

Recent Developments on Hamilton’s Ricci flow

Huai-Dong Cao, Bing-Long Chen, Xi-Ping Zhu

ABSTRACT. In 1982, Hamilton [41] introduced the Ricci flow to study compact three-manifolds with positive Ricci curvature. Through decades of works of many mathematicians, the Ricci flow has been widely used to study the topology, geometry and complex structure of manifolds. In particular, Hamilton’s fundamental works (cf. [12]) in the past two decades and the recent breakthroughs of Perelman [80, 81, 82] have made the Ricci flow one of the most intricate and powerful tools in geometric analysis, and led to the resolutions of the famous Poincaré conjecture and Thurston’s geometrization conjecture in three-dimensional topology.

In this survey, we will review the recent developments on the Ricci flow and give an outline of the Hamilton-Perelman proof of the Poincaré conjecture, as well as that of a proof of Thurston’s geometrization conjecture.

1. Analytic Aspect

1.1. Short-time Existence and Uniqueness. Let M be an n -dimensional manifold without boundary. The Ricci flow

$$\partial_t g = -2Ric$$

introduced by Hamilton [41] is a degenerate parabolic evolution system on metrics. In his seminal paper [41], Hamilton used the Nash-Moser implicit function theorem to prove the following short-time existence and uniqueness theorem for the Ricci flow on compact manifolds.

THEOREM 1.1 (Hamilton [41]). *Let $(M, g_{ij}(x))$ be a compact Riemannian manifold. Then there exists a constant $T > 0$ such that the Ricci flow $\partial_t g = -2Ric$, with $g_{ij}(x, 0) = g_{ij}(x)$, admits a unique smooth solution $g_{ij}(x, t)$ for all $x \in M$ and $t \in [0, T)$.*

The degeneracy of the system is caused by the diffeomorphism invariance of the equation. By composing the Ricci flow with a family of suitably chosen diffeomorphisms, one can obtain a strictly parabolic system. This is the De Turck trick. The resulting system is called Ricci-De Turck flow. By using this trick, De Turck [32] gave a simpler proof of the above short-time existence and uniqueness result.

In 1989, Shi [91] generalized the above short-time existence result to complete noncompact manifolds with bounded curvature.

THEOREM 1.2 (Shi [91]). *Let $(M, g_{ij}(x))$ be a complete noncompact Riemannian manifold of dimension n with bounded curvature. Then there exists a constant $T > 0$ such that the initial value problem*

$$\begin{cases} \frac{\partial}{\partial t} g_{ij}(x, t) = -2R_{ij}(x, t), \text{ on } M, t > 0, \\ g_{ij}(x, 0) = g_{ij}(x), \text{ on } M, \end{cases}$$

admits a smooth solution $g_{ij}(x, t)$, $t \in [0, T]$, with bounded curvature.

For the uniqueness of Ricci flow on complete noncompact manifolds, the situation is a little subtle. It is well known that, without extra growth conditions on the solutions, the uniqueness for the standard heat equation does not always hold. For example, even the simplest linear heat equation on \mathbb{R} with zero as initial data has a nontrivial solution which grows faster than $e^{a|x|^2}$ for any $a > 0$ whenever $t > 0$. The bounded curvature condition for the Ricci flow in some sense resembles the growth assumption $e^{a|x|^2}$ for the heat equation. Heuristically, it is natural to ask the uniqueness of Ricci flow in the class of bounded curvature solutions.

Recently, the last two authors proved the following uniqueness theorem.

THEOREM 1.3 (Chen-Zhu [24]). *Let $(M, g_{ij}(x))$ be a complete noncompact Riemannian manifold of dimension n with bounded curvature. Let $g_{ij}(x, t)$ and $\bar{g}_{ij}(x, t)$ be two solutions to the Ricci flow on $M \times [0, T]$ with $g_{ij}(x)$ as the initial data and with bounded curvatures. Then $g_{ij}(x, t) = \bar{g}_{ij}(x, t)$ for all $(x, t) \in M \times [0, T]$.*

We remark that Perelman [81] sketched a different proof of the above uniqueness result for a special rotationally symmetric initial metric on \mathbb{R}^3 . The detailed exposition of Perelman’s uniqueness result was given by Lu-Tian [63].

1.2. Shi’s Local Derivative Estimates. In the course of proving his short-time existence theorem in the noncompact case, Shi also obtained the following very useful local derivative estimates.

THEOREM 1.4 (Shi [91]). *There exist positive constants θ , $C_k, k=1, 2, \dots$, depending only on the dimension with the following property. Suppose that the curvature of a solution to the Ricci flow is bounded*

$$|Rm| \leq A, \text{ on } B_t(x_0, r_0) \times \left[0, \frac{\theta}{A}\right],$$

where $B_t(x_0, r_0)$ is compactly contained in the manifold, then we have

$$|\nabla^k Rm(p, t)|^2 \leq C_k A^2 \left(\frac{1}{r^{2k}} + \frac{1}{t^k} + A^k \right),$$

on $B_t(x_0, \frac{r_0}{2})$, $t \in [0, \frac{\theta}{A}]$, for $k=1, 2, \dots$

1.3. Advanced Maximum Principles. Maximum principle is a fundamental and powerful tool for studying heat equations. For Ricci flow, this principle was established by Hamilton [41, 42]. Roughly speaking, Hamilton's maximum principle states that if solutions to the corresponding ODE system always persist in some convex set C when they start from C , then the solutions to the original PDE system will also remain so as long as they stay in C at $t=0$. It turns out many key estimates, such as the Hamilton-Ivey pinching estimate, the Li-Yau-Hamilton estimate, are all proved by using this principle.

To introduce Hamilton's maximum principle, let us start with some basic set-up. We assume $(M, g_{ij}(x, t))$, $t \in [0, T]$, is a smooth complete solution to the Ricci flow with bounded curvature. Let V be an abstract vector bundle over M with a metric $h_{\alpha\beta}$, and connection $\nabla = \Gamma_{i\beta}^\alpha$ compatible with h . Now we may form the Laplace $\Delta\sigma = g^{ij}\nabla_i\nabla_j\sigma$ which acts on the sections $\sigma \in \Gamma(V)$ of V . Suppose $M_{\alpha\beta}(x, t)$ is a family of bounded symmetric bilinear forms on V satisfying the equation

$$(1.1) \quad \frac{\partial}{\partial t} M_{\alpha\beta} = \Delta M_{\alpha\beta} + u^i \nabla_i M_{\alpha\beta} + N_{\alpha\beta},$$

where $u^i(t)$ is a time-dependent uniform bounded vector field on the manifold M , and $N_{\alpha\beta} = \mathcal{P}(M_{\alpha\beta}, h_{\alpha\beta})$ is a polynomial in $M_{\alpha\beta}$ formed by contracting products of $M_{\alpha\beta}$ with itself using the metric $h = \{h_{\alpha\beta}\}$. In [41], Hamilton established the following weak maximum principle: *Let $M_{\alpha\beta}$ be a bounded solution to (1.1) and suppose $N_{\alpha\beta}$ satisfies the condition that*

$$N_{\alpha\beta} v^\alpha v^\beta \geq 0 \quad \text{whenever} \quad M_{\alpha\beta} v^\beta = 0,$$

for $0 \leq t \leq T$. If $M_{\alpha\beta} \geq 0$ at $t=0$, then it remains so for $0 \leq t \leq T$.

Hamilton [42] also established a strong maximum principle for solutions to equation (1.1): *Let $M_{\alpha\beta}$ be a bounded solution to (1.1) with $u^i=0$, and $N_{\alpha\beta}$ satisfies*

$$N_{\alpha\beta} \geq 0$$

whenever $M_{\alpha\beta} \geq 0$. Suppose $M_{\alpha\beta} \geq 0$ at $t=0$. Then there exists an interval $0 < t < \delta$ on which the rank of $M_{\alpha\beta}$ is constant and the null space of $M_{\alpha\beta}$ is invariant under parallel translation and invariant in time and also lies in the null space of $N_{\alpha\beta}$.

The evolution equation of the curvature operator $M_{\alpha\beta}$ of the Ricci flow satisfies

$$(1.2) \quad \frac{\partial M_{\alpha\beta}}{\partial t} = \Delta M_{\alpha\beta} + M_{\alpha\beta}^2 + M_{\alpha\beta}^\#$$

where $M_{\alpha\beta}^\# = C_\alpha^{\xi\gamma} C_\beta^{\eta\theta} M_{\xi\eta} M_{\gamma\theta}$ and $C_\alpha^{\beta\gamma} = \langle [\phi^\beta, \phi^\gamma], \phi^\alpha \rangle$ are Lie structural constants in a standard basis of the Lie algebra consisting of two-forms. Choosing an orthonormal frame such that $M_{\alpha\beta}$ is diagonal with eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots$, then

$$M_{11}^2 + M_{11}^\# = \lambda_1^2 + \sum_{\xi, \eta \geq 2} (C_1^{\xi\eta})^2 \lambda_\xi \lambda_\eta.$$

So $N_{\alpha\beta} = M_{\alpha\beta}^2 + M_{\alpha\beta}^\#$ satisfies the assumption in Hamilton’s strong maximum principle. Note also that if $M_{\alpha\alpha} = 0$ for $\alpha \leq k$, and $M_{\alpha\alpha} > 0$ for $\alpha > k$, then the condition $M_{\alpha\alpha}^\# = 0$ for $\alpha \leq k$ implies

$$C_\alpha^{\xi\gamma} = \langle \phi^\alpha, [\phi^\xi, \phi^\gamma] \rangle = 0, \quad \text{if } \alpha \leq k \text{ and } \xi, \gamma > k.$$

This says that the image of $M_{\alpha\beta}$ is a Lie subalgebra. So Hamilton’s strong maximum principle implies

THEOREM 1.5 (Hamilton [42]). *Suppose the curvature operator $M_{\alpha\beta}$ of the initial metric is nonnegative. Then, under the Ricci flow, for some interval $0 < t < \delta$ the image of $M_{\alpha\beta}$ is a Lie subalgebra of $so(n)$ which has constant rank and is invariant under parallel translation and invariant in time.*

This implies that under the Ricci flow, any compact manifold with nonnegative curvature operator which admits no strictly positive curvature operator has special holonomy group. This theorem may lead to complete topological classification of compact manifolds with nonnegative curvature operators.

Note that the nonnegative curvature operators form a convex cone. For general convex set, Hamilton [42] developed an advanced maximum principle as follows.

Let $V \rightarrow M$ be a vector bundle with a fixed bundle metric h_{ab} and

$$\nabla_t : \Gamma(V) \rightarrow \Gamma(V \otimes T^*M), \quad t \in [0, T]$$

be a smooth family of time-dependent connection compatible with h_{ab} . We may form the Laplacian

$$\Delta_t \sigma = g^{ij}(x, t)(\nabla_t)_i(\nabla_t)_j \sigma,$$

for $\sigma \in \Gamma(V)$.

Let $N : V \times [0, T] \rightarrow V$ be a fiber preserving map, i.e., $N(x, \sigma, t)$ is a time-dependent vector field defined on the bundle V and tangent to the fibers. Let K be a closed subset of V satisfying

(H1) K is invariant under parallel translation defined by the connection ∇_t for each $t \in [0, T]$;

(H2) the set $K_x \triangleq V_x \cap K$ is closed and convex in each fiber V_x .

Consider the following heat type equation

$$(1.3) \quad \frac{\partial}{\partial t} \sigma(x, t) = \Delta_t \sigma(x, t) + u^i (\nabla_t)_i \sigma(x, t) + N(x, \sigma(x, t), t)$$

where $u^i = u^i(t)$ is a time-dependent vector field on M which is uniformly bounded on $M \times [0, T]$, and $N(x, \sigma, t)$ is continuous in x, t and satisfies

$$|N(x, \sigma_1, t) - N(x, \sigma_2, t)| \leq C_B |\sigma_1 - \sigma_2|$$

for all $x \in M$, $t \in [0, T]$ and $|\sigma_1| \leq B$, $|\sigma_2| \leq B$, where C_B is a positive constant depending only on B . We will also consider the corresponding ODE system

$$(1.4) \quad \frac{d\sigma_x}{dt} = N(x, \sigma_x, t)$$

for $\sigma_x = \sigma_x(t)$ in each fiber V_x . Hamilton's advanced maximum principle is the following:

THEOREM 1.6 (Hamilton [42]). *Let K be a closed subset of V satisfying (H1) and (H2). Suppose that for any $x \in M$ and any initial time $t_0 \in [0, T]$, and any solution $\sigma_x(t)$ of the ODE (1.4) which starts in K_x at t_0 , the solution $\sigma_x(t)$ will remain in K_x for all later times. Then for any initial time $t_0 \in [0, T]$ the solution $\sigma(x, t)$ of the PDE (1.3) will remain in K for all later times if $\sigma(x, t)$ starts in K at time t_0 and the solution $\sigma(x, t)$ is uniformly bounded with respect to the bundle metric h_{ab} on $M \times [t_0, T]$.*

1.4. Hamilton-Ivey Pinching Estimate. The advanced maximum principle has several significant applications in Ricci flow. One of them is the following Hamilton-Ivey pinching estimate in dimension three.

THEOREM 1.7 ([47, 52, 49]). *Suppose we have a solution to the Ricci flow on a three-manifold which is complete with bounded curvature for each $t \geq 0$. Assume at $t=0$ the eigenvalues $\lambda \geq \mu \geq \nu$ of the curvature operator at each point is bounded below by $\nu \geq -1$ and the scalar curvature at each point is bounded below by $R = (\lambda + \mu + \nu) \geq -1$. Then at all points and all times $t \geq 0$ we have the pinching estimate*

$$(1.5) \quad R \geq (-\nu)[\log(-\nu) + \log(1 + t) - 3]$$

whenever $\nu < 0$.

This pinching estimate roughly says that if a solution to the Ricci flow on a three-manifold becomes singular at a time T , then the most negative sectional curvature will be small compared to the most positive sectional curvature. Thus after rescaling around the singularity, one will obtain a non-negatively curved limit. This fact will play a crucial role in the classification of singularities of the Ricci flow on three-dimensional manifolds.

1.5. Li-Yau-Hamilton inequalities. The Harnack inequality, comparing values of a positive solution at different points in space-time, is a very useful property of parabolic equations. In their seminal paper [61], Li-Yau studied the heat equation and found a fundamental important inequality for the gradient of positive solutions. Integrating this inequality along suitable paths, they obtained the usual Harnack inequality.

In 1993, Hamilton [44] discovered a highly nontrivial matrix version of Li-Yau type inequality for the Ricci flow on complete manifolds with positive curvature operator. This inequality is called the Li-Yau-Hamilton inequality. We now describe these inequalities in detail.

Let us begin with the original Li-Yau inequality for the heat equation

$$(1.6) \quad \left(\frac{\partial}{\partial t} - \Delta \right) u = 0.$$

THEOREM 1.8 (Li-Yau [61]). *Let (M, g_{ij}) be an n -dimensional complete Riemannian manifold with nonnegative Ricci curvature. Let $u(x, t)$ be any positive solution to the heat equation (1.6) for $t \in [0, \infty)$. Then we have*

$$(1.7) \quad \frac{\partial u}{\partial t} - \frac{|\nabla u|^2}{u} + \frac{n}{2t} u \geq 0 \text{ on } M \times (0, \infty).$$

The proof is a computation of $(\partial_t - \Delta) \left(\frac{\partial}{\partial t} \log u - |\nabla \log u|^2 \right)$ and an application of the maximum principle.

Next recall that under the Ricci flow on a Riemann surface the scalar curvature satisfies the following heat type equation

$$\left(\frac{\partial}{\partial t} - \Delta \right) R = R^2.$$

By the maximum principle, the positivity of the curvature is preserved by the Ricci flow. Hamilton considered the quantity $Q = \frac{\partial}{\partial t} \log R - |\nabla \log R|^2$ and computed

$$\frac{\partial}{\partial t} \left(Q + \frac{1}{t} \right) \geq \Delta \left(Q + \frac{1}{t} \right) + 2 \nabla \log R \cdot \nabla \left(Q + \frac{1}{t} \right) + \left(Q - \frac{1}{t} \right) \left(Q + \frac{1}{t} \right).$$

From the maximum principle, it follows

THEOREM 1.9 (Hamilton [43]). *Let $g_{ij}(x, t)$ be a complete solution to the Ricci flow with bounded curvature on a surface M . Assume the scalar*

curvature of the initial metric is positive. Then

$$\frac{\partial R}{\partial t} - \frac{|\nabla R|^2}{R} + \frac{R}{t} \geq 0.$$

For higher dimensions, the curvature operator is a matrix and satisfies a quite complicated evolution equation. Apparently, the first important thing is to find out the correct expression of the quantity we want to estimate. To this end, Hamilton observed a very useful fact: the Li-Yau inequality becomes an equality on the heat kernel $h(x, t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$ on \mathbb{R}^n which can be considered as an "expanding soliton." In fact, the Li-Yau inequality is equivalent to the nonnegativity of the following quadratic form:

$$(1.8) \quad \frac{\partial u}{\partial t} + 2\langle \nabla u, V \rangle + u|V|^2 + \frac{n}{2t}u \geq 0.$$

Substituting the optimal vector field $V = -\frac{\nabla u}{u}$, we recover (1.7). To illustrate the idea of forming the quantity (1.8), let us check the heat kernel $u(x, t) = (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$. Differentiating the function u , we get

$$(1.9) \quad \nabla_j u + uV_j = 0$$

where $V_j = \frac{x_j}{2t}$. Differentiating (1.9) again, we have

$$(1.10) \quad \nabla_i \nabla_j u + \nabla_i u V_j + \frac{u}{2t} \delta_{ij} = 0.$$

To make the expression in (1.10) symmetric in i, j , we multiply V_i to (1.9) and add it to (1.10)

$$(1.11) \quad \nabla_i \nabla_j u + \nabla_i u V_j + \nabla_j u V_i + uV_i V_j + \frac{u}{2t} \delta_{ij} = 0.$$

Taking the trace in (1.11), we obtain the Li-Yau expression

$$\frac{\partial u}{\partial t} + 2\nabla u \cdot V + u|V|^2 + \frac{n}{2t}u = 0$$

for the heat kernel u . Moreover, the above formulation suggests the matrix Li-Yau type inequality (1.11) discussed in [50].

Based on similar considerations, Hamilton found the matrix Li-Yau type expression for Ricci flow which vanishes on expanding gradient Ricci solitons, and established the following fundamental important inequality.

THEOREM 1.10 (Hamilton [44]). *Let $g_{ij}(x, t)$ be a complete solution with bounded curvature to the Ricci flow on a manifold M^n for $t \in (0, T)$ and suppose the curvature operator of $g_{ij}(x, t)$ is nonnegative. Then for any one-form W_a and any two-form U_{ab} we have*

$$M_{ab}W_aW_b + 2P_{abc}U_{ab}W_c + R_{abcd}U_{ab}U_{cd} \geq 0$$

for $x \in M$ and $t \in (0, T)$, where

$$\begin{aligned} M_{ab} &= \Delta R_{ab} - \frac{1}{2} \nabla_a \nabla_b R + 2R_{acbd} R_{cd} - R_{ac} R_{bc} + \frac{1}{2t} R_{ab}, \\ P_{abc} &= \nabla_a R_{bc} - \nabla_b R_{ac}. \end{aligned}$$

Consequently, for any one-form V_a , we have

$$(1.12) \quad \frac{\partial R}{\partial t} + \frac{R}{t} + 2\nabla_a R \cdot V_a + 2R_{ab} V_a V_b \geq 0.$$

Integrating (1.12) along suitable space-time paths, we obtain

COROLLARY 1.11 (Hamilton [44]). *Under the assumption of Theorem 1.10, for any $x_1, x_2 \in M$, $t_1 < t_2$, we have*

$$(1.13) \quad R(x_1, t_1) \leq \frac{t_2}{t_1} e^{d_{t_1}(x_1, x_2)^2/2(t_2-t_1)} \cdot R(x_2, t_2).$$

We remark that Brendle [5] has extended the Li-Yau-Hamilton inequality under certain curvature assumption which is weaker than nonnegative curvature operator but stronger than nonnegative sectional curvature. Also in the Kähler case, the above Li-Yau-Hamilton inequality has been generalized by the first author in [9] under the weaker assumption of nonnegative bisectional curvature.

Define the Lie bracket on $\Lambda^2 M \oplus \Lambda^1 M$ by

$$[U \oplus W, V \oplus X] = [U, V] \oplus (U \lrcorner X - V \lrcorner W),$$

and a degenerate inner product

$$\langle U \oplus V, W \oplus X \rangle = \langle U, W \rangle.$$

Hamilton [47] observed that the equation satisfied by the quantity Q can be formally written as

$$(1.14) \quad \frac{\partial}{\partial t} Q - \Delta Q = Q^2 + Q^\#$$

under appropriate space-time extensions of U and W . This fascinating structure led Hamilton to write in his survey [47]:

“The geometry would seem to suggest that the Harnack inequality is some sort of jet extension of positive curvature operator on some bundle including translations as well as rotation, this is somehow all related to solitons where the solution moves by translation. It would be very helpful to have a proper understanding of this suggestion.”

In [27], Chow and Chu verified this geometric interpretation of Hamilton, by showing that the Li-Yau-Hamilton quantity is in fact the curvature of a torsion free connection compatible with a degenerate metric on space-time.

2. Sphere Theorems and Uniformization Conjectures

2.1. Differential Sphere Theorems. One of the basic problems in Riemannian geometry is the classification of positively curved manifolds. The classical (Topological) Sphere Theorem due to Rauch [87], Berger [4], Klingenberg [57] (cf. [18]) states that *a simply connected Riemannian manifold with 1/4-pinched sectional curvature, in the sense that sectional curvatures at each point varying in the interval (1, 4], is homeomorphic to \mathbb{S}^n .* In 1951, Rauch [87] actually conjectured that such a Riemannian manifold is diffeomorphic to \mathbb{S}^n . This question is known as the *Differential Sphere Theorem*. The classical result of differential sphere theorem under δ -pinched assumption for δ close to 1 (with the best $\delta=0.87$) was obtained by Gromoll [36], Calabi, Sugimoto-Shiohama [94], Karcher [55], Ruh [88] etc (cf. [18]). Note that the well-known theorem of Cheeger-Gromoll-Meyer (cf. [18]) asserts that any complete *noncompact* Riemannian manifold of positive sectional curvature is diffeomorphic to Euclidean space \mathbb{R}^n .

The Ricci flow has profound application in proving various differential sphere theorems. In his 1982 seminal paper [41], Hamilton proved the following famous sphere theorem.

THEOREM 2.1 (Hamilton [41]). *A compact 3-manifold with positive Ricci curvature is diffeomorphic to a spherical space form, i.e., the three-sphere \mathbb{S}^3 or a quotient of it by a finite group of fixed point free isometries in the standard metric.*

The idea of the proof is to study the long-time behavior of the Ricci flow with the given metric of positive Ricci curvature as the initial data and to obtain spherical space forms as its asymptotic limit. A sketch of the proof, as shown in [46], can be described as follows. First of all, if we diagonalize the 3×3 curvature operator matrix $M_{\alpha\beta}$ with eigenvalues $\lambda \geq \mu \geq \nu$, the corresponding ODE system to the evolution equation of the curvature operator is given by

$$\begin{cases} \frac{d}{dt}\lambda = \lambda^2 + \mu\nu \\ \frac{d}{dt}\mu = \mu^2 + \lambda\nu \\ \frac{d}{dt}\nu = \nu^2 + \mu\lambda \end{cases}$$

It is easy to see $\lambda \geq \mu \geq \nu$ is preserved by this system. For any $0 < \delta \leq 2$, we compute

$$(2.1) \quad \frac{d}{dt}(\mu + \nu - \delta\lambda) \geq 0,$$

if $\mu + \nu = \delta\lambda$. This implies $\mu + \nu \geq \frac{\delta}{1+\delta}(\lambda + \mu + \nu)$ is preserved by the above ODE system. That is equivalent to say, for any $0 < \delta' \leq \frac{1}{3}$, the Ricci pinching

$Ric \geq \delta' Rg$ is preserved by the Ricci flow. Moreover, considering the convex set

$$K = \{M_{\alpha\beta} \mid \lambda - \nu \leq C(\lambda + \mu + \nu)^{1-\eta}\},$$

it is not hard to see

$$\frac{d}{dt} \log(\lambda - \nu) = \lambda + \nu - \mu,$$

and

$$\frac{d}{dt} \log(\lambda + \mu + \nu) \geq \left(1 + \left(\frac{\delta}{2(1+\delta)}\right)^2\right) (\lambda + \nu - \mu)$$

if $\mu + \nu \geq \frac{\delta}{1+\delta}(\lambda + \mu + \nu) > 0$. So $\frac{d}{dt} \log \frac{\lambda - \nu}{(\lambda + \mu + \nu)^{1-\eta}} \leq 0$ for some $\eta > 0$. Thus Hamilton's advanced maximum principle implies

$$\frac{\left| Ric - \frac{R}{3}g \right|}{R^{1-\eta}} \leq C.$$

where C is some positive constant. By a blow up argument and using the second Bianchi identity, we then obtain the gradient estimate with small coefficient, i.e., for any $\varepsilon > 0$ there is $C_\varepsilon > 0$ such that

$$\max_{t \leq \tau} \max_{x \in M} |\nabla Rm(x, t)| \leq \varepsilon \max_{t \leq \tau} \max_{x \in M} |Rm(x, t)|^{\frac{3}{2}} + C_\varepsilon.$$

On the other hand, we know that the solution to the Ricci flow exists only for a finite time and curvatures become unbounded. Now dilate the metrics around maximum curvature points so that the maximum curvature becomes one. Combining the gradient estimate, the pinching estimate, and the Bonnet-Meyers theorem, we know the diameter is bounded for the rescaled solution. Furthermore, by Klingenberg's injectivity radius estimate for $\frac{1}{4}$ -pinched manifold and Shi's derivative estimate, we may take a smooth convergent subsequence, whose limit is a round sphere \mathbb{S}^3 .

The combination of the above Hamilton's sphere theorem and Hamilton's strong maximum principle gives a complete classification of 3-dimensional compact manifolds with nonnegative Ricci curvature, see [42]. Actually, if the Ricci curvature can't be deformed to strictly positive, then the kernel of Ricci tensor gives rise to a parallel distribution of the tangent bundle. By De Rham splitting theorem, the universal cover is either flat or splits. Consequently, *a compact three-manifold with nonnegative Ricci curvature is diffeomorphic to \mathbb{S}^3 or a quotient of one of the spaces \mathbb{S}^3 or $\mathbb{S}^2 \times \mathbb{R}^1$ or \mathbb{R}^3 by a group of fixed point free isometries in the standard metrics.*

By using his advanced maximum principle in a similar way, Hamilton [42] also proved a 4-dimensional differentiable sphere theorem:

THEOREM 2.2 (Hamilton [42]). *A compact 4-manifold with positive curvature operator is diffeomorphic to the 4-sphere \mathbb{S}^4 or the real projective space $\mathbb{R}P^4$.*

One of the key steps in the proof is to show the existence of pinching sets associated to the cone $C = \{Rm > 0\}$ of positive curvature operators. Here a pinching set Z is a closed convex subset in the space of curvature operators that is invariant under the flow of ODE $\frac{dRm}{dt} = Rm^2 + Rm^\#$ (cf. equation (1.2)) and such that $|\tilde{Rm}| \leq c|Rm|^{1-\delta}$, for some constants $\delta > 0$ and $c > 0$ and all $Rm \in Z$, where $|\tilde{Rm}|$ denotes the traceless part of Rm . The latter condition implies that when $|Rm|$ is large, the rescaled curvature operator with maximal norm one becomes almost constant curvature. When $n = 4$, 2-forms can be written as the direct sum of self-dual 2-forms and anti self-dual forms, hence curvature operators admit block decompositions. Using this fact and by an elaborate argument, Hamilton [42] was able to construct pinching sets associated to the cone $C = \{Rm > 0\}$ such that any compact subset K of C is contained in some pinching set Z . Once this is established, it follows (as in the case of $n = 3$) every initial metric of positive curvature will evolve under the normalized Ricci flow to a round metric in the limit.

In [42], Hamilton also obtained the following classification theorem for four-manifolds with nonnegative curvature operator.

THEOREM 2.3 (Hamilton [42]). *A compact four-manifold with nonnegative curvature operator is diffeomorphic to one of the spaces \mathbb{S}^4 or $\mathbb{C}P^2$ or $\mathbb{S}^2 \times \mathbb{S}^2$ or a quotient of one of the spaces \mathbb{S}^4 or $\mathbb{C}P^2$ or $\mathbb{S}^3 \times \mathbb{R}^1$ or $\mathbb{S}^2 \times \mathbb{S}^2$ or $\mathbb{S}^2 \times \mathbb{R}^2$ or \mathbb{R}^4 by a group of fixed point free isometries in the standard metrics.*

We note that H. Chen [26] extended Theorem 2.2 to 2-positive curvature operator. Here *2-positive curvature operator* means the sum of the least two eigenvalues of the curvature operator is positive. Later, by using the Ricci flow, differential sphere theorems for higher dimensions under some suitable pointwise pinching conditions were obtained by Huisken [51] (see also Margerin [65, 66] and Nishikawa [79]).

Naturally, one would ask if a compact Riemannian manifold M^n , with $n \geq 5$, of positive curvature operator (or 2-positive curvature operator) is diffeomorphic to a space form. This was in fact conjectured so by Hamilton and proved only very recently by Böhm-Wilking [8].

THEOREM 2.4 (Böhm-Wilking [8]). *A compact Riemannian manifold of dimension $n \geq 5$ with a two-positive curvature operator is diffeomorphic to a spherical space form.*

In [8], Böhm-Wilking developed a powerful new method to construct closed convex sets, which are invariant under the Ricci flow, in the space of

curvature operators. They introduced a linear transformation $l_{a,b}$ ($a, b > 0$) on the space of curvature operators defined by $l_{a,b}(Rm) = Rm + aRm_I + bRm_0$ so that it increases the scalar curvature part Rm_I and the traceless Ricci part Rm_0 of Rm by factors of a and b respectively. A crucial property they found is that the associated transformation $D_{a,b}(Rm) = l_{a,b}^{-1}((l_{a,b}Rm)^2 + (l_{a,b}Rm)^\#) - Rm^2 - Rm^\#$ is independent of the Weyl curvature part of Rm . Based on this, they can construct new invariant cones from old ones. By choosing appropriate constants a 's and b 's, this construction gives rise to a (continuous) pinching family $C(s)$, $s \in [0, 1)$, of invariant closed convex cones such that $C(0)$ is the cone of 2-nonnegative curvature operators and, as $s \rightarrow 1$, $C(s)$ approaches $\{cI : c \in \mathbb{R}_+\}$, the set of constant curvature operators. From this pinching family $C(s)$, one can then construct a generalized pinching set F which is a certain special invariant convex set in the space of curvature operators, so that F contains the initial data and $F \setminus C(s)$ is compact for every $s \in [0, 1)$. Since the curvature operator of the evolving metric under the Ricci flow has to diverge to infinity in finite time, it must be contained in every $C(s)$ when the time is large after rescaling. Thus, the solution to the normalized Ricci flow converges to a round metric in the limit.

We have seen that the Ricci flow preserves the *positive curvature operator* condition in all dimensions and preserves the positive *Ricci curvature* condition in dimension 3. On the other hand, in [48] Hamilton also proved that the *positive isotropic curvature* (PIC) condition is preserved by the Ricci flow in dimension 4. We remark that in 1988, by using minimal surface theory, Micallef and Moore [67] were able to prove that any compact simply connected n -dimensional manifold with positive isotropic curvature is homeomorphic to the n -sphere \mathbb{S}^n , and the condition of positive isotropic curvature is weaker than both positive curvature operator and 1/4-pinched. Very recently Brendle-Schoen [6], and independently H. Nguyen [73], proved that the PIC condition is preserved by the Ricci flow in all dimensions $n \geq 4$.¹ More excitingly, Brendle and Schoen [6] showed that when the initial metric has (pointwise) 1/4-pinched sectional curvature (in fact under the weaker curvature condition that $M \times \mathbb{R}^2$ has PIC, see [6]), the Ricci flow will converge to a spherical space form. As a corollary, they proved the long-standing Differential Sphere Theorem.

THEOREM 2.5 (Brendle-Schoen [6]). *Let M be a compact manifold with (pointwise) 1/4-pinched sectional curvature. Then M is diffeomorphic to \mathbb{S}^n or a quotient of \mathbb{S}^n by a group of fixed point free isometries in the standard metrics.*

¹See also Andrews-Nguyen [1] for a proof that 1/4-pinched flag curvature is preserved for $n = 4$ which has some common features.

By generalizing the usual strong maximum principle to a powerful version, Brendle and Schoen [7] even obtained the following rigidity result, which extends the well-known rigidity result of Berger (cf. [18]).

THEOREM 2.6 (Brendle-Schoen [7]). *Let M be a compact manifold with (pointwise) weakly $1/4$ -pinched sectional curvature in the sense that $0 < \text{sect}(P_1) \leq 4\text{sect}(P_2)$ for all two-planes $P_1, P_2 \in T_p M$. If M is not diffeomorphic to a spherical space form, then it is isometric to a locally symmetric space.*

2.2. Kähler Manifolds with Nonnegative Holomorphic Bisectional Curvature. The classical uniformization theorem for Riemann surfaces implies that a complete simply connected Riemann surface with positive curvature is biholomorphic to either the Riemann sphere or the complex plane. The classification (in holomorphic category) of positively curved Kähler manifolds in higher dimensions is one of the most important problems in complex differential geometry. Corresponding to positive sectional curvature condition in Riemannian geometry, one usually considers the positive holomorphic bisectional curvature in complex differential geometry.

Let M^n be a complex n -dimensional compact Kähler manifold. The famous *Frankel conjecture* states that: *if M^n has positive holomorphic bisectional curvature, then it is biholomorphic to the complex projective space $\mathbb{C}P^n$.* This was independently proved by Mori [71] and Siu-Yau [100] by using different methods. After the work of Mori and Siu-Yau, it is natural to ask the similar question for the semi-positive case. This is often called the *generalized Frankel conjecture*. The complex three-dimensional case was first obtained by Bando [3]. When the curvature operator of M^n is assumed to be nonnegative, the result was proved by the first author and Chow [11]. The general case of the generalized Frankel conjecture is proved by Mok [69].

THEOREM 2.7 (Mok [69]). *Let (M^n, h) be a compact complex n -dimensional Kähler manifold of nonnegative holomorphic bisectional curvature and let (\tilde{M}^n, \tilde{h}) be its universal covering space. Then there exists nonnegative integers k, N_1, \dots, N_l, p and irreducible compact Hermitian symmetric spaces M_1, \dots, M_p of rank ≥ 2 such that (\tilde{M}^n, \tilde{h}) is isometrically biholomorphic to*

$$(\mathbb{C}^k, g_0) \times (\mathbb{C}P^{N_1}, \theta_1) \times \cdots \times (\mathbb{C}P^{N_l}, \theta_l) \times (M_1, g_1) \times \cdots \times (M_p, g_p)$$

where g_0 denotes the Euclidean metric on \mathbb{C}^k , g_1, \dots, g_p are canonical metrics on M_1, \dots, M_p , and $\theta_i, 1 \leq i \leq l$, is a Kähler metric on $\mathbb{C}P^{N_i}$ carrying nonnegative holomorphic bisectional curvature.

Mok’s method of proving the generalized Frankel conjecture in [69] depends on Mori’s theory of rational curves on Fano manifolds, so his proof is

not completely transcendental. Recently, by using the strong maximum principle of Brendle-Schoen in [7], H. L. Gu [39] gave a simpler and completely transcendental proof of the generalized Frankel conjecture. The above Mok’s theorem on generalized Frankel conjecture is indeed a factorization theorem for compact case. Based on the arguments in [39], we now formulate a new factorization theorem for noncompact cases as follows.

THEOREM 2.8. *Let (M, h) be a complete noncompact Kähler manifold with bounded and nonnegative holomorphic bisectional curvature. Then one of the following holds:*

- (i) *M admits a Kähler metric with bounded and positive bisectional curvature;*
- (ii) *The universal cover \hat{M} of M splits holomorphically, isometrically and nontrivially as*

$$\hat{M} = \mathbb{C}^k \times M_1 \times \cdots \times M_{l_1} \times N_1 \times \cdots \times N_{l_2}$$

where k, l_1, l_2 are nonnegative integers, \mathbb{C}^k is the complex Euclidean space with flat metric, $M_i, 1 \leq i \leq l_1$, are complete (compact or noncompact) Kähler manifolds with bounded and nonnegative bisectional curvature admitting a Kähler metric with bounded and positive bisectional curvature, $N_j, 1 \leq j \leq l_2$, are irreducible compact Hermitian symmetric spaces of rank ≥ 2 with the canonical metrics.

PROOF. We evolve the metric h by the Kähler Ricci flow:

$$\begin{cases} \frac{\partial}{\partial t} g_{i\bar{j}}(x, t) = -R_{i\bar{j}}(x, t), \\ g_{i\bar{j}}(x, 0) = h_{i\bar{j}}(x). \end{cases}$$

Then by Shi’s short-time existence theorem, we know that there is a $T > 0$ such that the Ricci flow has a smooth solution $g_{i\bar{j}}(t)$ with bounded curvature for $t \in [0, T)$. It is well-known (from [69] and [91]) that the solution $g_{i\bar{j}}(t)$ still has nonnegative holomorphic bisectional curvature. By lifting the solution to the universal cover \hat{M} of M , then the pull back evolving metric $\hat{g}_{i\bar{j}}(t)$ is a solution to the Ricci flow on \hat{M} . Clearly we may assume that the holomorphic bisectional curvature of the solution $\hat{g}_{i\bar{j}}(t)$ vanishes somewhere at each time $t \in [0, T)$; otherwise we will have case (i).

By applying the standard De Rham decomposition theorem, we know that the universal cover $(\hat{M}, \hat{g}_{i\bar{j}}(t)), t \in [0, \delta)$, can be isometrically and holomorphically splitted as

$$(\hat{M}, \hat{g}_{i\bar{j}}(t)) = (\mathbb{C}^k, \hat{g}^0) \times (\hat{M}_1, \hat{g}_{i\bar{j}}^1(t)) \times \cdots \times (\hat{M}_p, \hat{g}_{i\bar{j}}^p(t))$$

for some $\delta \in (0, T)$, where each $(\hat{M}_\alpha, \hat{g}_{i\bar{j}}^\alpha(t)), 1 \leq \alpha \leq p$, is irreducible and non-flat, \hat{g}^0 is the standard flat metric.

Consider each irreducible and non-flat factor $(\hat{M}_\alpha, \hat{g}_{i\bar{j}}^\alpha(t)), 1 \leq \alpha \leq p$. Suppose $(\hat{M}_\alpha, \hat{g}_{i\bar{j}}^\alpha(0))$ is not locally symmetric. We shall show that $(\hat{M}_\alpha, \hat{g}_{i\bar{j}}^\alpha(t)), t \in (0, \delta')$, has positive holomorphic bisectional curvature everywhere on $(0, \delta')$ for some $0 < \delta' < \delta$.

Since the smooth limit of locally symmetric space is also locally symmetric, we obtain that there exists $\delta' \in (0, T)$ such that $(\hat{M}_\alpha, \hat{g}_{i\bar{j}}^\alpha(t))$ is not locally symmetric for $t \in (0, \delta')$. Combining the Kählerity of $\hat{g}_{i\bar{j}}^\alpha(t)$ and Berger’s holonomy theorem, we know that the holonomy group $\text{Hol}(\hat{g}^\alpha(t))$ of $(\hat{M}_\alpha, \hat{g}^\alpha(t))$ is $U(n_\alpha)$, where $n_\alpha = \dim_C \hat{M}_\alpha$.

Recall the evolution equation of holomorphic bisectional curvature under an evolving orthonormal frame $\{e_i\}$ according to Hamilton [42]

$$\frac{\partial}{\partial t} \hat{R}_{i\bar{i}j\bar{j}}^\alpha = \Delta \hat{R}_{i\bar{i}j\bar{j}}^\alpha + \sum_{p,q} (\hat{R}_{i\bar{i}p\bar{q}}^\alpha \hat{R}_{q\bar{p}j\bar{j}}^\alpha - |\hat{R}_{i\bar{i}p\bar{q}}^\alpha|^2 + |\hat{R}_{i\bar{i}j\bar{p}q}^\alpha|^2).$$

Let P be the fiber bundle with the fixed metric h and the fiber P_x over $x \in \hat{M}^\alpha$ consisting of all 2-vectors $\{X, Y\} \subset T_x^{1,0}(\hat{M}^\alpha)$. Now define a function u on $P \times (0, \delta')$ by

$$u(\{X, Y\}, t) = \hat{R}^\alpha(X, \bar{X}, Y, \bar{Y}),$$

where \hat{R}^α denotes the pull-back of the curvature tensor of $\hat{g}_{i\bar{j}}^\alpha(t)$. For simplicity, we denote $R = \hat{R}^\alpha$. Since $(\hat{M}^\alpha, \hat{g}_{i\bar{j}}^\alpha(t))$ has nonnegative holomorphic bisectional curvature, we have $u \geq 0$. Let $N = \{(\{X, Y\}, t) | u(\{X, Y\}, t) = 0, X \neq 0, Y \neq 0\} \subset P \times (0, \delta')$. We will show in the following that if N is not empty then it is invariant under the parallel translation.

For fixed e_i , consider the Hermitian form $H_i(X, Y) = R(e_i, \bar{e}_i, X, \bar{Y})$ and let $\{E_p\}$ be an orthonormal basis associated to eigenvectors of H_i . In these basis we have

$$\sum_{p,q} R_{i\bar{i}p\bar{q}} R_{q\bar{p}j\bar{j}} = \sum_p R(e_i, \bar{e}_i, E_p, \bar{E}_p) R(E_p, \bar{E}_p, e_j, \bar{e}_j),$$

and

$$\sum_{p,q} |R_{i\bar{i}p\bar{q}}|^2 = \sum_{p,q} |R(e_i, \bar{E}_p, e_j, \bar{E}_q)|^2.$$

Moreover, we claim

$$\sum_{p,q} R_{i\bar{i}p\bar{q}} R_{q\bar{p}j\bar{j}} - \sum_{p,q} |R_{i\bar{i}p\bar{q}}|^2 \geq c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 u(\{e_i, e_j\}, t)(\xi, \xi)\}$$

for some constant $c_1 > 0$, where V denotes the vertical spaces of the fiber bundle P .

Indeed, inspired by Mok [69], for any given $\varepsilon_0 > 0$ and each fixed $q \in \{1, 2, \dots, n\}$, we consider the function

$$\tilde{G}_q(\varepsilon) = (R + \varepsilon_0 R_0)(e_i + \varepsilon E_q, \overline{e_i + \varepsilon E_q}, e_j + \varepsilon \sum_p C_p E_p, \overline{e_j + \varepsilon \sum_p C_p E_p}),$$

where R_0 is a curvature operator defined by $(R_0)_{i\bar{j}k\bar{l}} = g_{i\bar{j}}g_{k\bar{l}} + g_{i\bar{l}}g_{k\bar{j}}$ and C_p are complex constants to be determined later. For simplicity, we denote $\tilde{R} = R + \varepsilon_0 R_0$, then

$$\tilde{G}_q(\varepsilon) = \tilde{R}(e_i + \varepsilon E_q, \overline{e_i + \varepsilon E_q}, e_j + \varepsilon \sum_p C_p E_p, \overline{e_j + \varepsilon \sum_p C_p E_p}).$$

Then a direct computation gives

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \tilde{G}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \tilde{R}(E_q, \overline{E_q}, e_j, \overline{e_j}) + \sum_p |C_p|^2 \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_p}) \\ &\quad + 2\operatorname{Re} \sum_p \overline{C_p} \tilde{R}(e_i, \overline{E_q}, e_i, \overline{E_p}) + 2\operatorname{Re} \sum_p C_p \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_q}). \end{aligned}$$

Writing $C_p = x_p e^{i\theta_p}$, ($p \geq 1$) for some real numbers x_p, θ_p to be determined later, the above identity becomes:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \tilde{G}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \tilde{R}(E_q, \overline{E_q}, e_j, \overline{e_j}) + \sum_p |x_p|^2 \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_p}) \\ &\quad + 2 \sum_p x_p \cdot \operatorname{Re}(e^{-i\theta_p} \tilde{R}(e_i, \overline{E_q}, e_j, \overline{E_p}) + e^{i\theta_p} \tilde{R}(e_i, \overline{e_j}, E_p, \overline{E_q})). \end{aligned}$$

Following Mok [69], by setting $A_p = \tilde{R}(e_i, \overline{e_j}, E_p, \overline{E_q})$, $B_p = \tilde{R}(e_i, \overline{E_q}, e_j, \overline{E_p})$, we have:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \tilde{G}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \tilde{R}(E_q, \overline{E_q}, e_j, \overline{e_j}) + \sum_p |x_p|^2 \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_p}) \\ &\quad + \sum_p x_p (e^{-i\theta_p} B_p + e^{i\theta_p} \overline{B_p} + e^{i\theta_p} A_p + e^{-i\theta_p} \overline{A_p}) \\ &= \tilde{R}(E_q, \overline{E_q}, e_j, \overline{e_j}) + \sum_p |x_p|^2 \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_p}) \\ &\quad + \sum_p x_p \cdot (\overline{e^{i\theta_p} (A_p + \overline{B_p})} + e^{i\theta_p} (A_p + \overline{B_p})) \end{aligned}$$

By choosing θ_p such that $e^{i\theta_p} (A_p + \overline{B_p})$ is real and positive, the identity becomes:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \tilde{G}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \tilde{R}(E_q, \overline{E_q}, e_j, \overline{e_j}) + \sum_p |x_p|^2 \tilde{R}(e_i, \overline{e_i}, E_p, \overline{E_p}) \\ &\quad + 2 \sum_p x_p \cdot |A_p + \overline{B_p}|. \end{aligned}$$

If we change e_i by $e^{i\varphi}e_i$, then $A_p = \tilde{R}(e_i, \bar{e}_j, E_p, \bar{E}_q)$ is replaced by $e^{i\varphi}A_p$, and $\bar{B}_p = \tilde{R}(e_i, \bar{E}_q, e_j, \bar{E}_p)$ is replaced by $e^{-i\varphi}\bar{B}_p$. Then we have:

$$\begin{aligned} \frac{1}{2} \cdot \frac{d^2 \tilde{F}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} &= \tilde{R}(E_q, \bar{E}_q, e_j, \bar{e}_j) + \sum_p |x_p|^2 \tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p) \\ &\quad + 2 \sum_p x_p \cdot |e^{i\varphi}A_p + e^{-i\varphi}\bar{B}_p|, \end{aligned}$$

where

$$\tilde{F}_q(\varepsilon) = \tilde{R} \left(e^{i\varphi}e_i + \varepsilon E_q, \overline{e^{i\varphi}e_i + \varepsilon E_q}, e_j + \varepsilon \sum_p C_p E_p, \overline{e_j + \varepsilon \sum_p C_p E_p} \right).$$

Since the curvature operators R and R_0 have nonnegative and positive holomorphic bisectional curvature respectively, we know that the operator $\tilde{R} = R_0 + \varepsilon_0 R_0$ has positive holomorphic bisectional curvature. Now by choosing $x_p = -\frac{|e^{i\varphi}A_p + e^{-i\varphi}\bar{B}_p|}{\tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p)}$, for $p \geq 1$, it follows that

$$\frac{1}{2\pi} \int_0^{2\pi} \left(\frac{1}{2} \cdot \frac{d^2 \tilde{F}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} \right) d\varphi = \tilde{R}(E_q, \bar{E}_q, e_j, \bar{e}_j) - \sum_p \frac{|A_p|^2 + |B_p|^2}{\tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p)}$$

and then

$$\begin{aligned} &\tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q) \cdot \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{1}{2} \cdot \frac{d^2 \tilde{F}_q(\varepsilon)}{d\varepsilon^2} \Big|_{\varepsilon=0} \right) d\varphi \\ &= \tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q) \tilde{R}(E_q, \bar{E}_q, e_j, \bar{e}_j) \\ &\quad - \sum_p \frac{|A_p|^2 + |B_p|^2}{\tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p)} \tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q). \end{aligned}$$

Note that

$$\begin{aligned} \tilde{F}_q(\varepsilon) &= \tilde{R}(e^{i\varphi}e_i + \varepsilon E_q, \overline{e^{i\varphi}e_i + \varepsilon E_q}, e_j + \varepsilon \sum_p C_p E_p, \overline{e_j + \varepsilon \sum_p C_p E_p}) \\ &= \tilde{R}(e_i + \varepsilon e^{-i\varphi}E_q, \overline{e_i + \varepsilon e^{-i\varphi}E_q}, e_j + \varepsilon \sum_p C_p E_p, \overline{e_j + \varepsilon \sum_p C_p E_p}). \end{aligned}$$

Interchanging the roles of E_q and E_p , and then taking summation, we have

$$\begin{aligned} &\sum_q 2\tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q) \tilde{R}(E_q, \bar{E}_q, e_j, \bar{e}_j) \\ &\geq c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 \tilde{u}(\{e_i, e_j\}, t)(\xi, \xi)\} \\ &\quad + \sum_{p,q} (|A_p|^2 + |B_p|^2) \left(\frac{\tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q)}{\tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_i)} + \frac{\tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p)}{\tilde{R}(e_i, \bar{e}_i, E_q, \bar{E}_q)} \right) \\ &\geq c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 \tilde{u}(\{e_i, e_j\}, t)(\xi, \xi)\} + 2 \sum_{p,q} |\tilde{R}(e_i, \bar{E}_q, e_j, \bar{E}_p)|^2, \end{aligned}$$

where $\tilde{u}(\{X, Y\}, t) = \tilde{R}(X, \bar{X}, Y, \bar{Y}) = R(X, \bar{X}, Y, \bar{Y}) + \varepsilon_0 R_0(X, \bar{X}, Y, \bar{Y})$ and c_1 is a positive constant which depends on the bound of the curvature R , but does not depend on ε_0 .

Hence

$$\begin{aligned} & \sum_p \tilde{R}(e_i, \bar{e}_i, E_p, \bar{E}_p) \tilde{R}(E_p, \bar{E}_p, e_j, \bar{e}_j) - \sum_{p,q} |\tilde{R}(e_i, \bar{E}_p, e_j, \bar{E}_q)|^2 \\ & \geq c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 \tilde{u}(\{e_i, e_j\}, t)(\xi, \xi)\}. \end{aligned}$$

Since $\varepsilon_0 > 0$ is arbitrary, we can let $\varepsilon_0 \rightarrow 0$ and it follows that:

$$\sum_{p,q} R_{i\bar{i}p\bar{q}} R_{q\bar{p}j\bar{j}} - \sum_{p,q} |R_{i\bar{i}p\bar{q}}|^2 \geq c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 u(\{e_i, e_j\}, t)(\xi, \xi)\},$$

for some constant $c_1 > 0$. Therefore we proved our claim.

By the definition of u and the evolution equation of the holomorphic bisectional curvature, we know that

$$\begin{aligned} \frac{\partial}{\partial t} u(\{X, Y\}, t) &= \Delta u(\{X, Y\}, t) + \sum_{p,q} R(X, \bar{X}, e_p, \bar{e}_q) R(e_q, \bar{e}_p, Y, \bar{Y}) \\ &\quad - \sum_{p,q} |R(X, \bar{e}_p, Y, \bar{e}_q)|^2 + \sum_{p,q} |R(X, \bar{Y}, e_p, \bar{e}_q)|^2. \end{aligned}$$

Therefore, from the above inequality, we obtain that:

$$\frac{\partial u}{\partial t} \geq Lu + c_1 \cdot \min\{0, \inf_{|\xi|=1, \xi \in V} D^2 u(\xi, \xi)\},$$

where L is the horizontal Laplacian on P , V denotes the vertical subspaces. By Proposition 2 in [7] and note that the curvature is nonnegative and bounded, we know that the set

$$N = \{(\{X, Y\}, t) | u(\{X, Y\}, t) = 0, X \neq 0, Y \neq 0\} \subset P \times (0, \delta')$$

is invariant under parallel transport.

Next, we claim that $R_{i\bar{i}j\bar{j}} > 0$ for all $t \in (0, \delta')$. Indeed, suppose not. Then $R_{i\bar{i}j\bar{j}} = 0$ for some $t \in (0, \delta')$. Therefore

$$(\{e_i, e_j\}, t) \in N.$$

Combining $R_{i\bar{i}j\bar{j}} = 0$ with the evolution equation of the curvature operator and the first variation, we can obtain that

$$\begin{cases} \sum_{p,q} (R_{i\bar{i}p\bar{q}} R_{q\bar{p}j\bar{j}} - |R_{i\bar{i}p\bar{q}}|^2) = 0, \\ R_{i\bar{i}j\bar{p}q} = 0, \quad \forall p, q, \\ R_{i\bar{i}p\bar{j}} = R_{j\bar{j}p\bar{i}} = 0, \quad \forall p. \end{cases}$$

We define an orthonormal 2-frames $\{\tilde{e}_i, \tilde{e}_j\} \subset T_x^{1,0}(\hat{M}^\alpha)$ by

$$\tilde{e}_i = \sin \theta \cdot e_i - \cos \theta \cdot e_j,$$

$$\tilde{e}_j = \cos \theta \cdot e_i + \sin \theta \cdot e_j.$$

Then

$$\overline{\tilde{e}_i} = \sin \theta \cdot \overline{e_i} - \cos \theta \cdot \overline{e_j},$$

$$\overline{\tilde{e}_j} = \cos \theta \cdot \overline{e_i} + \sin \theta \cdot \overline{e_j}.$$

Since N is invariant under parallel transport and $(\hat{M}^\alpha, \hat{g}_{ij}^\alpha(t))$ has holonomy group $U(n_\alpha)$, we obtain that

$$(\{\tilde{e}_i, \tilde{e}_j\}, t) \in N,$$

that is,

$$R(\tilde{e}_i, \overline{\tilde{e}_i}, \tilde{e}_j, \overline{\tilde{e}_j}) = 0.$$

On the other hand,

$$\begin{aligned} R(\tilde{e}_i, \overline{\tilde{e}_i}, \tilde{e}_j, \overline{\tilde{e}_j}) &= \sin^2 \theta \cos^2 \theta R_{\tilde{i}\tilde{i}\tilde{i}\tilde{i}} + \sin^3 \theta \cos \theta R_{\tilde{i}\tilde{i}\tilde{j}\tilde{j}} + \sin^3 \theta \cos \theta R_{\tilde{i}\tilde{j}\tilde{i}\tilde{i}} \\ &\quad + \sin^4 \theta R_{\tilde{i}\tilde{j}\tilde{j}\tilde{j}} - \sin \theta \cos^3 \theta R_{\tilde{i}\tilde{j}\tilde{i}\tilde{i}} - \sin^2 \theta \cos^2 \theta R_{\tilde{i}\tilde{j}\tilde{j}\tilde{i}} \\ &\quad - \sin^2 \theta \cos^2 \theta R_{\tilde{i}\tilde{j}\tilde{j}\tilde{i}} - \sin^3 \theta \cos \theta R_{\tilde{i}\tilde{j}\tilde{j}\tilde{j}} - \cos^3 \theta \sin \theta R_{\tilde{j}\tilde{i}\tilde{i}\tilde{i}} \\ &\quad - \sin^2 \theta \cos^2 \theta R_{\tilde{j}\tilde{i}\tilde{i}\tilde{j}} - \sin^2 \theta \cos^2 \theta R_{\tilde{j}\tilde{i}\tilde{j}\tilde{i}} - \cos \theta \sin^3 \theta R_{\tilde{j}\tilde{i}\tilde{j}\tilde{j}} \\ &\quad + \cos^4 \theta R_{\tilde{j}\tilde{j}\tilde{i}\tilde{i}} + \cos^3 \theta \sin \theta R_{\tilde{j}\tilde{j}\tilde{i}\tilde{j}} + \cos^3 \theta \sin \theta R_{\tilde{j}\tilde{j}\tilde{j}\tilde{i}} \\ &\quad + \cos^2 \theta \sin^2 \theta R_{\tilde{j}\tilde{j}\tilde{j}\tilde{j}} \\ &= \cos^2 \theta \sin^2 \theta (R_{\tilde{i}\tilde{i}\tilde{i}\tilde{i}} + R_{\tilde{j}\tilde{j}\tilde{j}\tilde{j}}). \end{aligned}$$

So we have $R_{\tilde{j}\tilde{j}\tilde{j}\tilde{j}} + R_{\tilde{i}\tilde{i}\tilde{i}\tilde{i}} = 0$, if we choose θ such that $\cos^2 \theta \sin^2 \theta \neq 0$. But this contradicts with the fact that $(\hat{M}^\alpha, \hat{g}_{ij}^\alpha(t))$ has positive holomorphic sectional curvature. Hence we proved that $R_{\tilde{i}\tilde{i}\tilde{j}\tilde{j}} > 0$, for all $t \in (0, \delta')$.

This completes the proof of Theorem 2.8. \square

We remark that a (rough) factorization theorem, according to whether the manifold supports a strictly plurisubharmonic function, was obtained earlier by Ni and Tam [76] without assuming the curvature to be bounded.

Finally, by combining with the resolution of the Frankel conjecture, our (more precise) factorization Theorem 2.8 can reduce the classification of complete noncompact Kähler manifolds with bounded and nonnegative bisectional curvature to the case of strictly positive bisectional curvature. In the latter case there is a long standing conjecture due to Yau (Problem 34 in [101]):

Yau’s Conjecture (Yau [101]) *A complete noncompact Kähler manifold of positive holomorphic bisectional curvature is biholomorphic to a complex Euclidean space.*

In recent years, there have been many research activities in studying this conjecture of Yau. The Ricci flow has been found to be a useful tool to approach it. The following partial affirmative answer, due to Chen-Tang-Zhu [20] in complex dimension $n = 2$ and Chau-Tam [16] for all dimensions, was obtained via the Ricci flow.

THEOREM 2.9. *Let M be a complete noncompact n -dimensional Kähler manifold of positive and bounded holomorphic bisectional curvature. Suppose there exists a positive constant C_1 such that for a fixed base point x_0 , we have*

$$\text{Vol}(B(x_0, r)) \geq C_1 r^{2n} \quad 0 \leq r < +\infty,$$

then M is biholomorphic to \mathbb{C}^n .

We refer the readers to the survey article of A. Chau and L. F. Tam [17] in this volume for more information on works related to the Kähler-Ricci flow and Yau’s uniformization conjecture.

3. Perelman’s Noncollapsing Result

In the celebrated work [80], Perelman proved a remarkable (local) noncollapsing result for the Ricci flow on compact manifolds in all dimensions. This (local) noncollapsing result had been conjectured by Hamilton in his survey paper [47] and is crucial in applying Hamilton’s compactness theorem to understand the structure of singularities of the Ricci flow. Below, we follow Perelman [80] to give two approaches for deriving his noncollapsing result.

3.1. Perelman’s Conjugate Heat Equation Approach. For the Ricci flow on a compact manifold, Perelman [80] introduced a new functional

$$(3.1) \quad \mathcal{W}(g_{ij}, f, \tau) = \int_M [\tau(R + |\nabla f|^2) + f - n](4\pi\tau)^{-\frac{n}{2}} e^{-f} dV.$$

This functional has played a very important role in the Ricci flow; see also the more recent works by Feldman-Ilmanen-Ni [34], Cao-Hamilton-Ilmanen [13], Ma [64], Li [60], Zhang [107], Ye [102]–[105], X. Cao [15], Ling [62], etc.

Perelman proved that the \mathcal{W} -functional is monotone in time when the metric g evolves under the Ricci flow, the function f evolves under the backward heat equation

$$\frac{\partial f}{\partial \tau} = \Delta f - |\nabla f|^2 + R - \frac{n}{2\tau},$$

and $\frac{d\tau}{dt} = -1$. This entropy monotonicity can be interpreted as a Li-Yau type estimate for the conjugate heat equation

$$(3.2) \quad \square u := \frac{\partial u}{\partial \tau} - \Delta u + Ru = 0,$$

where $\tau = T - t$, and $g_{ij}(x, t)$, $0 \leq t < T$, is a solution to the Ricci flow. Note that $u = (4\pi\tau)^{-\frac{n}{2}} e^{-f}$ satisfies the conjugate heat equation if and only if f satisfies the above backward heat equation.

By considering the shrinking Ricci solitons, one can find the analogous Li-Yau expression for the conjugate heat equation to be

$$H = 2\Delta f - |\nabla f|^2 + R + \frac{f - n}{\tau}.$$

(We learned this argument from Hamilton. The details can be found in [14].) By direct computations, one has

$$\frac{\partial H}{\partial \tau} = \Delta H - 2\nabla f \cdot \nabla H - \frac{1}{\tau}H - 2 \left| R_{ij} + \nabla_i \nabla_j f - \frac{1}{2\tau}g_{ij} \right|^2.$$

Set

$$(3.3) \quad v = \tau H u = (\tau(R + 2\Delta f - |\nabla f|^2) + f - n)u,$$

then

$$(3.4) \quad \frac{\partial v}{\partial \tau} = \Delta v - Rv - 2\tau u \left| R_{ij} + \nabla_i \nabla_j f - \frac{1}{2\tau}g_{ij} \right|^2.$$

If u is a fundamental solution to (3.2), one can show $\lim_{\tau \rightarrow 0^+} \tau H \leq 0$ (see [75]). Then the maximum principle implies Perelman's Li-Yau type estimate for the conjugate heat equation:

$$H \leq 0$$

for all $\tau \in (0, T]$. Along any space-time path $(\gamma(\tau), \tau)$, $\tau \in [0, \bar{\tau}]$ with $\gamma(0) = p, \gamma(\bar{\tau}) = q$, there holds

$$\frac{d}{d\tau} (2\sqrt{\tau}f(\gamma(\tau), \tau)) \leq \sqrt{\tau}(R + |\dot{\gamma}(\tau)|_{g_{ij}(\tau)}^2).$$

If one defines

$$(3.5) \quad \mathcal{L}(\gamma) \triangleq \int_0^{\bar{\tau}} \sqrt{\tau}(R + |\dot{\gamma}(\tau)|_{g_{ij}(\tau)}^2) d\tau,$$

and

$$(3.6) \quad l(q, \bar{\tau}) \triangleq \inf_{\gamma} \frac{1}{2\sqrt{\bar{\tau}}} \mathcal{L}(\gamma),$$

where the inf is taken over all space curves $\gamma(\tau), 0 \leq \tau \leq \bar{\tau}$, joining p and q , then

$$f(q, \bar{\tau}) \leq l(q, \bar{\tau}).$$

This leads to a lower estimate for the fundamental solution u of the conjugate heat equation,

$$(3.7) \quad u(q, \bar{\tau}) \geq (4\pi\bar{\tau})^{-\frac{n}{2}} e^{-l(q, \bar{\tau})}.$$

Now since v happens to be the integrand of the \mathcal{W} -functional, by integrating (3.4), one obtains

$$(3.8) \quad \frac{d}{dt} \mathcal{W}(g_{ij}(t), f(t), \tau(t)) = \int_M 2\tau \left| R_{ij} + \nabla_i \nabla_j f - \frac{1}{2\tau} g_{ij} \right|^2 (4\pi\tau)^{-\frac{n}{2}} e^{-f} dV \geq 0.$$

Let

$$(3.9) \quad \mu(M, g, \tau) = \inf \left\{ W(g, f, \tau) \mid \int (4\pi\tau)^{-\frac{n}{2}} e^{-f} dv = 1 \right\},$$

then we have the monotonicity of Perelman’s entropy:

LEMMA 3.1 (Perelman [80]). $\mu(M, g(t), T - t)$ is nondecreasing along compact Ricci flow; moreover, the monotonicity is strict unless we are on a shrinking gradient soliton.

A direct consequence is the following important noncollapsing theorem of Perelman.

THEOREM 3.2 (Perelman [80]). Let $g_{ij}(x, t), 0 \leq t \leq T$, be a smooth solution to the Ricci flow on an n -dimensional compact manifold M . Then there exists a constant $\kappa > 0$ depending only on T and the initial metric such that the following holds: if $r_0 \leq \sqrt{T}$ and $|Rm|(x, t_0) \leq r_0^{-2}$ on $B_{t_0}(x_0, r_0)$, then

$$\text{vol}_{t_0}(B_{t_0}(x_0, r_0)) \geq \kappa r_0^n.$$

Indeed, let ξ be a smooth nonnegative non-increasing function, which is 1 on $(-\infty, \frac{1}{2}]$ and 0 on $[\frac{3}{4}, \infty)$. Substituting

$$u = (4\pi r_0^2)^{-\frac{n}{2}} e^{-f} = \frac{\xi\left(\frac{d_{t_0}(x_0, x)}{r_0}\right)}{\int_M \xi\left(\frac{d_{t_0}(x_0, x)}{r_0}\right) dv_{t_0}}$$

into (3.1), we have

$$W(g_{ij}(t_0), f, r_0^2) \leq C(n) + \log \frac{\text{vol}_{t_0}(B_{t_0}(x_0, r_0))}{r_0^n}.$$

By Lemma 3.1, we have

$$W(g_{ij}(t_0), f, r_0^2) \geq \mu(M, g_{ij}(0), r_0^2 + t_0).$$

Note that the right hand side is controlled by the log Sobolev constant of the initial metric (on the scales $\leq \sqrt{2T}$). This proves the noncollapsing theorem.

3.2. Perelman's Reduced Volume Approach. There is another way to get the noncollapsing result by establishing the comparison geometry to the \mathcal{L} -length introduced in (3.5). Moreover, this comparison geometric approach could be adapted to get the noncollapsing for *surgical* solutions. We now discuss this approach.

In [80], Perelman introduced the \mathcal{L} -length defined in (3.5) as a suitable renormalized distance function on potentially infinite dimensional space-time manifold, where the Ricci flow was embedded there. More explicitly, let $\frac{\partial g_{ij}}{\partial \tau} = 2Ric$ be a solution to the Ricci flow on M with $\tau = T - t$. Consider

$$\tilde{M} = M \times \mathbb{S}^N \times \mathbb{R}^+$$

with the following metric

$$\begin{aligned} \tilde{g}_{ij} &= g_{ij}, \\ \tilde{g}_{\alpha\beta} &= \tau g_{\alpha\beta}, \\ \tilde{g}_{oo} &= \frac{N}{2\tau} + R, \\ \tilde{g}_{i\alpha} &= \tilde{g}_{io} = \tilde{g}_{\alpha o} = 0, \end{aligned}$$

where i, j are coordinate indices on M , α, β are coordinate indices on \mathbb{S}^N and the coordinate τ on \mathbb{R}^+ has index o . The metric $g_{\alpha\beta}$ on \mathbb{S}^N has constant sectional curvature $\frac{1}{2N}$. This construction may be viewed as a “regularization” of what Chow-Chu did in [27]. Perelman twisted the sign of the time and coupled the space-time with a solution to the Ricci flow with positive curvatures on manifolds of very big dimensions. As we mentioned before, Chu and Chow [27] found a geometric interpretation of Li-Yau-Hamilton inequality. So the following proposition of Perelman is not surprising.

PROPOSITION 3.3. *The components of the curvature tensor of the metric \tilde{g} coincide (module N^{-1}) with the components of the Li-Yau-Hamilton quadratic.*

The key observation due to Perelman is the following

PROPOSITION 3.4 (Perelman [80]). *The Ricci curvature of the metric \tilde{g} is flat (module N^{-1}), i.e. $|\tilde{Ric}|_{\tilde{g}} = O\left(\frac{1}{N}\right)$.*

Recall that we have Bishop-Gromov volume comparison theorem on manifolds with Ricci curvature bounded from below. The above proposition 3.4 motivated an important monotonicity formula: the reduced volume.

Actually, by looking at the \tilde{g} length of a space time curve $\gamma(\tau)$, $0 \leq \tau \leq \bar{\tau}$,

$$\begin{aligned} & \int_0^{\bar{\tau}} \sqrt{\left(\frac{N}{2\tau} + R\right) + |\dot{\gamma}(\tau)|_{g_{ij}(\tau)}^2} d\tau \\ &= \sqrt{2N\bar{\tau}} + \frac{1}{\sqrt{2N}} \int_0^{\bar{\tau}} \sqrt{\tau}(R + |\dot{\gamma}(\tau)|_{g_{ij}}^2) d\tau + O(N^{-\frac{3}{2}}), \end{aligned}$$

we find the expression of \mathcal{L} distance $\int_0^{\bar{\tau}} \sqrt{\tau}(R + |\dot{\gamma}(\tau)|_{g_{ij}}^2) d\tau$. By computing the volumes of geodesic spheres of radii $\sqrt{2N\bar{\tau}}$ on \tilde{M} and \mathbb{R}^{n+N+1} , we have

$$\frac{Vol(S_{\tilde{M}}(\sqrt{2N\bar{\tau}}))}{Vol(S_{\mathbb{R}^{n+N+1}}(\sqrt{2N\bar{\tau}}))} \approx const \cdot N^{-\frac{n}{2}} \cdot \int_M (\bar{\tau})^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sqrt{\bar{\tau}}}L(x, \bar{\tau})\right\} dV_M.$$

Proposition 3.4 indicates that the quantity

$$\int_M (\bar{\tau})^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sqrt{\bar{\tau}}}L(x, \bar{\tau})\right\} dV_M$$

should be non-increasing in $\bar{\tau}$. This quantity is called *Perelman's reduced volume* and we denote it by $V(\bar{\tau})$.

The rigorous proof of this monotonicity can be obtained in the following way. One computes the first and second variation for the \mathcal{L} -length (3.5) to get

LEMMA 3.5 (Perelman [80]). *For the reduced distance $l(q, \bar{\tau})$ defined in (3.6), there hold*

$$(3.10) \quad \frac{\partial l}{\partial \bar{\tau}} = -\frac{l}{\bar{\tau}} + R + \frac{1}{2\bar{\tau}^{3/2}}K$$

$$(3.11) \quad |\nabla l|^2 = -R + \frac{l}{\bar{\tau}} - \frac{1}{\bar{\tau}^{3/2}}K$$

$$(3.12) \quad \Delta l \leq -R + \frac{n}{2\bar{\tau}} - \frac{1}{2\bar{\tau}^{3/2}}K.$$

where

$$K = \int_0^{\bar{\tau}} \tau^{\frac{3}{2}} Q(X) d\tau,$$

and

$$Q(X) = -R_\tau - \frac{R}{\tau} - 2\langle \nabla R, X \rangle + 2Ric(X, X)$$

is the trace Li-Yau-Hamilton quadratic. Moreover, the equality in (3.12) holds if and only if the solution along the \mathcal{L} minimal geodesic γ satisfies the gradient soliton equation

$$R_{ij} + \frac{1}{2\sqrt{\bar{\tau}}} \nabla_i \nabla_j L = \frac{1}{2\bar{\tau}} g_{ij}.$$

The combination of (3.10)-(3.12) gives

$$(3.13) \quad \left(\frac{\partial}{\partial \tau} - \Delta + R \right) ((4\pi\bar{\tau})^{-\frac{n}{2}} e^{-l}) \leq 0.$$

If the manifold is compact, then by integrating, the reduced volume satisfies

$$(3.14) \quad \frac{d}{d\bar{\tau}} V(\bar{\tau}) = \frac{d}{d\bar{\tau}} \int_M (4\pi\bar{\tau})^{-\frac{n}{2}} e^{-l} dV_{\bar{\tau}}(q) \leq 0,$$

and equality holds if and only if we are on a shrinking gradient soliton.

To carry the monotonicity to noncompact manifolds, Perelman [80] established a Jacobian comparison for the exponential map associated to the \mathcal{L} -length. From the \mathcal{L} -length, one defines an \mathcal{L} -exponential map (with parameter $\bar{\tau}$) $\mathcal{L}exp(\bar{\tau}) : T_p M \rightarrow M$ as follows: for any $X \in T_p M$, set $\mathcal{L}exp_X(\bar{\tau}) = \gamma(\bar{\tau})$, where γ is an \mathcal{L} -geodesic satisfying $\gamma(0) = p$ and $\lim_{\tau \rightarrow 0} \sqrt{\tau} \dot{\gamma}(\tau) = X$. Let $\mathcal{J}(\tau)$ be the Jacobian of the \mathcal{L} -exponential map along $\gamma(\tau)$, $0 \leq \tau \leq \bar{\tau}$. Then by the standard computation of Jacobi fields, we obtain

$$\frac{d}{d\tau} \log \mathcal{J}(\tau) = \Delta l + R$$

along any minimal \mathcal{L} -geodesic γ . Combining with equations (3.10)-(3.12) in the Lemma (3.5), this gives

THEOREM 3.6 (Perelman's Jacobian comparison [80]). *Along any minimal \mathcal{L} -geodesic γ , we have*

$$(3.15) \quad \frac{d}{d\tau} \{ (4\pi\tau)^{-\frac{n}{2}} \exp(-l(\tau)) \mathcal{J}(\tau) \} \leq 0.$$

Consequently, we obtain

THEOREM 3.7 (Monotonicity of the Perelman's reduced volume). *Let g_{ij} be a family of complete metrics evolving by the Ricci flow $\frac{\partial}{\partial \tau} g_{ij} = 2R_{ij}$ on a manifold M with bounded curvature. Fix a point p in M and let $l(q, \tau)$ be the reduced distance from $(p, 0)$. Then*

(i) *Perelman's reduced volume*

$$\tilde{V}(\tau) = \int_M (4\pi\tau)^{-\frac{n}{2}} \exp(-l(q, \tau)) dV_{\tau}(q)$$

is finite and nonincreasing in τ ;

- (ii) *the monotonicity is strict unless we are on a gradient shrinking soliton.*

Now we are going to use the reduced volume to derive a slightly weaker version of Theorem 3.2. The advantage of this new method is that it allows to be adapted to the case that the solutions are only locally defined. This will be extremely important in the analysis of surgical solutions.

DEFINITION 3.8. *We say a solution to the Ricci flow is κ -noncollapsed at (x_0, t_0) on the scale r for positive constants κ and r if it satisfies the following property: if $|Rm|(x, t) \leq r^{-2}$ for all $x \in B_{t_0}(x_0, r)$ and $t \in [t_0 - r^2, t_0]$, then we have*

$$\text{vol}_{t_0}(B_{t_0}(x_0, r)) \geq \kappa r^n.$$

THEOREM 3.9 (Perelman [80]). *Let (M^n, g_{ij}) be a complete Riemannian manifold with bounded curvature $|Rm| \leq k_0$ and with injectivity radius bounded from below by $\text{inj}(M, g_{ij}) \geq i_0$. Let $g_{ij}(x, t)$, $t \in [0, T)$ be a smooth solution to the Ricci flow with bounded curvature for each $t \in [0, T)$ and $g_{ij}(x, 0) = g_{ij}(x)$. Then there is a $\kappa > 0$ depending only on k_0, i_0 and T such that the solution is κ -noncollapsed on scales $\leq \sqrt{T}$.*

A sketch of the proof is given as follows. Argue by contradiction. Suppose $|Rm|(x, t) \leq r^{-2}$ for all $x \in B_{t_0}(x_0, r)$ and $t \in [t_0 - r^2, t_0]$, but

$$\frac{\text{vol}_{t_0}(B_{t_0}(x_0, r))}{r^n} = \varepsilon^n$$

is very small. Write

(3.16)

$$\begin{aligned} \tilde{V}(\varepsilon r^2) &= \int_M (4\pi\varepsilon r^2)^{-\frac{n}{2}} \exp(-l(q, \varepsilon r^2)) dV_{t_0 - \varepsilon r^2}(q) \\ &\leq \int_{\mathcal{L}exp_{\{|v| \leq \frac{1}{4}\varepsilon^{-1/2}\}}(\varepsilon r^2)} (4\pi\varepsilon r^2)^{-\frac{n}{2}} \exp(-l(q, \varepsilon r^2)) dV_{t_0 - \varepsilon r^2}(q) \\ &\quad + \int_{\mathcal{L}exp_{\{|v| > \frac{1}{4}\varepsilon^{-1/2}\}}(\varepsilon r^2)} (4\pi\varepsilon r^2)^{-\frac{n}{2}} \exp(-l(q, \varepsilon r^2)) dV_{t_0 - \varepsilon r^2}(q) \\ &\leq I + II. \end{aligned}$$

First of all, it can be shown $\mathcal{L}exp_{\{|X| \leq \frac{1}{4}\varepsilon^{-\frac{1}{2}}\}}(\varepsilon r^2) \subset B_{t_0}(x_0, r)$, and $l(q, \varepsilon r^2) \geq -C(n)\varepsilon$ on $B_{t_0}(x_0, r)$. This implies $I \leq C(n)\varepsilon^{-\frac{n}{2}}\varepsilon^n = C(n)\varepsilon^{\frac{n}{2}}$. By the monotonicity (3.15) of \mathcal{L} -Jacobian, one has

$$II \leq \int_{\{|X| \geq \frac{1}{4}\varepsilon^{-\frac{1}{2}}\}} (4\pi)^{-\frac{n}{2}} \exp(-|X|^2) dX \leq \varepsilon^{\frac{n}{2}},$$

hence

$$(3.17) \quad \tilde{V}(\varepsilon r^2) \leq C(n)\varepsilon^{\frac{n}{2}}.$$

On the other hand, by Lemma (3.5),

$$(3.18) \quad \frac{\partial \bar{L}}{\partial \tau} + \Delta \bar{L} \leq 2n$$

where $\bar{L}(q, \tau) = 4\tau l(q, \tau)$. It follows from maximum principle that there is a point $q_0 \in M$ such that $l(q_0, t_0 - \frac{1}{C(n)k_0}) \leq \frac{n}{2}$. Since the geometry is controlled on $B_0(q_0, 1) \times [0, \frac{1}{C(n)k_0}]$, one then has $l(q, t_0) \leq Const.$ on $B_0(q_0, 1)$, which implies

$$\tilde{V}(t_0) \geq \int_{B_0(P_0, 1)} (4\pi t_0)^{-\frac{n}{2}} e^{-l} dv \geq Const. > 0.$$

This contradicts with the monotonicity of the reduced volume when ε is small enough.

4. The Formation of Singularities

Given a compact Riemannian manifold (M, g) , we evolve the metric by the Ricci flow

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij}.$$

We say a solution $g_{ij}(x, t)$, $t \in [0, T)$ is a *maximal solution* to the Ricci flow with $g_{ij}(x, 0) = g_{ij}(x)$, if either $T = \infty$, or $T < \infty$ and the curvature becomes unbounded as $t \rightarrow T$. If T is finite, we say the solution develops singularities at the time T . In the early 90's, Hamilton initiated the program to investigate the formation of singularities.

4.1. Hamilton's Compactness Theorem. To understand the structure of a singularity, similar to the study of minimal surface theory and harmonic map theory, one tries to dilate the solution around the singularity and then take a limit of the rescaled sequence of solutions. In order to do this, a compactness theorem for solutions to the Ricci flow is needed.

The standard compactness theorems for Riemannian manifolds in $C^{1, \alpha}$ norm are available by the works of Gromov [38], Peters [83], Greene-Wu [35], etc. Thanks to Shi's derivative estimates (Theorem 1.4), we know that all the derivatives of curvature are guaranteed to be bounded once curvature is bounded for any solution to the Ricci flow. Based on this fact, Hamilton [46] established a C^∞ compactness theorem for solutions of the Ricci flow. A slight generalization of Hamilton's compactness theorem is given by below.

THEOREM 4.1 (Hamilton [46, 47]; see also [14]). *Let $(M_k, g_k(t), p_k)$, $t \in (A, \Omega]$ with $A < 0 \leq \Omega$, be a sequence of evolving marked complete Riemannian manifolds. Consider a sequence of geodesic balls $B_0(p_k, s_k) \subset M_k$ of*

radii $s_k (0 < s_k \leq +\infty)$, with $s_k \rightarrow s_\infty (\leq +\infty)$, around the base points p_k in the metrics $g_k(0)$. Suppose each $g_k(t)$ is a solution to the Ricci flow on $B_0(p_k, s_k) \times (A, \Omega]$. Suppose also

- (i) for every radius $r < s_\infty$ there exist positive constants $C(r)$ and $k(r)$ independent of k such that the curvature tensors $Rm(g_k)$ of the evolving metrics $g_k(t)$ satisfy the bound

$$|Rm(g_k)| \leq C(r)$$

on $B_0(p_k, r) \times (A, \Omega]$ for all $k \geq k(r)$, and

- (ii) there exists a constant $\delta > 0$ such that the injectivity radii of M_k at p_k in the metric $g_k(0)$ satisfy the bound

$$\text{inj}(M_k, p_k, g_k(0)) \geq \delta > 0$$

for all $k = 1, 2, \dots$

Then there exists a subsequence of $(B_0(p_k, s_k), g_k(t), p_k)$ over $t \in (A, \Omega]$ which converges in C_{loc}^∞ topology to a solution $(B_\infty, g_\infty(t), p_\infty)$ over $t \in (A, \Omega]$ to the Ricci flow, where, at the time $t=0$, B_∞ is a geodesic open ball centered at $p_\infty \in B_\infty$ with the radius s_∞ . Moreover the limiting solution is complete if $s_\infty = +\infty$.

4.2. Hamilton's Classification of Singularities. In [47], Hamilton divided all maximal solutions, according to the blow-up rate of maximal curvatures $K_{\max}(t) := \sup_{x \in M} |Rm|(x, t)$, into three types:

Type I: $T < +\infty$ and $\sup_{t \in [0, T)} (T - t)K_{\max}(t) < +\infty$;

Type II: (a) $T < +\infty$ but $\sup_{t \in [0, T)} (T - t)K_{\max}(t) = +\infty$;

(b) $T = +\infty$ but $\sup_{t \in [0, T)} tK_{\max}(t) = +\infty$;

Type III: (a) $T = +\infty$, $\sup_{t \in [0, T)} tK_{\max}(t) < +\infty$, and

$$\limsup_{t \rightarrow +\infty} tK_{\max}(t) > 0;$$

(b) $T = +\infty$, $\sup_{t \in [0, T)} tK_{\max}(t) < +\infty$, and

$$\limsup_{t \rightarrow +\infty} tK_{\max}(t) = 0.$$

To understand the structure of a singularity, one can follow Hamilton in [47] by first picking a sequence of space-time points (x_k, t_k) which approach the singularity, then rescaling the solution around these points so that the norm of the curvature of each rescaled solution in the sequence is bounded by 2 everywhere and equal to 1 at the chosen points. (Such space-time points (x_k, t_k) are called *almost maximal points* to the maximal solution). The noncollapsing theorem of Perelman in the previous section gives the desired injectivity radius estimate (ii) for the rescaled sequence of solutions.

Thus one can apply Hamilton's compactness theorem to take a limit and conclude that any rescaling limit must be one of the singularity models in the following sense.

DEFINITION 4.2 (Hamilton [47]). *A solution $g_{ij}(x, t)$ to the Ricci flow on the manifold M , where either M is compact or at each time t the metric $g_{ij}(\cdot, t)$ is complete and has bounded curvature, is called a **singularity model** if it is not flat and of one of the following three types:*

Type I: *The solution exists for $t \in (-\infty, \Omega)$ for some constant Ω with $0 < \Omega < +\infty$ and*

$$|Rm| \leq \Omega / (\Omega - t)$$

everywhere with equality somewhere at $t = 0$;

Type II: *The solution exists for $t \in (-\infty, +\infty)$ and*

$$|Rm| \leq 1$$

everywhere with equality somewhere at $t = 0$;

Type III: *The solution exists for $t \in (-A, +\infty)$ for some constant A with $0 < A < +\infty$ and*

$$|Rm| \leq A / (A + t)$$

everywhere with equality somewhere at $t = 0$.

In the special case of nonnegative curvatures, the singularity models of Type II and III can be further characterized as Ricci solitons.

THEOREM 4.3.

- (i) (Hamilton [45]) *Any Type II singularity model with nonnegative curvature operator and positive Ricci curvature must be a (steady) Ricci soliton.*
- (ii) (Chen-Zhu [21]) *Any Type III singularity model with nonnegative curvature operator and positive Ricci curvature must be a homothetically expanding Ricci soliton.*
- (iii) (Cao [10]) *Any Type II or III singularity model on a Kähler manifold with nonnegative holomorphic bisectional curvature and positive Ricci curvature must be a steady Kähler-Ricci soliton or an expanding Kähler-Ricci soliton.*

For Type I singularity models, N. Sesum [90] obtained the following characterization in the compact case.

THEOREM 4.4. *Let $(M, g_{ij}(x, t))$ be a compact Type I singularity model obtained as a rescaling limit of Type I maximal solution. Then $(M, g_{ij}(x, t))$ must be a (non-flat) gradient shrinking Ricci soliton.*

Very recently Naber [72] showed that a suitable rescaling limit of any Type I maximal solution is a gradient shrinking soliton. However, it is still an interesting question when the rescaling limit must be non-flat.

In recent years, there have been some research activities on the question what kind of singularity models can be realized by the Ricci flow. We have seen from the Differential Sphere Theorems obtained in [41, 42, 8], and [6] that manifolds with positive curvatures (positive Ricci curvature in dimension 3, positive or two-positive curvature operator and 1/4-pinch in dimensions greater than 3) always develop *spherical Type I* singularities in the sense the singularity model is the round sphere. Apart from the spherical Type I singularities, there should exist a *necklike Type I* singularity in the sense the singularity models are the round cylinders. The existence of necklike Type I singularities was first demonstrated by M. Simon [93] on noncompact warped product $\mathbb{R} \times_f \mathbb{S}^n$. Later, Feldman-Ilmanen-Knopf [33] also found such necklike Type I singularities on some noncompact Kähler manifold, the total space of certain holomorphic line bundle L^{-k} over the complex projective space $\mathbb{C}\mathbb{P}^n$. The existence of neckpinch Type I singularities on compact manifolds was recently proved by S. Angenent and D. Knopf [2] on \mathbb{S}^{n+1} with suitable rotationally symmetric metrics. It is also interesting to see if a Type II singularity could be really formed in the Ricci flow. In [31], Daskalopoulos and Hamilton showed that a Type II singularity can be developed by the Ricci flow on the noncompact \mathbb{R}^2 . The intuition of forming a Type II singularity on compact manifolds was described by Hamilton [47] (see also [28] and [99]) and the existence of a Type II singularity on compact manifolds was also proposed as an open question in the introduction of the book of Chow-Lu-Ni [29]. Most recently, Hui-Ling Gu and the last author [40] extended some arguments of Perelman to higher dimensions so as to show that a Type II singularity can be formed by the Ricci flow on \mathbb{S}^n with suitable rotationally symmetric metric for all $n \geq 3$.

4.3. Ancient κ -solutions. Once we have a basic understanding for those singularities developed by almost maximum points, we now want to consider those singularities which might not come from almost maximum points. If we are considering a general singularity developed by the Ricci flow in a finite time on a compact manifold, then any rescaling limit around the singularity will define at least on $(-\infty, 0)$, called an *ancient solution*. Moreover, by Perelman’s noncollapsing result, there is some positive constant κ so that the rescaling limit is κ -noncollapsing for all scales. So any rescaling limit for singularities developed by the Ricci flow on compact manifolds is κ -noncollapsing and defined at least on the time interval $(-\infty, 0)$. Up to now, all understandings to these rescaling limits are restricted on the class that have nonnegative curvature operators. That is, according to Perelman [80], we only consider *ancient κ -solutions*, i.e., each of them is defined on $(-\infty, 0)$, has bounded and nonnegative curvature operators and is κ -noncollapsing for all scales for some $\kappa > 0$. Notice, by Hamilton-Ivey

pinching estimate and Perelman’s noncollapsing result, that any rescaling limit of the Ricci flow on compact three-manifolds is an ancient κ -solution for some $\kappa > 0$.

The main purpose of this section is to review the properties of ancient κ -solutions and eventually to get a rather complete understanding to the ancient κ -solutions in certain low dimension cases. Firstly, two-dimensional ancient κ -solutions have been completely classified by Hamilton [47].

THEOREM 4.5. *Any two-dimensional κ -noncollapsing non-flat ancient solution must be either the round sphere \mathbb{S}^2 or the round real projective space \mathbb{RP}^2*

In fact, Hamilton [47] proved a somewhat stronger result: any two-dimensional complete non-flat ancient solution of bounded curvature must be the round sphere \mathbb{S}^2 , the round real project space \mathbb{RP}^2 , or the cigar soliton. Note that the cigar soliton does not satisfy the κ -noncollapsing property for large scales.

Three-dimensional ancient κ -solutions have not yet been completely classified. Nevertheless, Perelman obtained a complete classification to a special class of three-dimensional ancient κ -solutions – the shrinking gradient solitons.

LEMMA 4.6 (Perelman [81]). *Let $(M, g_{ij}(t))$ be a nonflat gradient shrinking soliton to the Ricci flow on a three-manifold. Suppose $(M, g_{ij}(t))$ has bounded and nonnegative sectional curvature and is κ -noncollapsed on all scales for some $\kappa > 0$. Then $(M, g_{ij}(t))$ is one of the followings:*

- (i) *the round three-sphere \mathbb{S}^3 , or its metric quotients;*
- (ii) *the round infinite cylinder $\mathbb{S}^2 \times \mathbb{R}$, or its \mathbb{Z}_2 quotients.*

Perelman’s proof is based on the investigation of the shrinking soliton equation

$$R_{ij} + f_{ij} + \frac{g_{ij}}{2t} = 0, \quad t < 0.$$

By applying Hamilton’s strong maximum principle, one can easily characterize the shrinking solitons as either the round three-sphere \mathbb{S}^3 , or the round infinite cylinder $\mathbb{S}^2 \times \mathbb{R}$ or a metric quotient of them, except the case when the soliton is noncompact and has positive sectional curvature everywhere. We now briefly describe Perelman’s arguments in excluding the possibility of such noncompact 3-dimensional solitons with positive sectional curvature.

Consider the metric at $t = -1$. By investigating the soliton equation and the second variation formula, we find that $f(x) \approx \frac{1}{4}d^2(x, x_0)$ and $|\nabla f|^2 \approx \frac{1}{4}d^2(x, x_0)$. From $\nabla_i R = 2R_{ij}\nabla_j f$, we know R is increasing along the integral curves of the potential function f . It is not hard to see that the solution at infinity splits off a line \mathbb{R} . By comparing the existence time of the Ricci flow on standard \mathbb{S}^2 , we find $\bar{R} = \limsup_{d_{-1}(x, x_0)} R(x, -1) \leq 1$. Consider the area

of level sets of f , we have

$$\begin{aligned} \frac{d}{dt} \text{area}\{f=a\} &= \int_{\{f=a\}} \text{div} \left(\frac{\nabla f}{|\nabla f|} \right) \\ &\geq \int_{\{f=a\}} \frac{1}{|\nabla f|} (1-R) \geq \frac{1-\bar{R}}{2\sqrt{a}} \text{area}\{f=a\}. \end{aligned}$$

This forces $\bar{R}=1$ and the area of $\{f=a\}$ is increasing to the area of \mathbb{S}^2 with constant curvature $\frac{1}{2}$. On the other hand, by the Gauss equation and the soliton equation, the intrinsic curvature of $\{f=a\} (a \gg 1)$ can be computed as

$$K = R_{1212} + \frac{\det(\text{Hess}(f))}{|\nabla f|^2} < \frac{1}{2},$$

which is a contradiction with the Gauss-Bonnet formula.

Remark: The above Perelman’s result has been improved by Ni-Wallach [77] and Naber [72] in which they dropped the assumption on κ -noncollapsing condition and replaced nonnegative sectional curvature by nonnegative Ricci curvature. In addition, Ni-Wallach [77] can allow the curvature to grow as fast as $e^{ar^2(x)}$, where $r(x)$ is the distance function and a is a suitable small positive constant. In particular, Ni-Wallach’s result implies that any 3-dimensional noncompact non-flat gradient shrinking soliton with nonnegative Ricci curvature and with curvature not growing faster than $e^{ar^2(x)}$ must be a quotient of the round infinite cylinder $\mathbb{S}^2 \times \mathbb{R}$. Now using the work of the second author in [19], we can further improve this latter result of Ni-Wallach as follows.

PROPOSITION 4.7. *Let (M^3, g_{ij}) be a 3-dimensional complete noncompact non-flat shrinking gradient soliton. Then (M^3, g_{ij}) is a quotient of the round neck $\mathbb{S}^2 \times \mathbb{R}$.*

PROOF. In view of the result of Ni-Wallach mentioned above, it suffices to show that our shrinking gradient soliton in fact has nonnegative Ricci curvature and satisfies the growth restriction on curvature.

First of all, by the work of the second author (see Corollary 2.4 of [19]), we know that the sectional curvature of g_{ij} must be nonnegative.

Next we claim that the scalar curvature, hence the curvature tensor, of g_{ij} grows at most quadratically in distance. Indeed, from the shrinking soliton equation

$$R_{ij} + f_{ij} - \frac{1}{2}g_{ij} = 0,$$

it is not hard to see

$$R_{jt} \nabla_t f = \frac{1}{2} \nabla_j R$$

and

$$R + |\nabla f|^2 - f = \text{Const.}$$

It then follows that $|\nabla f|^2 \leq f + \text{Const}$, because R is nonnegative. Thus, we obtain

$$|\nabla \sqrt{|f| + 1}| \leq \text{Const},$$

and hence

$$|f|(x) \leq C(d(x, x_0)^2 + 1).$$

Therefore,

$$R(x) \leq C(d(x, x_0)^2 + 1).$$

This completes the proof of the proposition. □

Clearly Perelman's argument using the Gauss-Bonnet formula imposes a restriction on the dimension. Thus an interesting open question is whether a similar classification of non-negatively curved shrinking solitons holds in higher dimensions. For $n = 4$, Ni and Wallach [78] showed that any 4-dimensional complete gradient shrinking soliton with nonnegative curvature operator and positive isotropic curvature, satisfying certain additional assumptions, is either a quotient of \mathbb{S}^4 or a quotient of $\mathbb{S}^3 \times \mathbb{R}$. Based on this result of Ni-Wallach, Naber [72] proved that

PROPOSITION 4.8 (Naber [72]). *Any 4-dimensional complete noncompact shrinking Ricci soliton with bounded nonnegative curvature operator is isometric to either \mathbb{R}^4 , or a finite quotient of $\mathbb{S}^3 \times \mathbb{R}$ or $\mathbb{S}^2 \times \mathbb{R}^2$.*

For higher dimensions, Gu and the last author [40] proved that any complete, *rotationally symmetric*, non-flat, n -dimensional ($n \geq 3$) shrinking Ricci soliton with κ -noncollapsing on all scales and with bounded and nonnegative sectional curvature must be the round sphere \mathbb{S}^n or the round cylinder $\mathbb{S}^{n-1} \times \mathbb{R}$. Subsequently, Kotschwar [59] proved a more general result that the only complete shrinking Ricci solitons (without curvature sign and bound assumptions) of rotationally symmetric metrics (on S^n , \mathbb{R}^n and $\mathbb{R} \times S^{n-1}$) are, respectively, the round, flat, and standard cylindrical metrics. Ni-Wallach [77] and Petersen-Wylie [86] also proved a classification result on gradient shrinking solitons with vanishing Weyl curvature tensor which includes all the rotationally symmetric ones. For additional recent results on shrinking or expanding Ricci solitons, see the works of Petersen and Wylie [84, 85].

Let us come back to the discussion on general ancient κ -solutions. Given a three-dimensional ancient κ -solution, one can pick a suitable sequence of space-time points (x_k, t_k) with $t_k \rightarrow -\infty$ as in [81] and take a rescaling limit, usually called a blow-down limit. By using the monotonicity of the reduced volume, Perelman [80] showed that the blow-down limit is necessarily a shrinking Ricci soliton. Then, based on the above classification lemma (Lemma 4.6) for three-dimensional shrinking Ricci solitons and imitating the argument as in proving his noncollapsing result, Perelman [81] obtained the following important universal noncollapsing property for all three-dimensional ancient κ -solutions.

PROPOSITION 4.9 (Universal Noncollapsing [81]). *There exists a positive constant κ_0 with the following property. Suppose we have a non-flat three-dimensional ancient κ -solution for some $\kappa > 0$. Then either the solution is κ_0 -noncollapsed on all scales, or it is a metric quotient of the round three-sphere.*

This universal noncollapsing property for three-dimensional ancient κ -solutions has been used indispensably by Perelman in [81] to prove the noncollapsing of surgical solutions to the Ricci flow with surgery. When extending the Hamilton-Perelman theory of three-dimensional Ricci flow with surgery to higher dimensions, one must meet the question how to verify the universal property to ancient κ -solutions. Due to the lack of complete classification of higher dimensional positively curved shrinking Ricci solitons, it is desirable to find an alternative way, without using the classification of shrinking Ricci solitons, to prove the universal noncollapsing property. Indeed, an alternative approach had been given by the last two authors in [25] to handle a class \mathcal{C} of ancient κ -solutions without any knowledge of classification of gradient shrinking solitons. Roughly speaking, the class \mathcal{C} contains all ancient κ -solutions where each of them at infinity splits as $\mathbb{S}^{n-1} \times \mathbb{R}$. In particular, all ancient three-dimensional κ -solutions and four-dimensional ancient κ -solutions with restrictive isotropic curvature pinching belong to this class \mathcal{C} . Here we say a four-dimensional ancient κ -solution satisfies *restricted isotropic curvature pinching* if there is some fixed $\Lambda > 0$ such that

$$a_3 \leq \Lambda a_1, \quad c_3 \leq \Lambda c_1, \quad b_3^2 \leq a_1 c_1,$$

where $Rm = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$ is the usual block decomposition of curvature operator in dimension 4 and a_i, b_i, c_i are eigenvalues of the corresponding matrixes A, B, C . By Hamilton's pinching estimate in [48], such four-dimensional ancient κ -solutions with restricted isotropic curvature pinching appears naturally as the singularity models of Ricci flow on compact four-manifolds with positive isotropic curvature.

Dimension reduction is a useful approach to understand the structure of singularities in the theory of minimal surfaces or harmonic maps. In his survey paper [47], Hamilton systemically developed the dimension reduction method for the Ricci flow. From Hamilton's classification to two-dimensional ancient solutions, one observes that any two-dimensional complete ancient solution of bounded curvature cannot be of maximal volume growth. Based on this observation and by applying a dimension reduction argument, Perelman [80] proved

PROPOSITION 4.10 (Non-maximal Volume Growth). *Let M be an n -dimensional complete noncompact Riemannian manifold. Suppose $g_{ij}(x, t)$, $x \in M$ and $t \in (-\infty, T)$ with $T > 0$, is a nonflat ancient solution of the Ricci flow with a nonnegative curvature operator and bounded*

curvature. Then the asymptotic volume ratio of the solution metric satisfies

$$\nu_M(t) = \lim_{r \rightarrow +\infty} \frac{\text{Vol}_t(B_t(O, r))}{r^n} = 0$$

for each $t \in (-\infty, T)$.

The same result for the Ricci flow on Kähler manifolds has been independently discovered by the last two authors in [23]. Moreover, for the Ricci flow on Kähler manifolds, it is proved by the last two authors and Tang [20] in complex dimension two and by Ni [74] for all dimensions that the nonnegative curvature operator condition can be replaced by the weaker condition of nonnegative holomorphic bisectional curvature.

By a standard rescaling argument, using the above non-maximal volume growth property, Perelman [80] got a local curvature bound of solutions in terms of local volume lower bound. Conversely, the noncollapsing estimate of Perelman says that local curvature bound can control the local volume lower bound (see Figure 1). Hence the combination of these two facts would imply an elliptic type estimate, which allows one to compare the values of the curvatures at different points at the same time. Such an estimate was first implicitly given by Perelman in [80]. The following version is taken from [14] and [25].

PROPOSITION 4.11 (Elliptic Type Estimate). *There exist a positive constant η and a positive function $\omega : [0, +\infty) \rightarrow (0, +\infty)$ with the following properties. Suppose that $(M, g_{ij}(t))$, $-\infty < t \leq 0$, is a 3-dimensional ancient κ -solution or a 4-dimensional ancient κ -solution with restricted isotropic curvature pinching, for some $\kappa > 0$. Then*

(i) *for every $x, y \in M$ and $t \in (-\infty, 0]$, there holds*

$$R(x, t) \leq R(y, t) \cdot \omega(R(y, t)d_t^2(x, y));$$

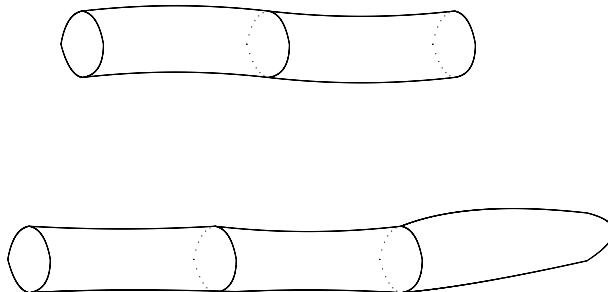


FIGURE 1. ϵ -neck and ϵ -cap.

(ii) for all $x \in M$ and $t \in (-\infty, 0]$, there hold

$$|\nabla R|(x, t) \leq \eta R^{\frac{3}{2}}(x, t) \text{ and } |R_t|(x, t) \leq \eta R^2(x, t).$$

Let us come back to consider three-dimensional ancient κ -solutions. In view of Hamilton’s dimension reduction, each noncompact three-dimensional ancient κ -solution splits off a line at infinity. Then by combining the classification of two-dimensional ancient κ -solutions, we see that each noncompact non-flat three-dimensional ancient κ -solution is asymptotic to a round cylinder at infinity. On the other hand, by applying the universal noncollapsing Proposition 4.9 and the above elliptic type estimate Proposition 4.11, we know that the space of non-flat three-dimensional ancient κ -solutions is compact modulo scalings and the quotients of the round sphere \mathbb{S}^3 . This compactness property and asymptotically cylindrical property allow us to use a standard rescaling argument to get a canonical neighborhood property, due to Perelman [81], for three-dimensional ancient κ -solutions.

Before stating the canonical neighborhood result, we introduce the terminologies of evolving ε -neck and ε -cap.

Fix $\varepsilon > 0$. Let $g_{ij}(x, t)$ be a non-flat ancient κ -solution on a three-manifold M for some $\kappa > 0$. We say that a point $x_0 \in M$ is the *center of an evolving ε -neck* at $t=0$, if the solution $g_{ij}(x, t)$ in the set $\{(x, t) \mid -\varepsilon^{-2}Q^{-1} < t \leq 0, d_t^2(x, x_0) < \varepsilon^{-2}Q^{-1}\}$, where $Q = R(x_0, 0)$, is, after scaling with factor Q , ε -close (in $C^{[\varepsilon^{-1}]}$ topology) to the corresponding set of the evolving round cylinder, having scalar curvature one at $t=0$. An evolving ε -cap is the time slice at the time t of an evolving metric on \mathbb{B}^3 or $\mathbb{RP}^3 \setminus \bar{\mathbb{B}}^3$ such that the region outside some suitable compact subset of \mathbb{B}^3 or $\mathbb{RP}^3 \setminus \bar{\mathbb{B}}^3$ is an evolving ε -neck.

THEOREM 4.12 (Canonical neighborhood theorem [81]). *For every sufficiently small $\varepsilon > 0$ one can find positive constants $C_1 = C_1(\varepsilon)$, $C_2 = C_2(\varepsilon)$ with the following property. Suppose we have a three-dimensional nonflat (compact or noncompact) ancient κ -solution $(M, g_{ij}(x, t))$. Then either the ancient solution is the round $\mathbb{RP}^2 \times \mathbb{R}$, or every point (x, t) has an open neighborhood B , with $B_t(x, r) \subset B \subset B_t(x, 2r)$ for some $0 < r < C_1 R(x, t)^{-\frac{1}{2}}$, which falls into one of the following three categories:*

- (a) B is an **evolving ε -neck**, or
- (b) B is an **evolving ε -cap**, or
- (c) B is a compact manifold (without boundary) with positive sectional curvature (thus it is diffeomorphic to the round three-sphere \mathbb{S}^3 or its metric quotients); furthermore, the scalar curvature of the ancient κ -solution in B at time t is between $C_2^{-1}R(x, t)$ and $C_2R(x, t)$, and its volume in cases (a) and (b) satisfies

$$(C_2R(x, t))^{-\frac{3}{2}} \leq Vol_t(B) \leq \varepsilon r^3.$$

Finally, we remark that this canonical neighborhood theorem has been extended by the last two authors [25] to all four-dimensional ancient κ -solutions with restrictive isotropic curvature pinching.

4.4. Singularity Structure Theorem. Let (M, g_{ij}) be a compact oriented three-manifold. Evolve the metric g_{ij} by the Ricci flow. Denote by $[0, T)$ the maximal time interval. Suppose $T < \infty$, then $\sup_{x \in M} |Rm|(x, t) \rightarrow \infty$ as $t \rightarrow T$. Let (x_k, t_k) be a sequence of almost maximal points, i.e. $\sup_{t \leq t_k} |Rm|(\cdot, t) \leq C |Rm|(x_k, t_k)$, $t_k \rightarrow T$, for some uniform constant C . Scale the solution around (x_k, t_k) with factor $Q_k = |Rm|(x_k, t_k)$ and shift the time t_k to 0. By applying Hamilton's compactness theorem, Perelman's local non-collapsing theorem, as well as Hamilton-Ivey pinching estimate, one can extract a convergent subsequence such that the limit is an oriented ancient κ -solution. Observe that $\mathbb{RP}^2 \times \mathbb{R}$ is excluded since it is not orientable. Consequently, for an arbitrarily given $\varepsilon > 0$, the solution around the points x_k and at times $t_k \rightarrow T$ have canonical neighborhoods which are either an ε -neck, or an ε -cap, or a compact positively curved manifold (without boundary). This gives the structure of singularities coming from a sequence of (almost) maximum points.

However the above argument does not work for singularities coming from a sequence of points (y_k, s_k) with $s_k \rightarrow T$ and $|Rm(y_k, s_k)| \rightarrow +\infty$ when $|Rm(y_k, s_k)|$ is not comparable with the maximum of the curvature at time s_k , since we cannot take a limit directly. To overcome this difficulty, Perelman [80] developed a refined blow up argument.

For convenience of stating the estimates, we may assume the initial data is normalized, namely, the norm of the curvature operator is less than $\frac{1}{10}$ and the volume of the unit ball is bigger than 1.

THEOREM 4.13 (Singularity structure theorem [80]). *Given $\varepsilon > 0$ and $T_0 > 0$, one can find $r_0 > 0$ with the following property. If $g_{ij}(x, t), x \in M$ and $t \in [0, T)$ with $1 < T \leq T_0$, is a solution to the Ricci flow on a compact oriented three-manifold M with normalized initial metric, then for any point (x_0, t_0) with $t_0 \geq 1$ and $Q = R(x_0, t_0) \geq r_0^{-2}$, the solution in $\{(x, t) \mid d_{t_0}^2(x, x_0) < \varepsilon^{-2} Q^{-1}, t_0 - \varepsilon^{-2} Q^{-1} \leq t \leq t_0\}$ is, after scaling by the factor Q , ε -close (in $C^{[\varepsilon^{-1}]}$ -topology) to the corresponding subset of some oriented ancient κ -solution (for some $\kappa > 0$).*

We now would like to give an outline of the proof. The proof is divided into four steps. The first three steps are basically following the line given by Perelman in [80]; while the last step is an alternative argument which is taken from [14] or [25].

The proof is an argument by contradiction. Suppose for some $\varepsilon > 0$, $T_0 > 1$, there exist a sequence of $r_k \rightarrow 0$, $1 < T_k \leq T_0$ and solutions $(M_k, g_k(\cdot, t))$, $t \in [0, T_k)$, satisfying the assumption of the theorem, but the conclusion of the theorem fails at some $x_k \in M_k$ and times $t_k \geq 1$

with $Q_k = R_k(x_k, t_k) \geq r_k^{-2}$. For each such solution, we adjust the point (x_k, t_k) so that the value of the curvature at (x_k, t_k) is as large as possible so that the conclusion of the theorem fails at (x_k, t_k) , but holds for any $(x, t) \in M_k \times [t_k - H_k Q_k^{-1}, t_k]$ satisfying $R_k(x, t) \geq 2Q_k$, where $H_k = \frac{1}{4} r_k^{-2} \rightarrow +\infty$ as $k \rightarrow +\infty$.

Let $(M_k, \tilde{g}_k(\cdot, t), x_k)$ be the rescaled solutions obtained by rescaling $(M_k, g_k(\cdot, t))$ around x_k with the factors $Q_k = R_k(x_k, t_k)$ and shifting the time t_k to the new time zero. Denote by \tilde{R}_k the rescaled scalar curvature. We will show that a subsequence of the rescaled solutions $(M_k, \tilde{g}_k(\cdot, t), x_k)$ converges in C_{loc}^∞ topology to an ancient κ -solution. This will be a contradiction.

The argument is divided into four steps.

Step 1. First of all, we need a local bound on curvature.

For each (\bar{x}, \bar{t}) with $t_k - \frac{1}{2} H_k Q_k^{-1} \leq \bar{t} \leq t_k$, we have

$$R_k(x, t) \leq 4\bar{Q}_k$$

whenever $\bar{t} - c\bar{Q}_k^{-1} \leq t \leq \bar{t}$ and $d_{\bar{t}}^2(x, \bar{x}) \leq c\bar{Q}_k^{-1}$, where $\bar{Q}_k = Q_k + R_k(\bar{x}, \bar{t})$ and $c > 0$ is a small universal constant.

This result is a simple consequence of the gradient estimate (ii) in Proposition 4.11. Indeed, since any ancient κ -solution satisfies the gradient estimate

$$(4.1) \quad |\nabla R^{-\frac{1}{2}}| + \left| \frac{\partial}{\partial t} R^{-1} \right| \leq 2\eta,$$

the desired curvature bound follows directly from integrating the gradient estimate along a space-time path.

Step 2. This step is to show that the curvature of rescaled solution is bounded at bounded distance at time $t = 0$. The detailed exposition to this step was first given by Kleiner-Lott in the first version of their notes [56]. The idea of the proof can be described as follows.

For all $\rho \geq 0$, set

$$M(\rho) = \sup\{\tilde{R}_k(x, 0) \mid k \geq 1, x \in M_k \text{ with } d_0(x, x_k) \leq \rho\}$$

and

$$\rho_0 = \sup\{\rho \geq 0 \mid M(\rho) < +\infty\}.$$

By Hamilton-Ivey’s pinching estimate, it suffices to show $\rho_0 = +\infty$. Still argue by contradiction. Suppose there is a sequence of points y_k so that the rescaled $\tilde{R}(y_k, 0) \rightarrow +\infty$ and $\tilde{d}_0(x_k, y_k) \rightarrow \rho_0 > 0$. Connecting x_k and y_k with a minimal geodesic γ_k . By Step 1, Hamilton’s compactness theorem 4.1 and Perelman’s noncollapsing theorem, there is a convergent subsequence such that the limit has nonnegative sectional curvature on the ball of radius ρ_0 . The curvature still blows up along the limiting geodesic γ_∞ by the gradient estimate for ancient κ -solutions. Then by the choice of the points x_k , one can

show each such point on the limiting geodesic has a neck-like neighborhood. So, by adding the end point q_∞ to the limit geodesic, the union of the limiting space and the added point q_∞ has nonnegative curvature in Alexandrov space sense. By blowing up the tangent cone at the q_∞ , we get a non-flat solution to the Ricci flow on the cone, which is a contradiction to Hamilton’s strong maximum principle.

Since the curvature is bounded at time 0, by gradient estimate (4.1) and Hamilton’s compactness theorem, one can show the limit solution is actually defined on the space-time open subset $\{(y, \tilde{t}) : y \in M_\infty, t \in [-\frac{1}{4}\eta^{-1} \tilde{R}_\infty(y)^{-1}, 0]\}$ containing $M_\infty \times \{0\}$.

Step 3. This step is to show that the limit $(M_\infty, \tilde{g}_\infty(\cdot, 0), x_\infty)$ at the time slice $\{t = 0\}$ has bounded curvature.

If the curvature is unbounded, by the virtue of Hamilton’s dimension reduction, we can choose a sequence of points $q_j \rightarrow \infty$, and take a rescaled limit around q_j to get infinite number of tiny ε -necks. But this is a contradiction with the following basic geometry lemma, which was written down by the last two authors in [25].

LEMMA 4.14. *There exists a constant $\varepsilon_0 = \varepsilon_0(n) > 0$ such that every complete noncompact Riemannian manifold (M^n, g_{ij}) of nonnegative sectional curvature has a positive constant r_0 such that any ε -neck of radius r on (M^n, g_{ij}) with $\varepsilon \leq \varepsilon_0$ must have $r \geq r_0$.*

Here we call an open subset $N \subset M^n$ to be an ε -neck of radius r if $(N, r^{-2}g_{ij})$ is ε -close, in $C^{[\varepsilon^{-1}]}$ topology, to a standard neck $\mathbb{S}^{n-1} \times (-\varepsilon^{-1}, \varepsilon^{-1})$ where \mathbb{S}^{n-1} has the scalar curvature 1.

As a consequence, the limit can be extended backward to some uniform interval $(-C, 0]$ for some $C > 0$.

Step 4. This step is to show the limit can be extended backward to $-\infty$.

Denote by

$$t' = \inf\{ \tilde{t} \mid \text{we can take a smooth limit on } (\tilde{t}, 0] \\ \text{from a subsequence of the rescaled solutions } \tilde{g}_k \}.$$

By the Li-Yau-Hamilton inequality, which must hold on the limit since the curvature is bounded by Step 1 and Step 3, and Hamilton’s compactness theorem, one can show that there is a subsequence of the rescaled solutions \tilde{g}_k which converges in C_{loc}^∞ topology to a smooth limit $(M_\infty, \tilde{g}_\infty(\cdot, t))$ on the maximal time interval $(t', 0]$. We next claim that $t' = -\infty$.

Suppose not, then the curvature of the limit $(M_\infty, \tilde{g}_\infty(\cdot, t))$ becomes unbounded as $t \rightarrow t' > -\infty$. By applying the maximum principle, we see that the infimum of the scalar curvature is nondecreasing in time. Thus

there exists some point $y_\infty \in M_\infty$ such that

$$\tilde{R}_\infty(y_\infty, t' + \frac{c}{3}) < \frac{3}{2}$$

where $c > 0$ is some universal small constant.

By using Step 1, we see that the limit $(M_\infty, \tilde{g}_\infty(\cdot, t))$ at a small neighborhood of the point $(y_\infty, t' + \frac{c}{3})$ extends backward to the time interval $[t' - \frac{c}{3}, t' + \frac{c}{3}]$. Moreover, one can show the distances at the time t and the time 0 are roughly equivalent in the following sense

$$(4.2) \quad d_t(x, y) \geq d_0(x, y) \geq d_t(x, y) - \text{const.}$$

This estimate ensures that the limit around the point y_∞ at any time $t \in (t', 0]$ is exactly the original limit around x_∞ at the time $t = 0$. By repeating the same arguments as in the above Step 2 and Step 3 to the solution around (y_k, t) for $t \in [t' - \frac{c}{3}, t' + \frac{c}{3}]$, we conclude the original limit $(M_\infty, \tilde{g}_\infty(\cdot, t))$ is actually well defined on the time slice $M_\infty \times \{t'\}$ and also has uniformly bounded curvature for all $t \in [t', 0]$. This is a contradiction.

Therefore the proof of the theorem is completed.

We remark that this singularity structure theorem had been extended by the last two authors in [25] to the Ricci flow on compact four-manifolds with positive isotropic curvature.

5. Ricci Flow with Surgery

In this section, we will discuss the surgery theory of the Ricci flow on three-dimensional manifolds. We also mention its extension to four-dimensional manifolds with positive isotropic curvature.

5.1. The Solution at the First Singular Time. Given any compact three-manifold M with an arbitrary Riemannian metric. By dilation, we may always assume that the metric is normalized so that the absolute values of the eigenvalues of its curvature operator at each point are bounded by $1/10$ and every geodesic ball of radius one has a volume of at least one. Let us evolve the normalized metric by the Ricci flow

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij},$$

and let $g(t)$, $t \in [0, T)$ be the maximal solution. If $T < \infty$, then curvature becomes unbounded as t tends to T , we say the maximal solution develops singularities as t tends to T and T is a singular time.

After obtaining the structure of points with suitably large curvature before the first singular time as in Theorem 4.13, we can give a clear picture of the solution near the singular time T as follows.

For the given $\varepsilon > 0$ and the maximal solution $(M, g_{ij}(\cdot, t))$ on $[0, T)$, with $T < \infty$, we can find $r_0 > 0$ depending only on T and ε such that each point (x, t) , with $R(x, t) \geq r_0^{-2}$, admits a canonical neighborhood which is either an ε -neck, or an ε -cap, or a compact positively curved manifold (without boundary). In the last case the solution, by the well-known theorem of Hamilton in [41] (see also Theorem 2.1), becomes extinct at time T and the manifold M is diffeomorphic to the round three-sphere \mathbb{S}^3 or a metric quotient of \mathbb{S}^3 .

Let Ω denote the set of all points in M where the curvature stays bounded as $t \rightarrow T$. If Ω is empty, then the solution becomes extinct at time T . In this case, either the manifold M is compact and positively curved, or it is entirely covered by ε -necks and ε -caps shortly before the maximal time T . So the manifold M is diffeomorphic to either \mathbb{S}^3 , or a metric quotient of the round \mathbb{S}^3 , or $\mathbb{S}^2 \times \mathbb{S}^1$, or $\mathbb{RP}^3 \# \mathbb{RP}^3$.

We now consider the case when Ω is nonempty. By using the local derivative estimates of Shi (Theorem 1.4), we see that as $t \rightarrow T$ the solution metric $g(t)$ has a smooth limit \bar{g} on Ω . Let \bar{R} denote the scalar curvature of \bar{g} . For any $0 < \rho < r_0$, let us consider the set

$$\Omega_\rho = \{x \in \Omega \mid \bar{R}(x) \leq \rho^{-2}\}.$$

First, we need some terminologies:

A metric on $\mathbb{S}^2 \times \mathbb{I}$, such that each point is contained in some ε -neck, is called an ε -**tube**, or an ε -**horn**, or a **double ε -horn**, if the scalar curvature stays bounded on both ends, or stays bounded on one end and tends to infinity on the other end, or tends to infinity on both ends, respectively (see Figure 2);

A metric on \mathbb{B}^3 or $\mathbb{RP}^3 \setminus \bar{\mathbb{B}}^3$ is called an **capped ε -horn** if each point outside some compact subset is contained in an ε -neck and the scalar curvature tends to infinity on the end (see Figure 3).

Now take any ε -neck in (Ω, \bar{g}) and consider a point x on one of its boundary components. If $x \in \Omega \setminus \Omega_\rho$, then there is either an ε -cap or an ε -neck, adjacent to the initial ε -neck. In the latter case we can take a point on the boundary of the second ε -neck and continue. This procedure can either terminate when we get into Ω_ρ or an ε -cap, or go on indefinitely, producing an ε -horn. The same procedure can be repeated for the other boundary component of the initial ε -neck. Therefore, taking into account that Ω has no compact components, we conclude that each ε -neck of (Ω, \bar{g}) is contained in a subset of Ω of one of the following types:

- (a) an ε -tube with boundary components in Ω_ρ , or
- (b) an ε -cap with boundary in Ω_ρ , or
- (c) an ε -horn with boundary in Ω_ρ , or
- (d) a capped ε -horn, or
- (e) a double ε -horn.

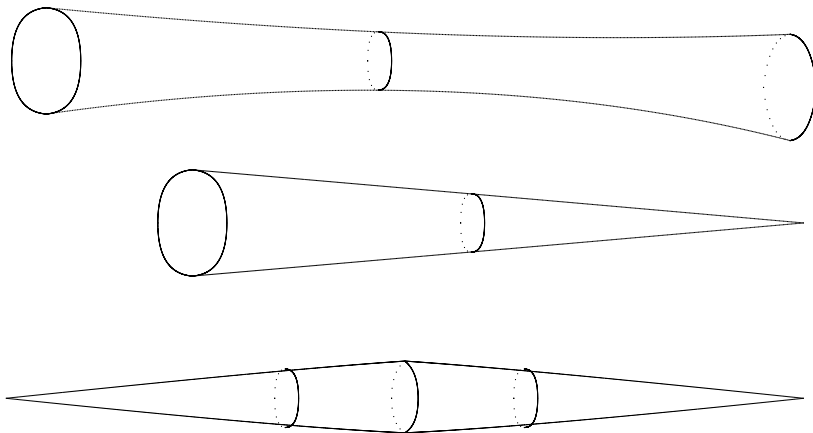


FIGURE 2. ε -tube, ε -horn and double ε -horn.

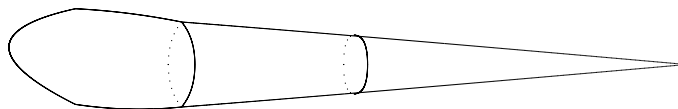


FIGURE 3. Capped ε -horn.

Similarly, each ε -cap of (Ω, \bar{g}_{ij}) is contained in a subset of Ω of either type (b) or type (d).

It is clear that there is a definite lower bound (depending on ρ) on the volume of subsets of type (a), (b), and (c). So there can be only a finite number of them. Thus we conclude that there is only a finite number of components of Ω , containing points of Ω_ρ , and every such component has a finite number of ends, each being an ε -horn. On the other hand, every component of Ω containing no points of Ω_ρ is either a capped ε -horn, or a double ε -horn. If we look at the solution $g(t)$ at a slightly earlier time, the above argument shows that each ε -neck or ε -cap of $(M, g(t))$ is contained in a subset of type (a) or (b), while the ε -horns, capped ε -horns and double ε -horns (at the maximal time T) are connected together to form ε -tubes and ε -caps at any time t shortly before T (see Figure 4).

Let us denote by Ω_j , $1 \leq j \leq m$, the connected components of Ω which contain points of Ω_ρ . Then *the initial three-manifold M is diffeomorphic to a connected sum of $\bar{\Omega}_j$, $1 \leq j \leq m$, with a finite number of copies of $\mathbb{S}^2 \times \mathbb{S}^1$*

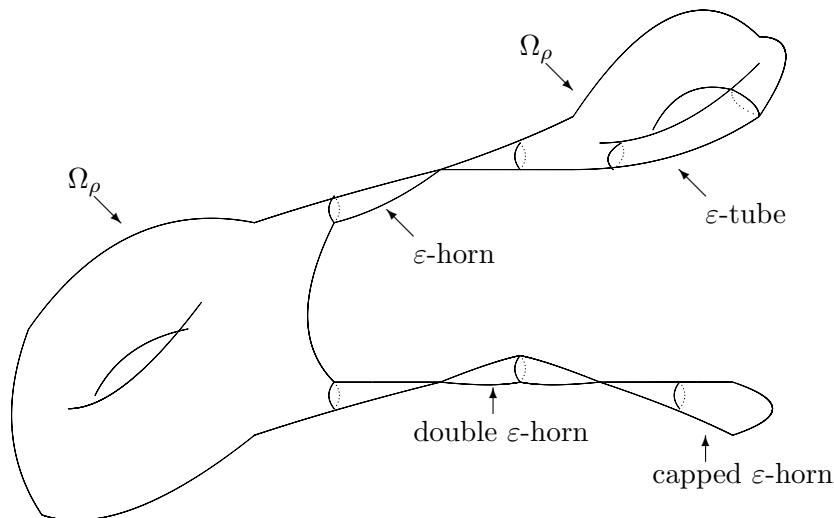


FIGURE 4. Solution at a maximal time.

(which correspond to gluing a tube to two boundary components of the same Ω_j), and a finite number of copies of $\mathbb{R}P^3$. Here $\bar{\Omega}_j$, $j = 1, 2, \dots, m$, is the compact manifold (without boundary) obtained from Ω_j by taking an ϵ -neck in every ϵ -horn of Ω_j , cutting it along the middle two-sphere, removing the horn-shaped end, and gluing back a cap (or more precisely, a differentiable three-ball).

5.2. Definition of Surgical Solutions. We have seen that when the Ricci flow develops singularities, it gives a natural way to split the underlying manifold M into pieces $\Omega_1, \dots, \Omega_m$ – the components of Ω containing points of Ω_ρ . Thus to capture the topology of M , one only needs to understand the topologies of the compact orientable three-manifolds $\bar{\Omega}_j$, $1 \leq j \leq m$, described above.

Let us evolve each $\bar{\Omega}_j$ by the Ricci flow again and, when the solution develops singularities, perform the above surgeries to get new compact orientable three-manifolds. By repeating this procedure, we will obtain a “weak” solution to the Ricci flow, called a solution to the *Ricci flow with surgery* or a *surgically modified solution* to the Ricci flow.

To get the topological information of the initial manifold M from the Ricci flow with surgery, we have to construct a surgically modified solution so that it has at most a finite number of surgeries at each finite interval and admits a well-understood long-time behavior. In this section, we only

consider the question of how to construct a surgically modified solution with at most a finite number of surgeries at each finite interval.

Let us look at the above construction for surgical solutions in more detail. Arbitrarily fix a small positive constant ε . On the given compact orientable three-manifold M with a normalized Riemannian metric, we evolve the normalized metric by the Ricci flow to obtain a maximal solution defined on the maximal time interval $[0, t_1)$ with $t_1 < +\infty$. By the theorem on the structure of singularity, there exists a small positive constant r_0 such that any point (x, t) at which (the norm of) the curvature is greater than r_0^{-2} has a canonical neighborhood. Then, according to the above discussions, we can cut off canonical neighborhoods to get a new compact orientable (not necessarily connected) three-manifold M_1 . Clearly, there are still some points, in the remaining parts near the surgery region, on M_1 at which the curvature are not less than r_0^{-2} and then we cannot expect that the metric of M_1 is still normalized. After evolving M_1 on a maximal time interval $[t_1, t_2)$ with $t_2 < +\infty$, we can only find canonical neighborhoods on the region where the curvature is at least r_0^{-4} (since, to apply the theorem on the structure of singularity, we have to dilate M_1 with a factor at least r_0^{-2}). By performing the surgery again, we get a compact orientable (not necessarily connected) three-manifold M_2 and there are still some points on M_2 with curvature not less than r_0^{-4} . By repeating this process, we will get a surgically modified solution on some time interval $[0, T_{\max})$ with the surgery times $0 < t_1 < \dots < t_k < \dots < T_{\max}$ such that at each $t_k, k = 1, 2, \dots$, the curvature is at least r_0^{-2k} somewhere. Intuitively, under this kind of surgery procedures, the curvatures would become higher and higher and the time intervals (t_{k-1}, t_k) become shorter and shorter. So, the surgery times of such constructed surgically modified solution are likely to accumulate in finite time. The trouble is basically caused by the inability to recognize the canonical neighborhoods on some fixed size of (high) curvature.

If one can improve the above surgery procedures so that there exists a uniform size on curvature to recognize canonical neighborhoods, then one will be able to cut down the solution so that its curvature never exceeds such a designed uniform size and hence each surgery will drop at least a fixed amount of volume. This, in turn, will prevent the surgery times from accumulating since one can easily show that the volume of the surgically modified solution can grow (in time) at most exponentially. So, what one really needs is to design a surgery procedure such that one can find a uniform positive function $r(t)$ on $[0, +\infty)$ so that any point (x, t) on the surgically modified solution at which the curvature is greater than $r(t)^{-2}$ has a canonical neighborhood. The theorem on the structure of singularity precisely ensures the existence of such a uniform function $r(t)$ for *smooth* solutions.

Thus, to prevent the accumulation of surgery times, we are led to construct surgically modified solutions which satisfy the following canonical neighborhood assumption (we refer the readers to Section 7.3 in [14] for precise definitions):

Canonical neighborhood assumption (with accuracy ε): There exists a nonincreasing positive function $r : [0, +\infty) \rightarrow (0, +\infty)$ such that at each time t , each point x , where the scalar curvature $R(x, t)$ is at least $r^{-2}(t)$, has a neighborhood B falling into one of the three categories:

- (a) B is a strong ε -neck, or
- (b) B is an ε -cap, or
- (c) B is a compact manifold (without boundary) of positive curvature.

The Hamilton-Ivey curvature pinching estimate is a special feature on three-dimension. It plays an important role in the proof of the theorem of structure of singularity. Thus one should also require the surgical solutions to satisfy the following Hamilton-Ivey pinching condition:

Pinching assumption: The eigenvalues $\lambda \geq \mu \geq \nu$ of the curvature operator Rm of the surgical solution at each point and each time satisfy

$$R \geq (-\nu)[\log(-\nu) + \log(1 + t) - 3]$$

whenever $\nu < 0$.

5.3. Long-Time Existence of Surgical Solutions. Let ε be an arbitrarily given small positive constant. We now describe how to use an inductive argument to construct a long time surgically modified solution satisfying the pinching assumption and the canonical neighborhood assumption (with accuracy ε).

Start with a (smooth) maximal solution $g(t)$, $t \in [0, T)$, to the Ricci flow on the compact, oriented three-manifold M with normalized initial metric. By the Hamilton-Ivey pinching estimate and Theorem 4.13 on the structure of singularity, we see that the maximal solution $g(t)$ satisfies the pinching assumption and the canonical neighborhood assumption on the maximal time interval $[0, T)$. If $T = +\infty$, we have the desired long time solution. Thus, without loss of generality, we may assume $T < +\infty$ and hence the solution goes singular at time T .

Suppose that we have a surgically modified solution on $[0, T)$ (with $T < +\infty$ and with the normalized metric as initial data) which satisfies the pinching assumption and the canonical neighborhood assumption (with accuracy ε), becomes singular at time T , and has only a finite number of surgery times on $[0, T)$. Let Ω denote the set of all points in M where the curvature stays bounded as $t \rightarrow T$. Then the solution $g(t)$ has a smooth limit \bar{g} , defined on Ω , as $t \rightarrow T$.

For some $\delta > 0$ to be chosen much smaller than ε , we let $\rho = \delta r(T)$, where $r(t)$ is the positive nonincreasing function in the definition of the canonical neighborhood assumption. We then consider the corresponding compact set

$$\Omega_\rho = \{x \in \Omega \mid \bar{R}(x) \leq \rho^{-2}\},$$

where \bar{R} is the scalar curvature of \bar{g} .

If Ω_ρ is empty, then the manifold (near the maximal time T) is entirely covered by ε -necks, ε -caps and compact components with positive curvature. As a consequence, the manifold is diffeomorphic to the union of a finite number of copies of \mathbb{S}^3 , or metric quotients of the round \mathbb{S}^3 , or $\mathbb{S}^2 \times \mathbb{S}^1$, or a connected sum of them. Thus when Ω_ρ is empty, the procedure stops here, and we say the **solution becomes extinct**.

We now assume Ω_ρ is not empty. As was explained before, we only need to consider those components Ω_j , $1 \leq j \leq m$, of Ω which contain points of Ω_ρ . We will perform surgical procedures, which have been roughly described before, by finding an ε -neck in all horns of Ω_j , $1 \leq j \leq m$, and then cutting it along the middle two-sphere, removing the horn-shaped end, and gluing back a cap.

However, in order to maintain the pinching assumption and the canonical neighborhood assumption *with the same accuracy* after surgery, we will need to find sufficiently “fine” necks in the ε -horns and to glue sufficiently “fine” caps. Note that $\delta > 0$ is to be chosen much smaller than $\varepsilon > 0$.

Actually, one can show (due to Perelman [81], see Lemma 7.3.2 [14]) that in every ε -horn of Ω_j , $1 \leq j \leq m$, there exists a δ -neck with its radius depending only on δ and $r(T)$. This gives us the “fine” necks in the ε -horns.

To construct “fine” caps, we consider the semi-infinite standard round cylinder $N_0 = \mathbb{S}^2 \times (-\infty, 4)$ with the metric g_0 of scalar curvature 1. Denote by z the coordinate of the second factor $(-\infty, 4)$. Let f be a smooth nondecreasing convex function on $(-\infty, 4)$ defined by

$$\begin{cases} f(z) = 0, & z \leq 0, \\ f(z) = ce^{-\frac{P}{z}}, & z \in (0, 3], \\ f(z) \text{ is strictly convex on } z \in [3, 3.9], \\ f(z) = -\frac{1}{2} \log(16 - z^2), & z \in [3.9, 4), \end{cases}$$

where the (small) constant $c > 0$ and (big) constant $P > 0$ will be determined later (see Figure 5).

Let us replace the standard metric g_0 on the portion $\mathbb{S}^2 \times [0, 4)$ of the semi-infinite cylinder by the conformal change $e^{-2f}g_0$. Then the resulting metric \hat{g} is smoothly defined on \mathbb{R}^3 obtained by adding a point to $\mathbb{S}^2 \times (-\infty, 4)$ at $z = 4$. We denote by $C(c, P) = (\mathbb{R}^3, \hat{g})$, and call it a **standard capped infinite cylinder** (see Figure 6). Clearly $C(c, P)$ has nonnegative sectional curvature and positive scalar curvature everywhere.

As a side remark, one might wonder whether we should also cut off all those ε -tubes and ε -caps in the surgery procedure. However, in general one may not be able to find a “fine” neck in an ε -tube or an ε -cap, and surgeries at “rough” ε -necks will certainly lose some accuracy. If one performs the surgeries at the necks with some fixed accuracy ε on the high curvature

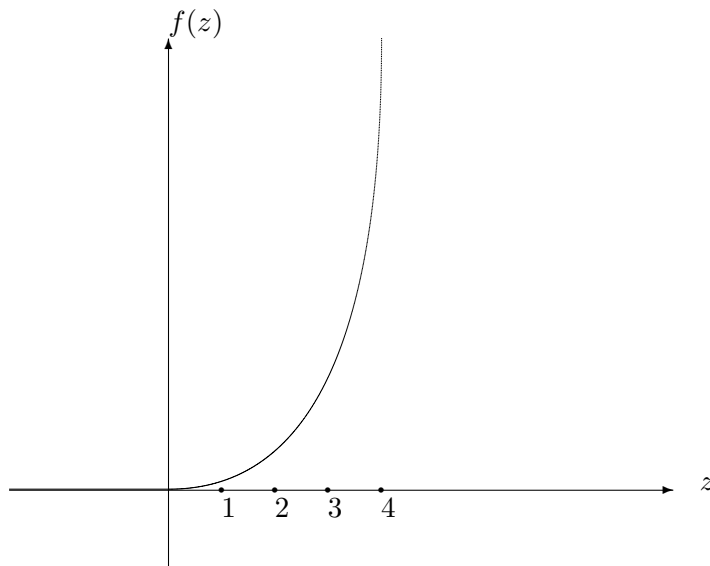


FIGURE 5. The function $f(z)$.



FIGURE 6. Standard capped infinite cylinder.

region at each surgery time, then it is possible that the errors of surgeries may accumulate to a certain amount so that at some later time one cannot recognize the structure of very high curvature region. This prevents us to carry out the whole process in finite time with finite steps. This is the reason why we will only perform the surgeries at ε -horns.

We can now perform Hamilton's geometric surgery procedure as follows. Take an ε -horn with boundary in Ω_ρ and take a δ -neck N of radius h , $0 < h < \delta\rho$, in the ε -horn. By definition, $(N, h^{-2}\bar{g})$ is δ -close (in the $C^{[\delta^{-1}]}$ topology) to the standard round neck $\mathbb{S}^2 \times \mathbb{I}$ of scalar curvature 1 with $\mathbb{I} = (-\delta^{-1}, \delta^{-1})$. The parameter $z \in \mathbb{I}$ induces a function on the δ -neck N .

Let us cut the δ -neck N along the middle (topological) two-sphere $N \cap \{z = 0\}$. Without loss of generality, we may assume that the right hand half portion $N \cap \{z \geq 0\}$ is contained in the horn-shaped end. Let φ be a smooth bump function with $\varphi = 1$ for $z \leq 2$, and $\varphi = 0$ for $z \geq 3$. Construct

a new metric \tilde{g} on a (topological) three-ball \mathbb{B}^3 as

$$\tilde{g} = \begin{cases} \bar{g}, & z = 0, \\ e^{-2f}\bar{g}, & z \in [0, 2], \\ \varphi e^{-2f}\bar{g} + (1 - \varphi)e^{-2f}h^2g_0, & z \in [2, 3], \\ h^2e^{-2f}g_0, & z \in [3, 4]. \end{cases}$$

The surgery, called a δ -cutoff surgery, is to replace the horn-shaped end by the cap (\mathbb{B}^3, \tilde{g}) (see Figure 7).

We remark that this type of surgery is topologically trivial. But it is geometrically significant: after suitable adjusting the parameters c , P and δ , the pinching assumption will survive under the surgeries. Indeed, we can prove

THEOREM 5.1 (see Lemma 7.3.4 in [14]). *There are universal positive constants δ_0 , c_0 and P_0 such that if one takes a δ -cutoff surgery at a δ -neck of radius h at time T with $\delta \leq \delta_0$ and $h^{-2} \geq 2e^2 \log(1 + T)$, then one can choose $c = c_0$ and $P = P_0$ in the definition of $f(z)$ such that after the surgery, the pinching condition*

$$\tilde{R} \geq (-\tilde{\nu})[\log(-\tilde{\nu}) + \log(1 + T) - 3]$$

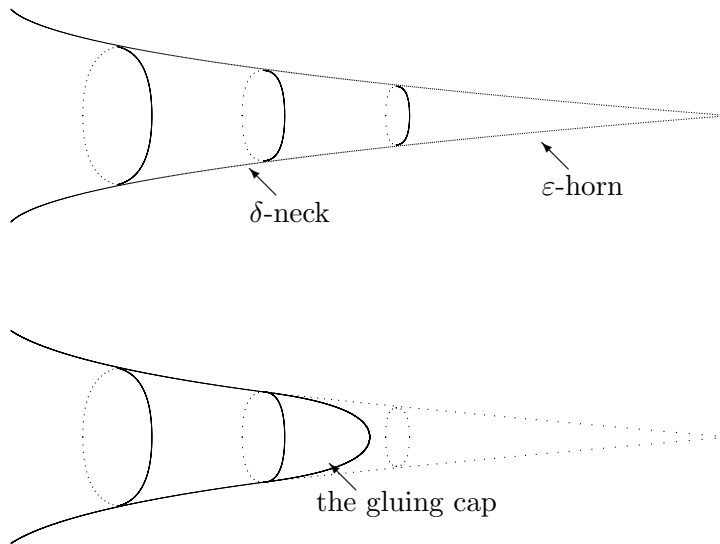


FIGURE 7. δ -cutoff surgery.

still holds whenever $\tilde{\nu} < 0$. Here \tilde{R} is the scalar curvature of \tilde{g} and $\tilde{\nu}$ is the least eigenvalue of the curvature operator of \tilde{g} .

Define the positive function $\bar{\delta}(t)$ on $[0, +\infty)$ by

$$\bar{\delta}(t) = \min \left\{ \frac{1}{2e^2 \log(1+t)}, \delta_0 \right\}.$$

From now on we will always assume $0 < \delta < \bar{\delta}(t)$ for any δ -cutoff surgery at a time $t > 0$ and take $c = c_0$ and $P = P_0$ so that the pinching assumption is preserved under the surgeries at T . After performing the δ -cutoff surgeries for all Ω_j , $1 \leq j \leq m$, we obtain the compact (without boundary), orientable three-manifolds $\bar{\Omega}_j$, $1 \leq j \leq m$. With these new compact manifolds as initial data, we can continue the solution under the Ricci flow until it becomes singular again at some later time $T' > T$. By the Hamilton-Ivey estimate (Theorem 1.7), we see that the solution still satisfies the pinching assumption on the extended time interval $[0, T')$.

By dilation and Theorem 4.13 on the structure of singularity, there always exists a nonincreasing positive function $r = r'(t)$, defined on $[0, +\infty)$, such that the canonical neighborhood assumption (with accuracy ε) holds on the extended time interval $[0, T')$ with the positive function $r = r'(t)$. Nevertheless, in order to prevent the surgery times from accumulating, the key is to choose the nonincreasing positive functions $r(t)$ uniformly. By a further restriction on the positive function $\bar{\delta}(t)$ we can verify the canonical neighborhood assumption with a uniform $r(t)$.

THEOREM 5.2 (Justification of the canonical neighborhood assumption [81]). *Given any small $\varepsilon > 0$, there exist decreasing sequences $0 < \tilde{r}_j < \varepsilon$ and $0 < \tilde{\delta}_j < \varepsilon^2$, $j = 1, 2, \dots$, with the following property. Define the positive function $\tilde{\delta}(t)$ on $[0, +\infty)$ by $\tilde{\delta}(t) = \tilde{\delta}_j$ for $t \in [(j-1)\varepsilon^2, j\varepsilon^2)$. Suppose there is a surgically modified solution, defined on $[0, T)$ with $T < +\infty$, to the Ricci flow which satisfies the following:*

- (1) *it starts on a compact orientable three-manifold with normalized initial metric, and*
- (2) *it has only a finite number of surgeries such that each surgery at a time $t \in (0, T)$ is a $\delta(t)$ -cutoff surgery with*

$$0 < \delta(t) \leq \min\{\tilde{\delta}(t), \bar{\delta}(t)\}.$$

Then on each time interval $[(j-1)\varepsilon^2, j\varepsilon^2] \cap [0, T)$, ($j = 1, 2, \dots$), the solution satisfies the canonical neighborhood assumption (with accuracy ε) with $r = \tilde{r}_j$.

This result was first given by Perelman in [81]. It extends the singularity structure theorem (Theorem 4.13) for smooth solutions to surgically modified solutions. However, when one tries to adapt the arguments of the smooth

case to the surgical case, they will encounter several difficulties: how to generalize the non local collapsing theorem of Perelman to surgical solutions to get the local injectivity radius bound; how to apply Hamilton's compactness theorem to surgically modified solutions; how to extend the rescaling limits backward in time without touching the surgical regions. Below we will give a brief description of the proof.

The proof is by induction: having constructed our sequences for $1 \leq j \leq m$, we make one more step, defining \tilde{r}_{m+1} and $\tilde{\delta}_{m+1}$. We follow a very clever idea of Perelman [81] by redefining $\tilde{\delta}_m = \tilde{\delta}_{m+1}$ in order to push the surgical regions to infinity in space.

We argue by contradiction. Suppose for some sequences of positive numbers $r^\alpha \rightarrow 0$ and $\tilde{\delta}^{\alpha\beta} \rightarrow 0$, there exist sequences of solutions $g_{ij}^{\alpha\beta}$ to the Ricci flow with surgery, with a compact orientable normalized three-manifold as initial data, so that

- (i) each δ -cutoff at time $t \in [(m-1)\varepsilon^2, (m+1)\varepsilon^2]$ satisfies $\delta \leq \tilde{\delta}^{\alpha\beta}$; and
- (ii) the solutions satisfy the statement of the proposition on $[0, m\varepsilon^2]$, but violate the canonical neighborhood assumption (with accuracy ε) with $r = r^\alpha$ on $[m\varepsilon^2, (m+1)\varepsilon^2]$.

For each solution $g_{ij}^{\alpha\beta}$, we choose $\bar{t}^{\alpha\beta}$ to be the nearly first time for which the canonical neighborhood assumption (with accuracy ε) is violated at some $(\bar{x}^{\alpha\beta}, \bar{t}^{\alpha\beta})$ but the canonical neighborhood assumption with accuracy parameter 2ε does hold on $t \in [m\varepsilon^2, \bar{t}^{\alpha\beta}]$.

Let $\tilde{g}_{ij}^{\alpha\beta}$ be the rescaled solutions around $(\bar{x}^{\alpha\beta}, \bar{t}^{\alpha\beta})$ with factors $R(\bar{x}^{\alpha\beta}, \bar{t}^{\alpha\beta}) (\geq (r^\alpha)^{-2} \rightarrow +\infty$ as $\alpha \rightarrow +\infty$) and shift $\bar{t}^{\alpha\beta}$ to zero. We hope to take a subsequential limit of the rescaled solutions as $\alpha, \beta \rightarrow \infty$ and show that the limit is an orientable ancient κ -solution, which will give the desired contradiction.

To do so, we first need to get a uniform lower bound for the injectivity radii of the rescaled sequence $\tilde{g}_{ij}^{\alpha\beta}$ at the marking points $(\bar{x}^{\alpha\beta}, \bar{t}^{\alpha\beta})$. Based on the fact that the canonical neighborhood assumption with accuracy parameter 2ε holds for $t \in [m\varepsilon^2, \bar{t}^{\alpha\beta}]$, we appeal the following lemma to show that the (unscaled) sequence $g_{ij}^{\alpha\beta}$ is κ -noncollapsed for some $\kappa > 0$ independent of α, β .

LEMMA 5.3 (Perelman [81], see also Lemma 7.4.2 in [14]). *Given $\varepsilon > 0$, suppose we have constructed the sequences satisfying the proposition for $1 \leq j \leq l$ (for some positive integer l). Then there exists $\kappa > 0$, such that for any r , $0 < r < \varepsilon$, one can find $\tilde{\delta} = \tilde{\delta}(r, \varepsilon)$, $0 < \tilde{\delta} < \varepsilon^2$, which may also depend on the already constructed sequences, with the following property. Suppose we have a solution with a compact oriented normalized three-manifold as initial data, to the Ricci flow with finite number of surgeries on a time interval $[0, T]$ with $l\varepsilon^2 \leq T < (l+1)\varepsilon^2$, satisfying the assumptions and the conclusions of Proposition 7.4.1 on $[0, l\varepsilon^2]$, and the canonical neighborhood*

assumption (with accuracy ε) with r on $[\varepsilon^2, \bar{T}]$, as well as $0 < \delta(t) \leq \tilde{\delta}$ for any δ -cutoff surgery with $\delta = \delta(t)$ at a time $t \in [(l-1)\varepsilon^2, \bar{T}]$. Then the solution is κ -noncollapsing on $[0, \bar{T}]$ for all scales less than ε .

The major observation in the proof of the lemma is that the space-time curves near the region of surgery carry large reduced distance, so one can find suitable surgically unaffected cone-like regions and to apply Perelman’s Jacobian comparison theorem (formula (3.15)) there as in the proof Theorem 3.9. The universal noncollapsing property of ancient κ -solutions is substantially used in this proof. We have to mention that the noncollapsing constant κ obtained in such a way does not depend on the canonical neighborhood parameter r in $[(l-1)\varepsilon^2, \bar{T}]$. This is the key point of the lemma².

The uniform noncollapsing estimate guarantees the desired injectivity radius bound for the rescaled sequence $\tilde{g}_{ij}^{\alpha\beta}$.

Next we need to get a uniform curvature bound for the rescaled sequence $\tilde{g}_{ij}^{\alpha\beta}$ on compact subsets around the marked points $\bar{x}^{\alpha\beta}$. Since the (unscaled) solutions satisfy the canonical neighborhood assumption with accuracy parameter 2ε on $[m\varepsilon^2, \bar{t}^{\alpha\beta}]$, we can use the gradient estimates in the canonical neighborhood assumption to get a uniform curvature estimate for the rescaled solutions $\tilde{g}_{ij}^{\alpha\beta}$ in some small space-time neighborhoods of $(\bar{x}^{\alpha\beta}, \bar{t}^{\alpha\beta})$. The sizes of these neighborhoods in space (at the new time zero) are uniform, but the time interval may vary due to the surgeries. If the time interval for the solution is too short, one can not apply Shi’s derivative estimates to obtain uniform higher derivatives estimates. This prevent us from applying Hamilton’s compactness theorem, which requires a uniform time interval for all the solutions in the sequence, to take a limit for the surgically modified rescaled solutions. To overcome this difficulty, in [14] and [25], the authors established three time-extension results:

The first assertion says that if we have curvature estimates for the renormalized solutions on a box $\tilde{B}_0(\bar{x}, A) \times [-b, 0]$, then the solution can be extended to a larger time interval.

Assertion 1. *For arbitrarily fixed α , $0 < A < +\infty$, $1 \leq C < +\infty$ and $0 \leq B < \frac{1}{2}\varepsilon^2(r^\alpha)^{-2} - \frac{1}{8}\eta^{-1}C^{-1}$, there is a $\beta_0 = \beta_0(\varepsilon, A, B, C)$ (independent of α) such that if $\beta \geq \beta_0$ and the rescaled solution $\tilde{g}_{ij}^{\alpha\beta}$ on the ball $\tilde{B}_0(\bar{x}, A)$ is defined on a time interval $[-b, 0]$ with $0 \leq b \leq B$ and the scalar curvature satisfies*

$$\tilde{R}(x, t) \leq C, \quad \text{on } \tilde{B}_0(\bar{x}, A) \times [-b, 0],$$

then the rescaled solution $\tilde{g}_{ij}^{\alpha\beta}$ on the ball $\tilde{B}_0(\bar{x}, A)$ is also defined on the extended time interval $[-b - \frac{1}{8}\eta^{-1}C^{-1}, 0]$.

²We have also learned the very recent works of Ye [105] and Zhang [108] on how to obtain a uniform Sobolev inequality, which is independent of the number of surgeries, and use it to derive κ -noncollapsing for surgically modified Ricci flow.

If the solution cannot be defined on the larger interval $[-b - \frac{1}{8}\eta^{-1}C^{-1}, 0]$, it must hit the surgery region. Since the surgery was done on the δ -neck, the solution on the surgery region is close to a standard solution by the uniqueness theorem of Ricci flow [24]. For standard solutions, we have canonical neighborhood decompositions. Note that once the solution is defined on $\tilde{B}_0(\bar{x}, A)$ for a subinterval $[-b - \nu, 0]$ of $[-b - \frac{1}{8}\eta^{-1}C^{-1}, 0]$, the curvature bound on this region follows directly from the gradient estimate on canonical neighborhoods with accuracy parameter 2ε . This curvature bound guarantees that the solution near the point \bar{x} is close to a standard solution until the time 0. Since the canonical neighborhood assumption with accuracy parameter ε is violated at (\bar{x}, \bar{t}) by assumption, this gives a contradiction.

Assertion 2. *For arbitrarily fixed α , $0 < A < +\infty$, $1 \leq C < +\infty$ and $0 \leq B < \frac{1}{2}\varepsilon^2(r^\alpha)^{-2} - \frac{1}{50}\eta^{-1}$, there is a $\beta_0 = \beta_0(\varepsilon, A, B, C)$ (independent of α) such that if $\beta \geq \beta_0$ and the rescaled solution $\tilde{g}_{ij}^{\alpha\beta}$ on the ball $\tilde{B}_0(\bar{x}, A)$ is defined on a time interval $[-b + \epsilon', 0]$ with $0 < b \leq B$ and $0 < \epsilon' < \frac{1}{50}\eta^{-1}$ and the scalar curvature satisfies*

$$\tilde{R}(x, t) \leq C \quad \text{on } \tilde{B}_0(\bar{x}, A) \times [-b + \epsilon', 0],$$

and there is a point $y \in \tilde{B}_0(\bar{x}, A)$ such that $\tilde{R}(y, -b + \epsilon') \leq \frac{3}{2}$, then the rescaled solution $\tilde{g}_{ij}^{\alpha\beta}$ at y is also defined on the extended time interval $[-b - \frac{1}{50}\eta^{-1}, 0]$ and satisfies the estimate

$$\tilde{R}(y, t) \leq 15$$

for $t \in [-b - \frac{1}{50}\eta^{-1}, -b + \epsilon']$.

Assertion 2 follows the same philosophy as in Assertion 1. Once we have the curvature estimates, and the solution hits the surgery, it must maintain the shape of standard solution until the time 0 by the uniqueness theorem in [24]. In practice, the point y will come from an almost minimal value of the scalar curvature, so its curvature is uniformly bounded.

The following Assertion 3 is based on the observation that the standard solution satisfies $R(x_1, t) \leq D''R(x_2, t)$ for any $t \in [0, \frac{1}{2}]$ and any two points x_1, x_2 , where D'' is a universal constant.

Assertion 3. *For arbitrarily fixed α , $0 < A < +\infty$, $1 \leq C < +\infty$, there is a $\beta_0 = \beta_0(\varepsilon, AC^{\frac{1}{2}})$ such that if any point (y_0, t_0) with $0 \leq -t_0 < \frac{1}{2}\varepsilon^2(r^\alpha)^{-2} - \frac{1}{8}\eta^{-1}C^{-1}$ of the rescaled solution $\tilde{g}_{ij}^{\alpha\beta}$ for $\beta \geq \beta_0$ satisfies $\tilde{R}(y_0, t_0) \leq C$, then either y_0 can be defined at least on $[t_0 - \frac{1}{16}\eta^{-1}C^{-1}, t_0]$ and the scalar curvature satisfies*

$$\tilde{R}(y_0, t) \leq 10C \quad \text{for } t \in [t_0 - \frac{1}{16}\eta^{-1}C^{-1}, t_0],$$

or we have

$$\tilde{R}(x_1, t_0) \leq 2D''\tilde{R}(x_2, t_0)$$

for any two points $x_1, x_2 \in \tilde{B}_{t_0}(y_0, A)$, where D'' is the above universal constant.

Based on these three time-extension results, we can adapt the arguments in the proof of Theorem 4.13 on the structure of singularity to the surgically modified solutions.

We next argue as in the second step of the proof of Theorem 4.13 to show that the curvatures of the rescaled solutions $\tilde{g}^{\alpha_m\beta_m}$ at new time zero (after shifting) stay uniformly bounded at bounded distances from \bar{x} as $m \rightarrow \infty$. More precisely, we will prove the following assertion:

Assertion 4. *For the rescaled solutions $\tilde{g}_{ij}^{\alpha_m\beta_m}$, we have that for any $L > 0$, there are constants $C(L) > 0$ and $m(L)$ such that for all $m \geq m(L)$ the rescaled solutions $\tilde{g}_{ij}^{\alpha_m\beta_m}$ satisfy*

- (i) $\tilde{R}(x, 0) \leq C(L)$ for all points x with $\tilde{d}_0(x, \bar{x}) \leq L$;
- (ii) the ball $\tilde{B}_0(\bar{x}, L)$ is defined at least on the time interval $[-\frac{1}{16}\eta^{-1}C(L)^{-1}, 0]$.

For all $\rho > 0$, set

$$M(\rho) = \sup\{\tilde{R}(x, 0) \mid \tilde{d}_0(x, \bar{x}) \leq \rho \text{ in the rescaled solutions } \tilde{g}_{ij}^{\alpha_m\beta_m}\}$$

and

$$\rho_0 = \sup\{\rho > 0 \mid M(\rho) < +\infty\}.$$

Note that the gradient estimate implies that $\rho_0 > 0$. For (i), it suffices to prove $\rho_0 = +\infty$.

Suppose $\rho_0 < +\infty$. By Assertion 3 or Assertion 1, we have for any $0 < \rho < \rho_0$, the rescaled solutions on the balls $\tilde{B}_0(\bar{x}, \rho)$ are defined on the time interval $[-\frac{1}{16}\eta^{-1}M(\rho)^{-1}, 0]$ for all large m . Once the solution is defined on this time interval, by gradient estimate and Shi's derivative estimate, we know that the covariant derivatives of the curvatures of all order on $\tilde{B}_0(\bar{x}, \rho - \frac{1}{m}) \times [-\frac{1}{32}\eta^{-1}M(\rho)^{-1}, 0]$ are also uniformly bounded. Hence Hamilton's compactness theorem is applicable now. Then we can apply the similar argument as in Step 2 of the proof of Theorem 4.13 to prove Assertion 4.

For any subsequence (α_m, β_m) of (α, β) with $r^{\alpha_m} \rightarrow 0$ and $\tilde{\delta}^{\alpha_m\beta_m} \rightarrow 0$ as $m \rightarrow \infty$, by Assertion 4, the κ -noncollapsing and Hamilton's compactness theorem, we can extract a C_{loc}^∞ convergent subsequence of $\tilde{g}_{ij}^{\alpha_m\beta_m}$ over some space-time open subsets containing $t = 0$. As in the proof of Singularity Structure Theorem 4.13, we can use Lemma 4.14 to show *any* such limit has bounded curvature at $t = 0$.

Choose $\alpha_m, \beta_m \rightarrow \infty$ so that $r^{\alpha_m} \rightarrow 0$, $\tilde{\delta}^{\alpha_m\beta_m} \rightarrow 0$, and Assertions 1-3 hold with $\alpha = \alpha_m, \beta = \beta_m$ for all $A \in \{p/q \mid p, q = 1, 2, \dots, m\}$, and $B, C \in \{1, 2, \dots, m\}$. By Assertion 4, we may assume the rescaled solutions $\tilde{g}_{ij}^{\alpha_m\beta_m}$ converge in C_{loc}^∞ topology at the time $t = 0$. Since the curvature of the limit at $t = 0$ is bounded, it follows from Assertion 1 and the choice of the sequence (α_m, β_m) that the limiting $(M_\infty, \tilde{g}_{ij}^\infty(\cdot, t))$ is defined at least on a backward time interval $[-a, 0]$ for some positive constant a and is a smooth solution to the Ricci flow there.

We can further extend the limit backward in time to infinity to get an ancient κ -solution. We omit the details here and refer the reader to consult [14] and [25]. The idea of time extension is used throughout our proof. We emphasize that, comparing to the no surgery case, this is a crucial point.

Summing up, for any $\varepsilon > 0$, there exist nonincreasing positive functions $\tilde{\delta}(t)$ and $\tilde{r}(t)$, defined on $[0, +\infty)$, such that for an arbitrarily given positive function $\delta(t)$ with $\delta(t) < \tilde{\delta}(t)$ on $[0, +\infty)$, the Ricci flow with surgery has a solution on $[0, T_{\max})$ obtained by evolving the Ricci flow and by performing δ -cutoff surgeries at a sequence of times $0 < t_1 < t_2 < \dots < t_i < \dots < T_{\max}$, with $\delta(t_i) \leq \delta \leq \tilde{\delta}(t_i)$ at each time t_i , so that the pinching assumption and the canonical neighborhood assumption (with accuracy ε) with $r = \tilde{r}(t)$ are satisfied. (At this moment we still do not know whether the surgery times t_i are discrete).

Each δ -cutoff surgery at time t_i cuts down the volume at least at an amount depending only on $\delta(t_i)$ and $\tilde{r}(t_i)$, while the volume of the surgically modified solution can be bounded by

$$V(t) \leq V(0)e^{-R_{\min}(0)t}.$$

Thus the surgery times t_i cannot accumulate in any finite interval. When the solution becomes extinct at some finite time T_{\max} , the solution at a time slightly before T_{\max} is entirely covered by canonical neighborhoods and then the initial manifold is diffeomorphic to a connected sum of a finite copies of $\mathbb{S}^2 \times \mathbb{S}^1$ and \mathbb{S}^3/Γ (the metric quotients of round three-sphere). So we have the following long-time existence result, which was proposed by Perelman in [81].

THEOREM 5.4 (Long-time Existence Theorem). *For any given small constant $\varepsilon > 0$, there exist nonincreasing (continuous) positive functions $\tilde{\delta}(t)$ and $\tilde{r}(t)$, defined on $[0, +\infty)$, such that for any arbitrarily given (continuous) positive function $\delta(t)$ with $\delta(t) \leq \tilde{\delta}(t)$ on $[0, +\infty)$, the Ricci flow with surgery, with an arbitrarily given compact orientable normalized three-manifold as initial data, has the following property: either*

- (i) *it is defined on a finite interval $[0, T_{\max})$ and obtained by evolving the Ricci flow and by performing a finite number of cutoff surgeries, with each δ -cutoff at a time $t \in (0, T_{\max})$ having $\delta = \delta(t)$, so that the solution becomes extinct at T_{\max} , and the initial manifold is diffeomorphic to a connected sum of a finite copies of $\mathbb{S}^2 \times \mathbb{S}^1$ and \mathbb{S}^3/Γ (the metric quotients of round three-sphere) ; or*
- (ii) *it is defined on $[0, +\infty)$ and obtained by evolving the Ricci flow and by performing at most a countably many cutoff surgeries, with each δ -cutoff at a time $t \in [0, +\infty)$ having $\delta = \delta(t)$, so that the pinching assumption and the canonical neighborhood assumption (with accuracy ε) with $r = \tilde{r}(t)$ are satisfied, and there exist at most a finite number of surgeries on every finite time interval.*

5.4. Topological Implications. As the first consequence of the above long-time existence result, one can obtain a complete classification to compact three-manifolds with nonnegative scalar curvature. Indeed, if the initial manifold has positive scalar curvature, then the solution becomes extinct in finite time and the manifold is either flat or diffeomorphic to a connected sum of a finite number of copies of $\mathbb{S}^2 \times \mathbb{S}^1$ and \mathbb{S}^3/Γ (the metric quotients of round three-spheres). This improves the well-known topological classification of Schoen-Yau [89] for compact three-manifolds with nonnegative scalar curvature.

The famous **Poincaré conjecture** states that *any simply connected compact three-manifold (without boundary) is homeomorphic to the three-sphere*. Recent works of Perelman [82] (see also detailed exposition given in Morgan-Tian [70]) and Colding-Minicozzi [30] proved that any surgically modified solution to the Ricci flow on a simply connected compact three-manifold must be extinct in finite time. Thus *The combination of the assertion (i) of the above long-time theorem 5.4 and the finite extinction result of Perelman and Colding-Minicozzi gives a complete proof to the Poincaré conjecture.*

The idea of proving the extinction result is adapted from Hamilton [49] where the argument was used to show the incompressibility of the boundary tori of hyperbolic pieces. This argument for the finite time extinction can be roughly described as follows. Suppose there would exist a surgical solution $g_{ij}(t)$ on the infinite time interval $[0, +\infty)$. Since the manifold is simply connected, a well-known result of J. P. Serre implies some higher homotopic group of the manifold is nontrivial. Then one can use the nontrivial homotopic group to construct a minimal surface $\Sigma(t)$ for each $t \in [0, +\infty)$. Denote by $A(t)$ the area of $\Sigma(t)$. By an argument of Schoen-Yau [89] of using the Gauss-Bonnet formula, one can bound the the time derivative of the area function

$$\frac{dA(t)}{dt} \leq -f(t)$$

for some positive function $f(t)$ which is nonintegrable on $[0, +\infty)$. Then it gives the desired contradiction.

To conclude this section, we mention the application of the Ricci flow with surgery to the classification of four-manifolds with positive isotropic curvature. Recall that a Riemannian four-manifold is said to have *positive isotropic curvature* if for every orthonormal four-frame the curvature tensor satisfies

$$R_{1313} + R_{1414} + R_{2323} + R_{2424} > 2R_{1234}.$$

An *incompressible space form* N^3 in a four-manifold M^4 is a three-dimensional submanifold diffeomorphic to \mathbb{S}^3/Γ (the quotient of the three-sphere by a group of isometries without fixed point) such that the fundamental group $\pi_1(N^3)$ injects into $\pi_1(M^4)$. The space form is said to be *essential* unless $\Gamma = \{1\}$, or $\Gamma = \mathbb{Z}_2$ and the normal bundle is non-orientable. In [25],

the last two authors obtained the following long-time existence result for the Ricci flow with surgery for a class of four-manifolds.

THEOREM 5.5 (Chen-Zhu [25]). *Let M^4 be a compact four-manifold with no essential incompressible space-form and with a metric g_{ij} of positive isotropic curvature. Then we have a finite collection of smooth solutions $g_{ij}^{(k)}(t)$, $k=0, 1, \dots, m$, to the Ricci flow, defined on $M_k^4 \times [t_k, t_{k+1})$, ($0=t_0 < \dots < t_{m+1}$) with $M_0^4 = M^4$ and $g_{ij}^{(0)}(t_0) = g_{ij}$, which go singular as $t \rightarrow t_{k+1}$, such that the following properties hold:*

- (i) *for each $k=0, 1, \dots, m-1$, the compact (possibly disconnected) four-manifold M_k^4 contains an open set Ω_k such that the solution $g_{ij}^{(k)}(t)$ can be smoothly extended to $t=t_{k+1}$ over Ω_k ;*
- (ii) *for each $k=0, 1, \dots, m-1$, $(\Omega_k, g_{ij}^{(k)}(t_{k+1}))$ and $(M_{k+1}^4, g_{ij}^{(k+1)}(t_{k+1}))$ contain compact (possibly disconnected) four-dimensional submanifolds with smooth boundary, which are isometric and then can be denoted by N_k^4 ;*
- (iii) *for each $k=0, 1, \dots, m-1$, $M_k^4 \setminus N_k^4$ consists of a finite number of disjoint pieces diffeomorphic to $S^3 \times \mathbb{I}$, \mathbb{B}^4 or $\mathbb{RP}^4 \setminus \mathbb{B}^4$, while $M_{k+1}^4 \setminus N_k^4$ consists of a finite number of disjoint pieces diffeomorphic to \mathbb{B}^4 ;*
- (iv) *for $k=m$, M_m^4 is diffeomorphic to the disjoint union of a finite number of S^4 , or \mathbb{RP}^4 , or $S^3 \times S^1$, or $S^3 \tilde{\times} S^1$, or $\mathbb{RP}^4 \# \mathbb{RP}^4$.*

As a direct consequence, it gives a complete proof to the following classification result of Hamilton [48].

COROLLARY 5.6. *A compact four-manifold with no essential incompressible space-form and with a metric of positive isotropic curvature is diffeomorphic to S^4 , or \mathbb{RP}^4 , or $S^3 \times S^1$, or $S^3 \tilde{\times} S^1$, or a connected sum of them.*

6. Geometrization of Three-manifolds

In the late 70's and early 80's, Thurston [95, 96, 97] proved a number of remarkable results on the existence of geometric structures on three-manifolds, especially the celebrated Haken manifold theorem. These results motivated him to formulate a profound conjecture

Thurston's Geometrization Conjecture *Let M be a compact, orientable and prime three-manifold. Then there is an embedding of a finite number of disjoint unions, possibly empty, of incompressible two-tori $\coprod_i T_i^2 \subset M$ such that every component of the complement admits a locally homogeneous Riemannian metric of finite volume.*

In dimension three, every locally homogeneous manifold with finite volume is modeled on one of the following eight homogeneous manifolds (see for example Theorem 3.8.4 of [98]):

- (1) \mathbb{S}^3 , the round three-sphere;
- (2) \mathbb{R}^3 , the Euclidean space;
- (3) \mathbb{H}^3 , the standard hyperbolic space;
- (4) $\mathbb{S}^2 \times \mathbb{R}$;
- (5) $\mathbb{H}^2 \times \mathbb{R}$;
- (6) Nil , the three-dimensional nilpotent Heisenberg group (consisting of upper triangular 3×3 matrices with diagonal entries 1);
- (7) $\widetilde{PSL}(2, \mathbb{R})$, the universal cover of the unit sphere bundle of \mathbb{H}^2 ;
- (8) Sol , the three-dimensional solvable Lie group.

According to Kneser [58] and Milnor [68], every compact orientable three-manifold admits a unique decomposition as a finite connected sum of orientable prime three-manifolds. Thus the geometrization conjecture is a complete classification to three-dimensional manifolds. In particular, the Poincaré conjecture can be deduced from Thurston's geometrization conjecture. Indeed, suppose that we have a compact simply connected three-manifold that satisfies the conclusion of the geometrization conjecture. If it were not diffeomorphic to the three-sphere \mathbb{S}^3 , there would be a prime factor in the prime decomposition of the manifold. Since the prime factor still has a vanishing fundamental group, the torus decomposition (by Jaco-Shalen [53] and Johannsen [54]) of the prime factor in the geometrization conjecture must be trivial. Thus the prime factor is a compact homogeneous manifold model. From the list of above eight models, we see that the only compact three-dimensional model is \mathbb{S}^3 . This is a contradiction. Consequently, the compact simply connected three-manifold is diffeomorphic to \mathbb{S}^3 .

The approach to prove the geometrization conjecture via the Ricci flow is to analyze long time behavior of surgically modified solutions. The argument of Perelman in [81] for the long-time behavior of surgical solutions is basically along the line given by Hamilton [49], in which Hamilton obtained the geometrization for a special class of solutions to the Ricci flow on three-manifolds, the so called nonsingular solutions. In the following, we present the long-time behavior analysis and sketch the proof of Thurston's geometrization conjecture.

Since we already have a complete (topological) classification to compact three-manifolds with nonnegative scalar curvature, we now assume that our initial manifold does not admit any metric with nonnegative scalar curvature and that *once we get a compact component with nonnegative scalar curvature, it is immediately removed*. Also by Theorem 5.4 (i), we only need to consider those solutions to the Ricci flow with surgery which exist for all time $t \geq 0$.

Let $g_{ij}(t)$, $0 \leq t < +\infty$, be a solution to the Ricci flow with δ -cutoff surgeries, constructed by the above long-time existence theorem (Theorem

5.2) with normalized initial data. Let $0 = t_0 < t_1 < t_2 < \dots < t_k < \dots$ be the surgery times. On each time interval (t_{k-1}, t_k) , the minimum of the scalar curvature $R_{\min}(t)$ at time t satisfies the differential inequality

$$\frac{d}{dt}R_{\min}(t) \geq \frac{2}{3}R_{\min}^2(t)$$

for $t \in (t_{k-1}, t_k)$, $k = 1, 2, \dots$. Since the surgeries occur only on regions of very large scalar curvature, it follows that

$$R_{\min}(t) \geq -\frac{3}{2} \cdot \frac{1}{t + \frac{3}{2}}$$

for all $t \geq 0$. Meanwhile, on each time interval (t_{k-1}, t_k) , the volume V satisfies the evolution equation

$$\frac{d}{dt}V = - \int R dV$$

and hence

$$\frac{d}{dt}V \leq \frac{3}{2} \cdot \frac{1}{(t + \frac{3}{2})}V.$$

Since the cutoff surgeries do not increase volume, the function $V(t)(t + \frac{3}{2})^{-\frac{3}{2}}$ is nonincreasing on $[0, +\infty)$. and there holds

$$\begin{aligned} \frac{V(t)}{(t + \frac{3}{2})^{\frac{3}{2}}} &\leq \frac{V(0)}{(\frac{3}{2})^{\frac{3}{2}}} \exp \left\{ - \int_0^t \left(R_{\min}(t) + \frac{3}{2(t + \frac{3}{2})} \right) dt \right. \\ &\quad \left. - \int_0^t \frac{1}{V} \int_M (R - R_{\min}(t)) dV dt \right\} \end{aligned}$$

for all $t > 0$. This inequality implies that whenever we have a rescaling limit along a sequence times $t^\alpha \rightarrow \infty$ and with factors $(t^\alpha)^{-1}$, the limit must be a hyperbolic manifold. Then by extending the elliptic type estimate in Proposition 4.11 to surgical solutions, one will be able to obtain the following important thick-thin decomposition (see Figure 8) for surgically modified solutions.

THEOREM 6.1 (Thick-thin decomposition theorem). *For any $w > 0$ and $0 < \varepsilon \leq \frac{1}{2}w$, there exists a positive constant $\rho = \rho(w, \varepsilon) \leq 1$ with the following property. Suppose $g_{ij}(t)$, $t \in [0, +\infty)$, is a surgically modified solution constructed by the above long-time existence theorem. Then for any arbitrarily fixed $\xi > 0$, for t large enough, the manifold M_t at time t admits a decomposition*

$$M_t = M_{thin}(w, t) \cup M_{thick}(w, t)$$

with the following properties:

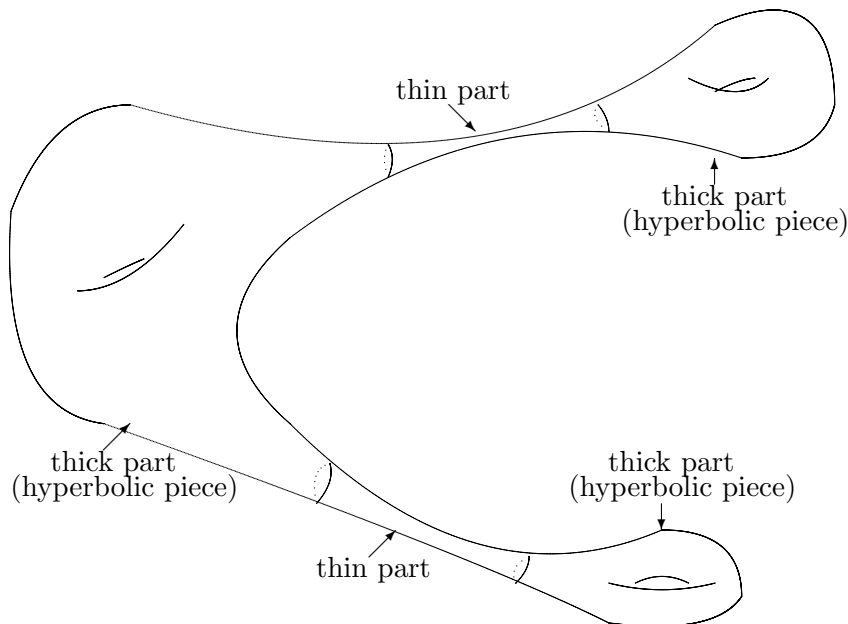


FIGURE 8. Thick-thin decomposition.

- (a) For every $x \in M_{thin}(w, t)$, there exists some $r = r(x, t) > 0$, with $0 < r\sqrt{t} < \rho\sqrt{t}$, such that

$$Rm \geq -(r\sqrt{t})^{-2} \text{ on } B_t(x, r\sqrt{t}), \text{ and}$$

$$Vol_t(B_t(x, r\sqrt{t})) < w(r\sqrt{t})^3;$$

- (b) For every $x \in M_{thick}(w, t)$, we have

$$|2tR_{ij} + g_{ij}| < \xi \text{ on } B_t(x, \rho\sqrt{t}), \text{ and}$$

$$Vol_t(B_t(x, \rho\sqrt{t})) \geq \frac{1}{10}w(\rho\sqrt{t})^3.$$

Moreover, if we take any sequence $t^\alpha \rightarrow +\infty$ and points $x^\alpha \in M_{thick}(w, t^\alpha)$, then a subsequence of the rescaled metrics of $g_{ij}(t^\alpha)$ around x^α with factor $(t^\alpha)^{-1}$ converge smoothly to a complete hyperbolic manifold of finite volume with constant sectional curvature $-\frac{1}{4}$.

This thick-thin decomposition theorem was implicitly given by Perelman in [81] without any restriction on the parameters ε and w . The above weaker version with the restriction is taken from [14]. The difference is because of our difficulty in understanding the original argument of Perelman [81].

Fortunately, this weaker version still allows us to complete the proof of the geometrization conjecture.

Without loss of generality, we may further assume that the initial manifold M is irreducible. By our surgery procedures, the solution manifold M_t at each time t consists of a finite number of components where one of the components, called the *essential component* and denoted by $M_t^{(1)}$, is diffeomorphic to the initial manifold M while the rest are diffeomorphic to the 3-sphere \mathbb{S}^3 .

Based on the thick-thin decomposition theorem, we can modify the arguments of Hamilton in [49] to obtain the following long-time behavior result.

THEOREM 6.2 (Long-time Behavior Theorem). *Let $w > 0$ and $0 < \varepsilon \leq \frac{1}{2}w$ be any small positive constants and let $(M_t, g_{ij}(t)), 0 < t < +\infty$, be a solution to the Ricci flow with surgery constructed by the long-time existence theorem. Then one of the following holds: either*

- (i) *for all sufficiently large t , we have $M_t = M_{thin}(w, t)$; or*
- (ii) *there exists a sequence of times $t^\alpha \rightarrow +\infty$ such that the scalings of the essential component $(M_{t^\alpha}^{(1)}, g_{ij}(t^\alpha))$, with factor $(t^\alpha)^{-1}$, converge in the C^∞ topology to a hyperbolic metric on the initial compact manifold M with constant sectional curvature $-\frac{1}{4}$; or*
- (iii) *we can find a finite collection of complete noncompact hyperbolic three-manifolds $\mathcal{H}_1, \dots, \mathcal{H}_m$, of finite volume, and compact subsets K_1, \dots, K_m of $\mathcal{H}_1, \dots, \mathcal{H}_m$ respectively obtained by truncating each cusp of the hyperbolic manifolds along constant mean curvature torus of small area, and for all t beyond some time $T < +\infty$ we can find diffeomorphisms $\varphi_l, 1 \leq l \leq m$, of K_l into M_t so that as long as t is sufficiently large, the metric $t^{-1}\varphi_l^*(t)g_{ij}(t)$ is as close to the hyperbolic metric as we like on the compact sets K_1, \dots, K_m ; moreover, the complement $M_t \setminus (\varphi_1(K_1) \cup \dots \cup \varphi_m(K_m))$ is contained in the thin part $M_{thin}(w, t)$, and the boundary tori of each K_l are incompressible in the sense that each φ_l injects $\pi_1(\partial K_l)$ into $\pi_1(M_t)$.*

If case (ii) holds, then it is clear that the initial manifold M is geometrizable. While if case (iii) holds, then it follows from Thurston’s theorem on Haken manifolds that the initial manifold M is also geometrizable. Thus it remains to consider case (i). For case (i), we will appeal to the following collapsing result, which was first announced (in a more general version) by Perelman [81] and proved by Shioya-Yamaguchi [92].

THEOREM 6.3 (Collapsing Theorem). *Suppose $(M^\alpha, g_{ij}^\alpha)$ is a sequence of compact orientable three-manifolds without boundary, and $w^\alpha \rightarrow 0$. Assume that for each point $x \in M^\alpha$ there exists a radius $\rho = \rho^\alpha(x)$, not exceeding the diameter of M^α , such that the volume of the ball $B(x, \rho)$ in the metric g_{ij}^α is*

at most $w^\alpha \rho^3$ and the sectional curvature on $B(x, \rho)$ is at least $-\rho^{-2}$. Then M^α , for sufficiently large α , are diffeomorphic to graph manifolds.

Now the proof of the Thurston geometrization conjecture can be completed in the following way.

Let M be a compact, orientable and irreducible three-manifold. Arbitrarily given a (normalized) Riemannian metric on the manifold M , we use it as initial data for the Ricci flow. Take an arbitrarily sequence of small positive constants $w^\alpha \rightarrow 0$ as $\alpha \rightarrow +\infty$. For each fixed α , we set $\varepsilon = w^\alpha/2 > 0$. Then we apply the long-time existence theorem (Theorem 5.2) to get a sequence of surgically modified solutions $(M_t^\alpha, g_{ij}^\alpha(t))$ on maximal time intervals $[0, T_{\max}^\alpha)$ satisfying the pinching assumption and the canonical assumption (with the accuracy parameter $\varepsilon = w^\alpha/2$). We may assume that the maximal time $T_{\max}^\alpha = +\infty$ for all α and the surgical solutions $(M_t^\alpha, g_{ij}^\alpha(t))$ always satisfy assertion (i) of the long-time behavior theorem. That is, for each α , $M_t^\alpha = M_{thin}(w^\alpha, t)$ when t is sufficiently large.

Clearly we only need to consider the essential component $(M_t^\alpha)^{(1)}$. We divide the discussion into two cases:

- (1) there is a positive constant $1 < C < +\infty$ such that for each α there is a sufficiently large time $t_\alpha > 0$ such that

$$r(x, t_\alpha)\sqrt{t_\alpha} < C \cdot \text{diam}((M_{t_\alpha}^\alpha)^{(1)})$$

for all $x \in (M_{t_\alpha}^\alpha)^{(1)} \subset M_{thin}(w^\alpha, t_\alpha)$;

- (2) there are a subsequence α_k and sequences of positive constants $C_k \rightarrow +\infty$ and $T_k < +\infty$ such that for each $t \geq T_k$, we have

$$r(x(t), t)\sqrt{t} \geq C_k \cdot \text{diam}((M_t^{\alpha_k})^{(1)})$$

for some $x(t) \in (M_t^{\alpha_k})^{(1)}$, $k = 1, 2, \dots$

In case (1), we apply the collapsing theorem of Shioya-Yamaguchi to conclude that $(M_{t_\alpha}^\alpha)^{(1)}$ are graph manifolds. So the initial manifold M is geometrizable.

For case (2), if there are subsequences α_k (still denoted by α_k) and $t_k \in (T_k, +\infty)$ such that

$$\text{Vol}_{t_k}((M_{t_k}^{\alpha_k})^{(1)}) < w'_k(\text{diam}((M_{t_k}^{\alpha_k})^{(1)}))^3$$

for some sequence $w'_k \rightarrow 0$, then we can apply the collapsing theorem again to conclude that $(M_{t_k}^{\alpha_k})^{(1)}$ are diffeomorphic to graph manifolds. On the other hand, if there is a positive constant w' such that

$$\text{Vol}_t((M_t^{\alpha_k})^{(1)}) \geq w'(\text{diam}((M_t^{\alpha_k})^{(1)}))^3$$

for each k and all $t \geq T_k$, we can obtain the curvature estimates and take a rescaling limit to conclude that the initial manifold M is diffeomorphic to a flat manifold. So the initial manifold M is also geometrizable in case (2).

Therefore, we see that the Thurston geometrization conjecture is true.

References

- [1] Andrews, B. and Nguyen, H., *Four-manifolds with 1/4-pinched flag curvatures*, Preprint, (2007).
- [2] Angenent, S. and Knopf, D., *An example of neckpinching for Ricci flow on S^{n+1}* , Math. Res. Lett. **11** (2004), no. 4, 493–518.
- [3] Bando, S., *On the classification of three-dimensional compact Kähler manifolds of nonnegative bisectional curvature*, J. Differential Geom. **19** (1984), no. 2, 283–297.
- [4] Berger, M., *Les variétés Riemanniennes 1/4-pincées*, Ann. Scuola Norm. Sup. Pisa, **14** (1960), 161–170.
- [5] Brendle, S., *A generalization of Hamilton’s Harnack inequality for the Ricci flow*, preprint, arXiv:math.DG/0707.2192
- [6] Brendle, S. and Schoen, R., *Manifolds with 1/4-pinched curvature are space forms*, preprint, arXiv:math.DG/0705.0766.
- [7] Brendle, S. and Schoen, R., *Classification of manifolds with weakly 1/4-pinched curvatures*, preprint, arXiv:math.DG/0705.3963.
- [8] Böhm, C., and Wilking, B., *Manifolds with positive curvature operator are space forms*, preprint, arxiv:math.DG/0606187.
- [9] Cao, H.-D., *On Harnack’s inequalities for the Kähler-Ricci flow*, Invent. Math. **109** (1992), no. 2, 247–263.
- [10] Cao, H.-D., *Limits of Solutions to the Kähler-Ricci flow*, J. Differential Geom., **45** (1997), 257–272.
- [11] Cao, H.-D. and Chow, B., *Compact Kähler manifolds with nonnegative curvature operator*, Invent. Math. **83** (1986), 553–556.
- [12] Cao, H.-D., Chow, B., Chu, S.-C., and Yau, S.-T. (editors), *Collected Papers on Ricci Flow, Series in Geometry and Topology*, **37**. Cambridge, MA: International Press, 2003.
- [13] Cao, H.-D., Hamilton, R. S., and Ilmanen, T., *Gaussian densities and stability for some Ricci solitons*, preprint, arXiv:math/0404165v1
- [14] Cao, H.-D. and Zhu, X.-P., *A complete proof of Poincare and geometrization conjectures—application of Hamilton-Perelman theory of Ricci flow*, Asian J. math. **102** (2006), 165–492.
- [15] Cao, X., *First Eigenvalues of Geometric Operators under the Ricci Flow*, preprint, arXiv:0710.3947.
- [16] Chau, A. and Tam, L. F., *On the complex structure of Kähler manifolds with non-negative curvature*, J. Differential Geom., **73** (2006), 491–530.
- [17] Chau, A. and Tam, L. F., *A survey on the Kähler-Ricci flow and Yau’s uniformization conjecture*, preprint, arXiv: math. DG/0702257.
- [18] Cheeger, J. and Ebin, D., *Comparison theorems in Riemannian geometry*, North-Holland (1975).
- [19] Chen, B. L., *Strong uniqueness of the Ricci flow*, preprint, arXiv:math.DG/0706.3081.
- [20] Chen, B. L., Tang, S. H. and Zhu, X. P., *A uniformization theorem of complete noncompact Kähler surfaces with positive bisectional curvature*, J. Diff. Geometry **67** (2004) 519–570.
- [21] Chen, B. L. and Zhu, X. P., *Complete Riemannian manifolds with pointwise pinched curvature*, Invent. Math. **140** (2000), no. 2, 423–452.
- [22] Chen, B. L. and Zhu, X. P., *On complete noncompact Kähler manifolds with positive bisectional curvature*, Math. Ann. **327** (2003) 1–23.
- [23] Chen, B. L. and Zhu, X. P., *Volume growth and curvature decay of positively curved Kähler manifolds*, Quarterly Journal of Pure and Applied Mathematics Vol. **1**. no. 1 68–108 (2005).

- [24] Chen, B. L. and Zhu, X. P., *Uniqueness of the Ricci flow on complete noncompact manifolds*, J. Diff. Geom. **74** (2006), 119–154.
- [25] Chen, B. L. and Zhu, X. P., *Ricci flow with Surgery on four-manifolds with positive isotropic curvature*, J. Diff. Geom. **74** (2006), 177–264.
- [26] Chen, H., *Pointwise quater-pinched 4 manifolds*, Ann. Global Anal. Geom. **9** (1991), 161–176.
- [27] Chow, B. and Chu, S. C., *A geometric interpretation of Hamilton's Harnack inequality for the Ricci flow*, Math. Res. Lett. **2** (1995) 701–718.
- [28] Chow, B. and Knopf, D., *The Ricci flow: An introduction*, Mathematical Surveys and Monographs, Amer. Math. Soc., Providence, RI, 2004.
- [29] Chow, B., Lu, P. and Ni, L., *Hamilton's Ricci Flow*, Lectures in Contemporary Mathematics, Amer. Math. Soc., Providence, RI, 2006.
- [30] Colding, T. H. and Minicozzi, W., P.II, *Estimates for the extinction time for the Ricci flow on certain 3-manifolds and a question of Perelman*, J. Amer. Math. Soc. **18** (2005), no. 3, 561–569.
- [31] Daskalopoulos, P. and Hamilton, R. S., *Geometric estimates for the logarithmic fast diffusion equation*, Communications in Analysis and Geometry, vol. 12, no. 1–2, pp. 143–164, 2004.
- [32] De Turck, D., *Deforming metrics in the direction of their Ricci tensors* J. Diff. Geom. **18** (1983), 157–162.
- [33] Feldman, M., Ilmanen, T., and Knopf, D., *Rotationally symmetric shrinking and expanding gradient Kähler-Ricci solitons*, J. Differential Geom. **65** (2003), 169–209.
- [34] Feldman, M., Ilmanen, T., and Ni, L., *Entropy and reduced distance for Ricci expanders*, J. Geom. Anal. **15** (2005), no. 1, 49–62.
- [35] Greene, R. E. and Wu, H., *Lipschitz convergence of Riemannian manifolds*, Pacific J. Math. **131** (1988), 119–141.
- [36] Gromoll, D., *Differenzierbare Strukturen und Metriken positiver Krümmung auf Sphären*, Math. Ann., **164**, (1966), 353–371.
- [37] Gromoll, D. and Meyer, W., *On complete open manifolds of positive curvature*, Ann. of Math., **90** (1969), 95–90.
- [38] Gromov, M., *Metric Structures for Riemannian and Non-Riemannian Spaces*, Edited by J.LaFontaine and P.Pansu English translation by Sean Michael Bates, 1998, Birkäuser.
- [39] Gu, H. L., *A simple proof for the generalized Frankel conjecture*, preprint, arXiv:math.DG/0707.0035.
- [40] Gu, H. L. and Zhu, X. P., *The Existence of Type II Singularities for the Ricci Flow on S^{n+1}* , preprint, arXiv:math.DG/0707.0033.
- [41] Hamilton, R. S., *Three manifolds with positive Ricci curvature*, J. Diff. Geom. **17** (1982), 255–306.
- [42] Hamilton, R. S., *Four-manifolds with positive curvature operator*, J. Differential Geom. **24** (1986), 153–179.
- [43] Hamilton, R. S., *The Ricci flow on surfaces*, Contemporary Mathematics **71** (1988) 237–261.
- [44] Hamilton, R. S., *The Harnack estimate for the Ricci flow*, J. Diff. Geom. **37** (1993), 225–243.
- [45] Hamilton, R. S., *Eternal solutions to the Ricci flow*, J. Diff. Geom. **38** (1993), no. 1, 1–11.
- [46] Hamilton, R. S., *A compactness property for solution of the Ricci flow*, Amer. J. Math. **117** (1995), 545–572.
- [47] Hamilton, R. S., *The formation of singularities in the Ricci flow*, Surveys in Differential Geometry (Cambridge, MA, 1993), **2**, 7–136, International Press, Combridge, MA, 1995.

- [48] Hamilton, R. S., *Four manifolds with positive isotropic curvature*, Commu. in Analysis and Geometry, **5** (1997), 1–92.
- [49] Hamilton, R. S., *Non-singular solutions to the Ricci flow on three manifolds*, Commu. Anal. Geom. **1** (1999), 695–729.
- [50] Hamilton, R. S., *A matrix Harnack estimate for the heat equation*, Commu. Anal. Geom. **1** (1993), 88–99.
- [51] Huisken, G., *Ricci deformation of the metric on a Riemannian manifold* J. Differential Geom. **21** (1985), 47–62.
- [52] Ivey, T., *Ricci solitons on compact three-manifolds*, Diff. Geom. Appl. **3** (1993), 301–307.
- [53] Jaco, W. and Shalen, P. *Seifert fibered spaces in 3-manifolds*, Mem. Amer. Math. Soc., **21** (1979), no. 220, viii+192 pp.
- [54] Johannson, K., *Homotopy equivalences of 3-manifolds with boundaries*, Lecture Notes in Mathematics, **761** (1979), Springer, Berlin, ii+303 pp.
- [55] Karcher, H., *Pinching implies strong pinching*, Comment. Math. Hevl. **46** (1971), 124–126.
- [56] Kleiner, B. and Lott, J., *Note on Perelman’s paper*, <http://www.math.lsa.umich.edu/research/ricciflow/perelman.html>.
- [57] Klingenberg, W., *Über Riemannsche Mannigfaltigkeiten mit positiver Krümmung*, Comment. Math. Helv. **35** (1961), 47–54.
- [58] Kneser, H., *Geschlossene Flächen in dreidimensionalen mannigfaltigkeiten*, Jahresbericht der Deut. Math. Verein., **38** (1929), 248–260.
- [59] Kotschwar, B., *On rotationally invariant shrinking gradient Ricci solitons*, preprint, arXiv:math/0702597.
- [60] Li, J.-F., *Eigenvalues and energy functionals with monotonicity formulae under Ricci flow*, Math. Ann. **338** (2007), no. 4, x927–946.
- [61] Li, P. and Yau, S. T., *On the parabolic kernel of the Schrödinger operator*, Acta Math. **156** (1986), no. 3–4, 153–201.
- [62] Ling, J., *Some Asymptotic Behavior of the first Eigenvalue along the Ricci Flow*, arXiv:0710.4326v1.
- [63] Lu, P. and Tian, G., *The uniqueness of standard solutions in the work of Perelman*, Available at www.math.lsa.umich.edu/lott/ricciflow/StanUniqWork2.pdf, 2005.
- [64] Ma, L., *Eigenvalue monotonicity for the Ricci-Hamilton flow*, Ann. Global Anal. Geom. **29** (2006), no. 3, 287–292.
- [65] Margerin, C., *Une caractrisation optimale de la structure différentielle standard de la sphère en terme de courbure pour (presque) toutes les dimensions. I. Les noncs*, (French) [A sharp curvature-type characterization of the smooth sphere in all but a few dimensions. I. Statement of the theorems] C. R. Acad. Sci. Paris Sr. I Math. **319** (1994), no. 6, 605–607.
- [66] Margerin, C., *A sharp characterization of the smooth 4-sphere in curvature terms*, Comm. Anal. Geom. **6** (1998), no. 1, 21–65.
- [67] Micallef, M. and Moore, J. D., *Minimal two-spheres and the topology of manifolds with positive curvature on totally isotropic two-planes*, Ann. of Math. (2) **127** (1988) 199–227.
- [68] Milnor, J., *A unique factorization theorem for 3-manifolds*, Amer. J. Math., **84** (1962), 1–7.
- [69] Mok, N., *The uniformization theorem for compact Kähler manifolds of nonnegative holomorphic bisectional curvature*, J. Differential Geom. **27** (1988), 179–214.
- [70] Morgan, J. and Tian, G., *Ricci flow and the Poincaré conjecture*, preprint, arXiv:math.DG/0605667.
- [71] Mori, S., *Projective manifolds with ample tangent bundles*, Ann. of Math. **100** (1979), 593–606.

- [72] Naber, A., *Noncompact Shrinking 4-Solitons with Nonnegative Curvature*, preprint, arXiv:math.DG/07105579.
- [73] Nguyen, H., Ph.D. thesis, the Australian National University.
- [74] Ni, L., *Ancient solution to Kahler-Ricci flow*, Math. Res. Lett. **12** (2005), 633–653.
- [75] Ni, L., *A note on Perelman's LYH inequality*, Comm. Anal. Geom. **14** (2006), 883–905.
- [76] Ni, L. and Tam, L. F., *Plurisubharmonic functions and the structure of complete Kähler manifolds with nonnegative curvature*, J. Differential Geom. **64** (2003), no. 3, 457–524.
- [77] Ni, L. and Wallach, N., *On a classification of the gradient shrinking solitons*, preprint, arXiv:math.DG/07103194
- [78] Ni, L. and Wallach, N., *On 4-dimensional gradient shrinking solitons*, arXiv:math.DG/07103195
- [79] Nishikawa, S., *Deformation of Riemannian metrics and manifolds with bounded curvature ratios*, Geometric measure theory and the calculus of variations (Arcata, Calif., 1984), 343–352, Proc. Sympos. Pure Math., 44, Amer. Math. Soc., Providence, RI, 1986
- [80] Perelman, G., *The entropy formula for the Ricci flow and its geometric applications*, arXiv:math.DG/0211159.
- [81] Perelman, G., *Ricci flow with surgery on three manifolds*, arXiv: math. DG/0303109.
- [82] Perelman, G., *Finite extinction time to the solutions to the Ricci flow on certain three manifolds*, arXiv:math.DG/0307245.
- [83] Peters, S., *Convergence of Riemannian manifolds*, Comp. Math. **62** (1987), 3–16.
- [84] Petersen, P. and Wylie, P., *Rigidity of gradient Ricci Solitons*, arXiv:math.DG/07103174
- [85] Petersen, P. and Wylie, P., *On gradient Ricci solitons with Symmetry*, arXiv:math.DG/07103595
- [86] Petersen, P. and Wylie, P., *On the classification of gradient Ricci solitons*, arXiv: math.DG/07121298
- [87] Rauch, H. E., *A contribution to differential geometry in the large*, Ann. of Math. **54**(1951), 38–55.
- [88] Ruh, E., *Curvature and differentiable structure on spheres*, Comment. Math. Hevl. 46, (1971), 127–136.
- [89] Schoen, R. and Yau, S. T., *Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature*, Ann. of Math. (2) **110** (1979), no. 1, 127–142.
- [90] Sesum, N., *Limiting behaviour of the Ricci flow*, arXiv:math.DG/0402194.
- [91] Shi, W. X., *Deforming the metric on complete Riemannian manifold*, J. Differential Geometry **30** (1989), 223–301.
- [92] Shioya, T. and Yamaguchi, T., *Volume collapsed three-manifolds with a lower curvature bound*, Math. Ann. **333** (2005), 131–155.
- [93] Simon, M., *A class of Riemannian manifolds that pinch when evolved by Ricci flow*, Manuscripta Math. **101** (2000), no. 1, 89–114.
- [94] Sugimoto, M. and Shiohama, K., *On the differentiable pinching problem*, Math. Ann. **195**, (1971), 1–16.
- [95] Thurston, W. P., *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 3, 357–381.
- [96] Thurston, W. P., *Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical manifolds*, Ann. of Math. (2) **124** (1986), no. 2, 203–246.
- [97] Thurston, W. P., *Hyperbolic structures on 3-manifolds II: Surface groups and 3-manifolds which fiber over the circle*, preprint, 1988.

- [98] Thurston, W. P., *Three-dimensional geometry and topology*, Vol. 1. Edited by Silvio Levy. Princeton Mathematical Series, 35. Princeton University Press, Princeton, NJ, 1997.
- [99] Topping, P., *Lectures on the Ricci flow*, L.M.S. Cambridge University Press, 2006.
- [100] Siu, Y. T. and Yau, S. T., *Complex Kähler manifolds of positive bisectional curvature*, *Invent. Math.* **59** (1980), 189–204.
- [101] Yau, S. T., *Problem Section in Seminar on Differential Geometry*, Edited by S. T. Yau, Princeton University Press, 1982.
- [102] Ye, R., *The logarithmic Sobolev inequality along the Ricci flow*, arXiv:math.DG/0707.2424
- [103] Ye, R., *The logarithmic Sobolev inequality along the Ricci flow: the case $\lambda_0(g_0) = 0$* , arXiv:math.DG/0708.2005.
- [104] Ye, R., *The Log Entropy Functional Along the Ricci Flow*, arXiv:math.DG/0708.2008.
- [105] Ye, R., *Sobolev Inequalities, Riesz Transforms and the Ricci Flow*, arXiv:math.DG/0709.0512.
- [106] Ye, R., *Entropy Functionals, Sobolev Inequalities and kappa-Noncollapsing Estimates along the Ricci Flow*, arXiv:math.DG/0709.2724
- [107] Zhang, Q., *A uniform Sobolev inequality under Ricci flow*, arXiv:math.DG/07061594.
- [108] Zhang, Q., *Strong non-collapsing and uniform Sobolev inequalities for Ricci flow with surgeries*, arXiv:math.DG/07121329