \in \mathcal{X}_{t_0}$ a point. We say that a metric flow pair $(\mathcal{X}', (v'_{x_{\max};t})_{t \in I})$ is a tangent flow of \mathcal{X} at x_0 if there is a sequence of scales $\lambda_k > 0$ with $\lambda_k \rightarrow \infty$ such that for any $T > 0$ the parabolic rescalings

$$(\mathcal{X}'_{[-T,0]}^{-t_0, \lambda_k}, (v'_{x_0;t}^{-t_0, \lambda_k})_{\lambda_k^{-2}t+t_0 \in I', t \in [-T,0]})$$

\mathbb{F} -converge to $(\mathcal{X}'_{[-T,0]}, (v'_{x_{\max};t})_{t \in [-T,0]})$.

Definition 6.14 (Metric soliton) A metric flow pair $(\mathcal{X}, (\mu_t)_{t \in I})$ is called a metric soliton if there is a tuple

$$(X, d, \mu, (v'_{x;t})_{x \in X; t \leq 0})$$

and a map $\phi : \mathcal{X} \rightarrow X$ such that the following holds:

1. For any $t \in I$, the map $\phi_t : (\mathcal{X}_t, d_t, \mu_t) \rightarrow (X, \sqrt{t}d, \mu)$ is an isometry between metric measure spaces.
2. For any $x \in \mathcal{X}_t$, $s \in I$ with $s \leq t$, we have $(\phi_s)_* v_{x;s} = v'_{\phi_t(x); \log(s/t)}$.

Roughly speaking, a metric soliton is a metric flow pair induced by a metric measure space in a shrinking way. In general, a tangent flow of a metric flow may not be a metric soliton. In the setting of Theorem 6.12, every tangent flow of $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ is also an \mathbb{F} -limit of a sequence of Ricci flows induced by Ricci shrinkers in $\mathcal{M}(A)$ (cf. [3, Theorem 6.58]).

Notice that the limit metric flow $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ in (6.3) always admits a regular-singular decomposition

$$\mathcal{X}_{t < 0}^\infty = \mathcal{R} \sqcup \mathcal{S},$$

so that \mathcal{R} is given by a Ricci flow spacetime (cf. [3, Definition 9.1]). The key point is to control the size of the singular part in the appropriate sense. To avoid the distance distortion at different time-slices, one can redefine the Hausdorff and Minkowski dimensions (denoted by \mathcal{H}^* and \mathcal{M}^* , respectively) by using the P^* -parabolic balls instead of the conventional ones; see [3, Definition 3.41, 3.42].

One can control the size of \mathcal{S} quantitatively. Let $(M^n, g(t))_{t \in I}$ be the Ricci flow induced by a Ricci shrinker in $\mathcal{M}(A)$. We fix a point $(x_0, t_0) \in M \times I$ and define $\tau = t_0 - t$ and $H(x_0, t_0, \cdot, \cdot) = (4\pi\tau)^{-\frac{n}{2}} e^{-b}$. We next recall the some definitions from [4, Definition 5.1, 5.5, 5.6, 5.7], which indicate the extent to which the local geometry around (x_0, t_0) is a Ricci shrinker, Ricci flat space or splitting off an \mathbb{R}^k .

Definition 6.15 (Almost self-similarity) Let $(M^n, g(t))_{t \in I}$ be the Ricci flow induced by a Ricci shrinker. The point $(x_0, t_0) \in M \times I$ is called (ϵ, r) -selfsimilar if the following holds:

$$\int_{t_0 - \epsilon^{-1}r^2}^{t_0 - \epsilon r^2} \int_M \tau \left| \text{Rc} + \nabla^2 b - \frac{1}{2\tau} g \right|^2 dv_{x_0, t_0; t} dt \leq \epsilon,$$

$$\int_M |\tau(2\Delta b - |\nabla b|^2 + R) + b - n - \mathcal{N}_{x_0, t_0}(r^2)| dv_{x_0, t_0; t} \leq \epsilon, \quad \forall t \in [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2].$$

Definition 6.16 (Almost static) The point (x_0, t_0) is called (ϵ, r) -static if the following holds:

$$r^2 \int_{t_0 - \epsilon^{-1}r^2}^{t_0 - \epsilon r^2} \int_M |\text{Rc}|^2 dv_{x_0, t_0; t} dt \leq \epsilon,$$

$$r^2 \int_M R dv_{x_0, t_0; t} \leq \epsilon, \quad \forall t \in [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2].$$

Definition 6.17 (Weak splitting) (x_0, t_0) is called weakly (k, ϵ, r) -split if there exists a vector-valued function $\vec{y} = (y_1, \dots, y_k) : M \times [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2] \rightarrow \mathbb{R}^k$ with the following properties for all $i, j = 1, \dots, k$:

(1) We have

$$r^{-1} \int_{t_0 - \epsilon^{-1}r^2}^{t_0 - \epsilon r^2} \int_M |\square y_i| dv_{x_0, t_0; t} dt \leq \epsilon.$$

(2) We have

$$r^{-2} \int_{t_0 - \epsilon^{-1}r^2}^{t_0 - \epsilon r^2} \int_M |\nabla y_i \cdot \nabla y_j - \delta_{ij}| dv_{x_0, t_0; t} dt \leq \epsilon.$$

Definition 6.18 (Strong splitting) (x_0, t_0) is called strongly (k, ϵ, r) -split if there exists a vector-valued function $\vec{y} = (y_1, \dots, y_k) : M \times [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2] \rightarrow \mathbb{R}^k$ with the following properties for all $i, j = 1, \dots, k$:

(1) y_i solves the heat equation $\square y_i = 0$ on $M \times [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2]$.

(2) We have

$$r^{-2} \int_{t_0 - \epsilon^{-1}r^2}^{t_0 - \epsilon r^2} \int_M |\nabla y_i \cdot \nabla y_j - \delta_{ij}| dv_{x_0, t_0; t} dt \leq \epsilon.$$

(3) For all $t \in [t_0 - \epsilon^{-1}r^2, t_0 - \epsilon r^2]$, we have

$$\int_M y_i dv_{x_0, t_0; t} = 0.$$

It can be proven (cf. [4, Proposition 12.1]) that if (x_0, t_0) is weakly (k, ϵ, r) -split, then it is strongly $(k, \delta(\epsilon), r)$ -split. With these definitions, one can consider the following quantitative stratification.

Definition 6.19 For $\epsilon > 0$ and $0 < r_1 < r_2 \leq \infty$, the effective strata

$$\tilde{\mathcal{S}}_{r_1, r_2}^{\epsilon, 0} \subset \tilde{\mathcal{S}}_{r_1, r_2}^{\epsilon, 1} \subset \tilde{\mathcal{S}}_{r_1, r_2}^{\epsilon, 2} \subset \dots \subset \tilde{\mathcal{S}}_{r_1, r_2}^{\epsilon, n+2} \subset M \times I$$

are defined as $(x', t') \in \tilde{\mathcal{S}}_{r_1, r_2}^{\epsilon, k}$ if and only if for all $r' \in (r_1, r_2)$, neither of the following two properties holds:

1. (x', t') is (ϵ, r') -selfsimilar and weakly $(k + 1, \epsilon, r')$ -split.
2. (x', t') is (ϵ, r') -selfsimilar, (ϵ, r') -static and weakly $(k - 1, \epsilon, r')$ -split.

By a delicate choice of the covering by P^* -parabolic balls, it can be proven, see [4, Proposition 11.2], that for any $0 < \sigma < \epsilon$, there are points $(x_1, t_1), \dots, (x_N, t_N) \in \tilde{\mathcal{S}}_{\sigma r, \epsilon r}^{\epsilon, k} \cap P^*(x_0, t_0; r)$ with $N \leq C(A, \epsilon)\sigma^{-k-\epsilon}$ and

$$\tilde{\mathcal{S}}_{\sigma r, \epsilon r}^{\epsilon, k} \cap P^*(x_0, t_0; r) \subset \bigcup_{i=1}^N P^*(x_i, t_i; \sigma r). \quad (6.11)$$

Notice that (6.11) can be regarded as a parabolic version of the covering in [12] by Cheeger and Naber.

On the complement of $\tilde{\mathcal{S}}_{\sigma r, \epsilon r}^{\epsilon, n-2}$, the following ϵ -regularity theorem is proven (cf. [4, Proposition 17.1]), which can be viewed as a parabolic analog of Cheeger-Naber's codimension 4 theorem in [13]. Roughly speaking, one needs to rule out the tangent flows which are Ricci-flat and split off an \mathbb{R}^{n-3} .

Proposition 6.20 There exists a constant $\epsilon = \epsilon(n, A) > 0$ such that the following holds. Let $(M^n, g(t))_{t \in I}$ be the Ricci flow induced by a Ricci shrinker in $\mathcal{M}(A)$. Suppose that (x_0, t_0) is strongly $(n-1, \epsilon, r)$ -split or strongly $(n-3, \epsilon, r)$ -split and (ϵ, r) -static. Then $r_{\text{Rm}}(x_0, t_0) \geq \epsilon r$.

There are many implications of Proposition 6.20. Notice that one has the following decomposition:

$$\mathcal{X}_{t < 0}^\infty = \mathcal{R}^* \sqcup \mathcal{S}^*,$$

where $\mathcal{R}^* \subset \mathcal{R}$ is the set of points where the convergence (6.3) is smooth as defined in [3, Section 9.4]. Since $\mathcal{S} \subset \mathcal{S}^*$, one can obtain the estimate of *-Minkowski dimension of \mathcal{S} by that of \mathcal{S}^* from (6.11) and Proposition 6.20 (cf. [4, Theorem 15.28 (a)]). Moreover, it can be proven that $\mathcal{S}^* \cap \mathcal{X}_t^\infty$ has measure 0 for any $t < 0$ (cf. [4, Theorem 15.28 (b)]). Therefore, Theorem 6.12 (1) is obtained.

Since \mathcal{S} has measure 0 on each time-slice, one can extend the definition of the Nash entropy on \mathcal{X}^∞ . Therefore, the Nash entropy at the base point x' of any tangent flow $(\mathcal{X}', (v_{x'; t})_{t \in I})$ of $(\mathcal{X}^\infty, (\mu_t^\infty)_{t \in I})$ is a constant. By the relation between the Nash entropy and the almost self-similarity (cf. [4, Proposition 7.1]), one concludes that $(\mathcal{X}', (v_{x'; t})_{t \in I})$ is a metric soliton since its regular part admits an incomplete Ricci shrinker and the tangent flow itself is determined by its regular part due to the high codimension of the singular part (cf. [4, Theorem 15.60, 15.69]). Moreover, the singular set on each time-slice of $(\mathcal{X}', (v_{x'; t})_{t \in I})$ has Minkowski dimension $\leq n-4$ (cf. [4, Theorem 2.16]). Furthermore, the fact that $x \in \mathcal{R}$ iff \mathcal{X}' is the Gaussian soliton follows from the ϵ -regularity theorem 5.15 and the convergence of the Nash entropies under (6.3) (cf. [4, Theorem 2.11, 2.14]). Notice that if $n = 4$, each time-slice of $(\mathcal{X}', (v_{x'; t})_{t \in I})$ is a smooth orbifold with isolated singularities, since each tangent flow at any singular point of $(\mathcal{X}', (v_{x'; t})_{t \in I})$ is a flat cone (cf. [4, Theorem 2.46]). Therefore, we obtain Theorem 6.12 (2).

For Theorem 6.12 (3), the inequality $d_t \leq d_{g_t}$ is clear. The opposite inequality is proven by showing that any $u \in C^0(\mathcal{R}_t)$ that is 1-Lipschitz with respect to d_{g_t} is also 1-Lipschitz with respect to d_t (cf. [4, Theorem 15.28 (c)]). The argument uses the high codimension of \mathcal{S} , the fact that \mathcal{X}^∞ is future continuous at t , and the fact that $\mathcal{R} = \mathcal{R}^*$, which can be proven by using the ϵ -regularity theorem and the convergence of the Nash entropies (cf. [4, Corollary 15.47]).

Once we know $\mathcal{R} = \mathcal{R}^*$, the diffeomorphisms in Theorem 6.12 (4) can be obtained by patching all local conventional Ricci flows into a Ricci flow spacetime by a center of mass construction (cf. [3, Theorem 9.31]). Notice that similar constructions are well-known for the

Cheeger-Gromov convergence (cf. Remark 7.7 of [25] and references therein). All assertions in Theorem 6.12 (4) are proven by smooth convergence. Therefore, Theorem 6.12 (4) is obtained.

As an application of the theory of \mathbb{F} -convergence, we have the following backward pseudolocality theorem; see [4, Theorem 2.47]. Earlier backward pseudolocality can be found in [35, Corollary 11.6(b)] [15, Lemma 4.2] [16, Theorem 4.7] [6, Theorem 1.5].

Theorem 6.21 (Backward pseudolocality theorem) For any $n \in \mathbb{N}$ and $\alpha > 0$ there is an $\epsilon(n, \alpha) > 0$ such that the following holds.

Let $(M^n, g(t))_{t \in I}$ be a Ricci flow induced by a Ricci shrinker. Given $(x_0, t_0) \in M \times I$ and $r > 0$, if

$$|B_{t_0}(x_0, r)| \geq \alpha r^n, \quad |\text{Rm}| \leq (\alpha r)^{-2} \quad \text{on } B_{t_0}(x_0, r),$$

then

$$|\text{Rm}| \leq (\epsilon r)^{-2} \quad \text{on } P(x_0, t_0; (1 - \alpha)r, -(\epsilon r)^2, 0).$$

Note that with the combination of the above theorem with the forward pseudolocality (cf. Theorem 24 of [29]), we arrive at the two-sided pseudolocality. Thus Theorem 1.6 is proven.

Combining Theorem 6.21 and (5.23), we have

Corollary 6.22 (Comparison of the curvature radii) There exists a constant $C(n, A) > 1$ such that the following holds.

Let $(M^n, g(t))_{t \in I'}$ be a Ricci flow induced by a Ricci shrinker in $\mathcal{M}(A)$. Then, for any $(x, t) \in M \times I'$,

$$r_{\text{Rm}}(x, t) \leq r_{\text{Rm}}^-(x, t) \leq r_{\text{Rm}}^s(x, t) \leq C r_{\text{Rm}}(x, t).$$

Another application is the following integral estimate using the quantitative stratification; see [4, Theorem 2.28].

Theorem 6.23 Let $(M^n, g(t))_{t < 1}$ be a Ricci flow associated with a Ricci shrinker in $\mathcal{M}(A)$. Then, for any $(x_0, t_0) \in M \times (-\infty, 1)$, $r > 0$ and $\epsilon > 0$,

$$\begin{aligned} & \int_{[t_0-r^2, t_0+r^2] \cap (-\infty, 1)} \int_{P^*(x_0, t_0; r) \cap M \times \{t\}} |\text{Rm}|^{2-\epsilon} dV_t dt \\ & \leq \int_{[t_0-r^2, t_0+r^2] \cap (-\infty, 1)} \int_{P^*(x_0, t_0; r) \cap M \times \{t\}} r_{\text{Rm}}^{-4+2\epsilon} dV_t dt \leq C(n, A, \epsilon) r^{n-2+2\epsilon}. \end{aligned} \tag{6.12}$$

As a corollary, we prove

Corollary 6.24 Let (M^n, g, f, p) be a Ricci shrinker in $\mathcal{M}(A)$. Then

$$\int_{d(p, \cdot) \leq r} |\text{Rm}|^{2-\epsilon} dV \leq \int_{d(p, \cdot) \leq r} r_{\text{Rm}}^{-4+2\epsilon} dV \leq C r^{n+2\epsilon-2}, \tag{6.13}$$

$$\int_{d(p, \cdot) \geq 1} \frac{|\text{Rm}|^{2-\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq \int_{d(p, \cdot) \geq 1} \frac{r_{\text{Rm}}^{-4+2\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq C \tag{6.14}$$

for any $\epsilon > 0$ and $r \geq 1$, where $r_{\text{Rm}}(\cdot) = r_{\text{Rm}}(\cdot, 0)$ and $C = C(n, A, \epsilon)$.

Proof We consider the Ricci flow $(M, g(t))_{t < 1}$ associated with the given Ricci shrinker. It follows from Proposition 5.7 that

$$Q(p, 0; 1, 0, 1) \subset P^*(p, 0; C_1, 0, 1) \subset P^*(p, 0; C_1)$$

for some constant $C_1 = C_1(n, A) > 1$. Therefore, it follows from Theorem 6.23 that with $(x_0, t_0) = (p, 0)$ and $r = C_1$

$$\int_0^1 \int_{d_t(p, \cdot) < 1} |\text{Rm}|^{2-\epsilon} dV_t dt \leq C(n, A, \epsilon). \quad (6.15)$$

Since $g(t) = (1-t)(\psi^t)^*g$ and ψ^t is defined by (2.3), we have $d_t(x, p) = \sqrt{1-t}d(x^t, p)$ and $|\text{Rm}|(x, t) = |\text{Rm}|(x^t)/(1-t)$, where $x^t = \psi^t(x)$. Therefore, we have

$$\int_{d_t(p, x) < 1} |\text{Rm}|^{2-\epsilon}(x, t) dV_t(x) = (1-t)^{\frac{n}{2}-2+\epsilon} \int_{d(x, p) < \frac{1}{\sqrt{1-t}}} |\text{Rm}|^{2-\epsilon}(x) dV(x). \quad (6.16)$$

By a change of variable with $t = 1 - r^{-2}$, it follows from (6.15) and (6.16) that

$$\int_1^\infty r^{1-2\epsilon-n} m(r) dr \leq C(n, A, \epsilon), \quad (6.17)$$

where

$$m(r) := \int_{d(\cdot, p) < r} |\text{Rm}|^{2-\epsilon} dV.$$

We claim that there exists a sequence $r_i \rightarrow \infty$ such that

$$\lim_{i \rightarrow \infty} \frac{m(r_i)}{r_i^{n+2\epsilon-2}} = 0. \quad (6.18)$$

Otherwise, there exists a constant $\delta > 0$ such that $m(r) \geq \delta r^{n+2\epsilon-2}$ for sufficiently large r . However, this contradicts (6.17).

We apply integration by parts to (6.17) from 1 to r_i and obtain

$$\int_{1 \leq d(p, \cdot) \leq r_i} \frac{|\text{Rm}|^{2-\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq C(n, A, \epsilon) + m(r_i) r_i^{2-2\epsilon+n}.$$

By letting $i \rightarrow \infty$, we have from (6.18) that

$$\int_{d(p, \cdot) \geq 1} \frac{|\text{Rm}|^{2-\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq C(n, A, \epsilon). \quad (6.19)$$

In addition, we have for any $r \geq 1$,

$$r^{2-2\epsilon-n} \int_{1 \leq d(p, \cdot) \leq r} |\text{Rm}|^{2-\epsilon} dV \leq \int_{1 \leq d(p, \cdot) \leq r} \frac{|\text{Rm}|^{2-\epsilon}}{d^{n+2\epsilon-2}(p, \cdot)} dV \leq C(n, A, \epsilon).$$

Therefore, for any $r \geq 1$,

$$\int_{d(p, \cdot) \leq r} |\text{Rm}|^{2-\epsilon} dV \leq C(n, A, \epsilon) r^{n+2\epsilon-2}, \quad (6.20)$$

since $m(1)$ is bounded by (6.17). In sum, the inequalities involving $|\text{Rm}|$ in (6.13) and (6.14) are proven.

Notice that for any $(x_0, t_0) \in M \times (-\infty, 1)$, it follows from the definition of ψ^t (2.3) that

$$\psi^{\theta(t)} \circ \psi^{t_0} = \psi^t,$$

where $\theta(t) := \frac{t-t_0}{1-t_0}$. Therefore, for any $t < 1$,

$$g(t) = (1-t)(\psi^t)^*g = (1-t_0)(1-\theta(t))(\psi^{t_0})^*(\psi^{\theta(t)})^*g = (1-t_0)(\psi^{t_0})^*g(\theta(t)).$$

Therefore,

$$r_{\text{Rm}}(x_0, t_0) = \sqrt{1 - t_0} r_{\text{Rm}}(x_0^{t_0}, 0).$$

Now, the conclusion regarding r_{Rm} can be proven similarly. \square

We end this section by proving a gap property for the volume ratio at infinity.

Corollary 6.25 Let (M^n, g, f, p) be a Ricci shrinker in $\mathcal{M}(A)$. Suppose

$$\liminf_{r \rightarrow \infty} \frac{|B(p, r)|}{r^n} = 0. \quad (6.21)$$

Then

$$|B(p, r)| \leq C r^{n-2+\epsilon} \quad (6.22)$$

for any $r \geq 1$ and some $C = C(n, A, \epsilon)$.

Proof We claim that $r_{\text{Rm}}(x) < 2$ for any x . Indeed, if $r_{\text{Rm}}(x) \geq 2$, we have

$$|\text{Rm}|(y, t) < 1$$

for any $y \in B(x, 1)$ and $t < 1$. By the same argument as in [29, Corollary 9], we obtain that

$$\psi^t(B(x, 1)) \subset B\left(p, \frac{c_1}{\sqrt{1-t}}\right) \setminus B\left(p, \frac{c_2}{\sqrt{1-t}}\right) \quad (6.23)$$

for $c_1 > c_2 > 0$, if t is sufficiently close to 1. From the standard distance distortion and Theorem 4.1, we obtain that

$$|\psi^t(B(x, 1))| \geq c_3(1-t)^{-\frac{n}{2}}. \quad (6.24)$$

However, (6.23) and (6.24) contradict (6.21). Thus, the desired inequality (6.22) follows immediately from (6.13). \square

Conflict of Interest The authors declare no conflict of interest.

References

- [1] Andrews B. Gradient and oscillation estimates and their applications in geometric PDE// International Congress of Chinese Mathematicians (ICCM2012), American Mathematical Society, USA, 2012: 3–19
- [2] Bamler R H. Entropy and heat kernel bounds on a Ricci flow background. arXiv:2008.07093v3
- [3] Bamler R H. Compactness theory of the space of super Ricci flow. Invent Math, 2023, **233**(3): 1121–1277
- [4] Bamler R H. Structure theory of non-collapsed limits of Ricci flow. arXiv:2009.03243v2
- [5] Bamler R H. On the fundamental group of non-collapsed ancient Ricci flows. arXiv:2110.02254
- [6] Bamler R H, Zhang Q S. Heat kernel and curvature bounds in Ricci flows with bounded scalar curvature. Adv Math, 2017, **319**: 396–450
- [7] Brendle S. Two-point functions and their applications in geometry. Bull Amer Math Soc, 2014, **51**(4): 581–596
- [8] Cao H D, Chen B L, Zhu X P. Recent developments on Hamilton's Ricci flow// Surveys in Differential Geometry, Vol. 12. Somerville, MA: International Press, 2008: 47–112
- [9] Cao H D, Zhou D. On complete gradient shrinking Ricci solitons. J Differ Geom, 2010, **85**(2): 175–186
- [10] Chan P, Ma Z, Zhang Y. Ancient Ricci flows with asymptotic solitons. arXiv: 2106.06904v1
- [11] Cheeger J, Colding T H. On the structure of space with Ricci curvature bounded below I. J Differ Geom, 1997, **46**: 406–480
- [12] Cheeger J, Naber A. Lower bounds on Ricci curvature and quantitative behavior of singular sets. Invent Math, 2013, **191**(2): 321–339
- [13] Cheeger J, Naber A. Regularity of Einstein manifolds and the codimension 4 conjecture. Ann Math, 2015, **182**(3): 1093–1165

- [14] Chen B L. Strong uniqueness of the Ricci flow. *J Differ Geom*, 2009, **82**(2): 363–382
- [15] Chen X, Wang B. Space of Ricci flows (I). *Comm Pure Appl Math*, 2012, **65**(10): 1399–1457
- [16] Chen X, Wang B. Space of Ricci flows (II)—Part B: Weak compactness of the flows. *J Differential Geom*, 2020, **116**(1): 1–123
- [17] Chen X, Wang B. Remarks of weak-compactness along Kähler Ricci flow// *Proceedings of the Seventh International Congress of Chinese Mathematicians*, ALM 44, 2016: 203–233
- [18] Chen C W, Zhang Z. On shrinking gradient Ricci solitons with positive Ricci curvature and small Weyl tensor. *Acta Math Sci*, 2019, **39B**(5): 1235–1239
- [19] Chow B, Chu S C, Glickenstein D, et al. *The Ricci Flow: Techniques and Applications. Part II. Analytic Aspects*. *Mathematical Surveys and Monographs*, vol 135. Providence RI: American Mathematical Society, 2008
- [20] Enders J, Müller R, Topping P. On Type-I singularities in Ricci flow. *Comm Anal Geom*, 2011, **19**(5): 905–922
- [21] Friedman A. *Partial Differential Equations of Parabolic Type*. Mineola, NY: Courier Dover Publications, 2008
- [22] Hamilton R S. The formation of singularities in the Ricci flow// *Surveys in Differential Geom*. Somerville, MA: International Press, 1995: 7–136
- [23] Haslhofer R, Müller R. A compactness theorem for complete Ricci shrinkers. *Geom Funct Anal*, 2011, **21**: 1091–1116
- [24] Hein H J, Naber A. New logarithmic Sobolev inequalities and an ϵ -regularity theorem for the Ricci flow. *Comm Pure Appl Math*, 2014, **67**(9): 1543–1561
- [25] Li H, Li Y, Wang B. On the structure of Ricci shrinkers. *J Funct Anal*, 2021, **280**(9): 108955
- [26] Li X, Ni L. Kähler-Ricci shrinkers and ancient solutions with nonnegative orthogonal bisectional curvature. *Math Pures Appl*, 2020, **138**: 28–45
- [27] Li X, Ni L, Wang K. Four-dimensional gradient shrinking solitons with positive isotropic curvature. *Int Math Res Not*, 2018, **2018**(3): 949–959
- [28] Li Y, Wang B. The rigidity of Ricci shrinkers of dimension four. *Trans Amer Math Soc*, 2019, **371**(10): 6949–6972
- [29] Li Y, Wang B. Heat kernel on Ricci shrinkers. *Calc Var Partial Differ Equ*, 2020, **59**: Art 194
- [30] Munteanu O, Wang J. Positively curved shrinking Ricci solitons are compact. *J Differ Geom*, 2017, **106**(3): 499–505
- [31] Ma Z, Zhang Y. Perelman’s entropy on ancient Ricci flows. *J Funct Anal* 2021, **281**(9): Art 109195
- [32] Naber A. Noncompact shrinking four solitons with nonnegative curvature. *J Reine Angew Math*, 2010, **645**: 125–153
- [33] Naff K. Shrinking Ricci solitons with positive isotropic curvature. arXiv: 1905.10305
- [34] Ni L, Wallach N. On a classification of gradient shrinking solitons. *Math Res Lett*, 2008, **15**(5): 941–955
- [35] Perelman G. The entropy formula for the Ricci flow and its geometric applications. arXiv:math.DG/0211159
- [36] Villani C. *Optimal Transport, Old and New*. *Grundlehren Math Wiss*, vol 338. Berlin: Springer, 2008

Appendix Integral Estimates for the Conjugate Heat Kernel

In this appendix, we generalize some integral estimates regarding the conjugate heat kernel from [4, Section 6] to Ricci flows associated with Ricci shrinkers. These estimates also hold for Ricci flows induced by Ricci shrinkers since they are scaling-invariant.

Throughout this appendix, let $(M^n, g(t))_{t < 1}$ be the Ricci flow associated with a Ricci shrinker in $\mathcal{M}(A)$. We fix a spacetime point $(x_0, t_0) \in M \times (-\infty, 1)$ and set $dv_t = dv_{x_0, t_0; t}$ and $\tau = t_0 - t$. Moreover, we define $w = w(x, t)$ and $b = b(x, t)$ by $w = H(x_0, t_0, x, t) = (4\pi(t_0 - t))^{-\frac{n}{2}} e^{-b}$.

Lemma A.1 There exists a constant $C = C(n, A) > 1$ such that for any $0 < \tau_0 < \tau_1$,

$$\int_{t_0 - \tau_1}^{t_0 - \tau_0} \int_M \{|\text{Rc}|^2 + |\nabla^2 b|^2\} dv_t dt \leq C\tau_0^{-1} \left(1 + \log \frac{\tau_1}{\tau_0}\right). \quad (\text{A.1})$$

Proof Without loss of generality, we assume $t_0 = 0$.

From Corollary 3.20, we have for any $\tau > 0$ that

$$\int_M |\nabla b|^2 + R dv_{-\tau} \leq \frac{n}{2\tau}. \tag{A.2}$$

Direct calculation shows that

$$\begin{aligned} \partial_t \int_M R w \phi^r dV_t &= \int_M \{(\square R)w\phi^r - R\square^*(w\phi^r)\} dV_t \\ &= \int_M 2|\text{Rc}|^2 w \phi^r + R \{w(\Delta\phi^r + \phi_t^r) + 2\langle \nabla w, \nabla \phi^r \rangle\} dV_t \\ &= \int_M \{2|\text{Rc}|^2 \phi^r + R(\Delta\phi^r + \phi_t^r) - 2R\langle \nabla b, \nabla \phi^r \rangle\} dv_t. \end{aligned}$$

Integrating the above equation from $-\tau_1$ to $-\tau_0$, we obtain

$$\begin{aligned} &\int_{-\tau_1}^{-\tau_0} \int_M 2|\text{Rc}|^2 \phi^r dv_t dt \\ &\leq \int_M R \phi^r dv_{-\tau_0} + \int_{-\tau_1}^{-\tau_0} \int_M \{R(|\Delta\phi^r| + |\phi_t^r|) + R^2|\nabla\phi^r| + |\nabla b|^2|\nabla\phi^r|\} dv_t dt. \end{aligned} \tag{A.3}$$

From (2.8) and Lemma 2.3, R increases at most quadratically. Combining (2.12), (2.13), (2.14) and (A.2), it follows that the last integral in (A.3) tends to 0 as $r \rightarrow \infty$. Therefore, we obtain

$$\int_{-\tau_1}^{-\tau_0} \int_M |\text{Rc}|^2 dv_t dt \leq \frac{1}{2} \int_M R dv_{-\tau_0} \leq \frac{n}{4\tau_0}. \tag{A.4}$$

On the other hand, it follows from (3.40) and Corollary 3.22 that

$$\int_{-\tau_1}^{-\tau_0} 2\tau \int_M \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 dv_t dt \leq -\mathcal{W}_{(x_0, t_0)}(\tau_1) \leq A. \tag{A.5}$$

By virtue of the elementary identity $(x - y)^2 \geq x^2/2 - y^2$, it follows from (A.5) that

$$\tau_0 \int_{-\tau_1}^{-\tau_0} \int_M |\text{Rc} + \nabla^2 b|^2 dv_t dt \leq \int_{-\tau_1}^{-\tau_0} \tau \int_M |\text{Rc} + \nabla^2 b|^2 dv_t dt \leq A + \frac{n}{2} \log \frac{\tau_1}{\tau_0}. \tag{A.6}$$

Combining (A.4) and (A.6), the conclusion follows immediately. □

Lemma A.2 There exists a constant $C = C(n, A) > 1$ such that the following estimates hold for any $t < t_0$ and $0 \leq s \leq 1/4$.

$$\int_M \{1 + |b| + \tau(|\Delta b| + |\nabla b|^2 + R)\} e^{sb} dv_t \leq C. \tag{A.7}$$

Proof We compute

$$\frac{d}{ds} \int_M e^{sb} dv_t = \int_M b e^{sb} dv_t. \tag{A.8}$$

Here, the differentiation under the integral sign is allowed by Theorem 4.16 and Remark 4.17. By the differential Harnack inequality [29, Theorem 21], we calculate

$$\begin{aligned} \int_M b e^{sb} dv_t &\leq \int_M (\tau(-2\Delta b + |\nabla b|^2 - R) + n) e^{sb} dv_t \\ &= \int_M (\tau((2s - 1)|\nabla b|^2 - R) + n) e^{sb} dv_t \leq n \int_M e^{sb} dv_t, \end{aligned} \tag{A.9}$$

where integration by parts in the equality can be justified similarly as to Remark 3.21. Combining (A.8) and (A.9), we obtain

$$\int_M e^{sb} dv_t \leq e^{ns}.$$

On the other hand, it follows from Theorem 4.4 that $b \geq -A$. Therefore, it follows from (A.9) and the above inequality that

$$\int_M \{|b| + \tau(|\nabla b|^2 + R)\} e^{sb} dv_t \leq C(n, A). \quad (\text{A.10})$$

Applying the differential Harnack inequality and integration by parts again, we obtain

$$\begin{aligned} \int_M 2\tau|\Delta b|e^{sb} dv_t &\leq \int_M \{|u| + \tau(|\nabla b|^2 + R) + |b| + n\} e^{sb} dv_t \\ &= \int_M \{-u + \tau(|\nabla b|^2 + R) + |b| + n\} e^{sb} dv_t \\ &\leq \int_M \{2s\tau|\nabla b|^2 + 2|b| + 2n\} e^{sb} dv_t \leq C(n, A), \end{aligned}$$

where $u = \tau(2\Delta b - |\nabla b|^2 + R) + b - n \leq 0$. It is clear that (A.7) follows from the combination of (A.10) and the above inequality. \square

Lemma A.3 There exists a constant $C = C(n) > 1$ such that the following estimates hold for any $t < t_0$ and $0 \leq s \leq 1/4$.

$$\int_M |\nabla b|^4 e^{sb} dv_t \leq C \int_M |\nabla^2 b|^2 e^{sb} dv_t. \quad (\text{A.11})$$

Proof In the proof, all constants $C > 1$ depend only on n , which may be different line by line.

We compute for $s \leq 1/4$ that

$$\begin{aligned} \int_M |\nabla b|^4 \phi^r e^{sb} dv_t &= (4\pi\tau)^{-\frac{n}{2}} \int_M |\nabla b|^4 \phi^r e^{(s-1)b} dV_t \\ &= (4\pi\tau)^{-\frac{n}{2}} (s-1)^{-1} \int_M |\nabla b|^2 \langle \nabla b, \nabla e^{(s-1)b} \rangle \phi^r dV_t \\ &= (4\pi\tau)^{-\frac{n}{2}} (1-s)^{-1} \int_M (2\nabla^2 b \langle \nabla b, \nabla b \rangle + |\nabla b|^2 \Delta b) \phi^r e^{(s-1)b} dV_t + Z \\ &\leq C(4\pi\tau)^{-\frac{n}{2}} (1-s)^{-1} \int_M |\nabla^2 b| |\nabla b|^2 \phi^r e^{(s-1)b} dV_t + Z \\ &\leq \frac{1}{4} \int_M |\nabla b|^4 \phi^r e^{sb} dv_t + C \int_M |\nabla^2 b|^2 \phi^r e^{sb} dv_t + Z, \end{aligned} \quad (\text{A.12})$$

where the remainder is

$$\begin{aligned} Z &:= (1-s)^{-1} \int_M |\nabla b|^2 \langle \nabla b, \nabla \phi^r \rangle e^{sb} dv_t \\ &\leq 2 \int_M |\nabla b|^3 |\nabla \phi^r| e^{sb} dv_t \\ &\leq \frac{1}{4} \int_M |\nabla b|^4 \phi^r e^{sb} dv_t + 4 \int_M |\nabla b|^2 |\nabla \phi^r|^2 (\phi^r)^{-1} e^{sb} dv_t. \end{aligned} \quad (\text{A.13})$$

Applying Lemma A.2 and (2.12), we conclude from (A.12) and (A.13) that

$$\int_M |\nabla b|^4 e^{sb} \phi^r dv_t \leq C \int_M |\nabla^2 b|^2 e^{sb} \phi^r dv_t + \epsilon(r), \quad (\text{A.14})$$

where $\epsilon(r) \rightarrow 0$ as $r \rightarrow \infty$. Thus we arrive at (A.11) by letting $r \rightarrow \infty$ in the above inequality. □

The main result of this section is the following spacetime integral estimate.

Proposition A.4 There exists a constant $C = C(n, A) > 1$ and $\bar{s} = \bar{s}(n) < 1$ such that the following estimates hold for any $r > 0, 0 < \theta < 1/2$ and $s \leq \bar{s}$.

$$\int_{t_0-\tau_0}^{t_0-\theta\tau_0} \int_M \tau(|\text{Rc}|^2 + |\nabla^2 b|^2 + |\nabla b|^4) e^{sb} dv_t dt \leq C \log \theta^{-1}. \tag{A.15}$$

Proof In the proof, all constants C depend on n , and C' depend on n and A . Moreover, we use $\epsilon(r)$ to denote a function independent of t such that $\epsilon(r) \rightarrow 0$ if $r \rightarrow \infty$. Those terms may be different line by line. Without loss of generality, we assume $t_0 = 0$.

We set $u = \tau(2\Delta b - |\nabla b|^2 + R) + b - n \leq 0$. Recall that from [35], we have the celebrated identity

$$\square^*(uw) = -2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 w.$$

Moreover, we have

$$\square b = -2\Delta b + |\nabla b|^2 - R + \frac{n}{2\tau} = \tau^{-1}(b - u - n/2).$$

Direct computation shows that

$$\begin{aligned} & \partial_t \int_M u w e^{sb} \phi^r dV_t \\ &= \int_M \{ \square(e^{sb} \phi^r) u w - e^{sb} \phi^r \square^*(u w) \} dV_t \\ &= \int_M \left\{ ((\square e^{sb}) \phi^r + e^{sb} (\square \phi^r) - 2\langle \nabla \phi^r, \nabla e^{sb} \rangle) u w + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 w e^{sb} \phi^r \right\} dV_t \\ &= \int_M \left\{ ((s \square b - s^2 |\nabla b|^2) e^{sb} \phi^r + e^{sb} (\square \phi^r) - 2\langle \nabla \phi^r, \nabla e^{sb} \rangle) u \right. \\ &\quad \left. + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 e^{sb} \phi^r \right\} w dV_t \\ &= \int_M \left\{ ((s\tau^{-1}(b - u - n/2) - s^2 |\nabla b|^2) \phi^r + \square \phi^r - 2s \langle \nabla \phi^r, \nabla b \rangle) u \right. \\ &\quad \left. + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 \phi^r \right\} e^{sb} dv_t. \end{aligned}$$

It follows that

$$\begin{aligned} & \partial_t \int_M u w e^{sb} \phi^r dV_t \\ & \geq \int_M \left\{ (s\tau^{-1}(b - u - n/2)) u \phi^r + (\square \phi^r - 2s \langle \nabla \phi^r, \nabla b \rangle) u + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 \phi^r \right\} e^{sb} dv_t \\ & \geq \int_M \left\{ -Cs\tau^{-1}(u^2 + b^2 + 1) \phi^r + (\square \phi^r - 2s \langle \nabla \phi^r, \nabla b \rangle) u + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 \phi^r \right\} e^{sb} dv_t \\ & \geq \int_M \left\{ -Cs(\tau((\Delta b)^2 + |\nabla b|^4 + R^2) + \tau^{-1}(b^2 + 1)) + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 \right\} \phi^r e^{sb} dv_t + X_t, \end{aligned} \tag{A.16}$$

where

$$X_t := \int_M (\square\phi^r - 2s\langle\nabla\phi^r, \nabla b\rangle) ue^{sb} dv_t.$$

Define $X' := \int_{-\tau_0}^{-\theta\tau_0} |X_t| dt$. Then it follows from Lemma A.2 and inequalities (2.12) to (2.15) that for any positive δ we have

$$\begin{aligned} |X'| &\leq \int_{-\tau_0}^{-\theta\tau_0} \int_M (|\square\phi^r| + 2s|\nabla\phi^r||\nabla b|)|u|e^{sb} dv_t dt \\ &\leq \epsilon(r) + \int_{-\tau_0}^{-\theta\tau_0} \int_M (\delta^{-1}|\nabla\phi^r|^2(\phi^r)^{-1}\tau|\nabla b|^2 + \delta\tau^{-1}u^2\phi^r) e^{sb} dv_t dt \\ &\leq \epsilon(r) + \delta \int_{-\tau_0}^{-\theta\tau_0} \int_M \tau^{-1}u^2\phi^r e^{sb} dv_t dt. \end{aligned} \quad (\text{A.17})$$

It is clear from the definition of u that $u^2 \leq C(\tau^2(|\Delta b|^2 + |\nabla b|^2 + R)) + b^2 + 1$. In addition, since $-A \leq b \leq -\tau(2\Delta b - |\nabla b|^2 + R) + n$, we have

$$b^2 \leq C'(\tau^2(|\Delta b|^2 + |\nabla b|^4 + R^2) + 1). \quad (\text{A.18})$$

Combining these facts with (A.14), we may choose δ in (A.17) sufficiently small such that

$$|X'| \leq \epsilon(r) + \frac{1}{10} \int_{-\tau_0}^{-\theta\tau_0} \int_M \tau(|\nabla^2 b|^2 + |\text{Rc}|^2)\phi^r e^{sb} dv_t dt + \log \theta^{-1}. \quad (\text{A.19})$$

Similarly, we compute

$$\begin{aligned} &\partial_t \int_M \tau R e^{sb} w \phi^r dV_t \\ &= \int_M \{\square(\tau R e^{sb})\phi^r w - \tau R e^{sb} \square^*(\phi^r w)\} dV_t \\ &= \int_M \{(\square(\tau R) e^{sb} + \tau R \square e^{sb} - 2\tau\langle\nabla R, \nabla e^{sb}\rangle)\phi^r w + \tau R e^{sb} ((\Delta\phi^r + \phi_t^r)w + 2\langle\nabla w, \nabla\phi^r\rangle)\} dV_t \\ &= \int_M \{(\square(\tau R) e^{sb} + \tau R \square e^{sb})\phi^r w + 2\tau R (\Delta e^{sb} - \langle\nabla e^{sb}, \nabla b\rangle)\phi^r w\} dV_t + Y_t \\ &= \int_M \{2\tau|\text{Rc}|^2 - R + \tau R (s\square b + (s^2 - 2s)|\nabla b|^2 + 2s\Delta b)\} \phi^r e^{sb} dv_t + Y_t \\ &\geq \int_M \{2\tau|\text{Rc}|^2 - R - Cs(\tau(R^2 + |\nabla b|^4 + (\Delta b)^2) + R)\} \phi^r e^{sb} dv_t + Y_t, \end{aligned} \quad (\text{A.20})$$

where

$$\begin{aligned} Y_t &:= \int_M \tau R e^{sb} ((\Delta\phi^r + \phi_t^r)w + 2\langle\nabla w, \nabla\phi^r\rangle + 2s\langle\nabla b, \nabla\phi^r\rangle) dV_t \\ &= \int_M \tau R (\Delta\phi^r + \phi_t^r + (2s - 2)\langle\nabla b, \nabla\phi^r\rangle) e^{sb} dv_t. \end{aligned}$$

We define similarly $Y := \int_{-\tau_0}^{-\theta\tau_0} |Y_t| dt$. Then it follows from Lemma A.2 and inequalities (2.12) to (2.15) that

$$\begin{aligned} |Y'| &\leq \epsilon(r) + C \int_{-\tau_0}^{-\theta\tau_0} \int_M (\delta^{-1}|\nabla\phi^r|^2(\phi^r)^{-1}\tau|\nabla b|^2 + \delta\tau R^2\phi^r) e^{sb} dv_t dt \\ &\leq \epsilon(r) + C \int_{-\tau_0}^{-\theta\tau_0} \int_M \delta\tau R^2\phi^r e^{sb} dv_t dt \end{aligned}$$

$$\leq \epsilon(r) + \frac{1}{10} \int_{-\tau_0}^{-\theta\tau_0} \int_M \tau(|\nabla^2 b|^2 + |\text{Rc}|^2) \phi^r e^{sb} dv_t dt, \tag{A.21}$$

for δ sufficiently small.

Combining (A.16) and (A.20), we obtain

$$\begin{aligned} & \partial_t \int_M (\tau R + u) e^{sb} \phi^r dv_t \\ & \geq \int_M \left\{ -Cs (\tau(|\nabla b|^4 + |\nabla^2 b|^2 + R^2) + \tau^{-1} + R) + 2\tau \left| \text{Rc} + \nabla^2 b - \frac{g}{2\tau} \right|^2 + 2\tau |\text{Rc}|^2 - R \right\} \\ & \quad \times \phi^r e^{sb} dv_t + X_t + Y_t \\ & \geq \int_M \left\{ -Cs (\tau(|\nabla b|^4 + |\nabla^2 b|^2 + R^2)) + \tau(|\nabla^2 b|^2 + |\text{Rc}|^2) \right\} \phi^r e^{sb} dv_t + X_t + Y_t - C'\tau^{-1}, \end{aligned} \tag{A.22}$$

where we have used Lemma A.2.

If s is sufficiently small, it follows from (A.22) and (A.14) that

$$\begin{aligned} & \partial_t \int_M (\tau R + u) e^{sb} \phi^r dv_t \\ & \geq \frac{1}{2} \int_M \tau(|\nabla^2 b|^2 + |\text{Rc}|^2) e^{sb} \phi^r dv_t + X_t + Y_t - C'\tau^{-1} + \epsilon(r). \end{aligned} \tag{A.23}$$

By integration from $-\tau_0$ to $-\theta\tau_0$, we obtain from (A.23), Lemma A.2, (A.19) and (A.21) that

$$\int_{-\tau_0}^{-\theta\tau_0} \int_M \tau(|\nabla^2 b|^2 + |\text{Rc}|^2) e^{sb} \phi^r dv_t dt \leq C' \log \theta^{-1} + \epsilon(r).$$

Letting $r \rightarrow \infty$, we obtain

$$\int_{-\tau_0}^{-\theta\tau_0} \int_M \tau(|\nabla^2 b|^2 + |\text{Rc}|^2) e^{sb} dv_t dt \leq C' \log \theta^{-1}. \tag{A.24}$$

Thus the inequality (A.15) follows from the combination of (A.24) and Lemma A.3. □