Acta Math., 228 (2022), 1–215
DOI: 10.4310/ACTA.2022.v228.n1.a1

Uniqueness and stability of Ricci flow through singularities

by

Richard H. Bamler
University of California, Berkeley
Berkeley, CA, U.S.A.

Bruce Kleiner
New York University
New York, NY, U.S.A.

Contents

The first author was supported by a Sloan Research Fellowship and NSF grant DMS-1611906. The research was in part conducted while the first author was in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2016 semester. The second author was supported by NSF grants DMS-1405899 and DMS-1406394, and a Simons Collaboration grant.

1. Introduction

1.1. Overview

The understanding of many aspects of Ricci flow has advanced dramatically in the last fifteen years. This has led to numerous applications, the most notable being Perelman's landmark proof of the geometrization and Poincaré conjectures. Nonetheless, from an analytical viewpoint, a number of fundamental questions remain, even for 3-dimensional Ricci flow. One of these concerns the nature of Ricci flow with surgery, a modification of Ricci flow that was central to Perelman's proof. Surgery, an idea initially developed by Hamilton, removes singularities as they form, allowing one to continue the flow. While Perelman's construction of Ricci flow with surgery was spectacularly successful, it is not entirely satisfying due to its ad hoc character and the fact that it depends on a number of non-canonical choices. Furthermore, from a PDE viewpoint, Ricci flow with surgery does not provide a theory of solutions to the Ricci flow PDE itself, since surgery violates the equation. In fact, Perelman himself was aware of these drawbacks and drew attention to them in both of his Ricci flow preprints [P1, p. 37] and [P2, p. 1], respectively:

> "*It is likely that by passing to the limit in this construction* [*of Ricci flow with surgery*] *one would get a canonically defined Ricci flow through singularities, but at the moment I do not have a proof of that.*"

> "*Our approach ... is aimed at eventually constructing a canonical Ricci flow ... a goal, that has not been achieved yet in the present work.*"

Motivated by the above, the paper [KL2] introduced a new notion of weak (or generalized) solutions to Ricci flow in dimension 3 and proved the existence within this class of solutions for arbitrary initial data, as well as a number of results about their geometric and analytical properties.

In this paper we show that the weak solutions of [KL2] are uniquely determined by their initial data (see Theorem 1.1 below). In combination with [KL2], this implies that the associated initial value problem has a *canonical* weak solution, thereby proving Perelman's conjecture (see Corollary 1.2). We also show that this weak solution depends continuously on its initial data, and that it is a limit of Ricci flows with surgery (see Corollary 1.4). In summary, our results provide an answer to the long-standing problem of finding a satisfactory theory of weak solutions to the Ricci flow equation in the 3-dimensional case.

From a broader perspective, it is interesting to compare the results in this paper with work on weak solutions to other geometric PDEs.

The theory of existence and partial regularity of such weak solutions has been studied

extensively. As with PDEs in general, proving existence of solutions requires a choice of objects and a topology that is strong enough to respect the equation, but weak enough to satisfy certain compactness properties. Establishing the finer structure of solutions (e.g. partial regularity) requires, generally speaking, a mechanism for restricting blow-ups. For minimal surfaces, harmonic maps and harmonic map heat flow, good notions of weak solutions with accompanying existence and partial regularity theorems were developed long ago [Al], [Si], [SU], [CS]. By contrast, the theory of weak solutions to mean curvature flow, the Einstein equation and Ricci flow, are at earlier stages of development. For mean curvature flow, for instance, different approaches to weak solutions (e.g. (enhanced) Brakke flows and level set flow) were introduced over the last forty years [Bra], [ES], [CGG], [I1]. Yet, in spite of deep results for the cases of mean convex or generic initial conditions [W1], [W3], [W4], [CM], to our knowledge, the best results known for flows starting from a general compact smooth surface in $\mathbb{R}^3$ are essentially those of [Bra], which are presumably far from optimal. For the (Riemannian) Einstein equation many results have been obtained in the Kähler case and on limits of smooth Einstein manifolds, but otherwise progress toward even a viable definition of weak solutions has been rather limited. Progress on Ricci flow has been limited to the study of specific models for an isolated singularity [FIK], [AK], [ACK] and the Kähler case, which has advanced rapidly in the last ten years after the appearance of [ST].

Regarding uniqueness of weak solutions, our focus in this paper, much less is known. The paper [I2] describes a mechanism for non-uniqueness, stemming from the dynamical instability of cones, which is applicable to a number of geometric flows. For example, for mean curvature flow of hypersurfaces in $\mathbb{R}^n$ this mechanism provides examples of non-uniqueness in high dimensions. Ilmanen and White [W2] found examples of non-uniqueness starting from compact smooth surfaces in $\mathbb{R}^3$. Examples for harmonic map heat flow are constructed in [GR], [GGM], and for Ricci flow in higher dimensions there are examples in [FIK], which suggest non-uniqueness. Since any discussion of uniqueness must refer to a particular class of admissible solutions, the interpretation of some of the above examples is not entirely clear, especially in the case of higher-dimensional Ricci flow, where a definition of weak solutions is lacking. In the other direction, uniqueness has been proven to hold in only a few cases: harmonic map heat flow with 2-dimensional domain [St], mean convex mean curvature flow [W3] and Kähler–Ricci flow [ST], [EGZ]. The proofs of these theorems rely on special features of these flows. In [St], the flow develops singularities only at a finite set of times, and at isolated points. The striking proof of uniqueness in [W3] is based on comparison techniques for scalar equations and a geometric monotonicity property specific to mean convex flow (see also the recent paper [HW], which localizes the mean convexity assumption). Lastly, Kähler–Ricci flow has

many remarkable features that play a crucial role in its uniqueness argument: the singularities, whose form is quite rigid, arise at a finite set of times determined by the evolution of the Kähler class; also, techniques specific to scalar equations play an important role.

The method of proving uniqueness used in this paper is completely different in spirit from earlier work. Uniqueness is deduced by comparing two flows with nearby initial condition and estimating the rate at which they diverge from one another. Due to the nature of the singularities, which might in principle occur at a Cantor set of times, the flows can only be compared after the removal of their almost singular regions. Since one knows nothing about the correlation between the almost singular parts of the two flows, the crux of the proof is to control the influence of effects emanating from the boundary of the truncated flows. This control implies a strong stability property, which roughly speaking states that both flows are close away from their almost singular parts if they are sufficiently close initially. A surprising consequence of our analysis is that this strong stability result applies not just to Ricci flows with surgery and the weak solutions of [KL2], but to flows whose almost singular parts are allowed to evolve in an arbitrary fashion, possibly violating the Ricci flow equation at small scales.

The main ideas of our proof may throw light on uniqueness problems in general. When distilled down to its essentials, our proof is based on the following ingredients:

(1) A structure theory for the almost singular part of the flow, which is based on a classification of all blow-ups, not just shrinking solitons.

(2) Uniform strict stability for solutions to the linearized equation, for all blow-ups.

(3) An additional quantitative rigidity property for blow-ups that makes it possible to fill in missing data to the evolution problem, after recently resolved singularities.

This list, which is not specific to Ricci flow, suggests a tentative criterion for when one might expect, and possibly prove, uniqueness for weak solutions to a given geometric flow. From a philosophical point of view, it is natural to expect (1) and (2) to be necessary conditions for uniqueness. However, implementation of even (1) can be quite difficult. Indeed, to date there are few situations where such a classification is known. It turns out that (3) is by far the most delicate part of the proof in our setting and it is responsible for much of the complexity in the argument (see the overview of the proof in §2 for more discussion of this point). Another context where the above criteria may be satisfied is the case of mean curvature flow of 2-spheres in $\mathbb{R}^3$, where uniqueness is conjectured to hold [W2].

We mention that our main result implies that weak solutions to Ricci flow behave well even when one considers continuous families of initial conditions. This continuous dependence leads to new results for diffeomorphism groups of 3-manifolds and spaces of metrics with positive scalar curvature, which will be discussed elsewhere [BK1]–[BK3].

1.2. Background and setup

In preparation for the statements of our main results, which will be presented in the next subsection, we now recall in greater detail some facts about Perelman's Ricci flow with surgery [P2], [KL1], [MT], [BBM+] and the weak solutions from [KL2], which will be needed for our setup. As these constructions are generally very technical, we will continue in a relatively informal style. The reader who is already familiar with this material may skip this subsection and proceed to the presentation of the main results in §1.3.

In his seminal paper [Ha1], Hamilton introduced the Ricci flow equation

$$\partial_t g(t) = -2\mathrm{Ric}(g(t)), \quad g(0) = g_0$$

and showed that any Riemannian metric g_0 on a compact manifold can be evolved into a unique solution $(g(t))_{t\in[0,T)}$. This solution may, however, develop a singularity in finite time. In [P1], Perelman analyzed such finite-time singularities in the 3-dimensional case and showed that those are essentially caused by two behaviors:

- *Extinction* (e.g. the flow becomes asymptotic to a shrinking round sphere).
- The development of *neck pinches* (i.e. there are regions of the manifold that become more and more cylindrical, $\approx S^2\times\mathbb{R}$, modulo rescaling, while the diameter of the cross-sectional 2-sphere shrinks to zero).

Based on this knowledge, and inspired by a program suggested by Hamilton, Perelman specified a surgery process in which the manifold is cut open along small cross-sectional 2-spheres, the high-curvature part of the manifold and extinct components are removed, and the resulting spherical boundary components are filled in with 3-disks endowed with a standard cap metric. This produces a new smooth metric on a closed manifold, from which the Ricci flow can be restarted. The process may then be iterated to yield a *Ricci flow with surgery*. More specifically, a Ricci flow with surgery is a sequence of conventional Ricci flows $(g_1(t))_{t\in[0,T_1]}$, $(g_2(t))_{t\in[T_1,T_2]}$, $(g_3(t))_{t\in[T_2,T_3]}$, ... on compact manifolds M_1, M_2, M_3, ..., where $(M_{i+1}, g_{i+1}(T_i))$ arises from $(M_i, g_i(T_i))$ by a surgery process, as described before.

As mentioned in §1.1, the construction of a Ricci flow with surgery depends on a variety of auxiliary parameters, for which there does not seem to be a canonical choice, such as the following:

- The scale of the cross-sectional 2-sphere along which a neck pinch singularity is excised; this scale is often called the *surgery scale*.
- The precise position and number of these 2-spheres.
- The standard cap metric that is placed on the 3-disks which are glued into the 2-sphere boundary components.

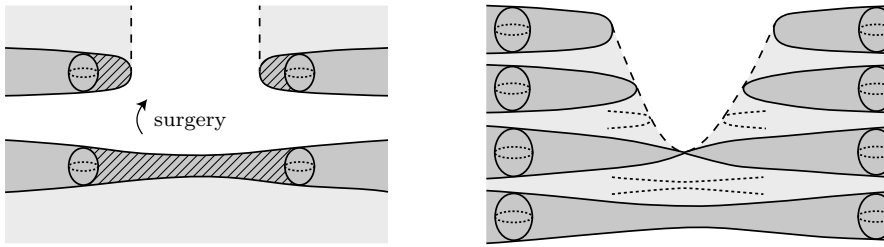

Figure 1. In a Ricci flow with surgery (left figure) surgeries are performed at a positive scale, whereas a singular Ricci flow (right figure) "flows through" a singularity at an infinitesimal scale. The hatched regions in the left figure mark the surgery points, i.e. the points that are removed or added during a surgery.

• The method used to interpolate between this metric and the metric on the nearby necks.

Different choices of these parameters may influence the future development of the flow significantly (as well as the space of future surgery parameters). Hence, a Ricci flow with surgery cannot be constructed in a *canonical* way or, in other words, a Ricci flow with surgery is not *uniquely* determined by its initial metric.

It is therefore a natural question whether a Ricci flow with surgery can be replaced by a more canonical object, which one may hope is uniquely determined by its initial data. This question was first addressed in [KL2], where the notion of a *singular Ricci flow*, a kind of weak solution to the Ricci flow equation, was introduced. In these flows, surgeries have been replaced by singular structure, i.e. regions with unbounded curvature, which may be thought of as "surgery at an infinitesimal scale" (see Figure 1).

In order to present the definition and summarize the construction of a singular Ricci flow, we need to introduce the spacetime picture of a Ricci flow or a Ricci flow with surgery. For this purpose, consider a Ricci flow with surgery consisting of the conventional Ricci flows $(M_1,(g_1(t))_{t\in[0,T_1]})$, $(M_2,(g_2(t))_{t\in[T_1,T_2]})$, ... and form the following 4-dimensional spacetime manifold (see Figure 2 for an illustration):

$$\mathcal{M} := (M_1\times[0,T_1]\bigcup_{\phi_1} M_2\times[T_1,T_2]\bigcup_{\phi_2} M_3\times[T_2,T_3]\bigcup_{\phi_3}\dots)\setminus\mathcal{S}. \tag{1.1}$$

Here $\mathcal{S}$ denotes the set of surgery points, i.e. the set of points that are removed or added during a surgery step and $\phi_i\colon M_i\supset U_i\to U_{i+1}\subset M_{i+1}$ are isometric gluing maps, which are defined on the complement of the surgery points in $M_i\times\{T_i\}$ and $M_{i+1}\times\{T_i\}$. The above construction induces a natural *time-function* $\mathfrak{t}\colon\mathcal{M}\to[0,\infty)$, whose level-sets are called *time-slices*, as well as a *time-vector field* $\partial_{\mathfrak{t}}$ on $\mathcal{M}$ with $\partial_{\mathfrak{t}}\cdot\mathfrak{t}=1$. The Ricci flows $(g_1(t))_{t\in[0,T_1]}$, $(g_2(t))_{t\in[T_1,T_2]}$, ... induce a metric g on the horizontal distribution

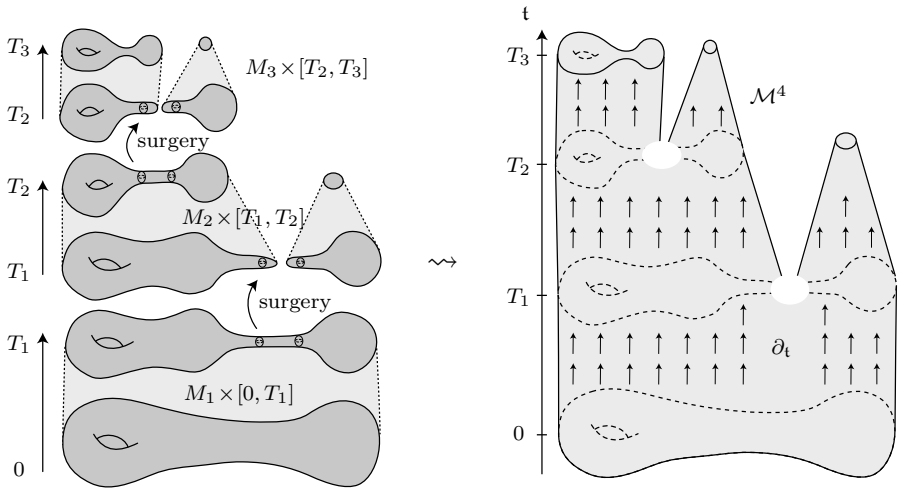

Figure 2. A Ricci flow with surgery (left) can be converted to a Ricci flow spacetime (right) by identifying pre and post-surgery time-slices and removing surgery points. The white circles in the right figure indicate that surgery points were removed at times T_1 and T_2.

$\{dt{=}0\} \subset T\mathcal{M}$, which satisfies the Ricci flow equation

$$\mathcal{L}_{\partial_t} g = -2\mathrm{Ric}(g).$$

The tuple $(\mathcal{M}, t, \partial_t, g)$ is called a *Ricci flow spacetime* (see Definition 5.1 for further details). We will often abbreviate this tuple by $\mathcal{M}$.

Note that a Ricci flow spacetime $\mathcal{M}$ that is constructed from a Ricci flow with surgery by the procedure above is incomplete (see Definition 5.3 for more details). More specifically, the time-slices corresponding to surgery times are incomplete Riemannian manifolds, because surgery points, consisting of necks near neck pinches or standard caps are not included in $\mathcal{M}$. So, these time-slices have "holes" whose "diameters" are $\leqslant C\delta$, where δ is the surgery scale and C is a universal constant. A Ricci flow with this property is called *$C\delta$-complete* (see again Definition 5.3 for further details).

In [KL2] it was shown that every Riemannian manifold is the initial time-slice of a Ricci flow spacetime $\mathcal{M}$ whose time-slices are *zero-complete*, which we also refer to as *complete* (see Figure 3 for an illustration). This means that the time-slices of $\mathcal{M}$ may be incomplete, but each time-slice can be completed as a metric space by adding a countable set of points. Note that since the curvature after a singularity is not uniformly bounded, we cannot easily control the time until a subsequent singularity arises. In fact, it is possible — although not known at this point — that the set of singular times on a finite time-interval is infinite or even uncountable. See [KL3] for a proof that this set has Minkowski dimension $\leqslant \frac{1}{2}$.

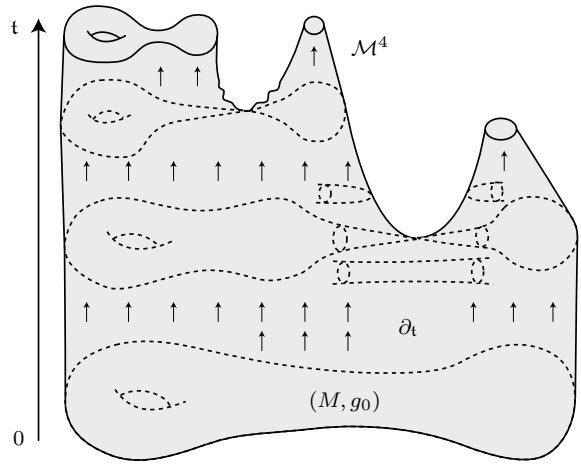

Figure 3. Example of a zero-complete Ricci flow spacetime with initial time-slice (M, g_0).

We briefly review the construction of the (zero-complete) Ricci flow spacetime $\mathcal{M}$ in [KL2]. Consider a sequence of Ricci flows with surgery with surgery scale $\delta_i \to 0$, starting from the same given initial metric, and construct the corresponding Ricci flow spacetimes $\mathcal{M}_{\delta_i}$ as in (1.1). Using a compactness argument, it was shown in [KL2] that, after passing to a subsequence, we have convergence

$$\mathcal{M}_{\delta_i} \to \mathcal{M} \tag{1.2}$$

in a certain sense. The Ricci flow spacetime $\mathcal{M}$ can then be shown to be zero-complete.

We remark that even though the surgery scale in this flow is effectively zero, which seems more canonical than in a Ricci flow with surgery, the entire flow may a priori not be canonical; i.e. the flow is a priori not uniquely determined by its initial data.

We also remind the reader that, while a Ricci flow spacetime describes a singular flow, the metric tensor field g on $\mathcal{M}$ is not singular itself, since the spacetime manifold $\mathcal{M}$ does not "contain the singular points". In other words, $\mathcal{M}$ describes the flow only on its regular part. A flow that includes singular points can be obtained, for example, by taking the metric completion of the time-slices. However, we do not take this approach, in order to avoid having to formulate the Ricci flow equation at the added singular points. This is in contrast to weak forms of other geometric flows, such as the Brakke flow (generalizing mean curvature flow), which is defined at singular points and therefore not smooth everywhere.

In lieu of an interpretation of the Ricci flow equation at the (non-existent) singular points of a Ricci flow spacetime, it becomes necessary to characterize the asymptotic

geometry in its almost singular regions. This is achieved via the *canonical neighborhood assumption*, which states that regions of high curvature are geometrically close to model solutions — *κ-solutions* — modulo rescaling (see Definition 5.6 for more details). Roughly speaking, this implies that these regions are either spherical, neck-like or cap-like. κ-solutions (see Definition 5.4 for more details) arise naturally as blow-up limits of conventional 3-dimensional Ricci flows and have also been shown to characterize high-curvature regions in Ricci flows with surgery. Moreover, the Ricci flow spacetimes constructed in [KL2] also satisfy the canonical neighborhood assumption in an even stronger sense (for more details see the discussion after Definition 5.6).

1.3. Statement of the main results

We now state the main results of this paper in their full generality. Some of the terminology used in the following was informally introduced in the previous subsection. For precise definitions and further discussions we refer the reader to §5.

Our first main result is the uniqueness of complete Ricci flow spacetimes that satisfy the canonical neighborhood assumptions. These spacetimes were also sometimes called "weak Ricci flows" in the previous two subsections.

THEOREM 1.1. (Uniqueness of Ricci flow spacetimes, general form) *There is a universal constant $\varepsilon_{\text{can}}>0$ such that the following holds.*

Let $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ and $(\mathcal{M}', \mathfrak{t}', \partial_{\mathfrak{t}'}, g')$ be two Ricci flow spacetimes that are both $(0,T)$-complete for some $T\in(0,\infty]$ and satisfy the ε_{can}-canonical neighborhood assumption at scales $(0,r)$ for some $r>0$. If the initial time-slices $(\mathcal{M}_0, g_0)$ and $(\mathcal{M}'_0, g'_0)$ are isometric, then the flows $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ and $(\mathcal{M}', \mathfrak{t}', \partial_{\mathfrak{t}'}, g')$ are isometric as well.

More precisely, assume that there is an isometry $\phi\colon (\mathcal{M}_0, g_0)\to(\mathcal{M}'_0, g'_0)$. Then, there is a unique smooth diffeomorphism $\hat{\phi}\colon \mathcal{M}_{[0,T]}\to\mathcal{M}'_{[0,T]}$ such that

$$\hat{\phi}^* g' = g, \quad \hat{\phi}|_{\mathcal{M}_0} = \phi, \quad \hat{\phi}_* \partial_{\mathfrak{t}} = \partial_{\mathfrak{t}'}, \quad \mathfrak{t}'\circ\hat{\phi} = \mathfrak{t}.$$

A Ricci flow spacetime is "$(0,T)$-complete" if the zero-completeness property holds up to time T (see Definition 5.3).

Both properties that are imposed on $\mathcal{M}$ and $\mathcal{M}'$ in Theorem 1.1 hold naturally for the Ricci flow spacetimes constructed in [KL2]. So, we obtain the following corollary.

COROLLARY 1.2. *There is a universal constant $\varepsilon_{\text{can}}>0$ such that the following holds.*

For every compact Riemannian manifold (M,g) there is a unique (i.e. canonical) Ricci flow spacetime $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ whose initial time-slice $(\mathcal{M}_0, g_0)$ is isometric to (M,g) and that is zero-complete, and such that for every $T>0$ the time-slab $\mathcal{M}_{[0,T)}$ satisfies the ε_{can}-canonical neighborhood assumption at scales $(0, r_T)$ for some $r_T>0$.

While we will not discuss this here, we remark that it is possible to modify the arguments in [Ba2]–[Ba6] to show that the flow $\mathcal{M}$ becomes non-singular past some time $T>0$ and we have a curvature bound of the form $|\mathrm{Rm}|<C/t$. So, the scale r_T in Corollary 1.2 can even be chosen independently of T.

Coming back to Theorem 1.1, we draw attention to the fact that the time-slices of $\mathcal{M}$ and $\mathcal{M}'$, including the initial time-slices, may have infinite diameter or volume. Also, they may have unbounded curvature even in bounded subsets, for instance when the flow starts from a manifold with finite-diameter cuspidal ends. We also emphasize that the constant $\varepsilon_{\mathrm{can}}$ is universal and does not depend on any geometric quantities.

Theorem 1.1 will follow from a stability result for Ricci flow spacetimes. We first present a slightly less general, but more accessible version of this stability result. In the following theorem, we only require the completeness and the canonical neighborhood assumption to hold above some small scale ε, i.e. where the curvature is $\lesssim\varepsilon^{-2}$. As such, the theorem can also be used to compare two Ricci flows with surgery or a Ricci flow with surgery and a Ricci flow spacetime, via the construction (1.1). Furthermore, we only require the initial time-slices of $\mathcal{M}$ and $\mathcal{M}'$ be close in the sense that there is a sufficiently precise bilipschitz map ϕ, which may only be defined on regions where the curvature is not too large. As a consequence, the two Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ can only be shown to be geometrically close. More specifically, the map $\hat{\phi}$ that compares $\mathcal{M}$ with $\mathcal{M}'$ can only shown to be bilipschitz and may not be defined on high-curvature regions. The map $\hat{\phi}$ is also not necessarily ∂_t-preserving (see Definition 6.12), but it satisfies the harmonic map heat flow equation (see Definition 6.13).

THEOREM 1.3. (Stability of Ricci flow spacetimes, weak form) *For every $\delta>0$ and $T<\infty$ there is an $\varepsilon=\varepsilon(\delta,T)>0$ such that the following holds.*

Consider two (ε,T)-complete Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ that each satisfy the ε-canonical neighborhood assumption at scales $(\varepsilon,1)$.

Let $\phi\colon U\to U'$ be a diffeomorphism between two open subsets $U\subset\mathcal{M}_0$ and $U'\subset\mathcal{M}'_0$. Assume that $|\mathrm{Rm}|\geqslant\varepsilon^{-2}$ on $\mathcal{M}_0\setminus U$ and

$$|\phi^*g'_0-g_0|\leqslant\varepsilon.$$

Assume also that the ε-canonical neighborhood assumption holds on U' at scales $(0,1)$.

Then, there is a time-preserving diffeomorphism $\hat{\phi}\colon\widehat{U}\to\widehat{U}'$ between two open subsets $\widehat{U}\subset\mathcal{M}_{[0,T]}$ and $\widehat{U}'\subset\mathcal{M}'_{[0,T]}$ that evolves by the harmonic map heat flow and that satisfies $\hat{\phi}=\phi$ on $U\cap\widehat{U}$ and

$$|\hat{\phi}^*g'-g|\leqslant\delta.$$

Moreover, $|\mathrm{Rm}|\geqslant\delta^{-2}$ on $\mathcal{M}_{[0,T]}\setminus\widehat{U}$.

We remark that the condition that the ε-canonical neighborhood assumption holds on U' at scales $(0,1)$ is automatically satisfied if the curvature scale on U' is $>\varepsilon$, which is implied by a bound of the form $|\mathrm{Rm}|<c\varepsilon^{-2}$ on U' (see Definition 5.6 for further details).

Theorem 1.3 is formulated using only C^0-bounds on the quantity $\phi^* g' - g$, which measures the deviation from an isometry. Using a standard argument involving local gradient estimates for non-linear parabolic equations, these bounds can be improved to higher-derivative bounds as follows:

Addendum to Theorem 1.3. *Let $m_0 \geqslant 1$ and $C<\infty$. If in Theorem 1.3 we additionally require that*

$$|\nabla^m(\phi^* g'_0 - g_0)| \leqslant \varepsilon$$

and

$$|\nabla^m \mathrm{Rm}| \leqslant C$$

on U for all $m=0,...,m_0+2$, and allow ε to depend on m_0 and C, then

$$|\nabla^m(\hat{\phi}^* g' - g)| \leqslant \delta$$

on $\widehat{U}$ for all $m=0,...,m_0$.

A similar addendum applies to Theorem 1.5 below.

Combining Theorem 1.3 with [KL2, Theorem 1.2] (see also [KL2, p. 6]) we obtain the following.

Corollary 1.4. *Let (M,g) be a compact Riemannian manifold, and consider a sequence of Ricci flows with surgery starting from (M,g), for a sequence of surgery scales $\delta_i \to 0$. Let $\mathcal{M}_{\delta_i}$ be the corresponding Ricci flow spacetimes, as defined in (1.1). Then, the $\mathcal{M}_{\delta_i}$ converge to a unique Ricci flow spacetime as in (1.2).*

We remark that in the case of mean curvature flow a similar result holds: In [He], [L] it was shown that the 2-convex mean curvature flow with surgery constructed in [HS] converges to the level set flow as the surgery parameter tends to zero. However, their proofs, which are remarkably elementary, are entirely different from ours: they use a quantitative variant of the barrier argument from White's uniqueness theorem [W3]. A similar convergence result holds for mean convex mean curvature flow with surgery in $\mathbb{R}^3$, as constructed in [BH], [HK].

Lastly, we state the stability theorem for Ricci flow spacetimes in its full generality. The following theorem is an improvement of Theorem 1.3 for the following reasons:

- It provides additional information on the bilipschitz constant and establishes a polynomial dependence on the curvature.

• It states that the precision of the canonical neighborhood assumption can be chosen independently of time and bilipschitz constant.

• In provides a condition under which the map $\hat{\phi}$ is almost surjective.

THEOREM 1.5. (Strong stability of Ricci flow spacetimes) *There is a constant $\underline{E}<\infty$ such that, for every $\delta>0$, $T<\infty$ and $\underline{E}\leqslant E<\infty$, there are constants $\varepsilon_{\text{can}}=\varepsilon_{\text{can}}(E)$ and $\varepsilon=\varepsilon(\delta,T,E)>0$ such that for all $0<r\leqslant 1$ the following holds.*

Consider two $(\varepsilon r,T)$-complete Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ that each satisfy the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon r,1)$.

Let $\phi\colon U\to U'$ be a diffeomorphism between two open subsets $U\subset\mathcal{M}_0$ and $U'\subset\mathcal{M}'_0$. Assume that $|\mathrm{Rm}|\geqslant(\varepsilon r)^{-2}$ on $\mathcal{M}_0\setminus U$ and

$$|\phi^* g'_0-g_0|\leqslant\varepsilon r^{2E}(|\mathrm{Rm}|+1)^E$$

on U. Assume moreover that the ε_{can}-canonical neighborhood assumption holds on U' at scales $(0,1)$.

Then, there is a time-preserving diffeomorphism $\hat{\phi}\colon\widehat{U}\to\widehat{U}'$ between two open subsets $\widehat{U}\subset\mathcal{M}_{[0,T]}$ and $\widehat{U}'\subset\mathcal{M}'_{[0,T]}$ that evolves by the harmonic map heat flow, satisfies $\hat{\phi}=\phi$ on $U\cap\widehat{U}$ and that satisfies

$$|\hat{\phi}^* g'-g|\leqslant\delta r^{2E}(|\mathrm{Rm}|+1)^E$$

on $\widehat{U}$. Moreover, we have $|\mathrm{Rm}|\geqslant r^{-2}$ on $\mathcal{M}_{[0,T]}\setminus\widehat{U}$.

If additionally $|\mathrm{Rm}|\geqslant(\varepsilon r)^{-2}$ on $\mathcal{M}'_0\setminus U'$, then we have $|\mathrm{Rm}|\geqslant r^{-2}$ on $\mathcal{M}'_{[0,T]}\setminus\widehat{U}'$.

1.4. A brief sketch of the proof, and further discussion

We now give a very brief and informal outline of the proof. See §2 for a more detailed overview.

Theorem 1.1, the main uniqueness theorem, is obtained from the strong stability Theorem 1.3 or 1.5 via a limit argument. In Theorems 1.3 and 1.5 we are given a pair of Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$, and an almost isometry $\phi\colon\mathcal{M}_0\supset U\to U'\subset\mathcal{M}'_0$ between open subsets of their initial conditions, and our goal is to construct an almost isometry $\hat{\phi}\colon\mathcal{M}\supset\widehat{U}\to\widehat{U'}\subset\mathcal{M}'$ that extends ϕ forward in time. The construction of $\hat{\phi}$ involves a procedure for choosing the domain $\widehat{U}$ of $\hat{\phi}$, and the map $\hat{\phi}$ on this domain. These two procedures interact in a complex way, and for this reason they are implemented by means of a simultaneous induction argument.

We now indicate some of the highlights in the two steps of the induction.

The domain $\widehat{U}$ is chosen to contain all points in $\mathcal{M}$ whose curvature $|\mathrm{Rm}|$ lies (roughly) below a certain threshold and is obtained from $\mathcal{M}$ by means of a delicate truncation argument. The truncation uses the fact that, roughly speaking, the part of $\mathcal{M}$ with large curvature looks locally either like a neck, or like a cap region. We cut along neck regions so that the time-slices of $\widehat{U}$ have spherical boundary. A critical complication stems from the occurrence of moments in time when the presence of cap regions interferes with the need to cut along neck regions. This occurrence necessitates modification of the domain by either insertion or removal of cap regions.

The map $\hat{\phi}$ is constructed by solving the harmonic map heat flow equation for its inverse $\hat{\phi}^{-1}$. There are many interrelated issues connected with this step, of which the three most important are:

- The distortion of the map $\hat{\phi}$ must be controlled under the harmonic map heat flow. For this, our main tool is an interior decay estimate, which may be applied away from the spacetime boundary of $\widehat{U}$.
- The presence of boundary in $\widehat{U}$ introduces boundary effects, which must be controlled. It turns out that the geometry of shrinking necks implies that the neck boundary recedes rapidly, which helps to stabilize the construction.
- The insertion of the cap regions alluded to above necessitates the extension of the map $\hat{\phi}$ over the newly added region. The implementation of this extension procedure relies on a delicate interpolation argument, in which the geometric models for the cap regions must be aligned with the existing comparison map $\hat{\phi}$ within tolerances fine enough to prolong the construction. This step hinges on several ingredients and their precise compatibility — rigidity theorems for the models of the cap regions [Ha2], [Bre2], quantitative asymptotics of the models [Bry], and strong decay estimates for the distortion of the map $\hat{\phi}$.

1.5. Acknowledgements

The first named author would like to thank Bennet Chow for pointing out his paper with Greg Anderson (see [AC]). Both authors would also like to thank Yongjia Zhang for many valuable corrections.

2. Overview of the proof

In this section we will describe the proof of the main theorem. Our aim here is to cover the most important ideas in an informal way, with many technicalities omitted. The first subsection of this overview provides an initial glimpse of the argument. It is intended

to be accessible to readers outside the field who would like to gain some sense of how the proof goes. The remaining subsections delve into the proof in greater detail and are primarily intended for people working in the area.

The main part of the paper is concerned with the proof of the strong stability theorem, Theorem 1.3 or 1.5, which asserts that two Ricci flow spacetimes are geometrically close, given that their initial data are geometrically close and we have completeness as well as the canonical neighborhood assumption in both spacetimes above a sufficiently small scale. All other results of this paper will follow from this theorem; in particular, the uniqueness theorem (Theorem 1.1) will follow from Theorem 1.3 or 1.5 via a limit argument.

In the strong stability theorem, we consider two Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$, whose initial time-slices, $(\mathcal{M}_0, g_0)$ and $(\mathcal{M}'_0, g'_0)$, are geometrically close or even isometric. Our goal is the construction of a map $\phi{:}\mathcal{M}\supset U\to\mathcal{M}'$, defined on a sufficiently large domain U, whose bilipschitz constant is sufficiently close to 1. In Theorems 1.3 and 1.5, this map is denoted by $\hat{\phi}$. However, in the main part of this paper, as well as in this overview, the hat will be omitted.

Our basic method for constructing ϕ, which goes back to DeTurck [DT], is to solve the harmonic map heat flow equation for the inverse ϕ^{-1}. In the non-singular case when both Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ may be represented by ordinary smooth Ricci flows on compact manifolds $(M, g(t))$ and $(M', g'(t))$, this reduces to finding a solution $\phi(t)\colon M\to M'$ to the equation $\partial_t(\phi^{-1})=\Delta(\phi^{-1})$. As DeTurck observed, the family of difference tensors $h(t):=(\phi(t))^*g'(t)-g(t)$, which quantify the deviation of $\phi(t)$ from being an isometry, then satisfies the *Ricci–DeTurck perturbation equation*:

$$\partial_t h(t) = \Delta_{g(t)} h(t) + 2\mathrm{Rm}_{g(t)}(h(t)) + \nabla h(t) * \nabla h(t) + h(t) * \nabla^2 h(t). \tag{2.1}$$

If $\phi(0)$ is an isometry, then $h(0)\equiv 0$. So, by the uniqueness of solutions to the strictly parabolic equation (2.1), one gets that $h(t)\equiv 0$ for all $t\geqslant 0$, and hence the two given Ricci flows are isometric. In our case we are given that $h(0)$ is small, and want to show that it remains small. Equation (2.1) has several properties that are important for maintaining control over of the size of the perturbation h, as the construction proceeds.

2.1. The construction process, an initial glimpse

In the general case, in which $\mathcal{M}$ and $\mathcal{M}'$ may be singular, the domain of the map ϕ will be the part of $\mathcal{M}$ that is not too singular, i.e. the set of points whose curvature is not too large. Note that this means that we will effectively be solving the harmonic map heat flow equation with a boundary condition.

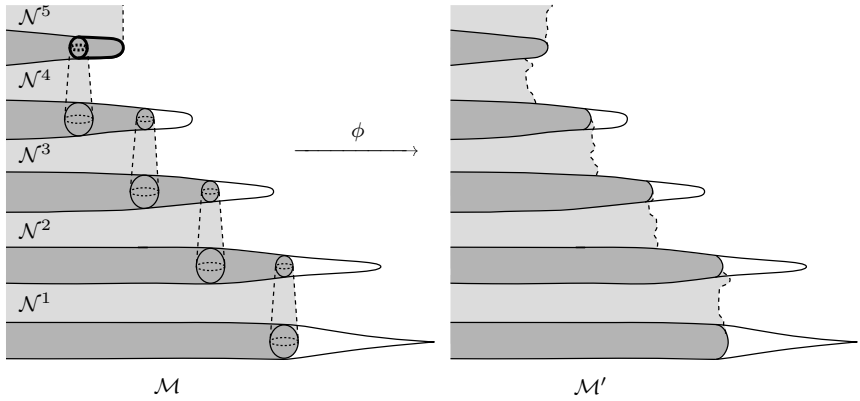

Figure 4. Comparison domain $\mathcal{N}\subset\mathcal{M}$ and comparison ϕ between $\mathcal{M}$ and $\mathcal{M}'$. The extension cap on the initial time-slice of $\mathcal{N}^5$ is outlined in bold.

The main objects of our construction are a subset $\mathcal{N}\subset\mathcal{M}$, called the *comparison domain*, and a time-preserving diffeomorphism onto its image $\phi:\mathcal{N}\to\mathcal{M}'$, called the *comparison* (*map*). We construct $\mathcal{N}$ and ϕ by a simultaneous induction argument using discrete time increments $[t_{j-1},t_j]$. The domain $\mathcal{N}$ is the union

$$\mathcal{N}=\mathcal{N}^1\cup\mathcal{N}^2\cup...\cup\mathcal{N}^J,$$

where $\mathcal{N}^j$ lies in the time-slab of $\mathcal{M}$ corresponding to the time-interval $[t_{j-1},t_j]$. The restriction of ϕ to each time-slab $\mathcal{N}^j$ is denoted by $\phi^j:\mathcal{N}^j\to\mathcal{M}'$. In the induction step, we enlarge $\mathcal{N}$ and ϕ in two stages: in the first we determine $\mathcal{N}^{J+1}$, and in the second we define the map $\phi^{J+1}:\mathcal{N}^{J+1}\to\mathcal{M}'$.

Before proceeding, we introduce the *curvature scale* ρ, which will be used throughout the paper. The precise definition may be found in §6.1, but for the purposes of this overview, ρ can be any function that agrees up to a fixed factor with $R^{-1/2}$ wherever $|\mathrm{Rm}|$ is sufficiently large. Here R denotes the scalar curvature. Note that ρ has the dimension of length.

We will now provide further details on the geometry of $\mathcal{N}$ and ϕ.

Fix a small *comparison scale* $r_{\mathrm{comp}}>0$. Our goal is to choose the comparison domain $\mathcal{N}$ such that it roughly contains the points for which $\rho\gtrsim r_{\mathrm{comp}}$. Then, we will have $R\leqslant Cr_{\mathrm{comp}}^{-2}$ on $\mathcal{N}$ and $R\geqslant cr_{\mathrm{comp}}^{-2}$ on $\mathcal{M}\backslash\mathcal{N}$ for some constants $C,c>0$. The constant r_{comp} will also determine the length of our time steps: we set $t_j=jr_{\mathrm{comp}}^2$, so that the time steps have duration r_{comp}^2.

Each time-slab $\mathcal{N}^j$ will be chosen to be a product domain on the time-interval $[t_{j-1},t_j]$. That is, the flow restricted to $\mathcal{N}^j$ can be described by an ordinary Ricci flow

parameterized by the time-interval $[t_{j-1}, t_j]$, on the initial time-slice of $\mathcal{N}^j$. We will sometimes denote this initial time-slice of $\mathcal{N}^j$ by $\mathcal{N}^j_{t_{j-1}}$ and the final time-slice by $\mathcal{N}^j_{t_j}$. Note that $\mathcal{N}^j_{t_{j-1}}$ and $\mathcal{N}^j_{t_j}$ are diffeomorphic, as $\mathcal{N}^j$ is a product domain. Each domain $\mathcal{N}^j$ will moreover be chosen in such a way that its time-slices $\mathcal{N}^j_t$ are bounded by 2-spheres of diameter $\approx r_{\text{comp}}$ that are central 2-spheres of sufficiently precise necks (i.e. cylindrical regions) in $\mathcal{M}$.

We now discuss the inductive construction of $\mathcal{N}$ and ϕ. For this purpose, assume that $\mathcal{N}^1, ..., \mathcal{N}^J$ and $\phi^1, ..., \phi^J$ have already been constructed. Our goal is now to construct $\mathcal{N}^{J+1}$ and ϕ^{J+1}.

We first outline the construction of $\mathcal{N}^{J+1}$. Our construction relies on the canonical neighborhood assumption, which guarantees that the large curvature part of the Ricci flow looks, roughly speaking, locally either neck-like or like a cap region diffeomorphic to a 3-ball. Using this geometric characterization, the final time-slice $\mathcal{N}^{J+1}_{t_{J+1}}$ of $\mathcal{N}^{J+1}$ is obtained by truncating the time-t_{J+1}-slice $\mathcal{M}_{t_{J+1}}$ along a suitable collection of central 2-spheres of necks of scale $\approx r_{\text{comp}}$. Due to the fact that a neck region shrinks substantially in a single time step and our neck regions have nearly constant scale, this process will ensure that the boundaries of successive time steps are separated by a distance $\gg r_{\text{comp}}$. So, our truncation process typically yields a rapidly receding "staircase" pattern (see Figure 4). However, it can happen that a cap region evolves in such a way that its scale increases slowly over a time-interval of duration $\gg r^2_{\text{comp}}$, so that at time t_J, this cap region is not contained in the final time-slice $\mathcal{N}^J_{t_J}$, but is contained in the initial time-slice $\mathcal{N}^{J+1}_{t_J}$. This behavior occurs, for instance, a short time after a generic neck pinch singularity. In such a situation, the comparison domain $\mathcal{N}$ is enlarged at time t_J by a cap region, which we call an *extension cap* (see again Figure 4). It then becomes necessary to extend the comparison map ϕ over the inserted region.

We now turn to the second stage of the induction step — the construction of the comparison map $\phi^{J+1}: \mathcal{N}^{J+1} \to \mathcal{M}'$.

As mentioned above, we will construct ϕ^{J+1} by solving the harmonic map heat flow equation for the inverse diffeomorphism $(\phi^{J+1})^{-1}$. For now, we will only provide a brief indication of a few of the obstacles that arise, leaving more detailed discussion to the subsequent subsections of this overview:

• (*Controlling* h, §2.2) Since our objective is to produce a map that is almost an isometry, one of the key ingredients in our argument is a scheme for maintaining control on the size of the metric *perturbation* $h = \phi^* g' - g$ as the map ϕ evolves. Our main tool for this is an interior decay estimate for $|h|$ with respect to a certain weight.

• (*Treatment of the boundary*, §2.3) The Ricci flow spacetime restricted to the product domain $\mathcal{N}^{J+1}$ is given by an ordinary Ricci flow on the manifold with boundary

$\mathcal{N}^{J+1}_{t_{J+1}}$. The process for solving the harmonic map heat flow equation must take this boundary into account and maintain control on any influence it may have on the rest of the evolution.

• (*Extending the comparison*, §2.4 and §2.5) As mentioned above, it may be necessary to extend the comparison map ϕ over an extension cap at time t_J. This requires a careful analysis of the geometry of $\mathcal{M}$ and $\mathcal{M}'$ in neighborhoods of the cap and its image, showing that both are well approximated by rescaled Bryant solitons. Then, the extension of ϕ is obtained by gluing the pre-existing comparison map with suitably normalized Bryant soliton "charts". This gluing construction is particularly delicate, since it must maintain sufficient control over the quality of the comparison map.

The actual construction of the comparison map ϕ is implemented using a continuity argument. The above issues interact with one another in a variety of different ways. For instance, both the treatment of the boundary and the procedure for extending ϕ over cap regions are feasible only under certain assumptions on the smallness of h, and both cause potential deterioration of h, which must be absorbed by the argument for controlling h. We defer further discussion of these interactions, and other points of a more technical nature, to §2.6.

2.2. Controlling the perturbation h

In order to control the perturbation $h=\phi^* g' - g$ in the inductive argument described above, we will consider the following weighted quantity:

$$Q \approx e^{-Ht} R^{-E/2} |h| \approx e^{-Ht} \rho^E |h|. \tag{2.2}$$

Here R denotes the scalar curvature and we have $H>0$ and $E>2$. We will show that this quantity satisfies an *interior decay estimate*, which may be thought of as a quantitative semi-local version of a maximum principle: rather than asserting that Q cannot attain an interior maximum, it roughly states that Q, evaluated at a point (x,t), must be a definite amount smaller than its maximum over a suitable parabolic neighborhood around (x,t) (see below for a more precise statement). This interior decay estimate will allow us to promote, and sometimes improve, a bound of the form $Q \leqslant \overline{Q}$ forward in time. We emphasize that the presence of the factor ρ^E, and the fact that E is strictly larger than 2, are both essential for the interior decay estimate. Moreover, the freedom to choose E large (>100, say) will be of crucial importance at a later point in our proof (see §2.5).

Before providing further details on this estimate, we want to illustrate the function of the weights in the definition of Q. The weight e^{-Ht} serves a technical purpose, which we will neglect in this overview. To appreciate the role of the weight $R^{-E/2}$, consider for

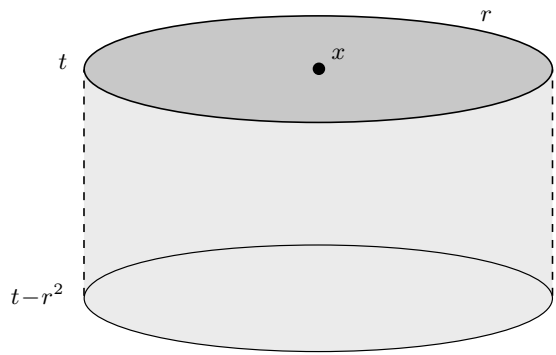

Figure 5. A parabolic ball $B(x,t,r)\times[t-r^2,t]$ of radius r.

a moment a classical Ricci flow $(M,g(t))$ with a perturbation $h(t)$ that evolves by (2.1). Suppose that $h(0)$ is bounded and supported in a region of large scalar curvature. So, due to the existence of the weight $R^{-E/2}$, the quantity Q is small at time zero. Our estimates will imply that Q remains small throughout the flow. Therefore, at any later time, the perturbation h must be small at points where the curvature is controlled. In the following, we will exploit this phenomenon, since, heuristically, we are considering two Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ whose initial data is either equal or very similar away from the almost singular regions, where the scalar curvature is large. So, even if $\mathcal{M}$ and $\mathcal{M}'$ were a-priori significantly different at those almost singular scales — resulting in a large perturbation $h(t)$ there — then Q would still be small, initially. Thus, the perturbation is expected to decay as we move forward in time and towards regions of bounded curvature, establishing an improved closeness there. More specifically, as remarked in the previous subsection, h may a priori only satisfy a rough bound near the neck-like boundary of each $\mathcal{N}^j$. However, as $R\approx r_{\text{comp}}^{-2}$ near such a boundary and r_{comp} is assumed to be small, our estimate suggests a significant improvement of this bound in regions where $R\approx 1$.

We now explain the statement of the interior decay estimate in more detail, in the case of a classical Ricci flow on $M\times[0,T)$. Assume that the perturbation h is defined on a sufficiently large backwards parabolic region $P\subset M\times[0,T)$ around some point (x,t). If H is chosen sufficiently large and $|h|\leqslant\eta_{\text{lin}}$ on P for some sufficiently small η_{lin}, where both H and η_{lin} depend on E, then our estimate states that

$$Q(x,t)\leqslant \tfrac{1}{100}\sup_P Q. \tag{2.3}$$

Here "P sufficiently large" means, roughly speaking, that the parabolic region P contains a product domain of the form $B(x,t,r)\times[t-r^2,t]$ (a parabolic ball), where $B(x,t,r)$ is

the r-ball centered at (x,t) in the time-t slice $M\times\{t\}$, and r is equal to a large constant times the scale $\rho(x,t)$ (see Figure 5).

In fact, the choice of the factor $\frac{1}{100}$ in (2.3) is arbitrary: for any $\alpha>0$ we have the estimate

$$Q(x,t)\leqslant\alpha\sup_P Q, \tag{2.4}$$

as long as we increase the size of the parabolic neighborhood P accordingly. An important detail here is that the constant η_{lin} in the bound $|h|\leqslant\eta_{\text{lin}}$ can be chosen independently of α.

The decay estimate (2.3) will be used to propagate a bound of the form

$$Q\leqslant\overline{Q} \tag{2.5}$$

throughout most parts of the comparison domain $\mathcal{N}$. Here we will choose the constant $\overline{Q}$ in such a way that (2.5) holds automatically near the neck-like boundary of the $\mathcal{N}^j$ and such that (2.5) implies $|h|\leqslant\eta_{\text{lin}}$ wherever $\rho\geqslant r_{\text{comp}}$. Note that at scales $\rho\gg r_{\text{comp}}$, the bound (2.5) implies a more precise bound on $|h|$, whose quality improves polynomially in ρ.

We will prove the interior decay estimate using a limit argument combined with a vanishing theorem for solutions of the linearized Ricci–DeTurck equation on κ-solutions, which uses an estimate of Anderson and Chow [AC]. See §9 for more details.

2.3. Treatment of the boundary

We now discuss aspects of the inductive construction of the map $\phi^{J+1}:\mathcal{N}^{J+1}\to\mathcal{M}'$ (sketched in §2.1) that are related to the presence of a boundary in the time slices $\mathcal{N}^{J+1}_t$. While the actual approach used in the body of the paper is guided by considerations that are beyond the scope of this overview, we will describe some of the main points in a form that is faithful to the spirit of the actual proof.

Recall from §2.1 that we wish to construct ϕ^{J+1} by solving the harmonic map heat flow equation (for the inverse $(\phi^{J+1})^{-1}$), in such a way that ϕ yields a perturbation $h=\phi^*g'-g$ satisfying the bound $Q\leqslant\overline{Q}$ near the neck-like boundary of $\mathcal{N}^j$, where Q is as in §2.2. Thus we need to specify boundary conditions so that the resulting evolution respects the bound $Q\leqslant\overline{Q}$.

Our strategy exploits the geometry of the boundary of $\mathcal{N}^{J+1}$. Recall from §2.1 that $\mathcal{N}^{J+1}$ is a product domain, and its boundary is collared by regions that look very close to shrinking round half-cylinders (half-necks) with scale comparable to r_{comp}. Under a smallness condition on h imposed in the vicinity of boundary components of $\mathcal{N}^{J+1}_{t_J}$,

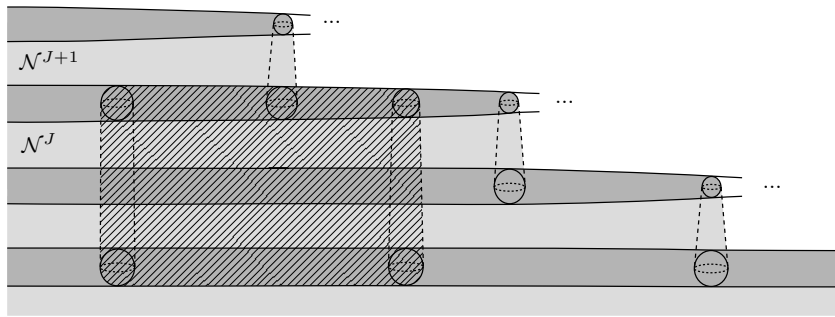

Figure 6. A parabolic neighborhood (hatched region) inside a comparison domain (shaded region). In order to apply the interior estimate at time t_J near the boundary of $\mathcal{N}^{J+1}$, a large parabolic neighborhood must fit underneath the staircase pattern.

we argue that at time t_J, our map ϕ must map the half-neck collar regions around the boundary to regions in the time-t_J slice $\mathcal{M}'_{t_J}$ that are nearly isometric to half-necks. Moreover, we will show that both half-necks evolve over the time-interval $[t_J, t_{J+1}]$ nearly like round half-necks. We then use this characterization and a truncation procedure to find an approximate product domain $\mathcal{N}'^{J+1} \subset \mathcal{M}'$ that serves as the domain for the evolving inverse map ϕ^{-1}. It turns out that if the half-neck regions in $\mathcal{M}$ and $\mathcal{M}'$ are sufficiently cylindrical, and ϕ is initially (at time t_J) sufficiently close to an isometry near the collar regions, then the map ϕ^{J+1} produced by harmonic map heat flow remains sufficiently close to an isometry near the boundary of $\mathcal{N}^{J+1}$, in the sense that $Q \leqslant \overline{Q}$.

The above construction is feasible only under improved initial control on $|h|$, which necessitates an improved bound of the form $Q \leqslant \alpha \overline{Q}$, for some $\alpha \ll 1$, near the boundary components of $\mathcal{N}^{J+1}_{t_J}$. To verify this improved bound, we apply the strong form of the interior decay estimate, (2.4), using parabolic regions that are large depending on α. This requires the geometry of the staircase pattern of the comparison domain to be "flat enough" to create enough space for such a parabolic region "under the staircase" (see Figure 6). Such flatness can be guaranteed, provided the half-neck collars are sufficiently precise.

2.4. Defining ϕ on extension caps

We recall from §2.1 that in the inductive construction of the time slab $\mathcal{N}^{J+1}$, we sometimes encounter *extension caps*, i.e. 3-disks $\mathcal{C}$ in the time-t_J slice $\mathcal{M}_{t_J}$ of $\mathcal{M}$ that belong to time slab $\mathcal{N}^{J+1}$, but that were not present in the preceding time slab $\mathcal{N}^J$. In this and the next subsection, we discuss how these extension caps are handled in the second stage of the induction step, in which ϕ^{J+1} is defined on $\mathcal{N}^{J+1}$.

Recall that we assume inductively that the map ϕ^J, as constructed in the previous step, restricts to an almost isometric map of the final time-slice $\mathcal{N}^J_{t_J}$ of $\mathcal{N}^J$ into $\mathcal{M}'_{t_J}$. We would like to proceed with the construction of ϕ^{J+1} on $\mathcal{N}^{J+1}$ using harmonic map heat flow, as described in the previous subsection. However, in order to do this, ϕ^{J+1} must be defined on the initial time-slice $\mathcal{N}^{J+1}_{t_J}$, whereas the previous induction step only determined ϕ^J on the complement of the extension caps. Thus we must first extend ϕ^J over the extension caps to an almost isometry defined on $\mathcal{N}^{J+1}_{t_J}$.

A priori, it is unclear why such an extension should exist; after all, since ϕ has thus far only been defined on $\mathcal{N}^J_{t_J}$, one might not expect an extension cap $\mathcal{C}\subset\mathcal{N}^{J+1}_{t_J}$ to be nearly isometric to a corresponding 3-ball region in $\mathcal{M}'$.

To obtain such an extension, we will need to combine several ingredients. The first is the canonical neighborhood assumption, which asserts that the geometry of $\mathcal{M}$ and $\mathcal{M}'$ near any point of large curvature is well approximated by a model Ricci flow — a κ-solution. For regions such as extension caps, the κ-soliton model is a Ricci flow on $\mathbb{R}^3$. Up to rescaling, the only known example of this type is the *Bryant soliton*, a rotationally symmetric steady gradient soliton, which can be expressed as a warped product

$$g_{\mathrm{Bry}} = dr^2 + a^2(r) g_{S^2},$$

where $a(r)\sim\sqrt{r}$ as $r\to\infty$. Bryant solitons commonly occur as singularity models of type-II blow-ups of singularity models, for example in the formation of a degenerate neck pinch [GZ], [AIK]. Moreover, they also occur in Ricci flow spacetimes when a singularity resolves. It is a well-known conjecture of Perelman that the Bryant soliton is the only κ-solution on $\mathbb{R}^3$, up to rescaling and isometry. This conjecture would imply that a Bryant soliton *always* describes singularity formation/resolution processes as above, and in particular the geometry of extension caps. Although this conjecture remains open, by using a combination of rigidity results of Hamilton and Brendle [Ha2], [Bre2], it is possible to show that Bryant solitons always describe the geometry at points where the curvature scale increases in time (i.e. where the scalar curvature decreases). Such points are abundant near a resolution of a singularity, as the curvature scale increases from zero to a positive value.

The above observation will be central to our treatment of extension caps. We will show that it is possible to choose the time slabs $\{\mathcal{N}^j\}$ so that each extension cap arises "at the right time", meaning at a time when the geometry near the extension cap in $\mathcal{M}$ and its counterpart in $\mathcal{M}'$ is sufficiently close to the Bryant soliton — at possibly different scales. The main strategy behind this choice of time will be to choose two different thresholds for the curvature scale on $\mathcal{N}$, specifying when an extension cap *may*, and when it *must*, be constructed. As curvature scales only grow slowly in time (with

respect to the time scale corresponding to the curvature), this extra play will produce sufficiently many time-steps during which an extension cap may, but need not necessarily be constructed. It can be shown that at one of these time-steps the geometry in both $\mathcal{M}$ and $\mathcal{M}'$ is in fact close to a Bryant soliton. This time-step will then be chosen as the "right time" for the construction of the extension cap.

The fact that the geometry near both the extension caps in $\mathcal{M}$ and the corresponding regions in $\mathcal{M}'$ can be described by the same singularity model (the Bryant soliton) is necessary in order to construct the initial time-slice of ϕ^{J+1}. However, it is not sufficient, as it is still not guaranteed that ϕ^J at time t_J extends over the extension caps almost isometrically, due to the following reasons:

- The scales of the approximate Bryant soliton regions in $\mathcal{M}$ and $\mathcal{M}'$ may differ, so that they are not almost isometric.
- Even if there is an almost isometry of the approximate Bryant soliton regions, in order to define a global map, there must be an almost isometry that is close enough to the existing almost isometry (given by ϕ^J) on the overlap, so that the two maps may be glued together to form an almost isometry.

These issues will be resolved by the *Bryant extension principle*, which will be discussed in the next subsection.

2.5. The Bryant extension principle

In the process of determining the initial data (at time t_J) of ϕ^{J+1} on or near the extension caps, as mentioned in the previous subsection, we are faced with the following task (see Figure 7 for an illustration). We can find two regions $W \subset \mathcal{M}_{t_J}$ and $W' \subset \mathcal{M}'_{t_J}$ in the time-t_J slices of $\mathcal{M}$ and $\mathcal{M}'$ that are each geometrically close to a Bryant soliton modulo rescaling by some constants λ and λ', respectively. Moreover, the region W contains an extension cap $\mathcal{C} \subset \mathcal{M}_{t_J}$. The map ϕ^J restricted to $W \setminus \mathcal{C}$ is an almost isometric map $W \setminus \mathcal{C} \to W'$. Our task is then to find another almost isometric map $\psi \colon W \to W'$, which is defined on the entire region W, and that coincides with ϕ^J away from some neighborhood of $\mathcal{C}$. Although in this overview we have largely avoided any mention of quantitative features of the proof, we point out that this step hinges on careful consideration of asymptotics, in order to make our construction independent of the diameters of W, W', and $\mathcal{C}$. In particular, it turns out to be of fundamental importance that we have the freedom to choose the exponent E in the definition of Q in (2.2) to be large.

We obtain ψ as follows. We use the fact that W and W' are approximate (rescaled) Bryant soliton regions to define an approximate homothety $\psi_0 \colon W \to W'$ that scales distances by the factor λ'/λ, possibly after shrinking W, W' somewhat. The map ψ_0 is

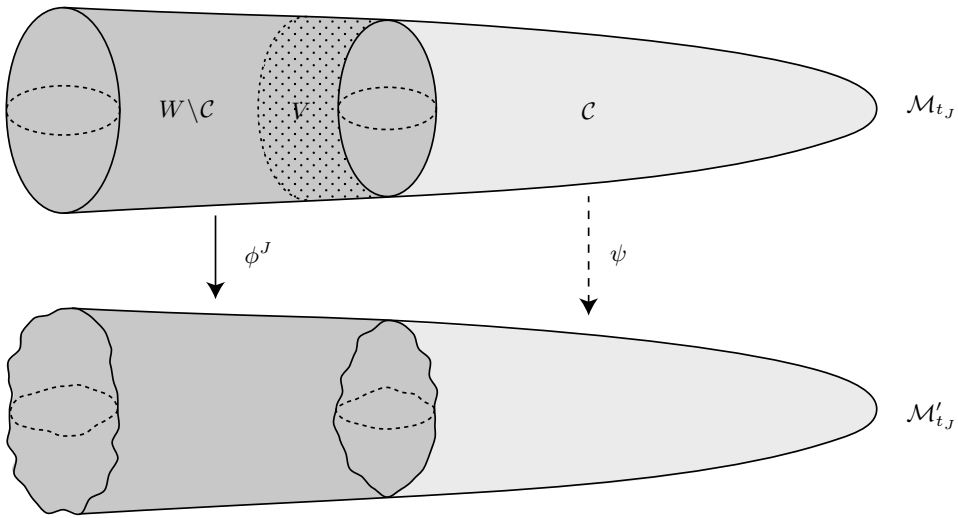

Figure 7. Extending the map $\phi^J_{W\backslash\mathcal{C}}$ over the extension cap $\mathcal{C}$ to an almost isometry $\psi\colon W\to W'$.

unique up to pre/post-composition with almost isometries, i.e. approximate rotations around the respective tips. We then compare ψ_0 with ϕ^J on $W\backslash\mathcal{C}$, and argue that ψ_0 may be chosen so that it may be glued to ϕ^J, to yield the desired map ψ. To do this, we must show that

- ψ_0 is an approximate isometry not just an approximate homothety, i.e. the ratio of the scales λ'/λ nearly equals 1.
- ψ_0 may be chosen to be sufficiently close to ϕ^J on a suitably chosen transition zone $V\subset W\backslash\mathcal{C}$.

However, we are only given information on the map ϕ^J far away from tip of the extension cap $\mathcal{C}$, where the metric is close to a round cylinder. Using this information, we must determine to within small error the scale of the tips and the discrepancy between the two maps. This aspect makes our construction quite delicate, because the only means of detecting the scale of the tip is to measure the deviation from a cylindrical geometry near V, which is decaying polynomially in terms of the distance to the tip. The crucial point in our construction is that we can arrange things so that this deviation can be measured to within an error that decays at a faster polynomial rate.

We now explain in some more detail the delicacy of the construction and our strategy for the case of showing that λ'/λ is nearly equal to 1. The problem of matching ϕ^J and ψ_0 on V will be handled similarly. The only means to compare λ and λ' is the almost isometry $\phi^J\colon W\backslash\mathcal{C}\to W'$. This almost isometry implies that the cross-sectional spheres

of $W\backslash\mathcal{C}$ have approximately the same diameter as their images in W'. Unfortunately, the closeness of these diameters does not imply a bound on λ'/λ, since these diameters vary — and even diverge — as we move away from the tips of W and W', and ϕ^J may map cross-sectional spheres of W to other almost-cross-sectional spheres in W' that are closer or farther from the tip. This fact requires us to estimate the *deviations* from the cylindrical geometry in W and W' by analyzing the precise asymptotics of the Bryant soliton. If the precision of ϕ^J is smaller than these deviations, then ϕ^J can be used to compare further geometric quantities on W and W', not just the diameters of the cross-sectional spheres. Combined with the almost preservation of the diameters of these spheres, this will imply that $\lambda'/\lambda\approx 1$.

The precision of the almost isometry ϕ^J is measured in terms of $|h|$. Using the bound $Q\leqslant\overline{Q}$, as discussed in §2.2, we obtain a bound of the form

$$|h|\lesssim R^{E/2}\lesssim\rho^{-E}. \tag{2.6}$$

Since W is an approximate rescaled Bryant soliton region and $\rho\to\infty$ as one goes to infinity on the Bryant soliton, the bound (2.6) improves as we move further away from the tip of W. If the exponent E is chosen large enough, then the precision of the almost isometry ϕ^J in the transition zone $V\subset W$ is good enough to compare the deviations from a cylinder in V and its image, to very high accuracy. As mentioned before, this will imply that $\lambda'/\lambda\approx 1$.

For more details, we refer to §10.

We mention that the mechanism that we are exploiting here can be illustrated using a cantilever: the longer the cantilever, and the less rigid it is, the more its tip may wiggle. However, the rigidity of a cantilever depends not only on its length, but also on the rigidity of the attachment at its base. A longer cantilever may be more stable than a short one, as long as the attachment at its base is chosen rigid enough to compensate for the increase in length. (Here, we are assuming the lever itself to be infinitely rigid.)

2.6. Further discussion of the proof

In this subsection we touch on a few additional features of the induction argument sketched in §2.1. Due to the complexity of the underlying issues, our explanations will be brief and relatively vague. For more details, we refer to §7.

We recall the bound $Q\leqslant\overline{Q}$ from §2.2, which enabled us to guarantee a bound of the form $|h|\leqslant\eta_{\text{lin}}$ in most parts of the comparison domain $\mathcal{N}$. As discussed in that subsection, this bound is propagated forward in time using the interior decay estimate. The bound $Q\leqslant\overline{Q}$, especially the factor ρ^E in its definition, was also crucial in the Bryant

extension process. In fact, it could be used to construct the initial time-slice of a new almost isometry ϕ^{J+1}, defined on the extension caps, that is precise enough such that a bound of the form $|h| \leqslant \eta_{\text{lin}}$ holds on or near each extension cap.

However, the bound $Q \leqslant \overline{Q}$, which is typically stronger than $|h| \leqslant \eta_{\text{lin}}$, may not remain preserved during the Bryant extension process; it may deteriorate by a fixed factor. In order to control the quality of the comparison map, measured by $|h|$, after the Bryant extension process has been performed, we will consider an additional bound of the form

$$Q^* \approx e^{H(T-t)} \frac{|h|}{R^{3/2}} \approx e^{H(T-t)} \rho^3 |h| \leqslant \overline{Q}^*.$$

The constant $\overline{Q}^*$ will be chosen such that this bound implies the bound $|h| \leqslant \eta_{\text{lin}}$ wherever it holds on $\mathcal{N}$. Due to the small exponent $3 \ll E$, which makes the bound $Q^* \leqslant \overline{Q}^*$ weaker than $Q \leqslant \overline{Q}$ at large scales, this bound still holds on and near each extension cap after the Bryant extension process has been carried out.

Both bounds, $Q \leqslant \overline{Q}$ and $Q^* \leqslant \overline{Q}^*$ will be propagated forward in time via the interior decay estimate from §2.2. The bound $Q \leqslant \overline{Q}$ will hold at all points on the comparison domain $\mathcal{N}$ that are sufficiently far (in space and forward in time) from an extension cap, while the bound $Q^* \leqslant \overline{Q}^*$ will hold sufficiently far (in space) from the neck-like boundary of $\mathcal{N}$. It will follow that, for a good choice of parameters, at least one of these bounds holds at each point of the comparison domain $\mathcal{N}$. This fact will enable us to guarantee that $|h| \leqslant \eta_{\text{lin}}$ everywhere on $\mathcal{N}$.

Even though the bound $Q \leqslant \overline{Q}$ may not hold in the near future of an extension cap, it may be important that it holds at *some time* in the future, thus allowing us to control the comparison map near a future neck-like boundary component, as described in §2.3. In order to guarantee this bound near such neck-like boundary components, in the future of extension caps, we first ensure that the neck-like boundary and the extension caps of the comparison domain are sufficiently separated (in space and time). Then, we use the strong form of the interior decay estimate to show that a weak bound of the form $Q \leqslant W\overline{Q}$, $W \gg 1$, which holds after a Bryant extension process, improves as we move forward in time and eventually implies $Q \leqslant \overline{Q}$. This interior decay estimate relies on the fact that $|h| \leqslant \eta_{\text{lin}}$, which is guaranteed by the bound $Q^* \leqslant \overline{Q}^*$.

3. Organization of the paper

The theorems stated in the introduction are proven in §13. They are all consequences of a more technical stability theorem, Theorem 13.1, which first appears in §13. This theorem asserts the existence of a comparison map between two Ricci flow spacetimes, satisfying

a (large) number of geometric and analytic bounds. As explained in the overview in the preceding section, Theorem 13.1 is proven using a simultaneous induction argument, in which the domain of this comparison map and the comparison map itself are constructed. The induction step consists of two stages; the first one is concerned with the comparison domain, and the second with the comparison map. These two stages are implemented in §11 and §12, respectively, and the induction hypotheses are collected beforehand as a set of a-priori assumptions, in §7. Both induction steps are formulated using objects and terminology that are introduced in preliminaries sections, §5 and §6. The arguments in §12 rely on two main ingredients: the interior decay estimate, which is discussed in §9, and the Bryant extension principle, which is presented in §10. We will also make use of a number of technical tools, which appear in §8.

To facilitate readability and verifiability, we have made an effort to make the proof modular and hierarchical. This eliminates unnecessary interdependencies, and minimizes the number of details the reader must bear in mind at any given stage of the proof. For instance, the two stages of the induction argument are formulated so as to be completely logically independent of each other. Also, within §12, which constructs the inductive extension of the comparison map, the argument is split into several pieces, which have been made as independent of one another as possible.

4. Conventions

4.1. Orientability

Throughout the paper we impose a blanket assumption that all 3-manifolds are orientable. The results remain true without this assumption — for instance Theorem 1.1 can be deduced from the orientable case by passing to the orientation cover. However, proving the main result without assuming orientability would complicate the exposition by increasing the number of special cases in many places. It is fairly straightforward, albeit time consuming, to modify the argument to obtain this extra generality.

4.2. Conventions regarding parameters

The statements of the a-priori assumptions in §7 involve a number of parameters, which will have to be chosen carefully. We will not assume these parameters to be fixed throughout the paper; instead, in each theorem, lemma or proposition we will include a list of restrictions on these parameters that serve as conditions for the hypothesis to hold. These restrictions state that certain parameters must be bounded from below or above by functions depending on certain other parameters. When we prove the main stability result,

Theorem 13.1, by combining our two main propositions, Propositions 11.1 and 12.1, we will need to verify that these restrictions are compatible with one another. This can be verified most easily via the parameter order, as introduced and discussed in §7.5. This parameter order is chosen in such a way that the required bounds from below/above on each parameter, if any, are given by a function depending on parameters that precede it. Hence, in order to verify the compatibility of all restrictions, it suffices to check that each parameter restriction is compatible with the parameter order in this way.

Throughout the entire paper, we will adhere to the convention that small (greek or arabic) letters stand for parameters that have to be chosen small enough and capital (greek or arabic) letters stand for parameters that have to be chosen sufficiently large. When stating theorems, lemmas or propositions, we will often express restrictions on parameters in the form

$$y \leqslant \bar{y}(x) \quad \text{and} \quad Z \geqslant \underline{Z}(x).$$

By this we mean that there are constants $\bar{y}$ and $\underline{Z}$, depending only on x such that if $0<y\leqslant\bar{y}$ and $Z\geqslant\underline{Z}$, then the subsequent statements hold. Furthermore, in longer proofs, we will introduce a restriction on parameters in the same form as a displayed equation. This makes it possible for the reader to check quickly that these restrictions are accurately reflected in the preamble of the theorem, lemma or proposition. Therefore, she/he may direct their full attention to the remaining details of the proof during the first reading.

5. Preliminaries I

In the following we define most of the notions that are needed in the statement of the main results of this paper, as stated in §1.3.

Definition 5.1. (Ricci flow spacetimes) A *Ricci flow spacetime* is a tuple $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ with the following properties:

(1) $\mathcal{M}$ is a smooth 4-manifold with (smooth) boundary $\partial\mathcal{M}$.

(2) $\mathfrak{t}\colon \mathcal{M}\to[0,\infty)$ is a smooth function without critical points (called *time function*). For any $t\geqslant 0$ we denote by $\mathcal{M}_t:=\mathfrak{t}^{-1}(t)\subset\mathcal{M}$ the *time-t slice* of $\mathcal{M}$.

(3) $\mathcal{M}_0=\mathfrak{t}^{-1}(0)=\partial\mathcal{M}$, i.e. the initial time-slice is equal to the boundary of $\mathcal{M}$.

(4) $\partial_{\mathfrak{t}}$ is a smooth vector field (the *time vector field*), which satisfies $\partial_{\mathfrak{t}}\mathfrak{t}\equiv 1$.

(5) g is a smooth inner product on the spatial sub-bundle $\ker(d\mathfrak{t})\subset T\mathcal{M}$. For any $t\geqslant 0$ we denote by g_t the restriction of g to the time-t-slice $\mathcal{M}_t$ (note that g_t is a Riemannian metric on $\mathcal{M}_t$).

(6) g satisfies the Ricci flow equation: $\mathcal{L}_{\partial_{\mathfrak{t}}} g=-2\,\mathrm{Ric}(g)$. Here, $\mathrm{Ric}(g)$ denotes the symmetric $(0,2)$-tensor on $\ker(d\mathfrak{t})$ that restricts to the Ricci tensor of $(\mathcal{M}_t, g_t)$ for all

$t \geqslant 0$.

For any interval $I \subset [0,\infty)$ we also write $\mathcal{M}_I = \mathfrak{t}^{-1}(I)$ and call this subset the *time-slab* of $\mathcal{M}$ over the time-interval I. Curvature quantities on $\mathcal{M}$, such as the Riemannian curvature tensor Rm, the Ricci curvature Ric, or the scalar curvature R will refer to the corresponding quantities with respect to the metric g_t on each time-slice. Tensorial quantities will be imbedded using the splitting $T\mathcal{M} = \ker(d\mathfrak{t}) \oplus \langle \partial_{\mathfrak{t}} \rangle$.

When there is no chance of confusion, we will sometimes abbreviate the tuple $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ by $\mathcal{M}$.

Ricci flow spacetimes were introduced by Lott and the second author (see [KL1]). The definition above is almost verbatim that of [KL1] with the exception that we require Ricci flow spacetimes to have initial time-slice at time zero and no final time-slice. This can always be achieved by applying a time-shift and removing the final time-slice from $\mathcal{M}$. Ricci flows with surgery, as constructed by Perelman in [P2], can be turned easily into Ricci flow spacetimes by removing a relatively small subset of surgery points. See (1.1) in §1.2 for further explanation.

We emphasize that, while a Ricci flow spacetime may have singularities — in fact the sole purpose of our definition is to understand flows with singularities — such singularities are not directly captured by a Ricci flow spacetime, as "singular points" are not contained in the spacetime manifold $\mathcal{M}$. Instead, the idea behind the definition of a Ricci flow spacetime is to understand a possibly singular flow by analyzing its asymptotic behavior on its regular part.

Any (classical) Ricci flow of the form $(g_t)_{t \in [0,T)}$, $0 < T \leqslant \infty$ on a 3-manifold M can be converted into a Ricci flow spacetime by setting $\mathcal{M} = M \times [0,T)$, letting $\mathfrak{t}$ be the projection to the second factor and letting $\partial_{\mathfrak{t}}$ correspond to the unit vector field on $[0,T)$. Vice versa, if $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is a Ricci flow spacetime with $\mathfrak{t}(\mathcal{M}) = [0,T)$ for some $0 < T \leqslant \infty$ and the property that every trajectory of $\partial_{\mathfrak{t}}$ is defined on the entire time-interval $[0,T)$, then $\mathcal{M}$ comes from such a classical Ricci flow.

We now generalize some basic geometric notions to Ricci flow spacetimes.

Definition 5.2. (Length, distance and metric balls in Ricci flow spacetimes) Let $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ be a Ricci flow spacetime. For any two points $x, y \in \mathcal{M}_t$ in the same time-slice of $\mathcal{M}$ we denote by $d(x,y)$ or $d_t(x,y)$ the *distance* between x and y within $(\mathcal{M}_t, g_t)$. The distance between points in different time-slices is not defined.

Similarly, we define the *length* $\mathrm{length}(\gamma)$ or $\mathrm{length}_t(\gamma)$ of a path $\gamma \colon [0,1] \to \mathcal{M}_t$ whose image lies in a single time-slice to be the length of this path when viewed as a path inside the Riemannian manifold $(\mathcal{M}_t, g_t)$.

For any $x \in \mathcal{M}_t$ and $r \geqslant 0$ we denote by $B(x,r) \subset \mathcal{M}_t$ the *r-ball* around x with respect

to the Riemannian metric g_t.

Our next goal is to characterize the (microscopic) geometry of a Ricci flow spacetime near a singularity or at an almost singular point. For this purpose, we will introduce a (*curvature*) *scale function* $\rho: \mathcal{M} \to (0, \infty]$ with the property that

$$C^{-1}\rho^{-2} \leqslant |\mathrm{Rm}| \leqslant C\rho^{-2} \tag{5.1}$$

for some universal constant $C < \infty$. The quantity ρ will be a (pointwise) function of the curvature tensor and therefore it can also be defined on (3-dimensional) Riemannian manifolds. For the purpose of this section, it suffices to assume that $\rho = |\mathrm{Rm}|^{-1/2}$. However, in order to simplify several proofs in subsequent sections, we will work with a slightly more complicated definition of ρ, which we will present in §6.1 (see Definition 6.1). Nonetheless, the discussion in the remainder of this subsection and the main results of the paper, as presented in §1.3, remain valid for any definition of ρ that satisfies (5.1).

We now define what we mean by completeness for Ricci flow spacetimes. Intuitively, a Ricci flow spacetime is called complete if its time-slices can be completed by adding countably many "singular points" and if no component appears or disappears suddenly without the formation of a singularity.

Definition 5.3. (Completeness of Ricci flow spacetimes) We say that a Ricci flow spacetime $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is (r_0, t_0)-*complete*, for some $r_0, t_0 \geqslant 0$, if the following holds: Consider a path $\gamma: [0, s_0) \to \mathcal{M}_{[0,t_0]}$ such that $\inf_{s \in [0,s_0)} \rho(\gamma(s)) > r_0$ for all $s \in [0, s_0)$ and

(1) its image $\gamma([0, s_0))$ lies in a time-slice $\mathcal{M}_t$ and the time-t length of γ is finite or

(2) γ is a trajectory of $\partial_{\mathfrak{t}}$ or of $-\partial_{\mathfrak{t}}$.

Then, the limit $\lim_{s \nearrow s_0} \gamma(s)$ exists.

If $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is (r_0, t_0)-complete for all $t_0 \geqslant 0$, then we also say that it is r_0-*complete*. Likewise, if $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is *zero-complete,* then we say that it is *complete*.

Note that the Ricci flow spacetimes constructed [KL2] are zero-complete, see [KL2, Proposition 5.11 (a) and Definition 1.8]. A Ricci flow with surgery and δ-cutoff, as constructed by Perelman in [P2], can be turned into a Ricci flow spacetime as in (1.1) that is $c\delta r$-complete for some universal constant $c > 0$, as long as the cutoff is performed in an appropriate way,[1] see [KL2, §3].

[1] As Perelman's objective was the characterization of the underlying topology, he allowed (but did not require) the removal of macroscopic spherical components during a surgery step. In contrast, Kleiner and Lott's version (cf. [KL1]) of the cutoff process does not allow this. However, both cutoff approaches allow some flexibility on the choice of the cutoff spheres inside the ε-horns. Some of these choices may result in the removal of points of scale larger than $c\delta r$; in such a case $c\delta r$-completeness cannot be guaranteed. Nevertheless, in both approaches it is always possible to perform the cutoff in such a way that the resulting Ricci flow spacetime is $c\delta r$-complete.

Lastly, we need to characterize the asymptotic geometry of a Ricci flow spacetime near its singularities. This is done by the canonical neighborhood assumption, a notion which is inspired by Perelman's work ([P2]) and which appears naturally in the study of 3-dimensional Ricci flows. The idea is to impose the same asymptotic behavior near singular points in Ricci flow spacetimes as is encountered in the singularity formation of a classical (smooth) 3-dimensional Ricci flow. The same characterization also holds in high-curvature regions of Perelman's Ricci flow with surgery that are far enough from "man-made" surgery points. Furthermore, an even stronger asymptotic behavior was shown to hold on Ricci flow spacetimes as constructed by Lott and the second author in [KL2].

The singularity formation in 3-dimensional Ricci flows is usually understood via singularity models called κ-solutions (see [P1, §11]). The definition of a κ-solution consists of a list of properties that are known to be true for 3-dimensional singularity models. Interestingly, these properties are sufficient to allow a qualitative (and sometimes quantitative) analysis of κ-solutions. We refer the reader to Appendix C and to [P2] and [KL1] for further details.

Let us recall the definition of a κ-solution.

Definition 5.4. (κ-solution) An ancient Ricci flow

$$(M, (g_t)_{t\in(-\infty,0]})$$

on a 3-dimensional manifold M is called a (3-*dimensional*) κ-*solution*, for $\kappa>0$, if the following holds:

(1) (M, g_t) is complete for all $t\in(-\infty, 0]$,

(2) $|\mathrm{Rm}|$ is bounded on $M\times I$ for all compact $I\subset(-\infty, 0]$,

(3) $\sec_{g_t}\geqslant 0$ on M for all $t\in(-\infty, 0]$,

(4) $R>0$ on $M\times(-\infty, 0]$,

(5) (M, g_t) is κ-non-collapsed at all scales for all $t\in(-\infty, 0]$.

(This means that for any $(x, t)\in M\times(-\infty, 0]$ and any $r>0$ if $|\mathrm{Rm}|\leqslant r^{-2}$ on the time-t ball $B(x, t, r)$, then we have $|B(x, t, r)|\geqslant\kappa r^n$ for its volume.)

We will compare the local geometry of a Ricci flow spacetime to the geometry of κ-solution using the following concept of pointed closeness.

Definition 5.5. (Geometric closeness) We say that a pointed Riemannian manifold (M, g, x) is ε-*close* to another pointed Riemannian manifold $(\overline{M}, \bar{g}, \bar{x})$ *at scale* $\lambda>0$ if there is a diffeomorphism onto its image

$$\psi\colon B^{\overline{M}}(\bar{x}, \varepsilon^{-1}) \longrightarrow M$$

such that $\psi(\bar{x})=x$ and

$$\|\lambda^{-2}\psi^*g-\bar{g}\|_{C^{[\varepsilon^{-1}]}(B^{\overline{M}}(\bar{x},\varepsilon^{-1}))}<\delta.$$

Here the $C^{[\varepsilon^{-1}]}$-norm of a tensor h is defined to be the sum of the C^0-norms of the tensors h, $\nabla^{\bar{g}}h$, $\nabla^{\bar{g},2}h$, ..., $\nabla^{\bar{g},[\varepsilon^{-1}]}h$ with respect to the metric $\bar{g}$.

We can now define the canonical neighborhood assumption. The main statement of this assumption is that regions of small scale (i.e. high curvature) are geometrically close to regions of κ-solutions.

Definition 5.6. (Canonical neighborhood assumption) Let (M,g) be a (possibly incomplete) Riemannian manifold. We say that (M,g) satisfies the *ε-canonical neighborhood assumption* at some point x if there is a $\kappa>0$, a κ-solution $(\overline{M},(\bar{g}_t)_{t\in(-\infty,0]})$ and a point $\bar{x}\in\overline{M}$ such that $\rho(\bar{x},0)=1$ and such that (M,g,x) is ε-close to $(\overline{M},\bar{g}_0,\bar{x})$ at some (unspecified) scale $\lambda>0$.

We say (M,g) satisfies the *ε-canonical neighborhood assumption* at scales (r_1,r_2), for some $0\leqslant r_1<r_2$, if every point $x\in M$ with $r_1<\rho(x)<r_2$ satisfies the ε-canonical neighborhood assumption.

We say that a Ricci flow spacetime $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ satisfies the *ε-canonical neighborhood assumption* at a point $x\in\mathcal{M}$ if the same is true at x in the time-slice $(\mathcal{M}_{\mathfrak{t}(x)},g_{\mathfrak{t}(x)})$. Moreover, we say that $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ satisfies the *ε-canonical neighborhood assumption* at scales (r_1,r_2) if the same is true for all its time-slices. Lastly, we say that a subset $X\subset\mathcal{M}$ satisfies the *ε-canonical neighborhood assumption* at scales (r_1,r_2), if the ε-canonical neighborhood assumption holds at all $x\in X$ with $\rho(x)\in(r_1,r_2)$.

Note that if $\mathcal{M}$ is a Ricci flow spacetime as constructed in [KL2], then $\mathcal{M}_{[0,T]}$ satisfies the ε-canonical neighborhood assumption at scales $(0,r)$, where $r=r(\varepsilon,T)>0$ [KL2, Theorem 1.3 and Proposition 5.30]. If $\mathcal{M}$ is the Ricci flow spacetime of a Ricci flow with surgery and δ-cutoff, as constructed by Perelman in [P2], then $\mathcal{M}$ satisfies the ε-canonical neighborhood assumption at $x\in\mathcal{M}$, provided the scale of x lies in the interval $(10h,r)$. Here $h=h(\varepsilon,t)$ and $r(\varepsilon,t)$ are decreasing functions of time, which appear in Perelman's construction, $h\leqslant\delta^2 r$, and $\delta=\delta(\varepsilon,t)$ may be chosen as small as desired.

Observe that we do not assume a global lower bound on κ in Definition 5.6. This slight generalization from other notions of the canonical neighborhood assumption does not create any serious issues, since by Perelman's work [P2], every 3-dimensional κ-solution is a κ_0-solution for some universal $\kappa_0>0$, unless it homothetic to a quotient of a round sphere (see assertion (a) of Lemma C.1 for further details).

We also remark that in Definition 5.5 we have put extra care in describing how the $C^{[\varepsilon^{-1}]}$-norm has to be understood. The reason for this is that the model metric $\bar{g}$ in Definition 5.6 is not fixed. So, it would be problematic, for example, to define the

$C^{[\varepsilon^{-1}]}$-norm using coordinate charts on $\overline{M}$, as the number and sizes of those coordinate charts may depend on the Riemannian manifold $(\overline{M}, \bar{g})$.

It may seem more standard to require spacetime closeness to a κ-solution on a backwards parabolic neighborhood — as opposed to closeness on a ball in a single time-slice — in the definition of the canonical neighborhood assumption. Such a condition would be stronger and, as our goal is to establish a uniqueness property, it would lead to a formally less general statement. We point out that spacetime closeness to a κ-solution is a rather straightforward consequence of time-slice closeness. The main purpose of the use of time-slice closeness in our work is because our uniqueness property also applies to Ricci flow spacetime with singular initial data. For this reason, the canonical neighborhood assumption also has to be applicable to the initial time-slice $\mathcal{M}_0$ or to time-slices $\mathcal{M}_t$ for small t.

6. Preliminaries II

In this section we present basic definitions and concepts that will be important for the proofs of the main results of this paper.

6.1. Curvature scale

As mentioned in §5, we will now define a notion of a curvature scale ρ that will be convenient for our proofs. The main objective in our definition will be to ensure that

$$\rho = \left(\tfrac{1}{3}R\right)^{-1/2}$$

wherever the sectional curvature is almost positive. For this purpose, observe that there is a constant $c_0>0$ such that the following holds. Whenever Rm is an algebraic curvature tensor with the property that its scalar curvature R is positive and all its sectional curvatures are bounded from below by $-\frac{1}{10}R$, then $c_0|\mathrm{Rm}|\leqslant\frac{1}{3}R$. We will fix c_0 for the remainder of this paper.

Definition 6.1. (Curvature scale) Let (M, g) be a 3-dimensional Riemannian manifold and $x\in M$ be a point. We define the (*curvature*) *scale* at x to be

$$\rho(x) = \min\left\{\left(\tfrac{1}{3}R_+(x)\right)^{-1/2}, (c_0|\mathrm{Rm}|(x))^{-1/2}\right\}. \tag{6.1}$$

Here, $R_+(x):=\max\{R(x), 0\}$ and we use the convention $0^{-1/2}=\infty$.

If $r_0>0$, then we set $\rho_{r_0}(x):=\min\{\rho(x), r_0\}$. Lastly, if $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ is a Ricci flow spacetime, then we define $\rho, \rho_{r_0}\colon\mathcal{M}\to\mathbb{R}$ such that they restrict to the corresponding scale functions on the time-slices.

LEMMA 6.2. *There is a universal constant $C<\infty$ such that*

$$C^{-1}\rho^{-2}(x) \leqslant |\mathrm{Rm}|(x) \leqslant C\rho^{-2}(x). \tag{6.2}$$

Moreover, there is a universal constant $\varepsilon_0>0$ such that if x satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption for some $\varepsilon_{\mathrm{can}}\leqslant\varepsilon_0$, then $R(x)=3\rho^{-2}(x)$.

Proof. The bound (6.2) is obvious. For the second part of the lemma observe that, for sufficiently small $\varepsilon_{\mathrm{can}}$, we have $R(x)>0$ and $\sec\geqslant-\frac{1}{10}R(x)$ at x. So,

$$\left(\tfrac{1}{3}R_+(x)\right)^{-1/2} \leqslant (c_0|\mathrm{Rm}|(x))^{-1/2}. \qquad \square$$

The normalization constant $\frac{1}{3}$ in front of the scalar curvature in (6.1) is chosen purely for convenience. More specifically, we will frequently consider the following round shrinking cylinder evolving by Ricci flow:

$$\left(S^2\times\mathbb{R}, \left(g_t=\left(\tfrac{2}{3}-2t\right)g_{S^2}+g_{\mathbb{R}}\right)_{t\in(-\infty,1/3]}\right).$$

The scale of this cylinder and the normalization of the curvature scale have been chosen in such a way that $\rho(\cdot,0)\equiv1$ and $\rho(\cdot,-1)\equiv2$ hold, which can be remembered easily; more generally, we have

$$\rho(\cdot,t)\equiv\sqrt{1-3t}.$$

Definition 6.3. ((Weakly) thick and thin subsets) Let X be a subset of a Riemannian manifold (M,g) or Ricci flow spacetime $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ and $r>0$ a number. We say that X is *r-thick* if $\rho(X)>r$ and *weakly r-thick* if $\rho(X)\geqslant r$. Similarly, we say that X is *r-thin* or *weakly r-thin* if $\rho(X)<r$ or $\rho(X)\leqslant r$, respectively.

6.2. Basic facts about the Bryant soliton

In the following, we will denote by $(M_{\mathrm{Bry}},(g_{\mathrm{Bry},t})_{t\in\mathbb{R}})$ the Bryant soliton and with tip $x_{\mathrm{Bry}}\in M_{\mathrm{Bry}}$ normalized in such a way that $\rho(x_{\mathrm{Bry}})=1$. The Bryant soliton was first constructed [Bry]. A more elementary construction can also be found in [Ap2]. Recall that $(M_{\mathrm{Bry}},(g_{\mathrm{Bry},t})_{t\in\mathbb{R}})$ is a steady gradient soliton all whose time-slices are rotationally symmetric with center x_{Bry}. More specifically, $(M_{\mathrm{Bry}},g_{\mathrm{Bry},t})$ can be expressed as a warped product of the form

$$g_{\mathrm{Bry},t}=d\sigma^2+w_t^2(\sigma)g_{S^2},$$

where $w_t(\sigma)\sim\sqrt{\sigma}$ for large σ. We refer to Lemma B.1 for a more extensive list of properties of the Bryant soliton that are being used in this paper. Note that, due to

the normalization of M_{Bry}, the definition of ρ and (B.2) of Lemma B.1, we have $\rho\geqslant 1$ on M_{Bry}. This fact will be important in this paper.

We will set $g_{\mathrm{Bry}}:=g_{\mathrm{Bry},0}$ for the time-zero slice of the Bryant soliton. Furthermore, we will denote by $M_{\mathrm{Bry}}(r):=B(x_{\mathrm{Bry}},r)$ the r-ball around the tip with respect to g_{Bry} and for $0<r_1<r_2$, we will denote by $M_{\mathrm{Bry}}(r_1,r_2)$ the open (r_1,r_2)-annulus around x_{Bry}.

6.3. Geometry of Ricci flow spacetimes

The goal of this subsection is to introduce several notions that we will frequently use in order to describe points or subsets in Ricci flow spacetimes.

Definition 6.4. (Points in Ricci flow spacetimes) Let $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ be a Ricci flow spacetime and $x\in\mathcal{M}$ be a point. Set $t:=\mathfrak{t}(x)$. Consider the maximal trajectory $\gamma_x\colon I\to\mathcal{M}$, $I\subset[0,\infty)$ of the time-vector field $\partial_{\mathfrak{t}}$ such that $\gamma_x(t)=x$. Note that then $\mathfrak{t}(\gamma_x(t'))=t'$ for all $t'\in I$. For any $t'\in I$ we say that x *survives until time* t' and we write

$$x(t'):=\gamma_x(t').$$

Similarly, if $X\subset\mathcal{M}_t$ is a subset in the time-t slice, then we say that X *survives until time* t' if this is true for every $x\in X$ and we set $X(t'):=\{x(t'):x\in X\}$.

We will also use the following two notions.

Definition 6.5. (Time-slice of a subset) Let $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ be a Ricci flow spacetime and let $X\subset\mathcal{M}$ be a subset. For any time $t\in[0,\infty)$ we define the *time-t slice* of X to be $X_t:=X\cap\mathcal{M}_t$ and for any interval $I\subset[0,\infty)$ we define the *I-time slab* of X to be $X_I:=X\cap\mathcal{M}_I$.

Definition 6.6. (Product domain) Let $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ be a Ricci flow spacetime and let $X\subset\mathcal{M}$ be a subset. We call X a *product domain* if there is an interval $I\subset[0,\infty)$ such that for any $t\in I$ any point $x\in X$ survives until time t and $x(t)\in X$.

Note that a product domain X can be identified with the product $X_{t_0}\times I$ for an arbitrary $t_0\in I$. If X_{t_0} is sufficiently regular (e.g. open or a domain with smooth boundary in $\mathcal{M}_{t_0}$), then the metric g induces a classical Ricci flow $(g_t)_{t\in I}$ on X_{t_0}. We will often use the metric g and the Ricci flow $(g_t)_{t\in I}$ synonymously when our analysis is restricted to a product domain.

Definition 6.7. (Parabolic neighborhood) Let $(\mathcal{M},\mathfrak{t},\partial_{\mathfrak{t}},g)$ be a Ricci flow spacetime. For any $y\in\mathcal{M}$ let $I_y\subset[0,\infty)$ be the set of all times until which y survives. Now consider

a point $x \in \mathcal{M}$ and two numbers $a \geqslant 0$, $b \in \mathbb{R}$. Set $t := \mathfrak{t}(x)$. Then, we define the *parabolic neighborhood* $P(x,a,b) \subset \mathcal{M}$ as follows:

$$P(x,a,b) := \bigcup_{y \in B(x,a)} \bigcup_{t' \in [t,t+b] \cap I_y} y(t').$$

If $b<0$, then we replace $[t,t+b]$ by $[t+b,t]$. We call $P(x,a,b)$ *unscathed* if $B(x,a)$ is relatively compact in $\mathcal{M}_t$ and if $I_y \supset [t,t+b]$ or $I_y \supset [t+b,t] \cap [0,\infty)$ for all $y \in B(x,a)$. Lastly, for any $r>0$ we introduce the simplified notation

$$P(x,r) := P(x,r,-r^2)$$

for the (*backward*) *parabolic ball* with center x and radius r.

Note that, if $P(x,a,b)$ is unscathed, then it is a product domain of the form $B(x,a) \times I_y$ for any $y \in B(x,a)$. We emphasize that $P(x,a,b)$ can be unscathed even if $t+b<0$, that is when it hits the initial time-slice earlier than expected. So, an unscathed parabolic neighborhood is not necessarily of the form $B(x,a) \times [t+b,t]$ if $b<0$.

6.4. Necks

Borrowing from Definition 5.5, we will introduce the notion of a δ-neck.

Definition 6.8. (δ-neck) Let (M,g) be a Riemannian manifold and $U \subset M$ an open subset. We say that U is a *δ-neck at scale $\lambda>0$* if there is a diffeomorphism

$$\psi \colon S^2 \times (-\delta^{-1}, \delta^{-1}) \longrightarrow U$$

such that

$$\big\| \lambda^{-2} \psi^* g - \big(\tfrac{2}{3} g_{S^2} + g_{\mathbb{R}} \big) \big\|_{C^{[\delta^{-1}]}(S^2 \times (-\delta^{-1}, \delta^{-1}))} < \delta.$$

We call the image $\psi(S^2 \times \{0\})$ a *central 2-sphere of U* and every point on a central 2-sphere a *center of U*.

Note that by our convention (see Definition 6.1) we have $\rho \equiv 1$ on $\big(S^2 \times \mathbb{R}, \frac{2}{3} g_{S^2} + g_{\mathbb{R}}\big)$. So, on a δ-neck at scale λ we have $\rho \approx \lambda$, where the accuracy depends on the smallness of δ. We also remark that a δ-neck U has infinitely many central 2-spheres, as we may perturb ψ slightly. This is why we speak of *a* central 2-sphere of U, as opposed to *the* central 2-sphere. Similarly, the centers of U are not unique, but form an open subset of U.

6.5. Ricci–DeTurck flow and harmonic map heat flow

In this subsection we recall some of the basic facts about the harmonic map heat flow and the Ricci–DeTurck flow equation in the classical setting, which were first observed by DeTurck [DT] and Hamilton [Ha4, §6]. More details, including precise statements of short-time existence and regularity of these flows, can be found in Appendix A.

Consider two n-dimensional manifolds M and M', each equipped with a smooth family of Riemannian metrics $(g_t)_{t\in[0,T]}$, $(g'_t)_{t\in[0,T]}$. Let moreover $(\chi_t)_{t\in[0,T]}$, $\chi_t\colon M'\to M$ be a smooth family of maps.

Definition 6.9. We say that the family $(\chi_t)_{t\in[0,T]}$ moves by *harmonic map heat flow between* (M',g'_t) *and* (M,g_t) if it satisfies the following evolution equation:

$$\partial_t\chi_t=\Delta_{g'_t,g_t}\chi_t=\sum_{i=1}^n(\nabla^{g_t}_{d\chi_t(e_i)}d\chi_t(e_i)-d\chi_t(\nabla^{g'_t}_{e_i}e_i)),\tag{6.3}$$

where $\{e_i\}_{i=1}^n$ is a local frame on M' that is orthonormal with respect to g'_t.

Assume now for the remainder of this subsection that $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$ evolve by the Ricci flow equations

$$\partial_t g_t=-2\mathrm{Ric}_{g_t}\quad\text{and}\quad\partial_t g'_t=-2\mathrm{Ric}_{g'_t}.$$

Furthermore, assume for the rest of this subsection that all the maps χ_t are diffeomorphisms and consider their inverses $\phi_t:=\chi_t^{-1}$. A basic calculation (see Appendix A for more details) reveals that the pullback $g^*_t:=\phi_t^*g'_t$ evolves by the *Ricci–DeTurck flow* equation

$$\partial_t g^*_t=-2\mathrm{Ric}_{g^*_t}-\mathcal{L}_{X_{g_t}(g^*_t)}g^*_t,\tag{6.4}$$

where the vector field $X_{g_t}(g^*_t)$ is defined by

$$X_{g_t}(g^*_t):=\Delta_{g^*_t,g_t}\mathrm{id}_M=\sum_{i=1}^n(\nabla^{g_t}_{e_i}e_i-\nabla^{g^*_t}_{e_i}e_i),\tag{6.5}$$

for a local frame $\{e_i\}_{i=1}^n$ that is orthonormal with respect to g^*_t.

The advantage of the Ricci–DeTurck flow equation over the Ricci flow equation is that it is a non-linear, strongly parabolic equation in the metric g^*_t. More specifically, if we express g^*_t in terms of the perturbation $h_t:=g^*_t-g_t$, then (6.4) becomes the *Ricci–DeTurck flow equation for perturbations*

$$\nabla_{\partial_t}h_t=\Delta_{g_t}h_t+2\,\mathrm{Rm}_{g_t}(h_t)+\mathcal{Q}_{g_t}[h_t].\tag{6.6}$$

Here we view g_t as a background metric. All curvature quantities and covariant derivatives are taken with respect to g_t. On the left-hand side of (6.6), we moreover use Uhlenbeck's trick:

$$(\nabla_{\partial_t} h_t)_{ij} = (\partial_t h_t)_{ij} + g_t^{pq}(\mathrm{Ric}^{g_t}_{pj}(h_t)_{iq} + \mathrm{Ric}^{g_t}_{ip}(h_t)_{qj}).$$

The expressions on the right-hand side of (6.6) are to be interpreted as follows:

$$(\mathrm{Rm}_{g_t}(h_t))_{ij} = g_t^{pq} R_{pij}{}^{u}(h_t)_{qu}$$

and $\mathcal{Q}_{g_t}[h_t]$ is an algebraic expression in g_t, h_t, ∇h_t, $\nabla^2 h_t$ of the form

$$\begin{aligned}\mathcal{Q}_{g_t}[h_t] = {} & (g_t+h_t)^{-1} * (g_t+h_t)^{-1} * \nabla h_t * \nabla h_t \\ & + (g_t+h_t)^{-1} * (g_t+h_t)^{-1} * \mathrm{Rm}_{g_t} * h_t * h_t + (g_t+h_t)^{-1} * (g_t+h_t)^{-1} * h_t * \nabla^2 h_t.\end{aligned}$$

See (A.10) in Appendix A for an explicit formula for $\mathcal{Q}_{g_t}$. The precise structure of the quantity $\mathcal{Q}_{g_t}$ will, however, not be of essence in this paper.

We remark that in the classical setting and in the compact case, the uniqueness of solutions to the Ricci flow equation follows from the existence of solutions to (6.3) and the uniqueness of solutions to (6.6). More specifically, for any two Ricci flows $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$ on M and M' for which there is an isometry $\chi\colon M' \to M$ with $\chi^* g_0 = g'_0$ one first constructs a solution $(\chi_t)_{t\in[0,\tau)}$ of (6.3), for some maximal $\tau < T$, with initial condition $\chi_0 = \chi$. The resulting perturbation $h_t = \phi_t^* g'_t - g_t$, for $\phi_t = \chi_t^{-1}$, solves (6.6), as long as it is well defined. As $h_0 \equiv 0$, we obtain by uniqueness that $h_t \equiv 0$, as long as it is defined. It then follows that χ_t is an isometry for all $t \in [0,T] = [0,\tau]$ and by (6.3) that $\partial_t \chi_t \equiv 0$.

In this paper we will mostly analyze solutions h_t to (6.6) of small norm. Via a limit argument, such solutions can be understood in terms the *linearized Ricci–DeTurck* equation

$$\nabla_{\partial_t} h'_t = \Delta_{g_t} h'_t + 2\,\mathrm{Rm}_{g_t}(h'_t).$$

For more details on this, see §9.

6.6. Maps between Ricci flow spacetimes

In this subsection consider two Ricci flow spacetimes $(\mathcal{M}, \mathfrak{t}, \partial_{\mathfrak{t}}, g)$ and $(\mathcal{M}', \mathfrak{t}', \partial_{\mathfrak{t}'}, g')$, which we will abbreviate in the following by $\mathcal{M}$ and $\mathcal{M}'$. Our goal will be to characterize maps between subsets of these spacetimes. Using the terminology introduced above, we will then generalize the notions introduced in the previous subsection to Ricci flow spacetimes.

Definition 6.10. (Time-preserving and time-equivariant maps) Let $X\subset\mathcal{M}$ be a subset and $\phi\colon X\to\mathcal{M}'$ be a map. We say that ϕ is *time-preserving* if $\mathfrak{t}'(\phi(x))=\mathfrak{t}(x)$ for all $x\in X$. We say that ϕ is *a-time-equivariant*, for some $a\in\mathbb{R}$, if there is some $t_0\in\mathbb{R}$ such that $\mathfrak{t}'(\phi(x))=a\mathfrak{t}(x)+t_0$ for all $x\in X$.

Observe that a time-preserving map is also 1-time-equivariant.

Definition 6.11. (Time-slices of a map) If $\phi\colon X\subset\mathcal{M}\to\mathcal{M}'$ is time-equivariant and $t\in[0,\infty)$ such that $X_t=X\cap\mathcal{M}_t\neq\varnothing$, then we denote by

$$\phi_t:=\phi|_{X_t}\colon X_t\longrightarrow\mathcal{M}'_{t'}\subset\mathcal{M}'$$

the *time-t slice* of ϕ. Here t' is chosen such that $\phi(X_t)\subset\mathcal{M}'_{t'}$.

Definition 6.12. ($\partial_{\mathfrak{t}}$-preserving maps) Let $\phi\colon X\to\mathcal{M}'$ be a differentiable map defined on a sufficiently regular domain $X\subset\mathcal{M}$. If $(d\phi)_*\partial_{\mathfrak{t}}=\partial_{\mathfrak{t}'}$, then we say that ϕ is *$\partial_{\mathfrak{t}}$-preserving.*

Note that the image of a product domain under a time-equivariant and $\partial_{\mathfrak{t}}$-preserving map is again a product domain.

Definition 6.13. (Harmonic map heat flow) Let $Y\subset\mathcal{M}'$ be a subset. We say that a map $\chi\colon Y\to\mathcal{M}$ evolves by *harmonic map heat flow* if it is 1-time-equivariant and if at all times $t,t'\in[0,\infty)$ with $Y_t\neq\varnothing$ and $\chi(Y_{t'})\subset\mathcal{M}'_t$ the identity

$$d\chi(\partial_{t'})=\partial_{\mathfrak{t}}+\Delta_{g'_t,g_t}\chi_t \tag{6.7}$$

holds on the interior of Y. The last term in this equation denotes the Laplacian of the map $\psi_t\colon(\mathcal{M}'_{t'},g'_{t'})\to(\mathcal{M}_t,g_t)$ (see (6.3) for further details).

It is not difficult to see that the notions of harmonic map heat flow in Definition 6.13 corresponds to Definition 6.9 in the case in which $\mathcal{M}$ and $\mathcal{M}'$ can be described in terms of classical Ricci flows $(M,(g_t)_{t\in I})$ and $(M',(g'_t)_{t\in I'})$, respectively. The same is true in the case in which χ is the inverse of a diffeomorphism $\phi\colon X\to Y\subset\mathcal{M}'$, where X is a product domain in $\mathcal{M}$ whose time-slices are domains with smooth boundary. In this case, which will be of main interest for us (see Definition 7.2), the equation (6.7) makes sense and holds, by continuity, on all of Y.

Next, we generalize the concept of Ricci–DeTurck flow to the setting of Ricci flow spacetimes.

Definition 6.14. Consider a smooth symmetric $(0,2)$-tensor field h on the sub-bundle $\ker(d\mathfrak{t})\subset T\mathcal{M}$ over a sufficiently regular domain $N\subset\mathcal{M}$ (in this paper we will only consider

the case in which N is a domain with smooth boundary or is a product domain whose time-slices are domains with smooth boundary). We say that h is a *Ricci–DeTurck perturbation* (on N) if

$$\mathcal{L}_{\partial_t}(g+h) = -2\mathrm{Ric}(g+h) - \mathcal{L}_{X_g(g+h)}(g+h), \tag{6.8}$$

where $X_g(g+h)$ is defined on each time-slice X_t as in (6.5).

If X is a product domain of the form $X' \times I$, and if we identify g and h with smooth families of the form $(g_t)_{t\in I}$ and $(h_t)_{t\in I}$, then (6.8) is equivalent to the classical Ricci–DeTurck equation (6.4).

The following lemma is an immediate consequence of our discussion from §6.5.

LEMMA 6.15. *Let $X \subset \mathcal{M}$ be open or a product domain whose time-slices are domains with smooth boundary and consider a diffeomorphism $\phi: X \to Y := \phi(X) \subset \mathcal{M}'$. Assume that the inverse map $\phi^{-1}: Y \to X$ evolves by harmonic map heat flow. Then, the perturbation $h := \phi^* g' - g$ is a Ricci–DeTurck perturbation in the sense of Definition 6.14.*

7. A-priori assumptions

In this section we introduce the objects and conditions that will be used to formulate and prove the main result (Theorem 13.1), which asserts the existence of a certain type of map between subsets of Ricci flow spacetimes. The domain of the map will be called a *comparison domain* (Definition 7.1), and the map itself a *comparison* (Definition 7.2). The comparison and its domain will be subject to a number of *a-priori assumptions* (Definitions 7.4 and 7.5). These definitions have been tailored to facilitate an existence proof by induction over time steps.

We recommend reading the overview in §2 prior to reading this section, because it provides motivation for the structures defined here, and gives some indication of the role they play in the proof. We refer the reader to §5 and §6 for the definitions relevant to this section.

7.1. Comparison domains

We begin with a definition that collects the qualitative features of the domain of our comparison map. Additional assumptions of a quantitative nature are imposed later, in the a-priori assumptions. Loosely speaking, a comparison domain is a sequence of product domains $\mathcal{N}^1, ..., \mathcal{N}^J$ defined on successive time-intervals, whose time-slices have spherical boundary (see Figure 8 for an illustration). One observes two types of behavior

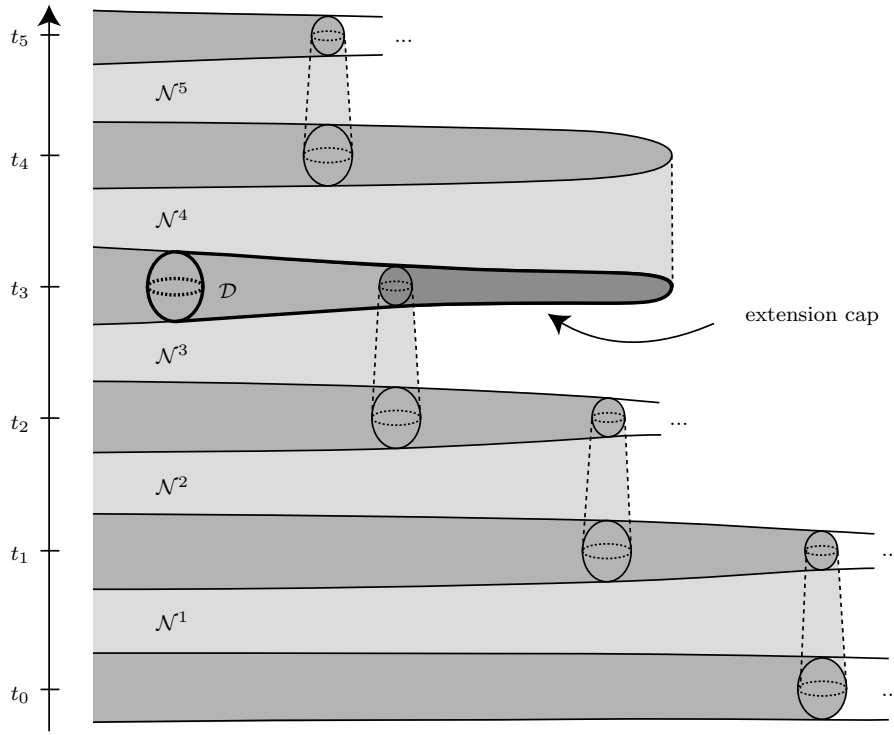

Figure 8. Example of a comparison domain defined over the time-interval $[0, t_5]$ and a cut. The dark shaded regions indicate the picture of the comparison domain at integral time-steps $t_0, ..., t_5$. The extension cap at time t_3 is shaded very dark. This extension cap is contained in a cut $\mathcal{D}$, which is outlined in bold. Note that cuts, such as $\mathcal{D}$, occur in the definition of a comparison, not of a comparison domain.

near the boundary as one transitions from one product domain to the next: boundary components can either "recede", or they can be filled in by 3-balls. In the main existence proof, the latter case corresponds to the situation when the comparison map is extended over a cap region lying in a subset that is approximated by a Bryant soliton; for this reason, we call the closures of such 3-balls extension caps.

Definition 7.1. (Comparison domain) A *comparison domain* (defined over the time-interval $[0, t_J]$) in a Ricci flow spacetime $\mathcal{M}$ is a triple $(\mathcal{N}, \{\mathcal{N}^j\}_{1 \leqslant j \leqslant J}, \{t_j\}_{j=0}^{J})$, where the following holds:

(1) The times $0 = t_0 < ... < t_J$ partition the time-interval $[0, t_J]$. Each $\mathcal{N}^j$ (for $1 \leqslant j \leqslant J$) is a subset of $\mathcal{M}_{[t_{j-1}, t_j]}$, and

$$\mathcal{N} = \bigcup_{1 \leqslant j \leqslant J} \mathcal{N}^j \subset \mathcal{M}_{[0, t_J]}.$$

(2) For all $1\leqslant j\leqslant J$ the subset $\mathcal{N}^j$ is closed in $\mathcal{M}$, and is a product domain, in the sense of Definition 6.6.

(3) For all $1\leqslant j<J$, we have $\partial\mathcal{N}^{j+1}_{t_j}\subset\operatorname{Int}\mathcal{N}^j_{t_j}$. Here $\operatorname{Int}\mathcal{N}^j_{t_j}$ denotes the interior of $\mathcal{N}^j_{t_j}$ inside $\mathcal{M}_{t_j}$. Consequently, the difference $\mathcal{N}^{j+1}_{t_j}\setminus\operatorname{Int}\mathcal{N}^j_{t_j}$ is a closed subset of $\mathcal{M}_{t_j}$ that is a domain with smooth boundary, with boundary contained in $\mathcal{N}^j_{t_j}$.

(4) For every $1\leqslant j<J$, the components of $\mathcal{N}^{j+1}_{t_j}\setminus\operatorname{Int}\mathcal{N}^j_{t_j}$ are 3-disks, which are called *extension caps*.

For any $t<t_J$, we define the forward time-t slice $\mathcal{N}_{t+}$ of $\mathcal{N}$ to be the set of accumulation points of $\mathcal{N}_{\bar t}$ as $\bar t\searrow t$, and if $t=t_J$ we define $\mathcal{N}_{t+}=\mathcal{N}_t$. We define the backward time-slices $\mathcal{N}_{t-}$ similarly, but taking accumulation points as $\bar t\nearrow t$, and when $t=0$, we put $\mathcal{N}_{0-}=\mathcal{N}_0$. Thus if $t\in(t_{j-1},t_j)$ then $\mathcal{N}_{t\pm}=\mathcal{N}^j_t$, $\mathcal{N}_{t_j-}=\mathcal{N}^j_{t_j}$ and $\mathcal{N}_{t_j+}=\mathcal{N}^{j+1}_{t_j}$ if $1\leqslant j<J$. Observe that $\mathcal{N}_t=\mathcal{N}_{t-}\cup\mathcal{N}_{t+}$.

In the case $J=0$ the comparison domain $(\mathcal{N}=\varnothing,\{\,\},\{t_0\})$ is called the *empty comparison domain*.

When there is no chance of confusion, we will sometimes abbreviate

$$(\mathcal{N},\{\mathcal{N}^j\}_{j=1}^J,\{t_j\}_{j=0}^J)$$

by $\mathcal{N}$.

7.2. Comparisons

Next, we collect the basic properties of our comparison maps between Ricci flow spacetimes. Roughly speaking, a comparison is a map between Ricci flow spacetimes that is defined on a comparison domain. Away from the transition times, the inverse of this map solves the harmonic map heat flow equation for the evolving metrics, or equivalently, the pullback metric satisfies the Ricci–DeTurck equation. At a transition time, the comparison is extended over the extension caps. In order to guarantee a good interpolation, it is necessary to adjust the comparison over a region that is much larger than the extension cap. As a consequence, the comparison, when viewed as a map between spacetimes, may have jump discontinuities near every extension cap. The discontinuity locus is contained in a disjoint union of closed disks, which we will call *cuts* (see Figure 8 for an illustration).

In the following definition, we allow a comparison to be defined on a shorter time-interval than the comparison domain. This is done for technical reasons having to do with a two part induction argument. More specifically, in §11, we will analyze a comparison that is defined on an entire comparison domain (over a time-interval $[0,t_J]$) and then extend the comparison domain by one time-step (to the time-interval $[0,t_{J+1}]$), without extending the comparison itself. So, we will end up with a comparison domain that is

defined up to some time t_{J+1}, while the comparison itself still remains defined only up to time t_J.

Definition 7.2. (Comparison) Let $\mathcal{M}$ and $\mathcal{M}'$ be Ricci flow spacetimes and consider a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ defined over the time-interval $[0, t_J]$ in $\mathcal{M}$.

A triple $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ is a *comparison from* $\mathcal{M}$ *to* $\mathcal{M}'$ defined on $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ (over the time-interval $[0, t_{J^*}]$) if the following holds:

(1) $J^* \leqslant J$.

(2) $\mathrm{Cut} = \mathrm{Cut}^1 \cup ... \cup \mathrm{Cut}^{J^*-1}$, where each Cut^j is a collection of pairwise disjoint 3-disks inside $\mathrm{Int}\,\mathcal{N}_{t_j+}$.

(3) Each $\mathcal{D} \in \mathrm{Cut}$ contains exactly one extension cap of the domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ and every extension cap of $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ that is contained in $\mathcal{M}_{[0,t_{J^*-1}]}$ is contained in one element of Cut.

(4) Each $\phi^j : \mathcal{N}^j \to \mathcal{M}'$ is a time-preserving diffeomorphism onto its image. More precisely, ϕ^j may be extended to a diffeomorphism onto its image defined on an open neighborhood of $\mathcal{N}^j$ in the manifold with boundary $\mathcal{M}_{[t_{j-1},t_j]}$.

(5) If $J^* \geqslant 1$, then $\phi : \bigcup_{j=1}^{J^*} \mathcal{N}^j \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D} \to \mathcal{M}'$ is a continuous map that is smooth on the interior of $\bigcup_{j=1}^{J^*} \mathcal{N}^j \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}$. If $J^* = 0$, then we assume that $\phi : \varnothing \to \varnothing$ is the trivial map.

(6) $\phi = \phi^j$ on the open time slab $\mathcal{N}^j_{(t_{j-1},t_j)}$ for all $j = 1, ..., J^*$.

(7) For all $j = 1, ..., J^*$, the inverse map $(\phi^j)^{-1} : \phi^j(\mathcal{N}^j) \to \mathcal{N}^j$ evolves by harmonic map heat flow (according to Definition 6.13).

We define ϕ_{t_j-} to be $\phi^j|_{\mathcal{N}^j_{t_j}}$ if $0 < j \leqslant J^*$ and ϕ^1_0 if $j = 0$. Similarly, we define ϕ_{t_j+} to be $\phi^j|_{\mathcal{N}^{j+1}_{t_j}}$ if $0 \leqslant j < J^*$ and $\phi^{J^*}|_{\mathcal{N}^{J^*}_{t_{J^*}}}$ if $j = J^*$.

We remark that Definition 7.2 implies that ϕ is injective, and that ϕ^{-1} satisfies the harmonic map heat flow equation everywhere it is defined.

Note that by Definition 7.2, the only comparison in the case $J^* = 0$ is the trivial comparison $(\mathrm{Cut} = \varnothing, \phi : \varnothing \to \varnothing, \varnothing)$.

As explained in §6.5, a map whose inverse is evolving by harmonic map heat flow induces a Ricci–DeTurck flow on its domain. We will now use this fact to define the Ricci–DeTurck perturbation associated with a comparison.

Definition 7.3. (Associated Ricci–DeTurck perturbation) Consider a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ in a Ricci flow spacetime $\mathcal{M}$ that is defined over the time-interval $[0, t_J]$ and a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ from $\mathcal{M}$ to $\mathcal{M}'$ defined on this domain over the time-interval $[0, t_{J^*}]$ for some $J^* \leqslant J$.

Define $h := \phi^* g' - g$ on $\mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}$ and $h^j := (\phi^j)^* g' - g$ on $\mathcal{N}^j$ for all $1 \leqslant j \leqslant J^*$.

Then, we say that $(h, \{h^j\}_{j=1}^{J^*})$ is the *associated Ricci–DeTurck perturbation* for

$$(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*}).$$

Moreover, for $1 \leqslant j \leqslant J^*$ we set $h_{t_j-} := h^j_{t_j}$, and define $h_{t_0-} := h^1_0$. Likewise, for $0 \leqslant j \leqslant J^*-1$ we set $h_{t_j+} := h^{j+1}_{t_j}$ and $h_{t_{J^*}+} = h^{J^*}_{t_{J^*}}$.

Note that by Lemma 6.15 the tensors h and h^j are Ricci–DeTurck perturbations in the sense of Definition 6.14.

7.3. A-priori assumptions I: the geometry of the comparison domain

Next, we introduce a-priori assumptions for a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ defined on a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J}, \{t_j\}_{j=0}^{J})$. We first state the first six a-priori assumptions, (APA1)–(APA6), which characterize the more geometric properties of the comparison domain and the comparison. These are the only a-priori assumptions needed to implement the first part of the main induction argument, in §11.

To make it easier to absorb the list of conditions, we make some informal preliminary remarks. The construction of the comparison domain and comparison involves a comparison scale r_{comp}. Most of the a-priori assumptions impose conditions at scales that are defined relative to r_{comp}. For instance, the final time-slice of each product domain $\mathcal{N}^j$ of the comparison domain is assumed to have boundary components that are central 2-spheres of necks at scale r_{comp}. Moreover, we assume the comparison domain to be $\lambda r_{\mathrm{comp}}$-thick and to contain all $\Lambda r_{\mathrm{comp}}$-thick points at integral time-slices. These and similar characterizations will be made in a-priori assumptions (APA1)–(APA3).

In addition, we impose two assumptions, (APA4) and (APA5), that restrict the situations when a component can be discarded or added, respectively. To appreciate the role of these two conditions, the reader may wish to imagine a scenario when a Bryant-like cap region in $\mathcal{M}$ evolves through a range of scales, initially well below $\lambda r_{\mathrm{comp}}$, then well above $\Lambda r_{\mathrm{comp}}$, possibly fluctuating between these over a time scale $\gg r^2_{\mathrm{comp}}$. Then, initially the cap region will lie outside the comparison domain, because its scale is too small, and later it will necessarily lie in the comparison domain, because it has scale $> \Lambda r_{\mathrm{comp}}$. A-priori assumptions (APA4) and (APA5) ensure that these events occurs when the tip of the cap has scale in the range approximately $(\lambda r_{\mathrm{comp}}, 10\lambda r_{\mathrm{comp}})$, and that they do not occur unnecessarily too often.

Finally, a-priori assumption (APA6) states that the comparison itself is an almost isometry of high enough precision.

We mention that a-priori assumptions (APA1)–(APA6) depend on a number of parameters, which will be chosen in the course of this paper. Also, as with Definition 7.2,

in the following definition we do not require a comparison to be defined on the entire comparison domain (see the discussion before Definition 7.2).

Definition 7.4. (A-priori assumptions (APA1)–(APA6)) Let $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ be a comparison domain in a Ricci flow spacetime $\mathcal{M}$ that is defined over the time-interval $[0, t_J]$ and consider a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ from $\mathcal{M}$ to $\mathcal{M}'$ on this domain to another Ricci flow spacetime $\mathcal{M}'$ that is defined over the time-interval $[0, t_{J^*}]$ for some $J^* \leqslant J$.

We say that $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ and $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ satisfy a-priori assumptions (APA1)–(APA6) with respect to the tuple of parameters

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}})$$

if the following holds:

(APA1) We have $t_j = j \cdot r_{\mathrm{comp}}^2$ for each $0 \leqslant j \leqslant J$.

(APA2) All points in $\mathcal{N}$ are $\lambda r_{\mathrm{comp}}$-thick.

(APA3) For every $1 \leqslant j \leqslant J$, the backward time-slice $\mathcal{N}_{t_j-} = \mathcal{N}_{t_j}^j$ has the following properties:

(a) The boundary components of $\mathcal{N}_{t_j-}$ are central 2-spheres of δ_{n}-necks at scale r_{comp}.

(b) $\mathcal{N}_{t_j-}$ contains all $\Lambda r_{\mathrm{comp}}$-thick points of $\mathcal{M}_{t_j}$.

(c) Each component of $\mathcal{N}_{t_j-}$ contains a $\Lambda r_{\mathrm{comp}}$-thick point.

(d) Each component of $\mathcal{M}_{t_j} \setminus \mathrm{Int}\, \mathcal{N}_{t_j-}$ with non-empty boundary contains a $10\lambda r_{\mathrm{comp}}$-thin point.

(e) The points on each cut $\mathcal{D} \in \mathrm{Cut}$ are $\Lambda r_{\mathrm{comp}}$-thin.

(APA4) (*Discarded disks become thin*) Suppose $1 \leqslant j \leqslant J$, and $\mathcal{C}$ is a component of $\mathcal{N}_{t_{j-1}-} \setminus \mathrm{Int}\, \mathcal{N}_{t_{j-1}+}$ (if $j \geqslant 2$) or $\mathcal{M}_0 \setminus \mathrm{Int}\, \mathcal{N}_{0+}$ (if $j = 1$) such that the following holds:

(a) $\mathcal{C}$ is diffeomorphic to a 3-disk.

(b) $\partial \mathcal{C} \subset \mathcal{N}_{t_{j-1}+}$.

Then, either $\mathcal{C}$ does not survive until time t_j (as in Definition 6.4), or for some time $t \in [t_{j-1}, t_j]$ we can find a weakly $\lambda r_{\mathrm{comp}}$-thin point on $\mathcal{C}(t)$ (recall the notation $\mathcal{C}(t)$ from Definition 6.4).

(APA5) (Geometry of extension caps) For each $1 \leqslant j \leqslant J^*$ and every component $\mathcal{C}$ of $\mathcal{M}_{t_j} \setminus \mathrm{Int}\, \mathcal{N}_{t_j-}$ the following holds: $\mathcal{C}$ is an extension cap of $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ if and only if there is a component $\mathcal{C}'$ of $\mathcal{M}'_{t_j} \setminus \phi_{t_j-}(\mathrm{Int}\, \mathcal{N}_{t_j-})$ such that the following holds:

(a) $\mathcal{C}$ and $\mathcal{C}'$ are 3-disks.
(b) $\partial\mathcal{C}' = \phi_{t_j-}(\partial\mathcal{C})$.
(c) There is a point $x \in \mathcal{C}$ such that $(\mathcal{M}_{t_j}, x)$ is δ_{b}-close to the pointed Bryant soliton $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at scale $10\lambda r_{\mathrm{comp}}$.
(d) There is a point $x' \in \mathcal{M}'_{t_j}$, at distance $\leqslant D_{\mathrm{cap}} r_{\mathrm{comp}}$ from $\mathcal{C}'$, such that $(\mathcal{M}'_{t_j}, x')$ is δ_{b}-close to the pointed Bryant soliton $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at some scale in the interval $[D_{\mathrm{cap}}^{-1} r_{\mathrm{comp}}, D_{\mathrm{cap}} r_{\mathrm{comp}}]$.
(e) $\mathcal{C}$ and $\mathcal{C}'$ have diameter $\leqslant D_{\mathrm{cap}} r_{\mathrm{comp}}$.

(APA6) Consider the Ricci–DeTurck perturbation $(h, \{h\}_{j=1}^{J^*})$ associated with the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$.

If $J^* \geqslant 1$, then $|h| \leqslant \eta_{\mathrm{lin}}$ on $\bigcup_{j=1}^{J^*} \mathcal{N}^j \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}$. Moreover, the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds at scales $(0, 1)$ on $\bigcup_{j=1}^{J^*} \phi^j(\mathcal{N}^j)$.

We point out that a-priori assumptions (APA1)–(APA4) are conditions on the comparison domain only. On the other hand, a-priori assumption (APA5) places restrictions on extension caps in terms of the comparison map and the local geometry of the image. This is to ensure that extension caps arise only when the geometry of the domain and target are nice enough to allow an extension of a comparison on that is a precise enough almost isometry.

7.4. A-priori assumptions II: analytic conditions on the comparison

Lastly, we introduce a further set of a-priori assumptions, (APA7)–(APA13), which characterize the behavior of the perturbation h and the geometry of the cuts more precisely. These assumptions will become important in §12, where we will extend the comparison by one time-step onto a larger comparison domain.

We now give a brief overview of a-priori assumptions (APA7)–(APA13). A-priori assumptions (APA7)–(APA10) impose global bounds on the Ricci–DeTurck perturbation h via two quantities Q and Q^*. These bounds essentially introduce a pointwise weight, which depends on the curvature scale ρ and time. A-priori assumption (APA7) imposes a bound on Q on the comparison domain, on the complement of forward parabolic neighborhoods of cuts. Similarly, a-priori assumption (APA9) imposes a bound on Q^* at points of the comparison domain that are far enough away from its neck-like boundary. For an illustration of the domains on which these bounds do or do not hold, see Figure 9. A-priori assumption (APA8) introduces a weaker bound on Q, which holds essentially everywhere on the comparison domain. Note that the constant W in this bound will be chosen to be large. Therefore, a-priori assumption (APA8) will not directly imply

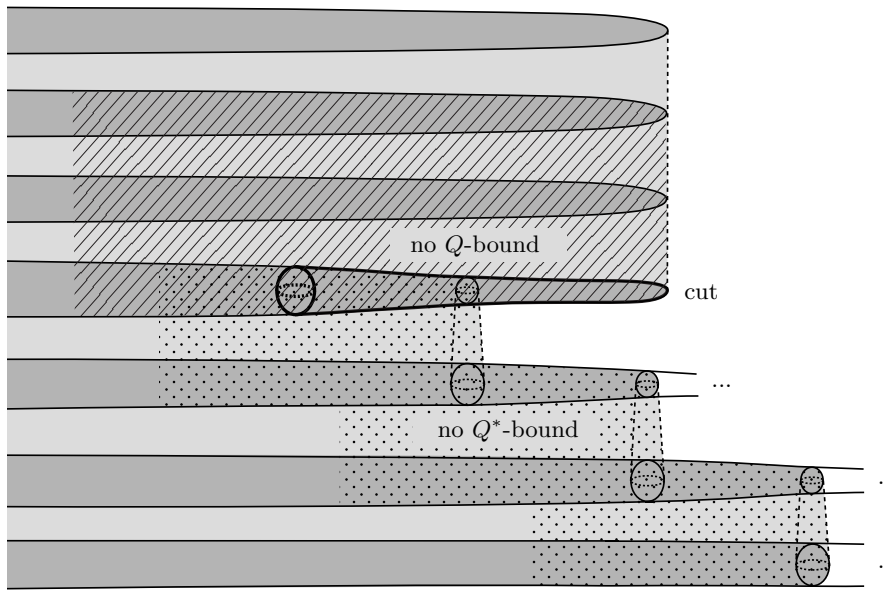

Figure 9. The bound $Q \leqslant \overline{Q}$ in (APA7) holds on all of $\mathcal{N}$ except for the hatched region. The bound $Q^* \leqslant \overline{Q}^*$ in (APA9) holds on all of $\mathcal{N}$ except for the dotted region.

a-priori assumption (APA7).

A-priori assumption (APA10) states that Q^* is small on each cut and a-priori assumption (APA12) guarantees a good bound on Q and Q^* on the initial time-slice. A-priori assumption (APA11) controls the geometry of the cuts. Lastly, a-priori assumption (APA13) imposes a bound on t_J.

Definition 7.5. (A-priori assumptions (APA7)–(APA13)) Let $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ be a comparison domain in a Ricci flow spacetime $\mathcal{M}$ that is defined on the time-interval $[0, t_J]$ and consider a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J^*})$ on this domain to another Ricci flow spacetime $\mathcal{M}'$ that is defined on the same time-interval $[0, t_J]$.

We say that $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ and $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^J)$ satisfy a-priori assumptions (APA7)–(APA13) with respect to the tuple of parameters

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}})$$

if the following holds. Define the functions

$$Q := e^{H(T-\mathfrak{t})}\rho_1^E|h| \quad \text{and} \quad Q^* := e^{H(T-\mathfrak{t})}\rho_1^3|h|$$

on $\mathcal{N} \backslash \bigcup_{\mathcal{D}\in\mathrm{Cut}} \mathcal{D}$, where $\mathfrak{t}\colon \mathcal{M} \to [0,\infty)$ is the time-function. On $\mathcal{N}_{t_j\pm}$ we denote by $Q_\pm$ and $Q^*_\pm$ the corresponding values for $h_{t_j\pm}$. We also set $Q_\pm := Q$ on $\mathcal{N} \backslash \bigcup_{\mathcal{D}\in\mathrm{Cut}} \mathcal{D}$. Set

$$\overline{Q} := 10^{-E-1}\eta_{\mathrm{lin}} r_{\mathrm{comp}}^E \quad \text{and} \quad \overline{Q}^* := 10^{-1}\eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3.$$

Then

(APA7) ($Q \leqslant \overline{Q}$ if no cuts in nearby past) For all $x \in \mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}$ for which

$$P(x, A\rho_1(x)) \cap \mathcal{D} = \varnothing \quad \text{for all } \mathcal{D} \in \mathrm{Cut},$$

we have

$$Q(x) \leqslant \overline{Q}.$$

Note that the bound is also required to hold if $P(x, A\rho_1(x)) \not\subset \mathcal{N}$.

(APA8) We have

$$Q \leqslant W \cdot \overline{Q} \quad \text{on } \mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}.$$

(APA9) ($Q^* \leqslant \overline{Q}^*$ away from time-slice boundary) For all $x \in \mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}} \mathcal{D}$ for which $B(x, A\rho_1(x)) \subset \mathcal{N}_{t(x)-}$, we have

$$Q^*(x) \leqslant \overline{Q}^*.$$

(APA10) On every cut $\mathcal{D} \in \mathrm{Cut}$, we have

$$Q^*_+ \leqslant \eta_{\mathrm{cut}} \overline{Q}^*.$$

(APA11) For every cut $\mathcal{D} \in \mathrm{Cut}$, $\mathcal{D} \subset \mathcal{M}_{t_j}$, the following holds: The diameter of $\mathcal{D}$ is less than $D_{\mathrm{cut}} r_{\mathrm{comp}}$ and $\mathcal{D}$ contains a $\frac{1}{10} D_{\mathrm{cut}} r_{\mathrm{comp}}$-neighborhood of the extension cap $\mathcal{C} = \mathcal{D} \setminus \operatorname{Int} \mathcal{N}_{t_j -}$.

(APA12) We have $Q \leqslant \nu \overline{Q}$ and $Q^* \leqslant \nu \overline{Q}^*$ on $\mathcal{N}_0$ (i.e. at time zero).

(APA13) We have $t_J \leqslant T$.

Note that a-priori assumptions (APA7)–(APA13) are vacuous if $J = 0$.

We briefly comment on the purpose of a-priori assumptions (APA7)–(APA9).

As explained in §2, a-priori assumption (APA7), the bound $Q \leqslant \overline{Q}$, serves as a main ingredient for the Bryant extension principle, as long as E is chosen large enough. It will also be used to ensure that $|h| \leqslant \eta_{\mathrm{lin}}$ at most points of the comparison domain.

Note however that $\overline{Q}$ is chosen such that the bound $Q \leqslant \overline{Q}$ only implies $|h| \leqslant \eta_{\mathrm{lin}}$ when $\rho_1 \gtrsim r_{\mathrm{comp}}$. So, it does not imply $|h| \leqslant \eta_{\mathrm{lin}}$ everywhere on the comparison domain. Unfortunately, we will not be able to remedy this issue by replacing $\overline{Q}$ in a-priori assumption (APA7) by a smaller constant, as our solution of the harmonic map heat flow will introduce an error of magnitude depending on δ_{n} near the neck-like boundary of $\mathcal{N}$.

More specifically, assuming that the bound $Q \leqslant \overline{Q}$ holds near the neck-line boundary, which has scale $\approx r_{\mathrm{comp}}$, then errors would force $\overline{Q} \gtrsim r_{\mathrm{comp}}^E \eta'$, where $\eta' = \eta'(\delta_{\mathrm{n}})$. On the other hand, since we would want the inequality $Q \leqslant \overline{Q}$ to enforce the bound $|h| \leqslant \eta_{\mathrm{lin}}$ everywhere in $\mathcal{N}$, and since $\mathcal{N}$ may contain points of scale $\approx \lambda r_{\mathrm{comp}}$, we would need to

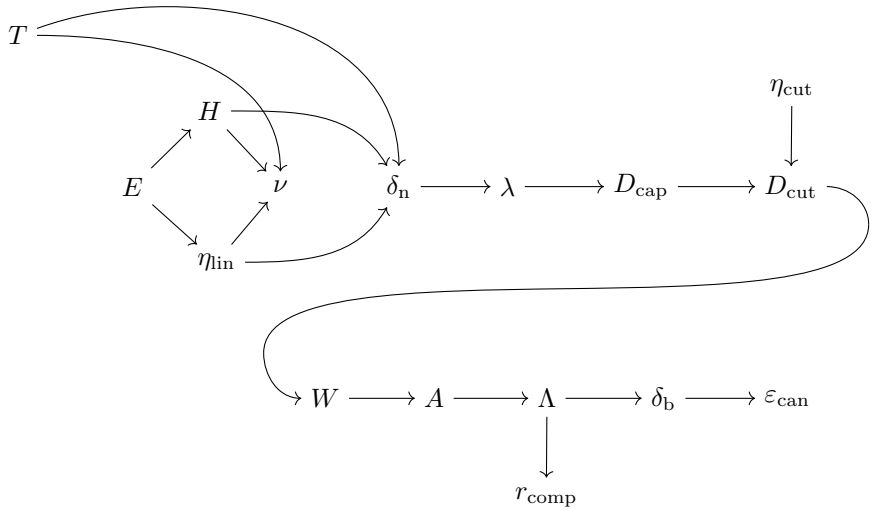

Figure 10. All restrictions on the parameters that will be imposed throughout this paper. Each parameter in this graph can be chosen depending only on the parameters that can be reached by following the arrows backwards. Note that the graph does not contain any oriented circles.

have $\overline{Q}\lesssim(\lambda r_{\mathrm{comp}})^E\eta_{\mathrm{lin}}$. Combining the two inequalities, we get $\eta'(\delta_{\mathrm{n}})\lesssim\lambda^E\eta_{\mathrm{lin}}$, so we end up with a condition of the form $\delta_{\mathrm{n}}\leqslant\delta_{\mathrm{n}}(\eta_{\mathrm{lin}},\lambda)$. However, to construct the comparison domain so that its boundary consists of (roughly) δ_{n}-necks, we need a condition of the form $\lambda\leqslant\lambda(\delta_{\mathrm{n}})$, which is incompatible. In summary, the constant $\overline{Q}$ cannot be chosen such that a-priori assumption (APA7) is both weak enough to hold near the boundary of $\mathcal{N}$ and strong enough to imply $|h|\leqslant\eta_{\mathrm{lin}}$ at all points of scale $\gtrsim\lambda r_{\mathrm{comp}}$.

The bound $Q^*\leqslant\overline{Q}^*$ in a-priori assumption (APA9), on the other hand, automatically implies $|h|\leqslant\eta_{\mathrm{lin}}$ everywhere on $\mathcal{N}$. However, we are not imposing it near the neck-like boundary of $\mathcal{N}$.

Lastly, note that the bound $Q\leqslant\overline{Q}$ may be violated after a Bryant extension construction. Therefore, we have not imposed it in a-priori assumption (APA7) at points that lie in the near future of cuts. At these points, the bound $Q^*\leqslant\overline{Q}^*$ will be used to guarantee $|h|\leqslant\eta_{\mathrm{lin}}$. Moreover, the bound $Q\leqslant W\overline{Q}$ from a-priori assumption (APA8) will be used to partially retain a-priori assumption (APA7) in the future of a cut. Using the interior decay from §9, this bound can in turn be improved to the bound $Q\leqslant\overline{Q}$ from a-priori assumption (APA7) after a sufficient time.

7.5. Parameter order

As mentioned earlier, the a-priori assumptions, as introduced in the last two subsections, involve several parameters, which will need to be chosen carefully in the course of this paper. Each step of our construction will require that certain parameters be chosen sufficiently small/large depending on certain other parameters. In order to show that these parameters can eventually be chosen such that all restrictions are met, we need to ensure that these restrictions are not circular. For this purpose, we introduce the following parameter order:

$$T,\ E,\ H,\ \eta_{\mathrm{lin}},\ \nu,\ \delta_{\mathrm{n}},\ \lambda,\ D_{\mathrm{cap}},\ \eta_{\mathrm{cut}},\ D_{\mathrm{cut}},\ W,\ A,\ \Lambda,\ \delta_{\mathrm{b}},\ \varepsilon_{\mathrm{can}},\ r_{\mathrm{comp}}.$$

In the entire paper, we will require each parameter to be chosen depending only on preceding parameters in this list. So, parameters can eventually be chosen successively in the order indicated by this list.

For a more detailed picture of all the parameter restrictions imposed in this paper see Figure 10. These restrictions also appear in the preamble of our main technical result, Theorem 13.1. Note that, as these restrictions are not completely linear, there are several admissible parameter orders. We have chosen the above parameter order, because we found it to be most intuitive.

We advise the first-time reader that it is not necessary to follow all parameter restrictions in detail when going through the proofs of this paper. Instead, it suffices to check that the above parameter order is obeyed in each step.

8. Preparatory results

In this section we collect a variety of technical results that will be needed in the proof of the main theorem. These are based on definitions from §§5–7. The reader may wish to skim (or skip) this section on a first reading.

8.1. Consequences of the canonical neighborhood assumption

The completeness and canonical neighborhood assumptions, as introduced in Definitions 5.3 and 5.6 lead in a straightforward way to local bounds on geometry, including local control on curvature and its derivatives, as well as control on neck and non-neck structure. We begin this subsection with a few such results (Lemmas 8.1–8.7), and then use them to deduce control on scale distortion of bilipschitz maps (Lemma 8.9), self-improvement of necks (Lemma 8.10) and scale bounds near necks (Lemma 8.11).

Our first two results are direct consequences of the definition of the canonical neighborhood assumption, and properties of κ-solutions.

LEMMA 8.1. (a) *For every $A<\infty$ there is a constant $C=C(A)<\infty$ such that, if*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(A)$$

and a Riemannian manifold M satisfies the ε_{can}-canonical neighborhood assumption at $x\in M$, then the following holds on the ball $B(x, A\rho(x))$ for all $0\leqslant m\leqslant A$:

$$\rho = \sqrt{3}R^{-1/2}, \quad C^{-1}\rho(x) \leqslant \rho \leqslant C\rho(x), \quad |\nabla^m \operatorname{Rm}| \leqslant C\rho^{-2-m}(x).$$

(b) *There is a $C<\infty$ such that, if*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}$$

and $\mathcal{M}$ is a Ricci flow spacetime that satisfies the ε_{can}-canonical neighborhood assumption at some point $x\in\mathcal{M}$, then

$$|\partial_t \rho^2|(x) = 3|\partial_t R^{-1}|(x) \leqslant C.$$

(c) *Given $\delta>0$, if*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta)$$

and $\mathcal{M}$ is a Ricci flow spacetime that satisfies the ε_{can}-canonical neighborhood assumption at some point $x\in\mathcal{M}$, then

$$\partial_t \rho^2(x) = 3\partial_t R^{-1}(x) \leqslant \delta.$$

Proof. Assertion (a) follows from the definition of the canonical neighborhood assumption, assertions (c) and (d) of Lemma C.1 and Lemma 6.2.

For assertions (b) and (c) we recall that in a Ricci flow the time derivative $\partial_t R(x)$ may be expressed as a universal continuous function $F(\operatorname{Rm}(x), \nabla \operatorname{Rm}(x), \nabla^2 \operatorname{Rm}(x))$ of spatial curvature derivatives. Now, assertions (b) and (c) follow from the definition of the canonical neighborhood assumption, and assertion (e) of Lemma C.1. □

LEMMA 8.2. *For every $\delta>0$ there is a constant $C_0=C_0(\delta)<\infty$ such that, if*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta),$$

then the following holds.

Assume that (M, g) is a Riemannian manifold that satisfies the ε_{can}-canonical neighborhood assumption at some point $x\in M$. Then, one of the following hold:

(a) *x is the center of a δ-neck at scale $\rho(x)$.*

(b) *There is a compact, connected domain $V\subset\mathcal{M}_t$ with connected (possibly empty boundary) such that the following hold*:

(1) $B(x,\delta^{-1}\rho(x))\subset V$.
(2) $\rho(y_1)<C_0\rho(y_2)$ *for all* $y_1,y_2\in V$.
(3) $\operatorname{diam} V<C_0\rho(x)$.
(4) *If* $\partial V\neq\varnothing$, *then*
   (i) ∂V *is a central 2-sphere of a* δ*-neck.*
   (ii) *Either* V *is a 3-disk or it is diffeomorphic to a twisted interval bundle over* $\mathbb{R}P^2$.
   (iii) *Any two points* $z_1,z_2\in\partial V$ *can be connected by a continuous path inside* ∂V *whose length is less than*

$$\min\{d(z_1,x),d(x,z_2)\}-100\rho(x).$$

(5) *If* V *is diffeomorphic to a twisted interval bundle over* $\mathbb{R}P^2$, *then*

$$\rho(y_1)<2\rho(y_2)\quad \text{for all } y_1,y_2\in V.$$

Proof. This follows immediately from Lemma C.2 using the definition of the canonical neighborhood assumption. □

LEMMA 8.3. *Suppose* $\mathcal{M}$ *is an* (r_0,t_0)*-complete Ricci flow spacetime. If for some* $r>r_0$ *we have* $\rho>r$ *on a parabolic neighborhood* $P(x,a,b)\subset\mathcal{M}_{[0,t_0]}$, *then it is unscathed.*

Proof. Let $t=\mathfrak{t}(x)$. From the (r_0,t_0)-completeness of $\mathcal{M}$, any unit speed geodesic in $\mathcal{M}_t$ starting at x can be extended up to a length of at least a. Therefore, the exponential map $\exp_x\colon T_x\mathcal{M}_t\supset\overline{B(0,a)}\to\mathcal{M}_t$ is well defined, and has compact image

$$\exp_x(\overline{B(0,a)})=\overline{B(x,a)}.$$

If $y\in B(x,a)$, then since $\rho>r$ on $P(x,a,b)$, it follows from (r_0,t_0)-completeness that $y(\bar{t})$ is defined on $[t,t+b]$ if $b>0$ or $[t+b,t]\cap[0,\infty)$ if $b<0$. □

Next, we derive a few results based on the bounds in Lemma 8.1.

LEMMA 8.4. (Scale nearly constant on small two-sided parabolic balls) *If* $L>1$ *and*

$$\eta\leqslant\bar{\eta}(L)\quad \text{and}\quad \varepsilon_{\text{can}}\leqslant\bar{\varepsilon}_{\text{can}},$$

then the following holds.

Suppose that $0<r\leqslant 1$ *and* $\mathcal{M}$ *is an* $(\varepsilon_{\text{can}}r,t_0)$*-complete Ricci flow spacetime satisfying the* ε_{can}*-canonical neighborhood assumption at scales* $(\varepsilon_{\text{can}}r,1)$. *If for some point* $x\in\mathcal{M}_t$ *with* $t\in[0,t_0]$ *we have* $\rho_1(x)\geqslant r$, *then the parabolic neighborhoods*

$$P_\pm:=P(x,\eta\rho_1(x),\pm(\eta\rho_1(x))^2)\cap\mathcal{M}_{[0,t_0]}$$

are unscathed and

$$L^{-1}\rho_1(x) \leqslant \rho_1 \leqslant L\rho_1(x) \tag{8.1}$$

on $P_+ \cup P_-$.

Proof. If

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}},$$

then, by Lemma 6.2 and assertions (a) and (b) of Lemma 8.1, there is a constant $C_0<\infty$ such that near any point that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption we have

$$|\nabla\rho|, |\partial_{\mathfrak{t}}\rho^2| \leqslant C_0. \tag{8.2}$$

Now choose a point $y\in P_\pm$, and let $\gamma\colon [0,a]\to\mathcal{M}$ be a curve from x to y that is a concatenation of curves γ_1 and γ_2, where γ_1 is a unit speed curve from x to $y(t)$ of length $<\eta\rho_1(x)$, and γ_2 is the integral curve of $\pm\partial_{\mathfrak{t}}$ from $y(t)$ to y. Then, by (8.2), we have

$$|(\rho_1\circ\gamma_1)'(s)| \leqslant C_0 \quad \text{and} \quad |(\rho_1^2\circ\gamma_2)'(s)| \leqslant C_0, \tag{8.3}$$

wherever the derivatives are defined and $\rho_1\circ\gamma_i(s)>\varepsilon_{\mathrm{can}}r$. Therefore, if

$$\eta \leqslant \bar{\eta}(L),$$

then (8.1) follows by integrating the derivative bound (8.3). The fact that $P_\pm$ are unscathed follows from Lemma 8.3. □

LEMMA 8.5. (Backward survival control) *If* $\delta>0$, $A<\infty$ *and*

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\delta, A),$$

then the following holds.

Suppose that $r>0$ *and* $\mathcal{M}$ *is an* $(\varepsilon_{\mathrm{can}}r, t_0)$*-complete Ricci flow spacetime satisfying the* $\varepsilon_{\mathrm{can}}$*-canonical neighborhood assumption at scales* $(\varepsilon_{\mathrm{can}}r, r)$. *Let* $x\in\mathcal{M}_t$ *with* $t\in[0,t_0]$ *and assume that* $\rho(x)\geqslant r$. *Then,* $x(t')$ *exists for all* $\bar{t}\in[t-Ar^2,t]\cap[0,\infty)$, *and we have*

$$\rho(x(\bar{t})) > (1-\delta)r.$$

Proof. Set $t:=\mathfrak{t}(x)$ and let $\delta_\#>0$ be a constant whose value we will choose at the end of the proof. Recall that $\rho_r=\min\{\rho, r\}$. By assertion (c) of Lemma 8.1, and assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\delta_\#),$$

we have

$$\frac{d}{d\bar{t}}(\rho_r^2(x(\bar{t}))) \leqslant \delta_{\#} \tag{8.4}$$

for all $\bar{t}\leqslant t$ for which both $x(\bar{t})$ and the derivative exist and $\rho_r^2(x(\bar{t}))>(\varepsilon_{\mathrm{can}}r)^2$. Therefore, if

$$\delta_{\#} \leqslant \bar{\delta}(\delta, A),$$

then we may integrate (8.4) to obtain that $\rho_r^2(x(\bar{t}))>(1-\delta)r$ for all $\bar{t}\leqslant t$ for which

$$t-\bar{t} \leqslant Ar^2$$

and $x(\bar{t})$ is defined. Assuming

$$\varepsilon_{\mathrm{can}} < 1-\delta,$$

we can use the $(\varepsilon_{\mathrm{can}}r, t_0)$-completeness to show that $x(\bar{t})$ is defined for all

$$\bar{t} \in [t-Ar^2, t]\cap[0,\infty). \qquad \square$$

LEMMA 8.6. (Bounded curvature at bounded distance) *For every $A<\infty$ there is a constant $C=C(A)<\infty$ such that if*

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(A),$$

then the following holds.

Let $0<r\leqslant 1$ and consider an $(\varepsilon_{\mathrm{can}}r, t_0)$-complete Ricci flow spacetime $\mathcal{M}$ that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}}r, 1)$. If $x\in\mathcal{M}_{[0,t_0]}$ and $\rho_1(x)\geqslant r$, then $P(x, A\rho_1(x))$ is unscathed and we have

$$C^{-1}\rho_1(x) < \rho_1 < C\rho_1(x) \quad \text{on } P(x, A\rho_1(x)). \tag{8.5}$$

Proof. We claim that there is a constant $C_1=C_1(A)<\infty$ such that

$$C_1^{-1}\rho_1(x) \leqslant \rho_1 \leqslant C_1\rho_1(x) \quad \text{on } B(x, A\rho_1(x)). \tag{8.6}$$

This is immediate if $\rho_1\equiv 1$ on $B(x, A\rho_1(x))$, so suppose $\rho_1(y)<1$ for some $y\in B(x, A\rho_1(x))$. By the continuity of ρ_1, we may choose y such that $\rho_1(y)\in\left(\frac{1}{2}\rho_1(x), 1\right)$. Applying assertion (a) of Lemma 8.1 to the ball $B(y, 4A\rho(y))\supset B(x, A\rho_1(x))$, we get (8.6).

If

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(A),$$

then using (8.6), we may apply Lemma 8.5 at any point $z\in B(x, A\rho_1(x))$ to conclude that $\gamma(\bar{t})$ is defined and

$$\rho_1(z(\bar{t})) \geqslant \tfrac{1}{2}C_1^{-1}\rho_1(x) > \varepsilon_{\mathrm{can}}r$$

for all $\bar{t}\in[t-A\rho_1(x)^2,t]\cap[0,\infty)$. Thus, $P(x,A\rho_1(x))$ is unscathed by Lemma 8.3.

Next, by assertion (a) of Lemma 8.1 there is a universal constant $C_2<\infty$ such that if

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(A),$$

then for all $\bar{t}\in[t-A\rho_1^2(x)]\cap[0,\infty)$ we have

$$\left|\frac{d}{d\bar{t}}(\rho_1^2(z(\bar{t}))\right|<C_2,$$

provided the derivative is defined. Integrating this bound yields

$$\rho_1^2(z(\bar{t}))\leqslant C^2\rho_1^2(z)$$

for $C=C(A)<\infty$. Thus (8.5) holds. □

In the next result we combine the bounded curvature at bounded distance estimate (Lemma 8.6) with a distance distortion estimate to find a parabolic neighborhood centered at a point x that contains all parabolic neighborhoods of the form $P(y,A_2\rho_1(y))$, where y varies over some parabolic neighborhood $P(x,A_1\rho_1(x))$.

LEMMA 8.7. (Containment of parabolic neighborhoods) *For any $A_1,A_2<\infty$ there is a constant $A'=A'(A_1,A_2)<\infty$ with $A'\geqslant A_1+A_2$ such that if*

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(A_1,A_2),$$

then the following holds.

Let $0<r\leqslant 1$ and consider an $(\varepsilon_{\mathrm{can}}r,t_0)$-complete Ricci flow spacetime $\mathcal{M}$ that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}}r,1)$. If $x\in\mathcal{M}_{[0,t_0]}$ and $\rho_1(x)\geqslant r$, then the parabolic neighborhood $P(x,A'\rho_1(x))$ is unscathed and we have

$$P(y,A_2\rho_1(y))\subset P(x,A'\rho_1(x)) \tag{8.7}$$

for all $y\in P(x,A_1\rho_1(x))$.

Proof. We first use Lemma 8.6, assuming

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(A_1),$$

to argue that $P(x,A_1\rho_1(x))$ is unscathed and

$$\rho_1<C_1(A_1)\rho_1(x)\quad\text{on } P(x,A_1\rho_1(x)) \tag{8.8}$$

for some $C_1=C_1(A_1)<\infty$.

The constant A' will be determined in the course of the proof. Again, by Lemma 8.6, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(A'), \tag{8.9}$$

we find that $P(x,2A'\rho_1(x))$ is unscathed and that $\rho_1>c_2(A')\rho_1(x)\geqslant c_2 r>\varepsilon_{\text{can}} r$ on it. At any point $z\in P(x,2A'\rho_1(x))$ with $\rho_1(z)<1$ the curvature operator is close to that of a κ-solution. Since κ-solutions have non-negative Ricci curvature, we can argue that

$$\text{Ric} \geqslant -c_2^2(A')\rho^{-2}(z) \geqslant -\rho_1^{-2}(x)$$

at z if we assume a bound of the form (8.9). On the other hand, at any $z\in P(x,2A'\rho_1(x))$ with $\rho_1(z)=1$ we have $\rho(z)\geqslant 1$, and therefore $\text{Ric}\geqslant -C_2$ at z for some universal constant C_2. So, in summary, we have

$$\text{Ric} \geqslant -C_2\rho_1^{-2}(x) \quad \text{on } P(x,2A'\rho_1(x)). \tag{8.10}$$

Now consider a point $y\in P(x,A_1\rho_1(x))$. Set $t_x:=\mathfrak{t}(x)$ and $t_y:=\mathfrak{t}(y)$. We first claim that for

$$A' \geqslant \underline{A}'(A_1,A_2)$$

we have

$$B(y,A_2\rho_1(y)) \subset P(x,A'\rho_1(x)). \tag{8.11}$$

Assume not and choose a smooth curve $\gamma\colon[0,1]\to(P(x,2A'\rho_1(x))_{t_y}$ between y and a point $z\in P(x,2A'\rho_1(x))\setminus P(x,A'\rho_1(x))$ such that $\ell_{t_y}(\gamma)<A_2\rho_1(y)$. Note that for all $t'\in[t_y,t_x]$ the curve $\gamma_{t'}\colon[0,1]\to\mathcal{M}_t$ with $\gamma_{t'}(s):=(\gamma(s))(t')$ is defined and its image is contained in $P(x,2A'\rho_1(x))$. So, by (8.10) and (8.8), we have

$$\begin{aligned} d_{t_x}(y(t_x),z(t_x)) &\leqslant \ell_{t_x}(\gamma_{t_x}) \\ &< \exp(C_2\rho_1^{-2}(x)A_1^2\rho_1^2(x))\cdot A_2\rho_1(y) \\ &\leqslant C_1A_2\exp(C_2A_1^2)\rho_1(x) \end{aligned}$$

So

$$A'\rho_1(x) \leqslant d_{t_x}(x,y)+d_{t_x}(y,z(t_x)) < A_1\rho_1(x)+C_1A_2\exp(C_2A_1^2)\rho_1(x).$$

Now set

$$A'(A_1,A_2) := A_1+C_1A_2\exp(C_2A_1^2)+\sqrt{A_1^2+A_2^2}.$$

Then, we obtain a contradiction and thus (8.11) holds. Since

$$(A')^2 \geqslant A_1^2+A_2^2,$$

we obtain (8.7). □

The next two results concern the behavior of the curvature scale ρ under nearly isometric mappings. We begin with a convergence lemma that shows that an immersion between Riemannian manifolds must nearly preserve the scale, provided it is nearly an isometry, and we have sufficient control on the curvature and possibly curvature derivatives on the domain and target. The main point is that the map is only assumed to be an almost isometry in the C^0-sense.

LEMMA 8.8. *Suppose that* $\{(Z_k^1, g_k^1, z_k^1\}_{k=1}^\infty$ *and* $\{(Z_k^2, g_k^2, z_k^2)\}_{k=1}^\infty$ *are sequences of pointed smooth Riemannian manifolds such that for some* $r_0>0$ *and for each* $i=1,2$ *the ball* $B(z_k^i, r_0)\subset Z_k^i$ *is relatively compact for all* k, *and one of the following holds*:

(i) $\sup_{B(z_k^i,r_0)} |\mathrm{Rm}|_{g_k^i} \to 0$ *as* $k\to\infty$.

(ii) $\limsup_{k\to\infty} \sup_{B(z_k^i,r_0)} |\nabla^j \mathrm{Rm}|_{g_k^i} < \infty$ *for* $0\leqslant j\leqslant 5$.

Let $\{\phi_k\colon Z_k^1 \to Z_k^2\}_{k=1}^\infty$ *be a sequence of smooth maps such that* $\phi_k(z_k^1)=z_k^2$ *and*

$$\sup_{B(z_k^1,r_0)} |(\phi_k^* g_k^2 - g_k^1)|_{g_k^1} \to 0 \quad \text{as } k\to\infty. \tag{8.12}$$

Then, after passing to a subsequence, the scale functions converge to the same limit:

$$\lim_{k\to\infty} \rho(z_k^1) = \lim_{k\to\infty} \rho(z_k^2) \in [0,\infty)\cup\{\infty\}.$$

Proof. We first prove the lemma under the additional conditions that the ϕ_ks are diffeomorphisms and the injectivity radii at z_k^1 satisfy

$$\liminf_{k\to\infty} \mathrm{InjRad}(Z_k^1, g_k^1, z_k^1) > 0. \tag{8.13}$$

Using standard injectivity radius estimates, conditions (i), (ii), (8.12), and (8.13) imply that for every $r<r_0$, and sufficiently large k, the injectivity radius is bounded uniformly from below on $B(z_k^i, r)\subset Z_k^i$. By standard compactness arguments, after passing to a subsequence, the sequence of pointed balls $\{(B(z_k^i, r_0), g_k^i, z_k^i)\}_{k=1}^\infty$ converges to a pointed C^4-Riemannian manifold $(Z_\infty^i, g_\infty^i, z_\infty^i)$, that is a proper r_0-ball (i.e. balls of radius $<r_0$ are relatively compact), and there is a basepoint-preserving map $\phi_\infty\colon (Z_\infty^1, z_\infty^1)\to(Z_\infty^2, z_\infty^2)$ that is an isometry of the Riemannian distance functions, where, for each $i=1,2$, the following holds:

- If $\{g_k^i\}$ satisfies (i), then the pointed convergence

$$(B(z_k^i, r_0), g_k^i, z_k^i) \to (Z_\infty^i, z_\infty^i)$$

is with respect to the Gromov–Hausdorff topology and Z_∞^i is flat.

- If $\{g_k^i\}$ satisfies (ii), then the pointed convergence

$$(B(z_k^i, r_0), g_k^i, z_k^i) \to (Z_\infty^i, g_\infty^i, z_\infty^i)$$

is with respect to the C^4-topology.

In view of the above, we have $\rho(z_k^i)\to\rho(z_\infty^i)\in[0,\infty)\cup\{\infty\}$ as $k\to\infty$ for $i=1,2$. Since ϕ_∞ is an isometry (of distance functions) between C^4 Riemannian manifolds, it is a C^3-isometry of Riemannian manifolds, and hence it preserves curvature tensors: $\phi_\infty^*(\mathrm{Rm}(z_\infty^2))=\mathrm{Rm}(z_\infty^1)$. It follows that $\rho(z_\infty^1)=\rho(z_\infty^2)$.

We now return to the general case. We may assume after shrinking r_0 that the conjugate radius of Z_k^1 at z_k^1 is $\geqslant 2r_0$. For $i=1,2$, let (W_k^i,w_k^i) be the ball $B(0,2r_0)\subset T_{z_k^1}Z_k^1$ with basepoint $w_k^i=0\in B(0,2r_0)$, and let $h_k^1:=\exp_{z_k^1}^* g_k^i$ and $h_k^2:=(\phi_k\circ\exp_{z_k^1})^* g_k^2$. Then, the injectivity radius at w_k^1 satisfies $\mathrm{InjRad}(W_k^1,h_k^1,w_k^1)\geqslant r_0$, and $B(w_k^i,r_0)\subset W_k^i$ is relatively compact. Therefore, applying the above argument to the identity maps $W_k^1\to W_k^2$, we obtain the lemma. □

LEMMA 8.9. (Scale distortion of bilipschitz maps) *There is a constant $10^3<C_{\mathrm{SD}}<\infty$ such that the following holds if*

$$\eta_{\mathrm{lin}}\leqslant\bar\eta_{\mathrm{lin}},\quad \delta_{\mathrm{n}}\leqslant\bar\delta_{\mathrm{n}},\quad \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}}\quad \textit{and}\quad r_{\mathrm{comp}}\leqslant\bar r_{\mathrm{comp}}.$$

Let $\mathcal{M}$ and $\mathcal{M}'$ be $(\varepsilon_{\mathrm{can}}r_{\mathrm{comp}},t_0)$-complete Ricci flow spacetimes. Consider a closed product domain $X\subset\mathcal{M}_{[0,t_0]}$ on a time-interval of the form $[t-r_{\mathrm{comp}}^2,t]$, $t\geqslant r_{\mathrm{comp}}^2$, such that the following holds:

(i) *∂X_t consists of embedded 2-spheres that are each centers of δ_{n}-necks at scale r_{comp}.*

(ii) *Each connected component of X_t contains a $2r_{\mathrm{comp}}$-thick point.*

Let $\bar t\in[t-r_{\mathrm{comp}}^2,t]$, $t'\geqslant 0$ and consider a diffeomorphism onto its image $\phi\colon X_{\bar t}\to\mathcal{M}'_{t'}$ such that $|\phi^*g'_{t'}-g_{\bar t}|\leqslant\eta_{\mathrm{lin}}$. We assume that $\mathcal{M}$ satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(0,1)$ on $X_{\bar t}$, and that $\mathcal{M}'$ satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(0,1)$ on $\phi(X_{\bar t})$.

Then, for any $x\in X_{\bar t}$, we have

$$C_{\mathrm{SD}}^{-1}\rho_1(x)<\rho_1(\phi(x))<C_{\mathrm{SD}}\rho_1(x). \tag{8.14}$$

This lemma will later be applied whenever a bound on the distortion of the scale function under a comparison (as defined in Definition 7.2) is needed. The product domain X in this lemma will later be taken to be a time-slab $\mathcal{N}^j$ of a comparison domain (as defined in Definition 7.1) and ϕ will denote the time-slice of a comparison. Assumptions (i) and (ii) correspond to a-priori assumptions (APA3) (a) and (APA3) (d), respectively (see Definition 7.4).

In order to avoid confusion, we point out that usually it is possible to derive stronger scale distortion bounds than (8.14), with C_{SD} replaced by a constant that can be chosen arbitrarily close to 1. These stronger bounds follow simply via local gradient estimates,

due to the parabolic nature of the comparison. This approach, however, fails if the point x lies close to the spatial or time-like boundary of X. This is why we have to work with a larger constant C_{SD} in this paper.

Proof. Assume the lemma was false. Then, there are sequences $\eta_{\mathrm{lin},k}\to 0$, $\delta_{\mathrm{n},k}\to 0$, $\varepsilon_{\mathrm{can},k}\to 0$, $r_{\mathrm{comp},k}\to 0$, $\{\mathcal{M}^k\}$, $\{\mathcal{M}'^k\}$, $\{X^k\}$, $\{x_k\}$, $\{t_k\}$, $\{\bar{t}_k\}$, $\{t'_k\}$ and $\phi_k\colon X^k_{\bar{t}_k}\to\mathcal{M}'^k_{t'_k}$ satisfying the assumptions of the lemma, such that

$$\frac{\rho_1(x_k)}{\rho_1(\phi_k(x_k))}\to 0 \quad \text{or} \quad \frac{\rho_1(x_k)}{\rho_1(\phi_k(x_k))}\to\infty \quad \text{as } k\to\infty. \tag{8.15}$$

To simplify notation, we let $M_k:=\mathcal{M}^k_{\bar{t}_k}$ and $M'_k:=\mathcal{M}'^k_{t'_k}$ denote the time-slices, with metrics g_k and g'_k, respectively, and let $Y_k:=X^k_{\bar{t}_k}\subset M_k$ be the relevant time-slice of the product domain X^k.

Let $r_k:=\min\{\rho_1(x_k),\rho_1(\phi_k(x_k))\}$. In view of (8.15), we have $r_k\to 0$. Note that, by our assumptions, for each of x_k and $\phi_k(x_k)$, either the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption holds, or we have $\rho_1(x_k)=1$ or $\rho_1(\phi_k(x_k))=1$, respectively. In the first case we may use the estimates on the derivatives of curvature in assertion (a) of Lemma 8.1, and we have

$$|\nabla^j\mathrm{Rm}|<C_1 r_k^{-2-j} \quad \text{on } B(x_k,r_k) \text{ or } B(\phi_k(x_k),r_k), \tag{8.16}$$

respectively, for some universal $C_1<\infty$ and large k and $0\leqslant j\leqslant 5$, and in the second case we may apply Lemma 8.6 to obtain

$$|\mathrm{Rm}|<C_2 \quad \text{on } B(x_k,r_k) \text{ or } B(\phi_k(x_k),r_k), \tag{8.17}$$

respectively, for some universal $C_2<\infty$ and large k.

Case 1. $\liminf_{k\to\infty} r_k^{-1}d(x_k,\partial Y_k)>0$.

If we let $\hat{g}_k:=r_k^{-2}g_k$ and $\hat{g}'_k:=r_k^{-2}g'_k$, then the assumptions of Lemma 8.8 hold for the sequence $\{\phi_k\colon(\mathrm{Int}\,Y_k,\hat{g}_k,x_k)\to(M'_k,\hat{g}'_k,\phi_k(x_k))\}$ by (8.15)–(8.17) and the fact that $r_k\to 0$. Hence, after passing to a subsequence,

$$\lim_{k\to\infty}\rho_{\hat{g}_k}(x_k)=\lim_{k\to\infty}\rho_{\hat{g}'_k}(\phi_k(x_k)).$$

Since for every k the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption holds at one of the points x_k, $\phi_k(x_k)$, the above limit must equal 1. This contradicts (8.15).

Case 2. $\liminf_{k\to\infty} r_k^{-1}d(x_k,\partial Y_k)=0$.

After passing to a subsequence, we may assume that

$$\lim_{k\to\infty} r_k^{-1}d(x_k,\partial Y_k)=0. \tag{8.18}$$

For each k we may choose a boundary component $\Sigma_k \subset X^k_{t_k}$ such that

$$\lim_{k\to\infty} r_k^{-1} d(x_k, \Sigma_k(\bar{t}_k)) = 0.$$

Let U_k be the $10r_{\mathrm{comp},k}$-neighborhood of Σ_k in $X^k_{t_k}$. If k is large, then

$$\tfrac{1}{2} r_{\mathrm{comp},k} < \rho < 2r_{\mathrm{comp},k} \quad \text{on } U_k.$$

So, by assumption (ii) and the fact that $\delta_{\mathrm{n},k} \to 0$, it follows that U_k does not fully contain the component of $X^k_{t_k}$ in which it lies, and moreover it does not intersect any other boundary components of $X^k_{t_k}$. Therefore, we can pick $y_k \in X^k_{t_k}$ with $d(y_k, \Sigma_k) = r_{\mathrm{comp},k}$. By Lemma 8.6, there is a universal constant $C_3 < \infty$ such that, for large k, we have

$$C_3^{-1} r_{\mathrm{comp},k} \leqslant \rho \leqslant C_3 r_{\mathrm{comp},k} \tag{8.19}$$

on $U_k(\bar{t}_k)$, in particular on $\Sigma_k(\bar{t}_k)$. By (8.16) or (8.17) and the fact that

$$B(x_k, r_k) \cap \Sigma_k(\bar{t}_k) \neq \varnothing$$

for large k, we get $r_k \leqslant C_4 r_{\mathrm{comp},k}$ for large k, where $C_4 < \infty$ is a universal constant. By (8.19), (8.18), and a distance distortion estimate, we have $x_k \in U_k(\bar{t}_k)$, and therefore $r_k \leqslant \rho_1(x_k) \leqslant C_3 r_{\mathrm{comp},k}$ for large k. Hence, $\lim_{k\to\infty} r^{-1}_{\mathrm{comp},k} d(x_k, \partial Y_k) = 0$. By a distance distortion estimate, there is a universal constant $C_5 < \infty$ such that, for large k,

$$C_5^{-1} r_{\mathrm{comp},k} \leqslant d(y_k(\bar{t}_k), x_k), d(y_k(\bar{t}_k), \partial Y_k) \leqslant C_5 r_{\mathrm{comp},k}. \tag{8.20}$$

So, using (8.20) and Case 1, we can find a uniform $C_6 < \infty$ such that

$$C_3^{-1} C_6^{-1} r_{\mathrm{comp},k} \leqslant C_6^{-1} \rho_1(y_k(\bar{t}_k)) \leqslant \rho_1(\phi_k(y_k(\bar{t}_k))) \leqslant C_6 \rho_1(y_k(\bar{t}_k)) \leqslant C_3 C_6 r_{\mathrm{comp},k}.$$

Since $d(\phi_k(y_k(\bar{t}_k)), \phi_k(x_k)) \leqslant 2C_3 r_{\mathrm{comp},k}$ for large k, Lemma 8.6 gives

$$C_7^{-1} r_{\mathrm{comp},k} \leqslant \rho_1(\phi_k(x_k)) \leqslant C_7 r_{\mathrm{comp},k}$$

for some uniform $C_7 < \infty$ and large k. This contradicts (8.15). □

In the following lemma we show that a region that is bilipschitz close to a cylinder contains a smaller region on which we have closeness to a cylinder in the C^m-sense, provided that the canonical neighborhood assumption holds. So, the smaller region is a neck of arbitrarily high accuracy, as long as the bilipschitz control on the larger region is strong enough.

LEMMA 8.10. (Self-improvement of necks) *If*

$$\delta_{\#}>0, \quad \delta\leqslant\bar{\delta}(\delta_{\#}) \quad \textit{and} \quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta_{\#}),$$

then the following holds.

Let (M,g) be a Riemannian manifold and $x\in M$ be a point that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption. Let $r>0$ be a constant and $\psi\colon S^2\times(-\delta^{-1},\delta^{-1})\to M$ be a diffeomorphism onto its image that satisfies $x\in\psi(S^2\times\{0\})$ and

$$\|r^{-2}\psi^*g-g^{S^2\times\mathbb{R}}\|_{C^0}<\delta,$$

where $g^{S^2\times\mathbb{R}}$ denotes the round cylindrical metric with $\rho\equiv 1$, and the C^0-norm is taken over the domain of ψ.

Then, x is a center of a $\delta_{\#}$-neck in M at scale r.

Proof. Without loss of generality, we may assume that $r=1$.

Assume that the lemma was false for some $\delta_{\#}>0$. Then, we can find sequences $\delta_k\to 0$ and $\varepsilon_{\mathrm{can},k}\to 0$, as well as a sequence $\{(M_k,g_k,x_k)\}$ of pointed Riemannian manifolds and a sequence $\{\psi_k\colon S^2\times(-\delta_k^{-1},\delta_k^{-1})\to M_k\}$ of diffeomorphisms onto their images such that, for all k, the following holds:

(1) (M_k,g_k) satisfies the $\varepsilon_{\mathrm{can},k}$ canonical neighborhood assumption at x_k;
(2) $x_k\in\psi_k(S^2\times\{0\})$;
(3) $\|\psi_k^*g_k-g^{S^2\times\mathbb{R}}\|_{C^0}<\delta_k\to 0$;
(4) x_k is not a center of a $\delta_{\#}$-neck at scale 1.

Let $\hat{r}_k:=\rho_1(x_k)$. Then, letting

$$(Z_k^1,g_k^1,z_k^1):=(S^2\times(-\delta_k^{-1},\delta_k^{-1}),\hat{r}_k^{-2}g^{S^2\times\mathbb{R}},\psi_k^{-1}(x_k)),$$
$$(Z_k^2,g_k^2,z_k^2):=(M_k,\hat{r}_k^{-2}g_k,x_k),$$

and $\phi_k:=\psi_k$, the assumptions of Lemma 8.8 hold by (3) above and assertion (a) of Lemma 8.1 together with the choice of $\hat{r}_k$. Therefore, we have $\rho(x_k)\to\rho(\psi_k^{-1}(x_k))=1$ as $k\to\infty$. It follows that (M_k,g_k,x_k) is $\varepsilon_{\mathrm{can},k}$-close at scale tending to 1 to the final time-slice $(\widehat{M}_k,\hat{g}_k,\hat{x}_k)$ of a κ_k-solution with $\rho(\hat{x}_k)=1$, as $k\to\infty$. Hence, $\mathrm{diam}(\widehat{M}_k,\hat{g}_k)\to\infty$. Since $\rho(\hat{x}_k)=1$, it follows that $(\widehat{M}_k,\hat{g}_k)$ cannot be a round metric for large k. So, by assertions (a) and (b) of Lemma C.1, after passing to a subsequence, the sequence $\{(M_k,g_k,x_k)\}$ converges in the pointed smooth topology to the final time-slice

$$(M_\infty,g_\infty,x_\infty)$$

of some κ-solution. However, by property (3) above, we conclude that (M_∞,g_∞) is isometric as a metric space to $(S^2\times\mathbb{R},g^{S^2\times\mathbb{R}})$ equipped with the induced length metric. So, (M_∞,g_∞) is isometric as a Riemannian manifold to $(S^2\times\mathbb{R},g^{S^2\times\mathbb{R}})$. Thus x_k is a center of a $\delta_{\#}$-neck at scale 1 for large k, contradicting (4). □

The next lemma gives control on the scale at bounded distance to a neck, assuming the canonical neighborhood assumption.

LEMMA 8.11. (Scale bounds near necks) *There is a constant $\delta_0>0$ such that for every $X<\infty$ there is a constant $Y=Y(X)<\infty$ such that, if*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(X),$$

then the following holds.

Let (M,g) be a (possibly incomplete) Riemannian manifold and let $\Sigma\subset M$ be a central 2-sphere of a δ_0-neck at scale 1 in M. Assume that M satisfies the ε_{can}-canonical neighborhood assumption at some point in Σ.

Consider a point $x\in M\setminus\Sigma$ and let $\mathcal{C}$ be the component of $M\setminus\Sigma$ containing x. If $d(x,\Sigma)\leqslant X$ and $\operatorname{diam}\mathcal{C}\geqslant Y$, then $\rho_1(x)>\frac{1}{10}$. Here the diameter is taken with respect to the distance function of (M,g).

The proof uses the geometry of non-negatively curved manifolds to bound neck scales from below. The argument is a variation on part of Perelman's proof of compactness of κ-solutions (see [P1]).

Proof. Fix $X<\infty$ and some small constant $\delta_0>0$. The precise conditions on the smallness of δ_0 will become clear in the course of the proof.

Assume that the statement of the lemma was false (for fixed X) and choose sequences $Y_k\to\infty$ and $\varepsilon_{\text{can},k}\to 0$. Then, we can find counterexamples (M_k,g_k), Σ_k, x_k, $\mathcal{C}_k\subset M_k\setminus\Sigma_k$ such that (M_k,g_k) satisfies the $\varepsilon_{\text{can},k}$-canonical neighborhood assumption at some point $y_k\in\Sigma_k$, $d(x_k,\Sigma_k)\leqslant X$, $\operatorname{diam}\mathcal{C}_k\geqslant Y_k$, but $\rho(x_k)\leqslant\frac{1}{10}$.

If

$$\delta_0 \leqslant \bar{\delta}_0,$$

then the injectivity radius at y_k is uniformly bounded from below by a positive constant. So, after passing to a subsequence, we may assume that

- The sequence of pointed Riemannian manifolds (M_k,g_k,y_k) converges to the pointed final time-slice $(M_\infty,g_\infty,y_\infty)$ of some κ-solution.
- The 2-spheres $\Sigma_k\subset M_k$ converge to a central 2-sphere Σ_∞ of a $2\delta_0$-neck $U_\infty\subset M_\infty$ at scale 1.
- The points x_k converge to a point $x_\infty\in M_\infty$ such that $\rho(x_\infty)\leqslant\frac{1}{10}$.
- $d(x_\infty,y_\infty)\geqslant\frac{1}{4}\delta_0^{-1}$, since we may assume that $\rho>\frac{1}{2}$ on the $2\delta_0$-neck U_∞.

As $\operatorname{diam}\mathcal{C}_k\geqslant Y_k\to\infty$, the κ-solution M_∞ must be non-compact. If

$$\delta_0 \leqslant \bar{\delta}_0,$$

then M_∞ cannot be isometric to a quotient of a round cylinder, because U_∞ is a $2\delta_0$-neck of scale 1, while $\rho(x_\infty)\leqslant\frac{1}{10}$. Therefore, M_∞ is diffeomorphic to $\mathbb{R}^3$, and the 2-sphere Σ_∞ bounds a compact domain, and a non-compact domain Z. We cannot have $x_\infty\in M_\infty\setminus Z$, since this would imply that $\operatorname{diam}\mathcal{C}_k\leqslant 2\operatorname{diam}(M_\infty\setminus Z)$ for large k, contradicting the fact that $\operatorname{diam}\mathcal{C}_k\to\infty$. So, $x_\infty\in Z$.

Let $\gamma\subset M_\infty$ be a minimizing geodesic ray starting from y_∞, and pick $z\in\gamma\cap Z$, to be determined later. Let $\overline{y_\infty z}$, $\overline{zx_\infty}$ and $\overline{x_\infty y_\infty}$ be minimizing geodesic segments between the corresponding pairs of points. Assuming

$$\delta_0\leqslant\bar{\delta}_0,$$

the segments $\overline{y_\infty z}$ and $\overline{y_\infty x_\infty}$ may intersect Σ_∞ at most once and are nearly parallel to the $\mathbb{R}$-factor of the neck U_∞. Therefore, both segments are contained in Z apart from the endpoint y_∞, and they form an angle of at most $\frac{1}{4}\pi$ at y_∞. By Toponogov's theorem, this implies that the comparison angle $\widetilde{\angle}_{y_\infty}(x_\infty,z)$ is at most $\frac{1}{4}\pi$. Provided that $d(z,y_\infty)$ is sufficiently large, we therefore have $\widetilde{\angle}_{x_\infty}(y_\infty,z)>\frac{1}{4}\pi$.

Fix some small $\delta_1>0$ whose value we will determine later. If

$$\delta_0\leqslant\bar{\delta}_0$$

and $d(z,y_\infty)$ is sufficiently large, then $\rho^{-1}(x_\infty)\min\{d(x_\infty,z),d(x_\infty,y_\infty)\}$ is large enough that we may apply [KL1, Corollary 49.1] to conclude that x_∞ is a center of a δ_1-neck, with central 2-sphere $\Sigma_{x_\infty}\subset M_\infty$. If $\delta_1\leqslant\bar{\delta}_1$, then the segments $\overline{x_\infty z}$ and $\overline{x_\infty y_\infty}$ intersect Σ_{x_∞} only at x_∞ and are nearly parallel to the $\mathbb{R}$-factor the neck at x_∞. Since their angle at x_∞ is $>\frac{1}{4}\pi$, it follows that y_∞ and z lie in distinct connected components of $M_\infty\setminus\Sigma_{x_\infty}$.

Let c_0 be the diameter of a central 2-sphere of a round cylinder of scale 1. If $\delta_0\leqslant\bar{\delta}_0$, we may choose a point $y'_\infty\in\Sigma_\infty$ such that $d(y'_\infty,y_\infty)\geqslant 0.99c_0$. Now consider geodesic segments $\overline{y_\infty z}$ and $\overline{y'_\infty z}$. If $\delta_0\leqslant\bar{\delta}_0$, both segments are contained in Z, and since Σ_{x_∞} separates y_∞ from z, both segments intersect Σ_{x_∞}. If $\delta_0\leqslant\bar{\delta}_0$ then

$$|d(z,y_\infty)-d(z,y'_\infty)|<0.01c_0,$$

as follows by applying the triangle inequality to points on $\overline{y_\infty z}$ and $\overline{y'_\infty z}$ at distance $\frac{1}{2}\delta_0^{-1}$. Therefore, after swapping the labels of y_∞ and y'_∞ if necessary, we may assume without loss of generality that there is a point $y''_\infty\in\overline{y'_\infty z}$ such that $d(z,y''_\infty)=d(z,y_\infty)$ and $d(y_\infty,y''_\infty)>.98c_0$. Similarly, if $\delta_1\leqslant\bar{\delta}_1$, there are points $w_\infty\in\overline{y_\infty z}$ and $w'_\infty\in\overline{y''_\infty z}$ such that

$$d(w_\infty,w'_\infty)<1.01c_0\rho(x_\infty)<\tfrac{1}{5}c_0,$$

$d(w_\infty, z) = d(w'_\infty, z)$ and one of w_∞ and w'_∞ lies on Σ_{x_∞}. By Toponogov's theorem (monotonicity of comparison angles) we have

$$\frac{d(y_\infty, y''_\infty)}{d(y_\infty, z)} \leqslant \frac{d(w_\infty, w'_\infty)}{d(w_\infty, z)}.$$

So, if $d(z, y_\infty)$ is sufficiently large, then

$$0.98 c_0 \leqslant d(y_\infty, y''_\infty) \leqslant 2 d(w_\infty, w'_\infty) < \tfrac{2}{5} c_0,$$

which is a contradiction. □

8.2. Promoting time-slice models to spacetime models

Our next two results show that under appropriate completeness and canonical neighborhood assumptions, if a time-slice of a Ricci flow spacetime is close to a neck or a Bryant soliton, then a parabolic region is also close to a neck or Bryant soliton, respectively. The proofs are standard convergence arguments based on a rigidity property of necks and Bryant solitons among κ-solutions.

LEMMA 8.12. (Time-slice necks imply spacetime necks) *If*

$$\delta_\# > 0, \quad 0 < \delta \leqslant \bar{\delta}(\delta_\#), \quad 0 < \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta_\#), \quad 0 < r \leqslant \bar{r},$$

then the following holds.

Assume that $\mathcal{M}$ is an $(\varepsilon_{\text{can}} r, t_0)$-complete Ricci flow spacetime that satisfies the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon_{\text{can}} r, 1)$. Let $a \in \left[-1, \frac{1}{4}\right]$ and consider a time $t \geqslant 0$ such that $t + ar^2 \in [0, t_0]$.

Assume that $U \subset \mathcal{M}_{t+ar^2}$ is a δ-neck at scale $r\sqrt{1-3a}$. So, there is a diffeomorphism

$$\psi_1 \colon S^2 \times (-\delta^{-1}, \delta^{-1}) \longrightarrow U$$

such that

$$\| r^{-2} \psi_1^* g_{t+ar^2} - g_a^{S^2 \times \mathbb{R}} \|_{C^{[\delta^{-1}]}} < \delta. \tag{8.21}$$

Here, $(g_t^{S^2 \times \mathbb{R}})_{t \in (-\infty, 1/3)}$ denotes the shrinking round cylinder with $\rho(\cdot, 0) = 1$ at time 0 and the $C^{[\delta^{-1}]}$-norm is taken over the domain of ψ_1.

Then, there is a product domain $U^* \subset \mathcal{M}_{[t-r^2, t+r^2/4] \cap [0, t_0]}$ and an r^2-time-equivariant and $\partial_{\mathfrak{t}}$-preserving diffeomorphism

$$\psi_2 \colon S^2 \times (-\delta_\#^{-1}, \delta_\#^{-1}) \times [t^*, t^{**}] \longrightarrow U^*,$$

with $t+t^*r^2=\max\{t-r^2,0\}$ *and* $t+t^{**}r^2=\min\{t+\frac{1}{4}r^2,t_0\}$, *such that*

$$\psi_2|_{S^2\times(-\delta_\#^{-1},\delta_\#^{-1})\times\{a\}}=\psi_1|_{S^2\times(-\delta_\#^{-1},\delta_\#^{-1})}$$

and

$$\|r^{-2}\psi_2^*g-g^{S^2\times\mathbb{R}}\|_{C^{[\delta_\#^{-1}]}}<\delta_\#.$$

Here the $C^{[\delta_\#^{-1}]}$*-norm is taken over the domain of* ψ_2.

Note that the lemma can be generalized to larger time-intervals. We have omitted this aspect, as it will not be important for us later. We also remark that one may prove a more general result to the effect that any parabolic region is close to a parabolic region in a κ-solution.

Proof. For the following proof, we may assume that $\bar{r}$ and $\bar{\varepsilon}_{\text{can}}$ are chosen small enough such that any point $x\in\mathcal{M}$ with $\frac{1}{10}r\leqslant\rho(x)\leqslant 10r$ satisfies the ε_{can}-canonical neighborhood assumption.

Assuming

$$\delta\leqslant\bar{\delta},$$

we have the following bound on the image of ψ_1:

$$\tfrac{1}{4}r\leqslant\tfrac{1}{2}r\sqrt{1-3a}<\rho<2r\sqrt{1-3a}\leqslant 4r. \tag{8.22}$$

Assume now that the statement of the lemma was false for some fixed $\delta_\#>0$. So, there are sequences $\varepsilon_{\text{can},k}\to 0$, $\delta_k\to 0$, $r_k\leqslant\bar{r}$, $t_k\geqslant r_k^2$, $a_k\in[-1,\frac{1}{4}]$, $t_{0,k}\geqslant 0$, $t_k^*\in[-1,0]$, $t_k^{**}\in[0,\frac{1}{4}]$, with $t_k+t_k^*r_k^2=\max\{t_k-r_k^2,0\}$ and $t_k+t_k^{**}r_k^2=\min\{t_k+\frac{1}{4}r_k^2,t_{0,k}\}$, as well as a sequence $\{\mathcal{M}^k\}$ of Ricci flow spacetimes that satisfy the $\varepsilon_{\text{can},k}$-canonical neighborhood assumption at scales $(\varepsilon_{\text{can},k}r_k,1)$ and maps $\psi_{1,k}$ belonging to δ_k-necks at time t_k and scale $r_k\sqrt{1-3a_k}$, but for which the conclusion of the lemma fails. After passing to a subsequence, we may assume that $t_\infty^*:=\lim_{k\to\infty}t_k^*$, $t_\infty^{**}:=\lim_{k\to\infty}t_k^{**}$ and $a_\infty:=\lim_{k\to\infty}a_k$ exist.

Choose $a_\infty^*\in[t_\infty^*,a_\infty]$ and $a_\infty^{**}\in[a_\infty,t_\infty^{**}]$ minimal and maximal, respectively, such that for any $d>0$ and any compact interval $[s_1,s_2]\subset(a_\infty^*,a_\infty^{**})$ the following holds for large k (possibly depending on d, s_1 and s_2): For all $x\in\psi_{1,k}(S^2\times(-d,d))$ and $t'\in[t_k+s_1r_k^2,t_k+s_2r_k^2]$ the point $x(t')$ is defined and we have

$$\tfrac{1}{10}r_k\leqslant\rho(x(t'))\leqslant 10r_k. \tag{8.23}$$

Note that, by the remark in the beginning of the proof, this implies that $x(t')$ satisfies the canonical $\varepsilon_{\text{can},k}$-canonical neighborhood assumption. By (8.22) and Lemma 8.4, we know that $a_\infty^*<a_\infty$ if $a_\infty>t_\infty^*$ and $a_\infty^{**}>a_\infty$ if $a_\infty<t_\infty^{**}$.

By the choices of $a^*_\infty, a^{**}_\infty$ we can find sequences $d_k \to \infty$, $a^*_k \in [-1, a_k]$ and $a^{**}_k \in \left[a_k, \frac{1}{4}\right]$ with $\lim_{k\to\infty} a^*_k = a^*_\infty$ and $\lim_{k\to\infty} a^{**}_k = a^{**}_\infty$ such that the set

$$P_k := (\psi_{1,k}(S^2 \times (-d_k, d_k)))([t_k + a^*_k r_k^2, t_k + a^{**}_k r_k^2])$$

is well defined and such that $\frac{1}{10} r_k \leqslant \rho \leqslant 10 r_k$ on P_k. For every k consider the parabolically rescaled flow $(g'_{k,s})_{s\in(a^*_k, a^{**}_k)}$ on $S^2 \times (-d_k, d_k)$ with

$$g'_{k,s} := r_k^{-2} \hat{g}_{t_k + s r_k^2}, \tag{8.24}$$

where $\hat{g}_{t_k+sr_k^2}$ denotes the pullback of $g_{t_k+sr_k^2}$ under the composition of $\psi_{1,k}$ with the map

$$\psi_{1,k}(S^2 \times (-d_k, d_k)) \longrightarrow P_k$$

that is given by the time $(s - a_k) r_k^2$-flow of $\partial_{\mathfrak{t}}$.

By (8.23) and the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption (see assertion (a) of Lemma 8.1), we obtain that the curvature of $(g'_{k,s})_{s\in(a^*_k, a^{**}_k)}$, along with its covariant derivatives, is uniformly bounded. Together with (8.21), these bounds imply uniform C^m-bounds on the tensor fields $(g'_{k,s})$ themselves. So, by passing to a subsequence, we obtain that the $(g'_{k,s})$ converge to a Ricci flow $(g'_{\infty,s})_{s\in(a^*_\infty, a^{**}_\infty)}$ on $S^2 \times \mathbb{R}$, which extends smoothly to the time-interval $[a^*_\infty, a^{**}_\infty]$.

The $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption implies that all time-slices of this limit are final time-slices of κ-solutions. By (8.21) we know that $g'_{\infty,a_\infty} = g^{S^2\times\mathbb{R}}_{a_\infty}$. Since $S^2 \times \mathbb{R}$ has two ends, $g'_{\infty,s}$ splits off an $\mathbb{R}$ factor for all $s \in (a^*_\infty, s^*_\infty)$ and must therefore be homothetic to a round cylinder. It follows that $g'_{\infty,s} = g^{S^2\times\mathbb{R}}_s$ for all $s \in [a^*_\infty, a^{**}_\infty]$. Since this limit is unique, we obtain that the $(g'_{k,s})$ converge to $(g'_{\infty,s})$ even without passing to a subsequence.

As $\frac{1}{2} \leqslant \rho \leqslant 2$ on $(S^2 \times \mathbb{R}) \times (a^*_\infty, a^{**}_\infty)$, we get that, for any $d > 0$ and $[s_1, s_2] \in (a^*_\infty, a^{**}_\infty)$, we have $\frac{1}{4} r_k \leqslant \rho \leqslant 4 r_k$ on $(\psi_{1,k}(S^2 \times (-d, d)))([s_1, s_2])$ for large k. So, by Lemma 8.4 and the minimal and maximal choices of a^*_∞ and a^{**}_∞, we have $a^*_\infty = t^*_\infty$ and $a^{**}_\infty = t^{**}_\infty$. Moreover, after adjusting the sequence $d_k \to \infty$, we may assume that $a^*_k = t^*_k$ and $a^{**}_k = t^{**}_k$ for large k.

For large k we now define $\psi_{2,k}$ by extending $\psi_{1,k}$ restricted to $S^2 \times (-\delta_\#^{-1}, \delta_\#^{-1})$ forward and backward using the flow of $r_k^2 \partial_{\mathfrak{t}}$. Then, we have

$$r_k^{-2} \psi^*_{2,k} g_k = g'_{k,s}$$

on $(S^2 \times (-\delta_\#^{-1}, \delta_\#^{-1})) \times [t^*_k, t^{**}_k]$. So, it suffices to show that $g'_{k,s}$ converges to $g^{S^2\times\mathbb{R}}$ on $(S^2 \times (-\delta_\#^{-1}, \delta_\#^{-1})) \times [t^*_k, t^{**}_k]$ uniformly in the $C^{[\delta_\#^{-1}]}$-sense. To see this, note that $g'_{k,s}$ from

(8.24) is uniformly bounded on $(S^2\times(-\delta_\#^{-1},\delta_\#^{-1}))\times\left[-1,\frac{1}{4}\right]$ in every C^m-norm and that we have uniform convergence of $g'_{k,s}$ to $g^{S^2\times\mathbb{R}}$ on every subset of the form

$$(S^2\times(-\delta_\#^{-1},\delta_\#^{-1}))\times[s_1,s_2]$$

for $[s_1,s_2]\subset(t^*_\infty,t^{**}_\infty)$, in every C^m-norm. □

For notation and facts about the Bryant soliton, see §6.2 and Appendix B. In the following result, it is important that $\rho\geqslant 1$ on the normalized Bryant soliton.

LEMMA 8.13. (Propagating Bryant-like geometry) *If*

$$\delta_\#>0,\quad T<\infty,\quad \delta\leqslant\bar\delta(\delta_\#,T),\quad \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}}(\delta_\#,T)\quad \textit{and}\quad r\leqslant\bar r,$$

then the following holds.

Assume that $\mathcal{M}$ is an $(\varepsilon_{\mathrm{can}}r,t_0)$-complete Ricci flow spacetime that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}}r,1)$. Let $t\in[0,t_0]$ and consider a diffeomorphism onto its image

$$\psi_1\colon M_{\mathrm{Bry}}(\delta^{-1})\times\{0\}\longrightarrow\mathcal{M}_t$$

with the property that

$$\|r^{-2}\psi_1^*g_t-g_{\mathrm{Bry}}\|_{C^{[\delta^{-1}]}(M_{\mathrm{Bry}}(\delta^{-1})\times\{0\})}<\delta. \tag{8.25}$$

Then, there is an r^2-time equivariant and ∂_t-preserving diffeomorphism onto its image

$$\psi_2\colon M_{\mathrm{Bry}}(\delta_\#^{-1})\times[t^*,t^{**}]\longrightarrow\mathcal{M}_{[t-Tr^2,t+Tr^2]\cap[0,t_0]},$$

where $t^*\leqslant 0\leqslant t^{**}$ are chosen such that

$$t+t^*r^2=\max\{t-Tr^2,0\}\quad \textit{and}\quad t+t^{**}r^2=\min\{t+Tr^2,t_0\}.$$

The map ψ_2 has the property that $\psi_2=\psi_1$ on $M_{\mathrm{Bry}}(\delta_\#^{-1})\times\{0\}$ and

$$\|r^{-2}\psi_2^*g-g_{\mathrm{Bry}}\|_{C^{[\delta_\#^{-1}]}}<\delta_\#,$$

where the norm is taken over the domain of ψ_2.

Proof. The proof is similar to the proof of Lemma 8.12.

In the following, we may assume that $\bar r$ and $\bar\varepsilon_{\mathrm{can}}$ are chosen small enough such that any point with $x\in\mathcal{M}$ with $\frac{1}{10}r\leqslant\rho(x)\leqslant 10r$ satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption.

Assuming

$$\delta \leqslant \bar{\delta},$$

we have

$$\tfrac{1}{4} r \leqslant \rho \quad \text{on } \operatorname{Im} \psi_1 \quad \text{and} \quad \tfrac{1}{4} r \leqslant \rho(\psi_1(x_{\mathrm{Bry}})) \leqslant 4r.$$

Assume now that the statement of the lemma was false for some fixed $\delta_{\#}>0$ and $T<\infty$. So, there are a sequences $\{\varepsilon_{\mathrm{can},k}\}$, $\{\delta_k\}$, $\{t_{0,k}\}$, $\{r_k\}$, $\{t_k\}$, $\{t_k^*\}$ and $\{t_k^{**}\}$ such that $\varepsilon_{\mathrm{can},k} \to 0$, $\delta_k \to 0$, as well as a sequence $\{\mathcal{M}^k\}$ of $(\varepsilon_{\mathrm{can},k} r_k, t_{0,k})$-complete Ricci flow spacetimes that satisfy the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can},k} r_k, 1)$ and a sequence of maps $\{\psi_{1,k}\}$ satisfying the hypotheses of the lemma, but for which the conclusion of the lemma fails for all k. By passing to a subsequence, we may assume that $t_\infty^* := \lim_{k\to\infty} t_k^*$ and $t_\infty^{**} := \lim_{k\to\infty} t_k^{**}$ exist.

Choose $a_\infty^* \in [t_\infty^*, 0]$, $a_\infty^{**} \in [0, t_\infty^{**}]$ minimal and maximal, respectively, such that for any $d>0$ and any compact interval $[s_1, s_2] \subset (a_\infty^*, a_\infty^{**})$ the following holds for large k: For every $x \in \psi_{1,k}(M_{\mathrm{Bry}}(d))$ and $t' \in [t_k + s_1 r_k^2, t_k + s_2 r_k^2]$ the point $x(t')$ is well defined and we have $\rho(x(t')) \geqslant \frac{1}{10} r_k$ and $\rho((\psi_{1,k}(x_{\mathrm{Bry}}))(t')) \leqslant 10 r_k$. Note that $a_\infty^* < 0$ if $t_\infty^* < 0$ and $a_\infty^{**} > 0$ if $t_\infty^{**} > 0$, due to Lemma 8.4.

As in the proof of Lemma 8.12, we can now find sequences $d_k \to \infty$ and $a_k^* \in [t_k^*, 0]$, $a_k^{**} \in [0, t_k^{**}]$ with $\lim_{k\to\infty} a_k^* = a_\infty^*$ and $\lim_{k\to\infty} a_k^{**} = a_\infty^{**}$, such that the product domains

$$P_k := (\psi_{1,k}(M_{\mathrm{Bry}}(d_k)))([t_k + a_k^* r_k^2, t_k + a_k^{**} r_k^2])$$

are well defined and such that

$$\rho \geqslant \tfrac{1}{10} r_k \quad \text{on } P_k,$$

and

$$\rho((\psi_{1,k}(x_{\mathrm{Bry}}))(t')) \leqslant 10 r_k \quad \text{for all } t' \in [t_k + a_k^* r_k^2, t_k + a_k^{**} r_k^2].$$

So, $(\psi_{1,k}(x_{\mathrm{Bry}}))(t')$ satisfies the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption for all

$$t' \in [t_k + a_k^* r_k^2, t_k + a_k^{**} r_k^2].$$

For every k consider the parabolically rescaled flow $(g'_{k,s})_{s \in [a_k^*, a_k^{**}]}$ on $M_{\mathrm{Bry}}(d_k)$ with $g'_{k,s} := r_k^{-2} \hat{g}_{t_k + s r_k^2}$, where $\hat{g}_{t_k + r_k^2 s}$ denotes the pullback of $g_{t_k + s r_k^2}$ under the composition of $\psi_{1,k}$ with the map

$$\mathcal{M}^k_{t_k} \supset \psi_{1,k}(M_{\mathrm{Bry}}(d_k) \times \{0\}) \longrightarrow P_k,$$

given by the time $r_k^2 s$-flow of $\partial_{\mathfrak{t}}$. By the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption at $(\psi_{1,k}(x_{\mathrm{Bry}}))(t')$ (see assertion (a) of Lemma 8.1) and a distance distortion estimate on P_k, we obtain that the curvature of this flow, along with its derivatives, is uniformly bounded

by a constant that may only depend on the spatial direction. Together with (8.25), these bounds imply uniform local C^m-bounds on the tensor fields $(g'_{k,s})$ themselves.

So, by passing to a subsequence, we obtain that $(g'_{k,s})$ converges to a Ricci flow $(g'_{\infty,s})_{s\in(a^*_\infty,a^{**}_\infty)}$ on M_{Bry} with uniformly bounded curvature, which extends smoothly to the time-interval $[a^*_\infty,a^{**}_\infty]$. By the $\varepsilon_{\mathrm{can},k}$-canonical neighborhood assumption at $(\psi_{1,k}(x_{\mathrm{Bry}}))(t')$ and the compactness of κ-solutions (see (a) and (b) of Lemma C.1), we find that all time-slices of this limit are final time-slices of κ-solutions. By (8.25), we furthermore know that $g'_{\infty,0}=g_{\mathrm{Bry},0}$.

We now claim that $g'_{\infty,s}=g_{\mathrm{Bry},s}$ for all $s\in[a^*_\infty,a^{**}_\infty]$. For $s\geqslant 0$, this follows from the uniqueness of Ricci flows with uniformly bounded curvature. To verify this in the case $s<0$, recall that there is a κ-solution $(g''_s)_{s\in(-\infty,0]}$ on M_{Bry} such that $g''_0=g'_{a^*_\infty}$. Now, set $g'''_s:=g'_{\infty,s}$ if $a^*_\infty\leqslant s\leqslant 0$ and $g'''_s:=g''_{s-a^*_\infty}$ if $s<a^*_\infty$. Then, $(g'''_s)_{s\in(-\infty,0]}$ is a smooth κ-solution (possibly after adjusting κ). Since $\partial_t R_{g'''}(x_{\mathrm{Bry}},0)=0$, it follows from Proposition C.3 that $(M_{\mathrm{Bry}},(g'''_s)_{s\in(-\infty,0]},x_{\mathrm{Bry}})$ is isometric to $(M_{\mathrm{Bry}},(g_{\mathrm{Bry},t})_{t\in(-\infty,0]},x_{\mathrm{Bry}})$. Thus $g'_{\infty,s}=g_{\mathrm{Bry},s}$ for all $s\in[a^*_\infty,a^{**}_\infty]$. As in the proof of Lemma 8.12, the uniqueness of the limit implies that the $(g'_{k,s})$ converge to $(g'_{\infty,s})$ even without passing to a subsequence.

So, $(g'_{\infty,s})_{s\in[a^*_\infty,a^{**}_\infty]}$ satisfies $\rho\geqslant 1$ everywhere and $\rho(x_{\mathrm{Bry}},s)=1$ for all $s\in[a^*_\infty,a^{**}_\infty]$. Therefore, by the minimal and maximal choices of a^*_∞ and a^{**}_∞ and Lemma 8.4, we obtain that $a^*_\infty=t^*_\infty$ and $a^{**}_\infty=t^{**}_\infty$. Moreover, after possibly adjusting the sequence $d_k\to\infty$, we may assume that $a^*_k=t^*_k$ and $a^{**}_k=t^{**}_k$ for large k. The claim now follows as in the proof of Lemma 8.12. □

8.3. Identifying approximate Bryant structure

In the next result, we exploit the rigidity theorems of Hamilton and Brendle to show that a large region must be well approximated by a Bryant soliton if the scale is nearly increasing at a point.

LEMMA 8.14. *If*

$$\delta_\#>0,\quad \delta\leqslant\bar\delta(\delta_\#)\quad \textit{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}}(\delta_\#),$$

then the following holds.

If $\mathcal{M}$ is a Ricci flow spacetime satisfying the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at $x\in\mathcal{M}$, and $\partial_t\rho^2(x)\geqslant-\delta$, then $(\mathcal{M}_t,x)$ is $\delta_\#$-close to $(M_{\mathrm{Bry}},g_{\mathrm{Bry}},x_{\mathrm{Bry}})$ at any scale $a\in((1-\delta)\rho(x),(1+\delta)\rho(x))$.

Note that $\partial_t\rho^2$ is scaling invariant.

Proof. Suppose the lemma was false for some $\delta_{\#}>0$. Then, there a sequence $\{\mathcal{M}^k\}$ of Ricci flow spacetimes satisfying the $(1/k)$-canonical neighborhood assumption at $x_k \in \mathcal{M}^k_{t_k}$, such that $\partial_t \rho^2(x_k) \geqslant -1/k$, but $(\mathcal{M}^k_{t_k}, x_k)$ is not $\delta_{\#}$-close to $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at some scale $a_k \in ((1-1/k)\rho(x_k), (1+1/k)\rho(x_k))$.

By the definition of the canonical neighborhood assumption, for every k there is a pointed κ_k-solution $(\overline{M}_k, (\bar{g}_{k,t})_{t\in(-\infty,0]}, \bar{x}_k)$ with $\rho(\bar{x}_k)=1$ and a diffeomorphism onto its image

$$\psi_k \colon B(\bar{x}_k, 0, k) \longrightarrow \mathcal{M}^k_{t_k}$$

with $\psi_k(\bar{x}_k)=x_k$ such that, for some $\lambda_k>0$ with $\lambda_k/\rho(x_k)\to 1$, we have

$$\|\lambda_k^{-2}\psi_k^* g_k - \bar{g}_k\|_{C^k(B(\bar{x}_k,k))} < k^{-1}.$$

So, we also have

$$\|a_k^{-2}\psi_k^* g_k - \bar{g}_k\|_{C^k(B(\bar{x}_k,k))} \to 0.$$

Hence, $\liminf_{k\to\infty} \partial_t \rho^2(\bar{x}_k) \geqslant 0$. Therefore, $(M_k, (\bar{g}_{k,t})_{t\in(-\infty,0]})$ cannot be a shrinking round spherical space form for large k. So, by assertions (a), (b) and (e) of Lemma C.1, after passing to a subsequence, $(\overline{M}_k, (\bar{g}_{k,t})_{t\in(-\infty,0]}, \bar{x}_k)$ converges in the pointed smooth topology to a κ-solution $(\overline{M}_\infty, (\bar{g}_{\infty,t})_{t\in(-\infty,0]}, \bar{x}_\infty)$ with $\partial_t \rho^2(\bar{x}_\infty)=0$. Now, by Proposition C.3, it follows that $(\overline{M}_\infty, (\bar{g}_{\infty,t})_{t\in(-\infty,0]}, \bar{x}_\infty)$ is isometric to a Bryant soliton. This is a contradiction. □

By combining Lemmas 8.14 and 8.13, we can deduce closeness to a Bryant soliton on a *parabolic* region.

LEMMA 8.15. (Nearly increasing scale implies Bryant-like geometry) *If*

$$\alpha, \delta > 0, \quad 1 \leqslant J < \infty, \quad \beta \leqslant \bar{\beta}(\alpha, \delta, J), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\alpha, \delta, J) \quad \textit{and} \quad r \leqslant \bar{r}(\alpha)$$

then the following holds.

Let $0<r\leqslant 1$. Assume that $\mathcal{M}$ is an $(\varepsilon_{\mathrm{can}} r, t_0)$-complete Ricci flow spacetime that satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}} r, 1)$.

Let $t \in [Jr^2, t_0]$ and $x \in \mathcal{M}_t$. Assume that x survives until time $t-r^2$ and that

$$\alpha r \leqslant \rho(x) \leqslant \alpha^{-1} r \quad \textit{and} \quad \rho(x(t-r^2)) \leqslant \rho(x) + \beta r.$$

Let $a \in [\rho(x(t-r^2)), \rho(x)+\beta r]$. Then, $(\mathcal{M}_{t'}, x(t'))$ is δ-close to $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at scale a for all $t' \in [t-r^2, t]$. Also, there is an a^2-time-equivariant and ∂_t-preserving diffeomorphism onto its image

$$\psi \colon \overline{M_{\mathrm{Bry}}(\delta^{-1})} \times [-J\cdot(ar^{-1})^{-2}, 0] \longrightarrow \mathcal{M}$$

such that $\psi(x_{\mathrm{Bry}},0)=x$ *and*

$$\|a^{-2}\psi^*g-g_{\mathrm{Bry}}\|_{C^{[\delta^{-1}]}}<\delta,$$

where the norm is taken over the domain of ψ.

Proof. Let $2>L>1$ and $\delta'>0$ be constants whose values will be determined in the course of this proof. By Lemma 8.5, and assuming that

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\alpha,L)\quad\text{and}\quad r\leqslant\bar{r}(\alpha),$$

we obtain that for all $t'\in[t-r^2,t]$ we have

$$\tfrac{1}{2}\alpha r\leqslant L^{-1}\rho(x)\leqslant\rho(x(t'))\leqslant L\rho(x(t-r^2)).$$

If moreover

$$\beta\leqslant\bar{\beta}(\alpha,L),$$

then

$$\rho(x(t-r^2))\leqslant\rho(x)+\beta r\leqslant\rho(x)+(L-1)\alpha r\leqslant L\rho(x).$$

Therefore, $L^{-1}\rho(x)\leqslant\rho(x(t'))\leqslant L^2\rho(x)$ and $a\in[L^{-1}\rho(x(t')),L^2\rho(x(t'))]$ for all $t'\in[t-r^2,t]$. We also obtain that $x(t')$ satisfies the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption for all $t'\in[t-r^2,t]$, assuming that

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}\quad\text{and}\quad r\leqslant\bar{r}.$$

By the mean value theorem, we can find a $t'\in[t-r^2,t]$ at which

$$\partial_t\rho^2(x(t'))\geqslant-2\rho(x(t'))\cdot\frac{(L^2-L^{-1})\rho(x)}{r^2}\geqslant-2\alpha^{-2}(L^2-L^{-1})L.$$

Therefore, if

$$L\leqslant1+\bar{L}(\delta')\quad\text{and}\quad\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta'),$$

then Lemma 8.14 implies that $(\mathcal{M}_{t'},x)$ is δ'-close to $(M_{\mathrm{Bry}},g_{\mathrm{Bry}},x_{\mathrm{Bry}})$ at scale a. Assuming

$$\delta'\leqslant\bar{\delta}'(\alpha,\delta,J),\quad\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\alpha,\delta,J)\quad\text{and}\quad r\leqslant\bar{r}(\alpha),$$

the claim now follows from Lemma 8.13. □

8.4. The geometry of comparison domains

The results in this subsection analyze the structure of comparison domains (and related subsets) of spacetimes that satisfy completeness and canonical neighborhood conditions, as well as some of the a-priori assumptions (APA1)–(APA6), as introduced in §7.

The first two results — the Bryant slice lemma (Lemma 8.16) and the Bryant slab lemma (Lemma 8.17) — describe the structure of comparison domains in approximate Bryant regions. These results are helpful in showing that neck-like boundaries of comparison domains and cuts are far apart (Lemma 8.19), and in facilitating the construction of the comparison domain in §11.

The Bryant slice lemma characterizes how a domain X in a time-slice $\mathcal{M}_t$ that is bounded by a central 2-sphere of a sufficiently precise neck intersects a domain $W \subset \mathcal{M}_t$ that is geometrically close to a Bryant soliton. The domain X will later be equal to either backward time-slice $\mathcal{N}_{t_j-}$ of a comparison domain or the domain Ω from §11.

LEMMA 8.16. (Bryant slice lemma) *If*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad 0 < \lambda < 1, \quad \Lambda \geqslant \underline{\Lambda} \quad \text{and} \quad \delta \leqslant \bar{\delta}(\lambda, \Lambda),$$

then the following holds for some $D_0 = D_0(\lambda) < \infty$.

Consider a Ricci flow spacetime $\mathcal{M}$ *and let* $r > 0$ *and* $t \geqslant 0$. *Consider a subset* $X \subset \mathcal{M}_t$ *such that the following holds*:

(i) X *is a closed subset and is a domain with smooth boundary.*

(ii) *The boundary components of* X *are central 2-spheres of* δ_{n}*-necks at scale* r.

(iii) X *contains all* Λr*-thick points of* $\mathcal{M}_t$.

(iv) *Every component of* X *contains a* Λr*-thick point.*

Consider the image W *of a diffeomorphism*

$$\psi \colon W^* := \overline{M_{\mathrm{Bry}}(d)} \longrightarrow W \subset \mathcal{M}_t,$$

such that $d \geqslant \delta^{-1}$ *and*

$$\| (10\lambda r)^{-2} \psi^* g_t - g_{\mathrm{Bry}} \|_{C^{[\delta^{-1}]}(W^*)} < \delta.$$

Then, $\psi(x_{\mathrm{Bry}})$ *is* $11\lambda r$*-thin. Moreover, if* $\mathcal{C} := W \setminus \operatorname{Int} X \neq \varnothing$, *then*

(a) $\mathcal{C}$ *is a 3-disk containing* $\psi(x_{\mathrm{Bry}})$.

(b) $\mathcal{C}$ *is a component of* $\mathcal{M}_t \setminus \operatorname{Int} X$, *and* $\partial \mathcal{C} \subset \partial X$.

(c) $\mathcal{C}$ *is* $9\lambda r$*-thick and* $1.1r$*-thin.*

(d) $\mathcal{C} \subset \psi(M_{\mathrm{Bry}}(D_0(\lambda))) \subset \operatorname{Int} W$.

Proof. Assuming

$$\delta \leqslant \bar{\delta}(\lambda, \Lambda),$$

it follows from the definition of W^* that ∂W is Λr-thick, W is $9\lambda r$-thick, and the image of the tip $\psi(x_{\mathrm{Bry}})$ is $11\lambda r$-thin. The fact that ∂W is Λr-thick and assumption (iii) imply that $\partial W \subset \operatorname{Int} X$.

Consider a boundary component $\Sigma \subset \partial X$ with $\Sigma \cap W \neq \varnothing$. Let $U_\Sigma \subset \mathcal{M}_t$ be a δ_{n}-neck at scale r that has Σ as a central 2-sphere. If

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}},$$

then we have $.99r < \rho < 1.01r$ on U_Σ. Assuming

$$\Lambda \geqslant 10,$$

we find that, then $U_\Sigma \cap \partial W = \varnothing$ and hence $U_\Sigma \subset W$.

Next, if

$$\delta \leqslant \bar{\delta}(\lambda),$$

then $.98(10\lambda)^{-1} < \rho < 1.02(10\lambda)^{-1}$ on $\psi^{-1}(U_\Sigma)$. Moreover, the 2-sphere $\Sigma^* := \psi^{-1}(\Sigma)$ is isotopic within the set $\{.98(10\lambda)^{-1} \leqslant \rho \leqslant 1.02(10\lambda)^{-1}\} \subset W^*$ to the 2-sphere

$$\widehat{\Sigma}^* = \{\rho = 1.02(10\lambda)^{-1}\}$$

in W^*.

By Alexander's theorem, Σ^* bounds a 3-disk $V^*_\Sigma \subset W^*$. By the previous paragraph, we have $V^*_\Sigma \subset \widehat{V}^*_\Sigma := \{\rho \leqslant 1.02(10\lambda)^{-1}\}$. Thus, if

$$\delta \leqslant \bar{\delta}(\lambda),$$

then $V_\Sigma := \psi(V^*_\Sigma) \subset \psi(\widehat{V}^*_\Sigma)$ is $1.1r$-thin and contains $\psi(x_{\mathrm{Bry}})$.

Lastly, suppose that Σ_1 and Σ_2 are distinct components of ∂X that intersect W. Let V_{Σ_1} and V_{Σ_2} be the corresponding 3-disk components, as defined in the discussion above. Since $\psi(x_{\mathrm{Bry}}) \in V_{\Sigma_1} \cap V_{\Sigma_2}$, we may assume (after reindexing) that $V_{\Sigma_1} \subset V_{\Sigma_2}$.

If X_0 is the component of X containing $\Sigma_1 = \partial V_{\Sigma_1}$, then it must be contained in V_{Σ_2} since every arc leaving V_{Σ_2} must intersect $\partial X \supset \partial V_{\Sigma_2}$. Thus, X_0 is $1.1r$-thin, contradicting assumption (iv) for

$$\Lambda > 1.1.$$

Thus W intersects at most one component of ∂X.

Now, suppose $\mathcal{C} := W \setminus \operatorname{Int} X$ is non-empty. Since $\partial W \subset \operatorname{Int} X$, we have $\mathcal{C} \neq W$. By the discussion above, we see that $\partial X \cap W$ consists of a single 2-sphere component Σ, where Σ bounds a 3-disk V_Σ which contains $\psi(x_{\mathrm{Bry}})$. Thus $\mathcal{C} = V_\Sigma$, and assertions (a)–(c) now follow immediately. For assertion (d), recall that $\mathcal{C} \subset \psi(\widehat{V}^*_\Sigma) = \psi(\{\rho \leqslant 1.02(10\lambda)^{-1}\})$, which can easily be converted into the desired bound. □

Next we consider a parabolic region $W \subset \mathcal{M}_{[t_0,t_1]}$ inside a *time-slab* of a Ricci flow spacetime that is geometrically close to an evolving Bryant soliton. Moreover, we consider two domains X_0 and X_1 that are contained in the initial and final time-slices $\mathcal{M}_{t_0}$ and $\mathcal{M}_{t_1}$ of this time-slab, respectively, and whose boundary components are central 2-spheres of sufficiently precise necks. The Bryant slab lemma describes the complements of these domains in W and characterizes their relative position.

LEMMA 8.17. (Bryant slab lemma) *If*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad 0 < \lambda < 1, \quad \Lambda \geqslant \underline{\Lambda} \quad \textit{and} \quad \delta \leqslant \bar{\delta}(\lambda, \Lambda),$$

then the following holds.

Consider a Ricci flow spacetime $\mathcal{M}$ and let $r>0$ and $t_0 \geqslant 0$. Set $t_1 := t_0 + r^2$. For $i = 0, 1$ let $X_i \subset \mathcal{M}_{t_i}$ be a closed subset that is a domain with boundary, satisfying conditions (i)–(iv) *from Lemma 8.16 and, in addition,*

(v) *$X_1(t)$ is defined for all $t \in [t_0, t_1]$, and $\partial X_1(t_0) \subset \operatorname{Int} X_0$.*

Consider a "δ-good Bryant slab" in $\mathcal{M}_{[t_0,t_1]}$, i.e. the image W of a map

$$\psi \colon W^* = \overline{M_{\mathrm{Bry}}(d)} \times [-(10\lambda)^{-2}, 0] \longrightarrow \mathcal{M}_{[t_0,t_1]},$$

where $d \geqslant \delta^{-1}$ and ψ is a $(10\lambda r)^2$-time equivariant and ∂_t-preserving diffeomorphism onto its image and

$$\|(10\lambda r)^{-2} \psi^* g - g_{\mathrm{Bry}}\|_{C^{[\delta^{-1}]}(W^*)} < \delta.$$

Set $\mathcal{C}_i := W_{t_i} \setminus \operatorname{Int} X_i \subset \mathcal{M}_{t_i}$ for $i = 0, 1$. Then, the following holds:

(a) *$\mathcal{C}_i(t)$ is well defined and $9\lambda r$-thick.*

(b) *If $\mathcal{C}_1 \neq \varnothing$, then $\mathcal{C}_0 \subset \mathcal{C}_1(t_0)$ and $\mathcal{C}_0 = \mathcal{C}_1(t_0) \setminus \operatorname{Int} X_0$.*

Proof. Assuming

$$0 < \lambda < 1, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}} \quad \text{and} \quad \delta \leqslant \bar{\delta}(\lambda, \Lambda),$$

Lemma 8.16 may be applied in the t_i-time-slice for $i = 0, 1$ and W is $9\lambda r$-thick. Since, by definition, $\mathcal{C}_i \subset W_{t_i}$, assertion (a) now follows from the fact that W is a product domain.

We now verify assertion (b). If $\mathcal{C}_1 \neq \varnothing$, then $\psi(x_{\mathrm{Bry}}, 0) \in \mathcal{C}_1$, by Lemma 8.16. So, since W is a product domain, we get that $\psi(x_{\mathrm{Bry}}, -(10\lambda)^{-2}) \in \mathcal{C}_1(t_0)$. If $\mathcal{C}_0 \neq \varnothing$, then both $\mathcal{C}_1(t_0)$ and $\mathcal{C}_0$ are 3-disks in W_{t_0} containing $\psi(x_{\mathrm{Bry}}, -(10\lambda)^{-2})$. By assumption (v) we have $\partial \mathcal{C}_1(t_0) \subset \partial X_1(t_0) \subset \operatorname{Int} X_0$ and hence $\partial \mathcal{C}_1(t_0)$ is disjoint from $\mathcal{C}_0 \subset \mathcal{M}_{t_0} \setminus \operatorname{Int} X_0$. Therefore, $\mathcal{C}_0 \subset \mathcal{C}_1(t_0)$. This gives $\mathcal{C}_0 = (W_{t_0} \setminus \operatorname{Int} X_0) \cap \mathcal{C}_1(t_0) = \mathcal{C}_1(t_0) \setminus \operatorname{Int} X_0$. □

We now show that a parabolic region $P(x,a,-b)$ lies in a comparison domain

$$(\mathcal{N},\{\mathcal{N}^j\}_{j=1}^J,\{t_j\}_{j=0}^J),$$

provided the ball $B(x,a,b)$ lies in $\mathcal{N}$ and $P(x,a,-b)$ "avoids the cuts"; see below for further discussion.

LEMMA 8.18. (Parabolic neighborhoods inside the comparison domain) *Consider a Ricci flow spacetime $\mathcal{M}$, a comparison domain $(\mathcal{N},\{\mathcal{N}^j\}_{j=1}^J,\{t_j\}_{j=0}^J)$ in $\mathcal{M}$ and a set $\mathrm{Cut}=\mathrm{Cut}^1\cup\ldots\cup\mathrm{Cut}^{J-1}$, where Cut^j is a collection of pairwise disjoint 3-disks inside $\mathcal{N}_{t_j+}$ in such a way that each extension cap of $\mathcal{N}$ is contained in some $\mathcal{D}\in\mathrm{Cut}$.*

Let $x\in\mathcal{M}$, $a,b>0$ and assume that $B(x,a)\subset\mathcal{N}$ and that $P(x,a,-b)\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}$. Then,

$$P(x,a,-b)\subset\mathcal{N}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}}\mathcal{D}. \tag{8.26}$$

As the notation suggests, the set Cut will later denote the set of cuts of a comparison, according to Definition 7.2. However, we will use Lemma 8.18 at a stage of the proof when this comparison will not have been fully constructed. More specifically, we will later consider a comparison domain defined over the time-interval $[0,t_{J+1}]$ and have to take Cut to be the union $\mathrm{Cut}\cup\mathrm{Cut}^J$. Here Cut is the set of cuts of a comparison that is only defined on the time-interval $[0,t_J]$ and Cut^J is a set of freshly constructed cuts at time t_J, which will not be part of a comparison yet. For this reason we have phrased Lemma 8.18 — and similarly Lemma 8.19 below — without using the terminology of a comparison and have instead only listed the essential properties of Cut.

Proof. Set $t:=\mathfrak{t}(x)$. Consider a point $y\in B(x,a)$ and choose j minimal with the property that $y(t)$ is defined and $y(t)\in\mathcal{N}$ for all $t\in[t_j,t]$. Assume that $t_j>0$ and $t_j>t-b$. Then, $y(t_j)\in\mathcal{N}_{t_j+}\setminus\mathcal{N}_{t_j-}$. So, $y(t_j)$ is contained in an extension cap and therefore $y(t_j)\in\mathcal{D}$ for some $\mathcal{D}\in\mathrm{Cut}$, in contradiction to our assumption. So, $t_j=0$ or $t_j\leqslant t-b$. It follows that $P(x,a,-b)\subset\mathcal{N}$. Combining this with the assumption of the lemma yields (8.26). □

The following result shows that any point near the neck-like boundary of a comparison domain is far from cuts, in the sense that there is a large backward parabolic region that is disjoint from the cuts. This result plays an important role in §12, where it allows us to isolate behavior occurring at the cuts from behavior that occurs near the neck-like boundary.

LEMMA 8.19. (Boundaries and cuts are far apart) *If*

$$\eta_{\mathrm{lin}}>0,\quad \delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda},\quad D_{\mathrm{cut}}>0,\quad A_0>0,\quad \Lambda\geqslant\underline{\Lambda},$$
$$\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda,D_{\mathrm{cut}},A_0,\Lambda),\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,D_{\mathrm{cut}},A_0,\Lambda),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

then the following holds.

Suppose that $0<T<\infty$*, and consider Ricci flow spacetimes* $\mathcal{M}$ *and* $\mathcal{M}'$ *that are* $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, T)$*-complete and that satisfy the* $\varepsilon_{\mathrm{can}}$*-canonical neighborhood assumption at scales* $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, 1)$*. Let* $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ *be a comparison domain on the time-interval* $[0, t_{J+1}]$*, and* $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ *be a comparison from* $\mathcal{M}$ *to* $\mathcal{M}'$ *defined on this comparison domain over the interval* $[0, t_J]$*. Assume that* $t_{J+1} \leqslant T$ *and that this comparison domain and comparison satisfy a-priori assumptions* (APA1)–(APA6) *for the tuple of parameters*

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}}).$$

Let Cut^J *be a set of pairwise disjoint 3-disks in* $\operatorname{Int} \mathcal{N}_{t_J+}$ *such that each* $\mathcal{D} \in \mathrm{Cut}^J$ *contains an extension cap of the comparison domain. Assume that the diameter of each* $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$ *is less than* $D_{\mathrm{cut}} r_{\mathrm{comp}}$*.*

Suppose that $x \in \mathcal{N}_t = \mathcal{N}_{t-} \cup \mathcal{N}_{t+}$ *and that*

$$P(x, A_0 \rho_1(x)) \cap \mathcal{D} \neq \varnothing$$

for some $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$*, where* $\mathcal{D} \subset \mathcal{M}_{t_k}$*.*

Then, $B(x, A_0 \rho_1(x)) \subset \mathcal{N}_{t+} \cap \mathcal{N}_{t-}$ *if* $t > t_k$*, and* $B(x, A_0 \rho_1(x)) \subset \mathcal{N}_{t+}$ *if* $t = t_k$*.*

As in Lemma 8.18 we have introduced a set Cut^J of "synthetic" cuts at time t_J in order to avoid complications due to the possible lack of a map ϕ^{J+1} that extends the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ past time t_J.

The sketch of the proof is as follows. The cut $\mathcal{D}$ contains an extension cap, which by a-priori assumption (APA5) and Lemma 8.13 implies that a large future parabolic region is Bryant-like. Then, the Bryant slice and slab lemmas, applied inductively on time steps, imply that the comparison domain contains this Bryant-like region for many time steps, which excludes neck-like boundary in the vicinity.

Proof. Pick $y \in P(x, A_0 \rho_1(x)) \cap \mathcal{D}$. By Lemma 8.6 and a-priori assumption (APA2), if

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, A_0),$$

then

$$C_1^{-1} \rho_1(x) \leqslant \rho_1(y) \leqslant C_1 \rho_1(x), \tag{8.27}$$

for some $C_1 = C_1(A_0) < \infty$.

By Definition 7.2 (3) and our assumptions regarding Cut^J, we know that $\mathcal{D}$ contains an extension cap $\mathcal{C}$. A-priori assumption (APA5) implies that there is a point $z \in \mathcal{C}$ such that $(\mathcal{M}_{t_j}, z)$ is δ_{b}-close to a Bryant soliton at scale $10\lambda r_{\mathrm{comp}}$.

Let $\delta_{\#}>0$ be a constant that will be chosen at the end of the proof.

Choose $l\in\{2,...,J+1\}$ such that $t\in[t_{l-1},t_l)$ or $[t_{l-1},t_l]$ if $l=J+1$. Since $\mathcal{D}\subset\mathcal{M}_{t_k}$, we have $k\leqslant l-1$. By a-priori assumption (APA3) (e) and (8.27), we have

$$t_l-t_k\leqslant(A_0\rho_1(x))^2+r_{\mathrm{comp}}^2\leqslant((A_0C_1\Lambda)^2+1)r_{\mathrm{comp}}^2. \tag{8.28}$$

Assuming

$$\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda,A_0,\Lambda,\delta_{\#}),\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,A_0,\Lambda,\delta_{\#})\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

we can use (8.28), a-priori assumption (APA2) and Lemma 8.13 to find a $(10\lambda r_{\mathrm{comp}})^2$-time equivariant and ∂_t-preserving diffeomorphism

$$\psi\colon W^*:=\overline{M_{\mathrm{Bry}}(\delta_{\#}^{-1})}\times[0,(t_l-t_k)\cdot(10\lambda r_{\mathrm{comp}})^{-2}]\longrightarrow\mathcal{M}$$

onto its image, such that $\psi(x_{\mathrm{Bry}},0)=z$ and

$$\|(10\lambda r_{\mathrm{comp}})^{-2}\psi^*g-g_{\mathrm{Bry}}\|_{C^{[\delta_{\#}^{-1}]}(W^*)}<\delta_{\#}.$$

Let $W:=\psi(W^*)$.

In the following we will apply the Bryant slice lemma (Lemma 8.16) at time t_j for $X=\mathcal{N}_{t_j-}$, using the time-slice W_{t_j}, where $k\leqslant j\leqslant l$. We will also apply the Bryant slab lemma (Lemma 8.17) for $X_0=\mathcal{N}_{t_{j-1}-}$ and $X_1=\mathcal{N}_{t_j-}$, using the time slab $W_{[t_{j-1},t_j]}$, where $k+1\leqslant j\leqslant l$. Note that assumptions (i)–(iv) of the Bryant slice lemma hold due to a-priori assumptions (APA3) (a)–(c) and assumption (v) of the Bryant slab lemma holds due to Definition 7.1 (3). If

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad 0<\lambda<1,\quad \Lambda\geqslant\underline{\Lambda}\quad\text{and}\quad \delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,A_0,\Lambda),$$

then the remaining assumptions of both the Bryant slice and the Bryant slab lemma are satisfied. This means, in particular, that the time-slice W_{t_j} and the slab $W_{[t_{j-1},t_j]}$ satisfy the assumptions of the Bryant slice/slab lemma for all $k\leqslant j\leqslant l$ and all $k+1\leqslant j\leqslant l$, respectively.

Claim. (a) *$W_{t_j}\subset\mathcal{N}_{t_j-}$ for all $k+1\leqslant j\leqslant l$.*
(b) *$W_t\subset\mathcal{N}_{t+}\cap\mathcal{N}_{t-}$ if $t>t_k$, and $W_t\subset\mathcal{N}_{t+}$ if $t=t_k$.*

Proof. Let $\mathcal{C}_j:=W_{t_j}\setminus\operatorname{Int}\mathcal{N}_{t_j-}$ for $k\leqslant j\leqslant l$. By assertion (a) of the Bryant slice lemma, we know that $\mathcal{C}_j$ is either empty or is a 3-disk in $\operatorname{Int}W_{t_j}$ for all $k\leqslant j\leqslant l$. Furthermore, assertion (b) of the Bryant slice lemma implies that $\mathcal{C}_k=\mathcal{C}$.

We will now show by induction that $\mathcal{C}_j=\varnothing$ for all $k+1\leqslant j\leqslant l$. This will imply assertion (a).

To see this, observe first that if $\mathcal{C}_{k+1}\neq\varnothing$, then by the Bryant slab lemma we have $\mathcal{C}=\mathcal{C}_k\subset\mathcal{C}_{k+1}(t_k)$. However, since $\mathcal{C}$ is an extension cap, we have $\mathcal{C}\subset\mathcal{N}_{t_{k^+}}$, in contradiction to the fact that $\operatorname{Int}\mathcal{C}_{k+1}(t_k)\subset\mathcal{M}_{t_k}\setminus\mathcal{N}_{t_{k^+}}$.

Next, assume that $k+2\leqslant j\leqslant l$ and that $\mathcal{C}_{j-1}=\varnothing$, but $\mathcal{C}_j\neq\varnothing$. Then, by the Bryant slab lemma, $\mathcal{C}_j(t)$ is defined and $9\lambda r_{\mathrm{comp}}$-thick for all $t\in[t_{j-1},t_j]$. Since $\mathcal{C}_{j-1}=\varnothing$, W is a product domain and $\mathcal{C}_j\subset\operatorname{Int}W_{t_j}$, we have $\mathcal{C}_j(t_{j-1})\subset\operatorname{Int}W_{t_{j-1}}\subset\mathcal{N}_{t_{j-1}-}$. So, $\mathcal{C}_j(t_{j-1})$ is a component of $\mathcal{N}_{t_{j-1}-}\setminus\operatorname{Int}\mathcal{N}_{t_{j-1}+}$ and $\partial\mathcal{C}_j(t_{j-1})\subset\mathcal{N}_{t_{j-1}+}$. This, however, contradicts a-priori assumption (APA4), finishing the induction.

To see assertion (b), observe that by assertion (a) for $j=l\geqslant k+1$ we have

$$W_t=W_{t_l}(t)\subset\mathcal{N}_{t_l-}(t).$$

As $t<t_l$ if $l\neq J+1$ and $\mathcal{N}_{t+}=\mathcal{N}_{t-}$ if $l=J+1$, this implies that $W_t\subset\mathcal{N}_{t+}$. Assume now that $t>t_k$. If $t>t_{l-1}$, then we trivially have $W_t\subset\mathcal{N}_{t+}=\mathcal{N}_{t-}$. Lastly, if $t=t_{l-1}>t_k$, then $l-1\geqslant k+1$ and therefore assertion (a) yields that $W_{t_{l-1}}\subset\mathcal{N}_{t_{l-1}-}$. □

We will now show that $B(x,A_0\rho_1(x))\subset W_t$. In combination with assertion (b) of the claim, this completes the proof of the lemma.

By the assumption of the lemma, we have $d_{t_k}(y,z)<D_{\mathrm{cut}}r_{\mathrm{comp}}$. So, if

$$\delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,D_{\mathrm{cut}}),$$

then $y\in\mathcal{D}\subset W_{t_k}$. Recall that $\mathrm{Ric}>0$ on $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$. So, g_{Bry} is decreasing in time. Therefore, if

$$\delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,D_{\mathrm{cut}}),$$

then

$$d_t(y(t),z(t))\leqslant 2d_{t_k}(y,z)\leqslant 2D_{\mathrm{cut}}r_{\mathrm{comp}}.$$

Now, by (APA3) (e),

$$\begin{aligned}d_t(x,z(t))&\leqslant d_t(x,y(t))+d_t(y(t),z(t))\\&<A_0\rho_1(x)+2D_{\mathrm{cut}}r_{\mathrm{comp}}\\&\leqslant A_0C_1\rho_1(y)+2D_{\mathrm{cut}}r_{\mathrm{comp}}\\&<(A_0C_1\Lambda+2D_{\mathrm{cut}})r_{\mathrm{comp}}.\end{aligned}$$

Therefore, assuming

$$\delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,D_{\mathrm{cut}},A_0,\Lambda),$$

we have $B(x,A_0\rho_1(x))\subset W_t$, as desired. □

The next lemma characterizes parabolic neighborhoods whose initial time-slices intersect a cut of a comparison. It states that points that belong to such an initial time-slice, but not to the corresponding cut, must have large scale if certain parameters are chosen appropriately. We also obtain that such an initial time-slice must be far from cuts that occur at earlier times. The first assertion will follow from the fact that the geometry on and near a cut is geometrically sufficiently close to a Bryant soliton and the second assertion will be a consequence of Lemma 8.19.

The results of the following lemma are specific for the proof in §12.4. As in the previous lemmas, we will use a set Cut^J of "synthetic" cuts in time-t_J-slice. Instead, we have listed the relevant properties of the cuts as assumptions of the lemma.

LEMMA 8.20. *For all $C_\#<\infty$, if*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad D_{\mathrm{cut}} \geqslant \underline{D}_{\mathrm{cut}}(\lambda, C_\#), \quad \Lambda \geqslant \underline{\Lambda},$$
$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, C_\#, D_{\mathrm{cut}}, A_0, \Lambda), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A_0, \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(C_\#),$$

then the following holds.

Suppose that $0<T<\infty$, and consider Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ that are $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, T)$-complete and that satisfy the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, 1)$. Let $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ be a comparison domain on the time-interval $[0, t_{J+1}]$, and $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ be a comparison from $\mathcal{M}$ to $\mathcal{M}'$ defined on this comparison domain over the time-interval $[0, t_J]$. Assume that $t_{J+1} \leqslant T$ and that this comparison domain and comparison satisfy a-priori assumptions (APA1)–(APA6) *for the tuple of parameters*

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}}).$$

Let Cut^J be a set of pairwise disjoint 3-disks in $\mathcal{N}_{t_J+}$ such that each $\mathcal{D} \in \mathrm{Cut}^J$ contains exactly one extension cap of the comparison domain.

Assume that the diameter of every $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$ is less than $D_{\mathrm{cut}} r_{\mathrm{comp}}$ and that the $\frac{1}{10} D_{\mathrm{cut}} r_{\mathrm{comp}}$-neighborhood of every extension cap is contained in some $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$.

Let $x \in \mathcal{N}$ and $t := \mathfrak{t}(x)$. Let $B_{t-T_0} := (B(x, A_0\rho_1(x)))(t-T_0)$ be the initial time-slice of the parabolic neighborhood $P(x, A_0\rho_1(x), -T_0)$ for some $0 \leqslant T_0 \leqslant (A_0\rho_1(x))^2$ and assume that $B_{t-T_0} \cap \mathcal{D}_0 \neq \varnothing$ for some $\mathcal{D}_0 \in \mathrm{Cut} \cup \mathrm{Cut}^J$.

Then,

$$\rho_1 \geqslant C_\# r_{\mathrm{comp}} \quad \text{on } B_{t-T_0} \setminus \mathcal{D}_0.$$

Moreover, for all $y \in B_{t-T_0}$, we have

$$P(y, A_0\rho_1(y)) \cap \mathcal{D} = \varnothing$$

for all $\mathcal{D} \in \mathrm{Cut}$ with $\mathcal{D} \subset \mathcal{M}_{[0,\mathfrak{t}(y))}$.

Proof. Let $t:=\mathfrak{t}(x)$ and choose $j\in\{1,...,J\}$ such that $t_j=t-T_0$, so $B_{t-T_0}\cup\mathcal{D}_0\subset\mathcal{M}_{t_j}$. Let $\mathcal{C}_0$ be the extension cap that is contained in $\mathcal{D}_0$.

By Lemma 8.6 and a-priori assumption (APA2), and assuming

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,A_0),$$

we find that the parabolic neighborhood $P(x,A_0\rho_1(x))$ is unscathed and that

$$C_1^{-1}\rho_1(x)\leqslant\rho_1\leqslant C_1\rho_1(x) \tag{8.29}$$

on $P(x,A_0\rho_1(x))$, where $C_1=C_1(A_0)<\infty$. By a distance distortion estimate this implies that $B_{t-T_0}\subset B(x(t_j),A_1\rho_1(x))$ for some $A_1=A_1(A_0)<\infty$.

Choose a point $z\in\partial\mathcal{C}_0\subset\mathcal{N}_{t_j-}\cap\mathcal{D}_0$. By a-priori assumption (APA3) (a) and assuming

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},$$

we have $\frac{1}{2}r_{\mathrm{comp}}\leqslant\rho_1(z)\leqslant 2r_{\mathrm{comp}}$. So, again by Lemma 8.6, and assuming

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(D_{\mathrm{cut}}),$$

we obtain that

$$C_2^{-1}r_{\mathrm{comp}}\leqslant\rho_1\leqslant C_2r_{\mathrm{comp}}\quad\text{on }\mathcal{D}_0$$

for some $C_2=C_2(D_{\mathrm{cut}})<\infty$. Combining this bound with (8.29) and the fact that

$$B_{t-T_0}\cap\mathcal{D}_0\neq\varnothing,$$

we obtain that

$$C_1^{-2}C_2^{-1}r_{\mathrm{comp}}\leqslant\rho_1\leqslant C_1^2C_2r_{\mathrm{comp}}\quad\text{on }B_{t-T_0}. \tag{8.30}$$

Therefore, for all $y\in B_{t-T_0}$,

$$d_{t_j}(y,z)\leqslant(2C_1^2C_2A_1+D_{\mathrm{cut}})r_{\mathrm{comp}}\leqslant C_3r_{\mathrm{comp}}, \tag{8.31}$$

for some $C_3=C_3(D_{\mathrm{cut}},A_0)<\infty$.

By a-priori assumption (APA5) (c), there is a diffeomorphism

$$\psi\colon M_{\mathrm{Bry}}(\delta_{\mathrm{b}}^{-1})\longrightarrow W\subset\mathcal{M}_{t_j}$$

such that $\psi(x_{Bry})\subset\mathcal{C}_0$ and

$$\|(10\lambda r_{\mathrm{comp}})^{-2}\psi^*g_{t_j}-g_{\mathrm{Bry}}\|_{C^{\lceil\delta_{\mathrm{b}}^{-1}\rceil}(M_{\mathrm{Bry}}(\delta_{\mathrm{b}}^{-1}))}<\delta_{\mathrm{b}}.$$

So, by (8.31), and the fact that $z\in\partial\mathcal{C}_0$ and that the diameter of $\mathcal{C}_0\subset\mathcal{D}_0$ is bounded by $D_{\mathrm{cut}}r_{\mathrm{comp}}$, we have

$$B_{t-T_0}\subset W, \tag{8.32}$$

assuming that

$$\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A_0).$$

Choose

$$D_{\#}=D_{\#}(\lambda, C_{\#})<\infty$$

such that $\rho>20\lambda C_{\#}$ on $M_{\mathrm{Bry}}\setminus M_{\mathrm{Bry}}(D_{\#})$ (see Lemma B.1). So, if

$$\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda, C_{\#})\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}(C_{\#}),$$

then

$$\rho_1\geqslant C_{\#}r_{\mathrm{comp}}\quad\text{on } W\setminus\psi(M_{\mathrm{Bry}}(D_{\#})). \tag{8.33}$$

If

$$D_{\mathrm{cut}}\geqslant\underline{D}_{\mathrm{cut}}(\lambda, D_{\#}(\lambda, C_{\#}))\quad\text{and}\quad\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}},$$

then $M_{\mathrm{Bry}}(D_{\#})\subset\mathcal{D}_0$. Together with (8.32) and (8.33), this implies the first assertion of this lemma.

For the second assertion note that by (8.30) and (8.31) we have

$$B(y, C_1^2C_2C_3\rho_1(y))\not\subset\mathcal{N}_{t_j-}$$

for all $y\in B_{t-T_0}$. So, the second assertion follows from Lemma 8.19, assuming

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad\lambda\leqslant\bar{\lambda},\quad\Lambda\geqslant\underline{\Lambda},\quad\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A_0, \Lambda),$$
$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A_0, \Lambda),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}.$$

This finishes the proof. □

9. Semilocal maximum principle

In this section we will show that small Ricci–DeTurck perturbations satisfy a uniform decay estimate when weighted by a suitable function of time and scale. More precisely, we show that quantities of the form

$$Q:=e^{H(T-\mathrm{t})}\rho_1^E|h|$$

satisfy a semi-local maximum principle as long as the Ricci–DeTurck perturbation h is small enough, and the Ricci flow background satisfies appropriate geometric assumptions.

The estimates of this section are based on a vanishing theorem for solutions h of the linearized Ricci–DeTurck equation on a κ-solution background, for which $|h|R^{-1-\chi}$ is uniformly bounded, where $\chi>0$ (see Theorem 9.5). The most important ingredient for the proof of this vanishing theorem is a maximum principle due to Chow and Anderson (see [AC]).

We first present the two main results of this section, Propositions 9.1 and 9.2. The first result states that a Ricci–DeTurck perturbation decays by a factor of at least 100 in the interior of a large enough neighborhood, in a weighted sense, as long as the solution is small enough. The factor 100 is chosen arbitrarily here and can be replaced by any number >1.

PROPOSITION 9.1. (Semi-local maximum principle) *If*

$$E>2, \quad H\geqslant \underline{H}(E), \quad \eta_{\text{lin}}\leqslant \bar{\eta}_{\text{lin}}(E), \quad \varepsilon_{\text{can}}\leqslant \bar{\varepsilon}_{\text{can}}(E),$$

then there are constants $L=L(E), C=C(E)<\infty$ such that the following holds.

Let $\mathcal{M}$ be a Ricci flow spacetime and pick $x\in\mathcal{M}_t$. Let us assume that $\mathcal{M}$ is $(\varepsilon_{\text{can}}\rho_1(x),t)$-complete and satisfies the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon_{\text{can}}\rho_1(x),1)$.

Then, the parabolic neighborhood $P:=P(x,L\rho_1(x))$ is unscathed and the following is true. Let h be a Ricci–DeTurck perturbation on P. Assume that $|h|\leqslant\eta_{\text{lin}}$ everywhere on P and define the scalar function

$$Q:=e^{H(T-\mathfrak{t})}\rho_1^E|h| \tag{9.1}$$

on P, where $T\geqslant t$ is some arbitrary number.

Then, in the case $t>(L\rho_1(x))^2$ (i.e. if P does not intersect the time-zero slice) we have

$$Q(x)\leqslant \tfrac{1}{100}\sup_P Q.$$

In the case $t\leqslant(L\rho_1(x))^2$ (i.e. if P intersects the time-zero slice) we have

$$Q(x)\leqslant \tfrac{1}{100}\sup_P Q+C\sup_{P\cap\mathcal{M}_0} Q.$$

Note that the parabolic neighborhood P may be defined on a time-interval of size less than $(L\rho_1(x))^2$ if P intersects the initial time-slice $\mathcal{M}_0$. By performing a time shift, Proposition 9.1 can be generalized to the case in which P is defined on a time-interval of size less than $(L\rho_1(x))^2$ that does not necessarily intersect $\mathcal{M}_0$. This fact will be used in §12 when P intersects a cut, i.e. a discontinuity locus of h, at some positive time.

We also remark that the constant T in Proposition 9.1 does not have any mathematical significance and could be eliminated from the statement. It appears in Proposition 9.1 only to conform with the notation later in the paper where it is used.

In the next result, we improve the interior estimate and replace the factor 100 by an arbitrary factor. As a trade-off, we need to choose the parabolic neighborhood on which h and Q are defined large enough; note however that we do not need to change the bound on $|h|$ appearing in the assumptions.

PROPOSITION 9.2. (Interior decay) *If*

$$E>2, \quad H\geqslant \underline{H}(E), \quad \eta_{\text{lin}}\leqslant \bar{\eta}_{\text{lin}}(E), \quad \alpha>0,$$
$$A\geqslant \underline{A}(E,\alpha), \quad \varepsilon_{\text{can}}\leqslant \bar{\varepsilon}_{\text{can}}(E,\alpha),$$

then there is a constant $C=C(E)<\infty$ such that the following holds.

Let $\mathcal{M}$ be a Ricci flow spacetime and $x\in\mathcal{M}_t$. Assume $\mathcal{M}$ is $(\varepsilon_{\text{can}}\rho_1(x),t)$-complete and satisfies the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon_{\text{can}}\rho_1(x),1)$.

Consider the parabolic neighborhood $P:=P(x,A\rho_1(x))$ and let h be a Ricci–DeTurck perturbation on P such that $|h|\leqslant\eta_{\text{lin}}$ everywhere. Define Q as in (9.1).

Then, in the case $t>(Ar)^2$ (i.e. if P does not intersect the time-zero slice) we have

$$Q(x)\leqslant \alpha \sup_P Q.$$

In the case $t\leqslant(Ar)^2$ (i.e. if P intersects the time-zero slice) we have

$$Q(x)\leqslant \alpha \sup_P Q+C\sup_{P\cap\mathcal{M}_0} Q.$$

We remark that it follows from the proof that the parabolic neighborhood $P(x,\underline{A}\,\rho_1(x))$ is unscathed, although we cannot guarantee this for $P(x,A\rho_1(x))$. Due to the way the proposition will be applied later, it is more convenient to state the conditions using the possibly larger scale A.

The proofs of Propositions 9.1 and 9.2 are based on the following strong maximum principle for solutions of the linearized Ricci–DeTurck flow. This maximum principle is a special case of a result of Anderson and Chow (cf. [AC]). The proof of Anderson and Chow's result simplifies in this special case, which is why we have decided to include it in this paper.

LEMMA 9.3. (Strong maximum principle of Anderson–Chow) *Let $(M,(g_t)_{t\in(-T,0]})$, $T>0$, be a Ricci flow on a connected 3-manifold M such that (M,g_t) has non-negative sectional curvature for all $t\in(-T,0]$.*

Consider a solution $(h_t)_{t\in[-T,0]}$ of the linearized Ricci–DeTurck equation on M, i.e.

$$\partial_t h_t = \Delta_{L,g_t} h_t \quad \Longleftrightarrow \quad \nabla_{\partial_t} h_t = \Delta_{g_t} h_t + 2\,\mathrm{Rm}_{g_t}(h_t).$$

Assume that

$$|h| \leqslant CR \quad \text{on } M\times(-T,0]$$

for some $C>0$ and that $|h|(x_0,0)=CR(x_0,0)$ for some $x_0\in M$. Then,

$$|h| = CR \quad \text{on } M\times(-T,0].$$

Proof. Using Kato's inequality it is not hard to see that, wherever $|h|\neq 0$, we have

$$\partial_t |h| \leqslant \Delta_{g_t}|h| + 2\frac{\mathrm{Rm}(h,h)}{|h|^2}\cdot|h|.$$

On the other hand, whenever $R>0$, we have

$$\partial_t(CR) = \Delta_{g_t}(CR) + 2\frac{|\mathrm{Ric}|^2}{R}\cdot CR.$$

So, the claim follows by the strong maximum principle applied to $|h|-CR$, if we can show that, for any symmetric 2-tensor h,

$$\frac{\mathrm{Rm}(h,h)}{|h|^2} \leqslant \frac{|\mathrm{Ric}|^2}{R}. \tag{9.2}$$

To see (9.2), let $h_{ij}\neq 0$ be a non-zero 3-dimensional symmetric 2-tensor and Rm_{ijkl} be a 3-dimensional algebraic curvature tensor with non-negative sectional curvature. We denote by Ric_{ij} and R its Ricci and scalar curvatures. Without loss of generality, we may assume that $|h|=1$ and that Ric_{ij} is diagonal. Then, Rm_{ijkl} is only non-zero if $\{i,j,k,l\}$ has cardinality 2. Set $a_1:=\mathrm{Rm}_{2332}$, $a_2:=\mathrm{Rm}_{1331}$, $a_3:=\mathrm{Rm}_{1221}$ and $x_i:=h_{ii}$. Then,

$$\begin{aligned}\mathrm{Rm}(h,h) &= \mathrm{Rm}_{ijkl}\,h_{il}h_{jk}\\ &= -2a_1h_{23}^2 - 2a_2h_{13}^2 - 2a_3h_{12}^2 + 2a_1h_{22}h_{33} + 2a_2h_{11}h_{33} + 2a_3h_{11}h_{22}\\ &\leqslant 2(a_1x_2x_3 + a_2x_1x_3 + a_3x_1x_2).\end{aligned}$$

On the other hand,

$$|\mathrm{Ric}|^2 = (a_2+a_3)^2 + (a_1+a_3)^2 + (a_1+a_2)^2$$

and

$$R = 2(a_1+a_2+a_3).$$

Since $x_1^2+x_2^2+x_3^2\leqslant|h|^2=1$, the next lemma implies (9.2). □

LEMMA 9.4. *If* $x_1^2+x_2^2+x_3^2\leqslant 1$ $a_1,a_2,a_3\geqslant 0$ *and* $a_1+a_2+a_3>0$, *then*

$$a_1x_2x_3+a_2x_1x_3+a_3x_1x_2\leqslant \frac{(a_2+a_3)^2+(a_1+a_3)^2+(a_1+a_2)^2}{4(a_1+a_2+a_3)}, \tag{9.3}$$

Proof. Let $\lambda_1\leqslant\lambda_2\leqslant\lambda_3$ be the eigenvalues of the symmetric matrix

$$A:=\frac{1}{2}\begin{pmatrix} 0 & a_3 & a_2\\ a_3 & 0 & a_1\\ a_2 & a_1 & 0\end{pmatrix}$$

and denote by $v_1,v_2,v_3\in\mathbb{R}^3$ the corresponding orthonormal basis of eigenvectors. The left-hand side of (9.3) is bounded from above by λ_3.

Since the trace of A vanishes and its determinant equals $\frac{1}{4}a_1a_2a_3\geqslant 0$, we must have $\lambda_1,\lambda_2\leqslant 0$ and $\lambda_3\geqslant 0$. In the case $\lambda_3=0$, we are done. So, assume from now on that $\lambda_3>0$. Consider the vector

$$u:=\begin{pmatrix}1\\1\\1\end{pmatrix}=c_1v_1+c_2v_2+c_3v_3.$$

Since

$$Au=\frac{1}{2}\begin{pmatrix}a_2+a_3\\ a_1+a_3\\ a_1+a_2\end{pmatrix}\quad\text{and}\quad u^TAu=a_1+a_2+a_3,$$

we obtain

$$\frac{c_1^2\lambda_1^2+c_2^2\lambda_2^2+c_3^2\lambda_3^2}{c_1^2\lambda_1+c_2^2\lambda_2+c_3^2\lambda_3}=\frac{(a_2+a_3)^2+(a_1+a_3)^2+(a_1+a_2)^2}{4(a_1+a_2+a_3)}.$$

Since $\lambda_1,\lambda_2<0$ and numerator and denominator of the first fraction are both positive, we obtain

$$\lambda_3=\frac{c_3^2\lambda_3^2}{c_3^2\lambda_3}\leqslant\frac{(a_2+a_3)^2+(a_1+a_3)^2+(a_1+a_2)^2}{4(a_1+a_2+a_3)}.$$

This is what we wanted to show. □

THEOREM 9.5. (Vanishing theorem) *Consider a 3-dimensional* κ*-solution*

$$(M,(g_t)_{t\in(-\infty,0]})$$

and a smooth, time-dependent tensor field $(h_t)_{t\in(-\infty,0]}$ *on* M *that satisfies the linearized Ricci–DeTurck equation*

$$\partial_t h_t=\Delta_{L,g_t}h_t.$$

Assume that there are numbers $\chi>0$ *and* $C<\infty$ *such that*

$$|h|\leqslant CR^{1+\chi}\quad\text{on } M\times(-\infty,0]. \tag{9.4}$$

Then, $h\equiv 0$ *everywhere.*

Proof. Assume that $h_0 \not\equiv 0$. Since $(M, (g_t)_{t\in(-\infty,0]})$ has uniformly bounded curvature, we have

$$|h| \leqslant C' R$$

for some $C' < \infty$. Choose a sequence $(x_k, t_k) \in (-\infty, 0] \times M$ such that

$$\lim_{k\to\infty} \frac{|h|(x_k, t_k)}{R(x_k, t_k)} = \sup_{M\times(-\infty,0]} \frac{|h|}{R}.$$

It follows from (9.4) that

$$CR^{\chi}(x_k, t_k) \geqslant \frac{|h|(x_k, t_k)}{R(x_k, t_k)}.$$

So, there is a $c > 0$ such that $R(x_k, t_k) > c$ for all k. Consider the sequence of pointed flows $(M, (g_{t+t_k})_{t\in(-\infty,0]}, x_k)$. After passing to a subsequence, this sequence converges to a pointed κ-solution $(M_\infty, (g_{\infty,t})_{t\in(-\infty,0]}, x_\infty)$. Similarly, consider the sequence of time-dependent tensor fields $h_k(\cdot, t+t_k)$. After passing to another subsequence, these tensor fields converge to a solution $(h_{\infty,t})_{t\in(-\infty,0]}$ of the linearized Ricci–DeTurck flow on $M_\infty \times (-\infty, 0]$. The bound (9.4) carries over in the limit to

$$|h_\infty| \leqslant CR^{1+\chi} \tag{9.5}$$

and by the choice of the points (x_k, t_k) we obtain the extra property that

$$\frac{|h_\infty|(x_\infty, 0)}{R(x_\infty, 0)} = \sup_{M_\infty\times(-\infty,0]} \frac{|h_\infty|}{R} = \sup_{M\times(-\infty,0]} \frac{|h|}{R} =: C' > 0.$$

We can now apply the strong maximum principle, Lemma 9.3, and obtain that

$$|h_\infty| \equiv C' R \quad \text{on} \quad M_\infty \times (-\infty, 0].$$

Combining this with (9.5) yields that, on $M_\infty \times (-\infty, 0]$,

$$C' R \leqslant CR^{1+\chi}.$$

So, R is uniformly bounded from below on $M_\infty \times (-\infty, 0]$. It follows that

$$(M, (g_{\infty,t})_{t\in(-\infty,0]})$$

cannot be the round shrinking cylinder or a quotient thereof. If M_∞ was non-compact, then we can obtain the round shrinking cylinder as a pointed limit of $(M, (g_{\infty,t})_{t\in(-\infty,0]})$, which contradicts the positive lower bound on R. If, on the other hand, M_∞ was compact, then the maximum principle applied to the evolution equation of R would imply that $\min_{M_\infty} R(\cdot, t) \to 0$ as $t \to -\infty$, again contradicting the positive lower bound on R. □

Proof of Proposition 9.1. Fix some $E>2$ for the remainder of the proof. By linearity of the desired bounds, we may assume for simplicity that $T=\mathfrak{t}(x)$.

Next, observe that, by bounded curvature at bounded distance, Lemma 8.6, for any choice of $L<\infty$ we may choose $\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(L)$ small enough such the parabolic neighborhood $P(x,L\rho_1(x))$ is unscathed and such that $\rho_1>c_0(L)\rho_1(x)$ on this parabolic neighborhood for some $c_0=c_0(L)>0$.

Assume now that the statement was false (for fixed $E>2$). Choose sequences $\eta_{\mathrm{lin},k},\varepsilon_{\mathrm{can},k}\to 0$ and $H_k,L_k,C_k\to\infty$ such that $\varepsilon_{\mathrm{can},k}$ is small enough depending on L_k, as discussed in the preceding paragraph. For each k we can choose a Ricci flow spacetime $\mathcal{M}_k$, points $x_k\in\mathcal{M}_{k,t_k}$, an (unscathed) parabolic neighborhood $P_k:=P(x_k,L_k\rho_1(x_k))$ and a Ricci–DeTurck perturbation h_k on P_k such that $|h_k|\leqslant\eta_{\mathrm{lin},k}$ on P_k, which violate the conclusion of the proposition. Thus, setting

$$Q_k(y):=e^{H_k(t_k-\mathfrak{t}(y))}\rho_1^E(y)|h_k|(y)\quad\text{for } y\in P_k,$$

either $t_k:=\mathfrak{t}(x_k)>(L_k\rho_1(x_k))^2$ and

$$Q_k(x_k)>\tfrac{1}{100}\sup_{P_k}Q_k \tag{9.6}$$

or $t_k=\mathfrak{t}(x_k)\leqslant(L_k\rho_1(x_k))^2$ and

$$Q_k(x_k)>\tfrac{1}{100}\sup_{P_k}Q_k+C_k\sup_{P_k\cap\mathcal{M}_{k,0}}Q_k. \tag{9.7}$$

Let us rephrase the bounds (9.6) and (9.7) in a more convenient form. To do this, let $\alpha_k:=|h_k(x_k)|\leqslant\eta_{\mathrm{lin},k}\to 0$ and consider the tensor field $h_k':=\alpha_k^{-1}h_k$. Then, h_k' is a solution to the rescaled Ricci–DeTurck equation (A.11) for $\alpha=\alpha_k$:

$$|h_k'|(x_k)=1 \tag{9.8}$$

and, on P_k,

$$|h_k'|=\frac{|h_k|}{|h_k|(x_k)}=e^{-H_k(t_k-\mathfrak{t})}\left(\frac{\rho_1}{\rho_1(x_k)}\right)^{-E}\cdot\frac{Q_k}{Q_k(x_k)}.$$

So, by (9.6) and (9.7), we have

$$|h_k'|\leqslant 100e^{-H_k(t_k-\mathfrak{t})}\left(\frac{\rho_1}{\rho_1(x_k)}\right)^{-E}\quad\text{on } P_k \tag{9.9}$$

and, if $P_k\cap\mathcal{M}_{k,0}\neq\varnothing$, then

$$|h_k'|\leqslant C_k^{-1}e^{-H_k(t_k-\mathfrak{t})}\left(\frac{\rho_1}{\rho_1(x_k)}\right)^{-E}\quad\text{on } P_k\cap\mathcal{M}_{k,0}. \tag{9.10}$$

We now distinguish two cases.

Case 1. $t_k \geqslant c\rho_1^2(x_k)$ *for all* k *and some* $c>0$.

The metric g_k restricted to P_k can be expressed in terms of a classical Ricci flow $(g_{k,t})_{t\in[t_k-\Delta t_k,t_k]}$ on $B_k:=B(x_k, L_k\rho_1(x_k))$, where

$$\Delta t_k := \min\{t_k, (L_k\rho_1(x_k))^2\}.$$

Let $r_k:=\rho_1(x_k)$ and

$$T_\infty := \limsup_{k\to\infty} r_k^{-2}\Delta t_k \geqslant c > 0.$$

Consider the parabolically rescaled flows

$$(g'_{k,t} := r_k^{-2} g_{k,r_k^2 t+t_k})_{t\in[-r_k^{-2}\Delta t_k,0]}.$$

By bounded curvature at bounded distance, Lemma 8.6, and since $\varepsilon_{\mathrm{can},k}\to 0$, for any $s<\infty$ and $T'<T_\infty$, for sufficiently large k we find uniform bounds on the curvature of $g'_{k,0}$ on the $g'_{k,0}$-ball $B(x_k,0,s)$ over the time-interval $[-T',0]$.

Case 1*a*. *We have* $\liminf_{k\to\infty}\rho_1(x_k)>0$, *and the injectivity radius satisfies*

$$\liminf_{k\to\infty} \mathrm{InjRad}(g'_{k,0}, x_k) > 0.$$

After passing to a subsequence, we may extract a smooth limiting pointed flow $(M_\infty, (g_{\infty,t})_{t\in(-T_\infty,0]}, x_\infty)$. Due to (9.9) and the local gradient estimates from Lemma A.2, the reparameterized tensor fields $(r_k^{-2}h'_{k,r_k^2t+t_k})_{t\in[-r_k^{-2}\Delta t_k,0]}$ converge, after passing to another subsequence, to a smooth solution $(h'_{\infty,t})_{t\in(-T_\infty,0]}$ on M_∞ of the linearized Ricci–DeTurck equation with background metric $(g_{\infty,t})_{t\in(-T_\infty,0]}$ (see (A.12)), such that

$$|h'_\infty|(x_\infty,0) = 1. \tag{9.11}$$

Since $\lim_{k\to\infty} H_k\rho_1^2(x_k)=\infty$, we can use the exponential factor in (9.9) to show that $h'_\infty\equiv 0$ on $M_\infty\times(-T_\infty,0)$, which implies $h'_\infty(x_\infty,0)=0$. This contradicts (9.11).

Case 1*a*′. *We have* $\liminf_{k\to\infty}\rho_1(x_k)>0$, *and the injectivity radius satisfies*

$$\liminf_{k\to\infty} \mathrm{InjRad}(g'_k, x_k) = 0.$$

For some $\hat r>0$, we may pull back g'_k to the $\hat r$-ball in the tangent space at x_k via the exponential map to reduce to Case 1a. Note that in Case 1a it was not important that the time-slices of the limiting flow $(M_\infty, (g_{\infty,t})_{t\in(-T_\infty,0]}, x_\infty)$ were complete.

Case 1b. $\liminf_{k\to\infty}\rho_1(x_k)=0$, *and the injectivity radius satisfies*

$$\liminf_{k\to\infty}\operatorname{InjRad}(g'_k(0),x_k)>0.$$

As explained in the beginning of Case 1a, by passing to a subsequence, we may assume that the pointed flows $(B_k,(g'_{k,t})_{k,t},x_k)$ converge to a smooth pointed flow $(M_\infty,(g_{\infty,t})_{t\in(-T_\infty,0]},x_\infty)$ and, moreover, the tensor fields $(r_k^{-2}h'_{k,r_k^2t+t_k})_{t\in[-r_k^{-2}\Delta t_k,0]}$ converge to a smooth solution $(h'_{\infty,t})_{t\in(-T_\infty,0]}$ on M_∞ of the linearized Ricci–DeTurck equation with background metric $(g_{\infty,t})_{t\in(-T_\infty,0]}$ (see (A.12)), such that (9.11) holds.

Using Lemma 8.6 and the canonical neighborhood assumption, it follows that $R>0$ everywhere on $M_\infty\times(-T_\infty,0]$. By assertion (a) of Lemma C.1, there is a $\kappa_0>0$ such that every κ-solution is either a shrinking round spherical space form or is a κ_0-solution. Therefore, in view of the injectivity radius bound, there is a $\kappa_1>0$ such that, by the canonical neighborhood assumption, every time-slice $(M_\infty,g_{\infty,t})$, $t\in(-T_\infty,0]$, is isometric to the final time-slice of a κ_1-solution. Since, by assertion (e) of Lemma C.1, we have $\partial_tR\geqslant0$ on κ-solutions, we get that $(M_\infty,(g_{\infty,t})_{t\in(-T_\infty,0]},x_\infty)$ has bounded curvature, so it is a κ-solution if $T_\infty=\infty$.

Passing (9.9) to the limit yields

$$|h'_\infty|\leqslant100\rho^{-E}\leqslant(C')^{E/2}R^{E/2}\quad\text{on } M_\infty\times(-T_\infty,0],$$

for some universal constant $C'<\infty$.

If $T_\infty=\infty$, then the vanishing theorem (Theorem 9.5) yields that $h'_\infty\equiv0$, in contradiction to (9.11).

Now suppose that $T_\infty<\infty$. We will show that, for some constant $C''<\infty$, we have

$$|h'_\infty(x,t)|\leqslant C''(t+T_\infty).\tag{9.12}$$

for all $x\in M_\infty$, $t\in(-T_\infty,0]$.

As $(M_\infty,g_{\infty,0})$ is isometric to the final time-slice of a κ_1-solution, and therefore has uniformly bounded curvature, we can find a constant $a_1>0$ such that, for any L', we have

$$\rho>a_1\rho_1(x_k)\quad\text{on } B(x_k,L'\rho_1(x_k)),$$

as long as k is chosen large enough. So, by bounded curvature at bounded distance, Lemma 8.6, there is a constant $a_2>0$ such that for any $L'<\infty$ we have

$$\rho>a_2\rho_1(x_k)\quad\text{on}\quad P(x_k,L'\rho_1(x_k),-t_k)$$

for large k. By (9.9), (9.10) and Proposition A.12, we find a sequence $c_k \to 0$ and a constant $C'' < \infty$ such that, for any $L' < \infty$, we have

$$|h'_k| < C'' \rho_1^{-2}(x_k) \cdot \mathfrak{t} + c_k \quad \text{on } P(x_k, L'\rho_1(x_k), -t_k) \tag{9.13}$$

for large k. Passing this bound to the limit implies (9.12).

Since $\sup |h'_\infty| < \infty$, this forces $h'_\infty \equiv 0$, again contradicting (9.11).

Case 1b'. $\liminf_{k\to\infty} \rho_1(x_k) = 0$, *and the injectivity radius satisfies*

$$\liminf_{k\to\infty} \operatorname{InjRad}(g'_k(0), x_k) = 0.$$

After passing to a subsequence, we may assume that $\operatorname{InjRad}(g'_k(0), x_k) \to 0$ as $k \to \infty$. By Lemma C.1, the universal covers of the flows $(M_k, g'_{k,t})$ converge to shrinking round spheres on the time-interval $(-\infty, 0]$. We may now pull back the tensor fields h_k to the universal covers and reduce to Case 1b.

Case 2. $\liminf_{k\to\infty} \rho_1^{-2}(x_k) t_k = 0$.

In this case, by combining the curvature bounds from Lemma 8.6 with (9.9) and (9.10), we can apply Proposition A.12 to show that there is a sequence $c_k \to 0$ and constants $C'', L' < \infty$ such that (9.13) holds for large k. It follows that

$$\lim_{k\to\infty} |h'_k|(x_k) = 0,$$

in contradiction to (9.8). □

Proof of Proposition 9.2. The bound follows by iterating the bound from Proposition 9.1.

Assume that

$$E > 2, \quad H \geqslant \underline{H}(E), \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}(E), \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(E),$$

and set $C = 2C(E)$ and $L = L(E)$ according to Proposition 9.1. So, Proposition 9.2 holds if $\alpha \geqslant \frac{1}{100}$. Assume now by induction that $\alpha_0 < \frac{1}{100}$ and that Proposition 9.2 holds for $\alpha = 100\alpha_0$ under an assumption of the form

$$A \geqslant A' := \underline{A}(E, 100\alpha_0) \quad \text{and} \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(E, 100\alpha_0).$$

Consider the point $x \in \mathcal{M}$. By Lemma 8.7 we can find a constant

$$A'' = A''(L(E), A'(E, 100\alpha_0)) < \infty$$

such that, if

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L(E), A'(E, 100\alpha_0)),$$

then the parabolic neighborhood $P(x, A'\rho_1(x))$ is unscathed and we have

$$P(y, A'\rho_1(y)) \subset P(x, A''\rho_1(x)) \quad \text{for all } y \in P(x, L\rho_1(x)).$$

Also, by bounded curvature at bounded distance, Lemma 8.6, assuming $\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L(E))$, we know that $\rho_1 \geqslant c\rho_1(x)$ on $P(x, L\rho_1(x))$ for some $c = c(L(E)) > 0$.

Assume now that $A \geqslant A''$ and apply Proposition 9.2 at each $y \in P(x, L\rho_1(x))$ for $\alpha = 100\alpha_0$. Note that in order to do this, we need to ensure that $\mathcal{M}$ is $(\varepsilon_{\mathrm{can}}\rho_1(y), \mathfrak{t}(y))$-complete and satisfies the canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}}\rho_1(y), 1)$. This can always be guaranteed if we assume that $\varepsilon_{\mathrm{can}} \leqslant c(L(E))\varepsilon_{\mathrm{can}}(E, 100\alpha_0)$. Proposition 9.2 for $\alpha = 100\alpha_0$ gives us

$$\sup_{P(x, L\rho_1(x))} Q \leqslant 100\alpha_0 \sup_{P(x, A''\rho_1(x))} Q + C \sup_{P(x, A''\rho_1(x)) \cap \mathcal{M}_0} Q.$$

Applying Proposition 9.1 then implies (recall that we have replaced C by $2C$)

$$\begin{aligned} Q(x) &\leqslant \frac{100\alpha_0}{100} \sup_{P(x, L\rho_1(x))} Q + \left(\frac{C}{100} + \frac{1}{2}C\right) \sup_{P(x, A''\rho_1(x)) \cap \mathcal{M}_0} Q \\ &\leqslant \alpha_0 \sup_{P(x, A''\rho_1(x))} Q + C \sup_{P(x, A''\rho_1(x)) \cap \mathcal{M}_0} Q. \end{aligned}$$

This finishes the induction. □

10. Extending maps between Bryant solitons

In this section we consider two regions that are close to Bryant solitons, at possibly different scales, and an almost isometry between annular subdomains inside these regions. We will then prove that the scales of both Bryant soliton regions are almost equal and that the given almost isometry can be extended to an almost isometry, of possibly lesser accuracy, over the entire Bryant soliton regions. An important aspect of the main result of this section is that the accuracy that is required from the given almost isometry depends only polynomially on the local scale — or on the distance from the tip.

Our main result, the Bryant extension proposition (Proposition 10.1), will be needed in the proof of Proposition 12.2 in §12. In this proposition, we extend an almost isometry between two Ricci flow spacetime time-slices over an extension cap. By assumption, the accuracy of this almost isometry improves at a large polynomial rate as we move away from the extension cap. As long as this polynomial rate is sufficiently large, we can use

Proposition 10.1 to construct an extension of the almost isometry over the extension cap whose accuracy still improves at a large polynomial rate. This enables us to retain the fine geometric bounds needed to prolong our comparison.

In this section, we will use the notation $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ for the pointed Bryant soliton with $\rho(x_{\mathrm{Bry}})=1$; for this and other notation related to the geometry of the Bryant soliton, we refer to §6.2. We will also frequently use the curvature scale function

$$\rho\colon M_{\mathrm{Bry}} \longrightarrow (0,\infty)$$

as introduced in Definition 6.1. Recall that $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}})$ is an $O(3)$-invariant gradient steady soliton diffeomorphic to $\mathbb{R}^3$ and $\rho(x)\to\infty$ as $x\to\infty$.

We first present a version of the Bryant comparison result in a form that is most useful for its application in the proof of Proposition 12.2.

Proposition 10.1. (Bryant extension) *If*

$$E \geqslant \underline{E}, \quad C>0, \quad \beta>0, \quad D \geqslant \underline{D}(E,C,\beta),$$
$$0<b\leqslant C, \quad 0<\delta\leqslant \bar{\delta}(E,C,\beta,D,b),$$

then the following holds for any $D'>0$.

Let g and g' be Riemannian metrics on $M_{\mathrm{Bry}}(D)$ and $M_{\mathrm{Bry}}(D')$, respectively, such that for some $\lambda\in[C^{-1},C]$,

$$\|g-g_{\mathrm{Bry}}\|_{C^{[\delta^{-1}]}(M_{\mathrm{Bry}}(D))} \quad \text{and} \quad \|\lambda^{-2}g'-g_{\mathrm{Bry}}\|_{C^{[\delta^{-1}]}(M_{\mathrm{Bry}}(D'))}<\delta. \tag{10.1}$$

Consider a diffeomorphism onto its image

$$\phi\colon M_{\mathrm{Bry}}\big(\tfrac{1}{2}D, D\big) \longrightarrow M_{\mathrm{Bry}}(D')$$

such that for $h:=\phi^*g'-g$ we have, for all $m=0,...,4$,

$$\rho_g^E|\nabla_g^m h|_g \leqslant b \quad \text{on } M_{\mathrm{Bry}}\big(\tfrac{1}{2}D, D\big),$$

where ρ_g denotes the scale function with respect to the metric g. Then, there is a diffeomorphism onto its image $\tilde{\phi}\colon M_{\mathrm{Bry}}(D)\to M_{\mathrm{Bry}}(D')$ such that the following holds:

(a) $\tilde{\phi}=\phi$ *on* $M_{\mathrm{Bry}}(D-1,D)$.

(b) *For* $\tilde{h}:=\tilde{\phi}^*g'-g$ *we have*

$$\rho_g^3|\tilde{h}|_g \leqslant \beta b \quad \text{on } M_{\mathrm{Bry}}(D).$$

We remark that there are several ways in which one could strengthen or sharpen this proposition. We chose the statement above, because it is adequate for our purposes and keeps the complications in the proof to a minimum. For example, the constant $\underline{E}$ in this proposition could be taken to be equal to 100, or even smaller. Also, the choice of the exponent 3 in assertion (b) is arbitrary. This exponent is needed in the proof of Proposition 12.2, but it could be replaced by any other number, assuming that E is chosen sufficiently large.

The Bryant extension Proposition 10.1 is a consequence of the following simpler result, on which we will focus for the larger part of this section. A proof that Proposition 10.2 implies Proposition 10.1 is provided at the end of this section.

PROPOSITION 10.2. (Bryant extension, simple form) *There is a constant $C<\infty$ such that, if*

$$0<\alpha<1, \quad E\geqslant \underline{E} \quad \text{and} \quad D\geqslant \underline{D}(\alpha),$$

then the following holds. Assume that

(i) $g_1=g_{\mathrm{Bry}}$ *and* $g_2=\lambda_2^2 g_{\mathrm{Bry}}$ *is a rescaled Bryant soliton metric.*

(ii) $\lambda_2\in[\alpha,\alpha^{-1}]$.

(iii) $\phi\colon M_{\mathrm{Bry}}\big(\frac{1}{2}D,D\big)\to M_{\mathrm{Bry}}$ *is a diffeomorphism onto its image.*

(iv) *For* $h=\phi^*g_2-g_1$ *and for some* $b\leqslant\alpha^{-1}$ *we have, for all* $m=0,...,4$,

$$|\nabla_{g_1}^m h|_{g_1}\leqslant bD^{-E} \quad \text{on } M_{\mathrm{Bry}}\big(\tfrac{1}{2}D,D\big).$$

Then, there a diffeomorphism onto its image $\tilde{\phi}\colon M_{\mathrm{Bry}}(D)\to M_{\mathrm{Bry}}$ *such that the following holds*:

(a) $\tilde{\phi}=\phi$ *on* $M_{\mathrm{Bry}}(D-1,D)$.

(b) *For* $\tilde{h}:=\tilde{\phi}^*g_2-g_1$ *we have*

$$|h|_{g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C} \quad \text{on } M_{\mathrm{Bry}}(D).$$

The strategy of the proof is as follows. We first show that ϕ almost preserves the curvature operator and its first covariant derivative, up to an error that decays polynomially in D. As the scale of a Bryant soliton can be expressed in terms of the curvature and its derivative, this will imply that the scale λ_2 of g_2 is close to the scale 1 of g_1, up to an error that decays polynomially in D. Similarly, we can argue that ϕ preserves the distance function to the tip x_{Bry} up to a polynomially decaying error. Using this extra information, we can in turn argue that ϕ is sufficiently close to an isometric rotation of $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$ around the tip x_{Bry}, again up to an error that decays polynomially in D. By an interpolation argument, we eventually extend ϕ to a map on $M_{\mathrm{Bry}}(D)$ that is equal to this isometric rotation sufficiently far away from the boundary.

The proof will use some standard properties geometric properties of the Bryant soliton, which are reviewed in Appendix B. Recall that $x_{\mathrm{Bry}} \in M_{\mathrm{Bry}}$ denotes the tip, i.e. the center of rotational symmetry, of M_{Bry}. In the following we furthermore denote by $\sigma := d_{g_{\mathrm{Bry}}}(\cdot, x_{\mathrm{Bry}})$ the distance function from the tip.

The remainder of this section will be devoted to the proof of Proposition 10.2. Until the end of the section, we will let $g_1 = g_{\mathrm{Bry}}$, g_2, λ_2, etc., be as in the statement of this proposition. Let $g_3 = \phi^* g_2$. We begin with some estimates on the difference between geometric quantities for g_1 and g_3.

We will use the convention that $1 < C < \infty$ denotes a generic universal constant, which may change from line to line.

LEMMA 10.3. *If*

$$E \geqslant \underline{E} \quad \text{and} \quad D \geqslant \underline{D}(\alpha),$$

then the following holds.

Let $\mathcal{D} = \nabla_{g_3} - \nabla_{g_1}$ be the difference tensor for the Levi-Civita connections of g_3 and g_1, respectively. Then, we have

$$|T|_{g_1} \leqslant b \cdot C D^{-E} \quad \text{on } M_{\mathrm{Bry}}\left(\tfrac{1}{2}D, D\right),$$

where T is any tensor field from the following list:

$$\{\nabla_{g_1}^k (g_3 - g_1)\}_{0 \leqslant k \leqslant 4}, \quad \{\nabla_{g_1}^k \mathcal{D}\}_{0 \leqslant k \leqslant 3}$$
$$\{\nabla_{g_1}^k (\mathrm{Rm}_{g_3} - \mathrm{Rm}_{g_1}), \nabla_{g_1}^k (\mathrm{Ric}_{g_3} - \mathrm{Ric}_{g_1}), \nabla_{g_1}^k (R_{g_3} - R_{g_1})\}_{0 \leqslant k \leqslant 2}.$$

The bound also holds if we view Ric_{g_i}, $i = 1, 3$, *as a* $(1,1)$*-tensor.*

Proof. Consider a point $x \in M_{\mathrm{Bry}}\left(\frac{1}{2}D, D\right)$ and identify $T_x M_{\mathrm{Bry}}$ with $\mathbb{R}^3$ such that $g_{1,x}$ corresponds to the Euclidean inner product. The tensors $h_x, \nabla_{g_1} h_x, ..., \nabla_{g_1}^4 h_x$ and $\mathrm{Rm}_{g_1,x}, ..., \nabla_{g_1}^2 \mathrm{Rm}_{g_1,x}$ and T_x can be viewed as tensors on $\mathbb{R}^3$. As T can be written in the form of an algebraic expression involving the tensors $g_1, (g_1 + h)^{-1}, h, ..., \nabla_{g_1}^4 h$ and $\mathrm{Rm}_{g_1}, ..., \nabla_{g_1}^2 \mathrm{Rm}_{g_1}$, there is a smooth tensor-valued function F such that

$$T_x = F(h_x, ..., \nabla_{g_1}^4 h_x, \mathrm{Rm}_{g_1,x}, ..., \nabla_{g_1}^2 \mathrm{Rm}_{g_1,x}).$$

Note that

$$F(0, ..., 0, \mathrm{Rm}_{g_1,x}, ..., \nabla_{g_1}^2 \mathrm{Rm}_{g_1,x}) = 0.$$

So, by (B.8) we have

$$|T_x|_{g_1} \leqslant C |(h_x, ..., \nabla_{g_1}^4 h_x)|_{g_1} \leqslant C(|h_x|_{g_1} + ... + |\nabla_{g_1}^4 h_x|_{g_1}) \leqslant C b D^{-E},$$

as long as $E \geqslant \underline{E}$ and $D \geqslant \underline{D}(\alpha)$. □

We now prove that the scales of g_1 and g_2 are close, up to an error that decays polynomially in D.

LEMMA 10.4. (Scale detection) *If*

$$E \geqslant \underline{E} \quad \textit{and} \quad D \geqslant \underline{D}(\alpha),$$

then we have

$$|\lambda_2 - 1| \leqslant b \cdot C \alpha^{-C} D^{-E+4}.$$

Proof. Set $\lambda_1 := 1$. Then, $g_i = \lambda_i^2 g_{\mathrm{Bry}}$ for $i=1,2$, and, by rescaling (B.2) and (B.3) by λ_i, we obtain that, for $i=1,2$,

$$R_{g_i} + |\nabla_{g_i} f|^2_{g_i} \equiv R_{g_i}(x_{\mathrm{Bry}}) = \lambda_i^{-2} R_{g_{\mathrm{Bry}}}(x_{\mathrm{Bry}}) \quad \text{and} \quad dR_{g_i} = 2\,\mathrm{Ric}_{g_i}(\nabla_{g_i} f, \cdot). \tag{10.2}$$

In the following, we will express these equations in terms of the metrics g_1 and g_2, by combining the difference estimates from the previous lemma with some estimates on the geometry of the normalized Bryant soliton from Lemma B.1. It will then follow that λ_1 and λ_2 are close.

In the following, we will work on the annulus $M_{\mathrm{Bry}}\left(\frac{1}{2}D, D\right)$ and assume $D > 2C_B$, where C_B is the constant from Lemma B.1. Therefore, $\sigma > \frac{1}{2}D > C_B$ on $M_{\mathrm{Bry}}\left(\frac{1}{2}D, D\right)$ and thus the bounds of Lemma B.1 apply for g_1. We may also assume that $E \geqslant \underline{E}$ and $D \geqslant \underline{D}(\alpha)$ have been chosen large enough so that g_1 and g_3 are 2-bilipschitz on $M_{\mathrm{Bry}}\left(\frac{1}{2}D, D\right)$.

From (B.5) in Lemma B.1 we obtain the following bound for the Ricci tensor, viewed as a quadratic form on T^*M:

$$\mathrm{Ric}_{g_1} > C_B^{-1} D^{-2} g_1.$$

Therefore, assuming D large enough, the inverse $\mathrm{Ric}_{g_1}^{-1}$, viewed as a map $T^*M \to T^*M$, is well defined and satisfies

$$|\mathrm{Ric}_{g_1}^{-1}|_{g_1} < C C_B D^2. \tag{10.3}$$

Hence, by Lemma 10.3, if $E \geqslant \underline{E}$ and $D \geqslant \underline{D}(\alpha)$, then

$$|\mathrm{Ric}_{g_1}^{-1}(\mathrm{Ric}_{g_3} - \mathrm{Ric}_{g_1})|_{g_1} \leqslant |\mathrm{Ric}_{g_1}^{-1}|_{g_1} |\mathrm{Ric}_{g_3} - \mathrm{Ric}_{g_1}|_{g_1} \leqslant b \cdot C D^{-E+2}.$$

So, if $E \geqslant \underline{E}$ and $D \geqslant \underline{D}(\alpha)$, then the inverse of

$$I + \mathrm{Ric}_{g_1}^{-1}(\mathrm{Ric}_{g_3} - \mathrm{Ric}_{g_1}) = \mathrm{Ric}_{g_1}^{-1}\,\mathrm{Ric}_{g_3}$$

exists and we have

$$|\mathrm{Ric}_{g_3}^{-1}\,\mathrm{Ric}_{g_1} - I|_{g_1} \leqslant b \cdot C D^{-E+2}.$$

Therefore, again by (10.3),

$$|\mathrm{Ric}_{g_3}^{-1}-\mathrm{Ric}_{g_1}^{-1}|_{g_1}\leqslant|\mathrm{Ric}_{g_3}^{-1}\,\mathrm{Ric}_{g_1}-I|_{g_1}|\mathrm{Ric}_{g_1}^{-1}|_{g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+4}. \tag{10.4}$$

Using the second relation in (10.2) we find that

$$\begin{aligned} d(f\circ\phi)-df&=2\,\mathrm{Ric}_{g_3}^{-1}(dR_{g_3})-2\,\mathrm{Ric}_{g_1}^{-1}(dR_{g_1})\\ &=2(\mathrm{Ric}_{g_3}^{-1}-\mathrm{Ric}_{g_1}^{-1})(dR_{g_3})+2\,\mathrm{Ric}_{g_1}^{-1}(dR_{g_3}-dR_{g_1}). \end{aligned} \tag{10.5}$$

So, as $|dR_{g_3}|_{g_1}\leqslant C|dR_{g_3}|_{g_3}\leqslant C\lambda_2^{-3}\leqslant C\alpha^{-3}$, we obtain by (10.3)–(10.5) and Lemma 10.3 that

$$|d(f\circ\phi)-df|_{g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+4}.$$

It follows using (B.7) that

$$\begin{aligned} \big||d(f\circ\phi)|_{g_3}^2-|df|_{g_1}^2\big|&\leqslant\big||d(f\circ\phi)|_{g_3}^2-|df|_{g_3}^2\big|+\big||df|_{g_3}^2-|df|_{g_1}^2\big|\\ &\leqslant|d(f\circ\phi)-df|_{g_3}\cdot|d(f\circ\phi)+df|_{g_3}+C|h|_{g_1}|df|_{g_1}^2\\ &\leqslant b\cdot C\alpha^{-C}D^{-E+4}(|d(f\circ\phi)|_{g_3}+|df|_{g_3})+b\cdot C\alpha^{-C}D^{-E}\\ &\leqslant b\cdot C\alpha^{-C}D^{-E+4}(|d(f\circ\phi)-df|_{g_3}+2|df|_{g_1})+b\cdot C\alpha^{-C}D^{-E}\\ &\leqslant b\cdot C\alpha^{-C}D^{-E+4}. \end{aligned}$$

Combining this with (10.2) and Lemma 10.3 yields

$$|\lambda_2^{-2}-\lambda_1^{-2}|\cdot R_{g_{\mathrm{Bry}}}(x_{\mathrm{Bry}})\leqslant|R_{g_3}-R_{g_1}|+\big||d(f\circ\phi)|_{g_3}^2-|df|_{g_1}^2\big|\leqslant b\cdot C\alpha^{-C}D^{-E+4}.$$

So, the bound on $|\lambda_2-1|$ follows for large enough D, as $\lambda_1=1$. □

Next, we prove that ϕ nearly preserves the radial distance function σ, up to an error that decays polynomially in D.

LEMMA 10.5. (ϕ nearly preserves σ) *If*

$$E\geqslant\underline{E}\quad\textit{and}\quad D\geqslant\underline{D}(\alpha), \tag{10.6}$$

then we have, for $k=0,1,2$,

$$|\nabla_{g_1}^k(\sigma\circ\phi-\sigma)|_{g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C}. \tag{10.7}$$

Proof. Let $F\colon(0,\infty)\to(0,\infty)$ be the function with the property that $R=F\circ\sigma$ on $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$. Consider the constant C_B from Lemma B.1. By (B.4), (B.6) and (B.8), we have, for $s>C_B$,

$$C_B^{-1}s^{-1}<F(s)<C_Bs^{-1},\quad C_B^{-1}s^{-2}<-F'(s)<C_B,\quad |F''(s)|,|F'''(s)|<C_B. \tag{10.8}$$

So, there is a $c_0>0$ such that $F^{-1}((0,c_0))=(C_B,\infty)$ and such that there is an inverse $H\colon(0,c_0)\to(C_B,\infty)$ of $F|_{(C_B,\infty)}$. A straightforward application of the chain rule gives

$$|H'(r)|<C_Br^{-2},\quad |H''(r)|<Cr^{-6}\quad\text{and}\quad |H'''(r)|<Cr^{-10}.$$

(Note that these bounds are not optimal.)

Assume now that E and D have been chosen large enough, in the sense of (10.6), that $\frac{1}{2}<\lambda_2<2$ by Lemma 10.4 and that by (10.8) and Lemma 10.3 we have, for $i=1,3$,

$$(10C_B)^{-1}D^{-1}<R_{g_i}<10C_BD^{-1}<\tfrac{1}{10}c_0$$

on $M_{\mathrm{Bry}}\big(\frac{1}{2}D,D\big)$. Then, on $M_{\mathrm{Bry}}\big(\frac{1}{2}D,D\big)$

$$\sigma\circ\phi-\sigma=P(R_{g_1},R_{g_3}-R_{g_1},\lambda_2),\tag{10.9}$$

where

$$P(r_1,r_2,\lambda):=H(\lambda^2(r_1+r_2))-H(r_1).$$

Note that $P(r,0,1)=0$ for all $r\in(0,c_0)$ and that on $((10C_B)^{-1}D^{-1},10C_BD^{-1})^2\times\big(\frac{1}{2},2\big)$ we have

$$|\partial^kP|\leqslant CD^{10}$$

for $k=1,2,3$. So, for $k=0,1,2$, we have

$$|\partial_1^kP|(r_1,r_2,\lambda)\leqslant CD^{10}(|r_2|+|\lambda-1|)\quad\text{on }((10C_B)^{-1}D^{-1},10C_BD^{-1})^2\times\big(\tfrac{1}{2},2\big).$$

So, (10.7) follows by differentiating (10.9) and using Lemmas 10.3 and 10.4. □

Recall that the Bryant soliton metric is a warped product $g_{\mathrm{Bry}}=d\sigma^2+w^2g_{S^2}$ on $M_{\mathrm{Bry}}\setminus\{x_{\mathrm{Bry}}\}$ and that $C_B^{-1}\sqrt{s}<w(s)<C_B\sqrt{s}$ for large s (see Lemma B.1 for more details). Fix some D that is sufficiently large such that $D>w(D)$. We now let $g_4=d\sigma_4^2+w_4^2g_{S^2}$ be a warped product metric on $M_{\mathrm{Bry}}\big(D-\frac{3}{4}w(D),D+\frac{1}{4}w(D)\big)$ with

$$\sigma_4=\frac{\sigma-D}{w(D)}$$

and the warping function

$$w_4=w_4(\sigma)=1+\sigma_4=1+\frac{\sigma-D}{w(D)}.$$

Note that there is an isometry

$$\Phi\colon M_{\mathrm{Bry}}\big(D-\tfrac{3}{4}w(D),D+\tfrac{1}{4}w(D)\big)\longrightarrow A_{1/4,5/4}\subset\mathbb{R}^3$$

to a Euclidean annulus such that $1+\sigma_4(x)=|\Phi(x)|_{\mathbb{R}^3}$. So

$$\Phi\big(M_{\mathrm{Bry}}(D-\tfrac{1}{2}w(D),D)\big)=A_{1/2,1}.$$

Due to Lemma 10.5, we may assume in the following that

$$\phi\big(M_{\mathrm{Bry}}(D-\tfrac{1}{2}w(D),D)\big)\subset M_{\mathrm{Bry}}\big(D-\tfrac{3}{4}w(D),D+\tfrac{1}{4}w(D)\big).$$

So, ϕ induces a map

$$\Phi\circ\phi\circ\Phi^{-1}\colon A_{1/2,1}\longrightarrow A_{1/4,5/4}.$$

We now show that ϕ restricted to $M_{\mathrm{Bry}}\big(D-\frac{1}{2}w(D),D\big)$ almost preserves the metric g_4 and the function σ_4. This is equivalent to saying that $\Phi\circ\phi\circ\Phi^{-1}$ almost preserves the Euclidean metric and the radial distance function on $\mathbb{R}^3$.

LEMMA 10.6. *If*

$$E\geqslant \underline{E}\quad \textit{and}\quad D\geqslant \underline{D}(\alpha), \tag{10.10}$$

then, for $k=0,1$,

$$|\nabla^k_{g_4}(\sigma_4\circ\phi-\sigma_4)|_{g_4}\leqslant b\cdot C\alpha^{-C}D^{-E+C}, \tag{10.11}$$

$$|\nabla^k_{g_4}(\phi^*g_4-g_4)|_{g_4}\leqslant b\cdot C\alpha^{-C}D^{-E+C}, \tag{10.12}$$

on $M_{\mathrm{Bry}}(D-\frac{1}{2}w(D),D)$.

Proof. Let us first consider the rescaled metric $\bar{g}_1:=w^{-2}(D)g_1$. This metric is a warped product of the form

$$\bar{g}_1=d\sigma_4^2+\bar{w}^2g_{S^2},$$

where

$$\bar{w}=\frac{w}{w(D)}.$$

Note that for large D the metric $\bar{g}_1$ on $M_{\mathrm{Bry}}\big(D-\frac{3}{4}w(D),D+\frac{1}{4}w(D)\big)$ is geometrically close to $S^2\times\big(-\frac{3}{4},\frac{1}{4}\big)$ equipped with the standard cylindrical metric. More precisely, if we express $\bar{w}=\bar{w}(\sigma_4)$ as a function in σ_4, then, by (B.10) in Lemma B.1, we have the following bounds, when $\sigma_4\in(-\frac{3}{4},\frac{1}{4})$:

$$|\bar{w}-1|\leqslant CD^{-1/2},\quad \left|\frac{d\bar{w}}{d\sigma_4}\right|\leqslant CD^{-1/2}\quad\text{and}\quad \left|\frac{d^2\bar{w}}{d\sigma_4^2}\right|\leqslant CD^{-1/2}. \tag{10.13}$$

Let us now consider the map ϕ. We have

$$\phi^*\bar{g}_1-\bar{g}_1=\lambda_2^{-2}w^{-2}(D)(\phi^*g_2-g_1+(1-\lambda_2^2)g_1).$$

Combining this with the scale detection Lemma 10.4 gives us the following bound for $k=0,1$, assuming an estimate of the form (10.10):

$$|\nabla^k_{\bar g_1}(\phi^*\bar g_1-\bar g_1)|_{\bar g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C}. \tag{10.14}$$

Note that here we have taken the covariant derivative with respect to $\bar g_1$, as opposed to g_1. This change produces a factor of the order of $O(D^{k/2})$, which can be absorbed in the right-hand side. Similarly, by rescaling (10.7) in Lemma 10.5 and assuming an estimate of the form (10.10), we obtain that, for $k=0,1,2$,

$$|\nabla^k_{\bar g_1}(\sigma_4\circ\phi-\sigma_4)|_{\bar g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C}. \tag{10.15}$$

This implies (10.11) for $k=0$ immediately and for $k=1$ after observing that $\bar g_1$ and g_4 are uniformly bilipschitz for large D.

So, it remains to show (10.12). The bound (10.15) implies that, for $k=0,1$,

$$|\nabla^k_{\bar g_1}(\phi^*d\sigma_4^2-d\sigma_4^2)|_{\bar g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C}. \tag{10.16}$$

Combining (10.14) and (10.16), one gets

$$\begin{aligned}|\nabla^k_{\bar g_1}(\phi^*(\bar w^2 g_{S^2})-\bar w^2 g_{S^2})|_{\bar g_1}&=|\nabla^k_{\bar g_1}((\phi^*\bar g_1-\bar g_1)-(\phi^*d\sigma_4^2-d\sigma_4^2))|_{\bar g_1}\\&\leqslant b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned} \tag{10.17}$$

Set now

$$\chi:=\frac{w_4^2}{\bar w^2}=\left(\frac{1+\sigma_4}{\bar w}\right)^2.$$

Let us first express $\chi(\sigma_4)$ as a function of σ_4. Then, by (10.13) we have, for $k=0,1,2$, as long as $-\frac34<\sigma_4<\frac14$,

$$|\chi|\leqslant C,\quad \left|\frac{d\chi}{d\sigma_4}\right|\leqslant C\quad\text{and}\quad\left|\frac{d^2\chi}{d\sigma_4^2}\right|\leqslant C. \tag{10.18}$$

It follows using (10.15) that

$$|\chi\circ\sigma_4\circ\phi-\chi\circ\sigma_4|\leqslant CD^{-1/2}|\sigma_4\circ\phi-\sigma_4|\leqslant b\cdot C\alpha^{-C}D^{-E+C}, \tag{10.19}$$

and

$$\begin{aligned}|\nabla_{\bar g_1}(\chi\circ\sigma_4\circ\phi-\chi\circ\sigma_4)|_{\bar g_1}&\leqslant|(\chi'\circ\sigma_4\circ\phi)\phi^*\,d\sigma_4-(\chi'\circ\sigma_4)\,d\sigma_4|_{\bar g_1}\\&\leqslant|\chi'\circ\sigma_4\circ\phi|\cdot|\phi^*d\sigma_4-d\sigma_4|_{\bar g_1}+|(\chi'\circ\sigma_4\circ\phi)-\chi'(\sigma_4)|\cdot|d\sigma_4|_{\bar g_1}\\&\leqslant b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned}$$

So, assuming a bound of the form (10.10), we get using (10.17) that, for $k=0,1$,

$$\begin{aligned}
&|\nabla^k_{\bar g_1}(\phi^*(w_4^2 g_{S^2})-w_4^2 g_{S^2})|_{\bar g_1}\\
&\quad=|\nabla^k_{\bar g_1}(\phi^*((\chi\circ\sigma_4)\bar w^2 g_{S^2})-(\chi\circ\sigma_4)\bar w^2 g_{S^2})|_{\bar g_1}\\
&\quad\leqslant|\nabla^k_{\bar g_1}((\chi\circ\sigma_4\circ\phi-\chi\circ\sigma_4)\phi^*(\bar w^2 g_{S^2}))|_{\bar g_1}\\
&\qquad+|\nabla^k_{\bar g_1}((\chi\circ\sigma_4)(\phi^*(\bar w^2 g_{S^2})-\bar w^2 g_{S^2}))|_{\bar g_1}\\
&\quad\leqslant b\cdot C\alpha^{-C}D^{-E+C}.
\end{aligned}$$

Combining this again with (10.16) gives us that, for $k=0,1$,

$$|\nabla^k_{\bar g_1}(\phi^* g_4-g_4)|_{\bar g_1}\leqslant b\cdot C\alpha^{-C}D^{-E+C}$$

This implies (10.12) for $k=0$, as g_4 and $\bar g_1$ are uniformly bilipschitz for large D. To see (10.12) for $k=1$, note that due to (10.18), we have $|\nabla_{g_4}-\nabla_{\bar g_1}|_{\bar g_1}\leqslant C$. □

In the following lemma, we extend the map $\Phi\circ\phi\circ\Phi^{-1}\colon A_{1/2,1}\to A_{1/4,5/4}$ to a map $\hat\phi$ on the unit ball $B_1\subset\mathbb{R}^3$.

LEMMA 10.7. *If*

$$E\geqslant \underline{E}\quad\text{and}\quad D\geqslant \underline{D}(\alpha),\tag{10.20}$$

then there is a diffeomorphism onto its image $\hat\phi\colon B_1\to\mathbb{R}^3$ such that the following holds:

(a) $|\hat\phi^* g_4-g_4|_{g_4}\leqslant b\cdot C\alpha^{-C}D^{-E+C}$.
(b) $|\sigma_4\circ\hat\phi-\sigma_4|, |\hat\phi^* d\sigma_4-d\sigma_4|_{g_4}\leqslant b\cdot C\alpha^{-C}D^{-E+C}$.
(c) $\hat\phi\equiv\Phi\circ\phi\circ\Phi^{-1}$ *on* $A_{5/6,1}$.
(d) $\hat\phi\equiv\psi$ *on* $B_{4/6}$ *for an orthogonal map* $\psi\in O(3)$ *of* $\mathbb{R}^3$.

Proof. By Lemma 10.6 and the fact that $g_4\equiv\Phi^* g_{\mathbb{R}^3}$ we have, for $k=0,1$,

$$|\nabla^k_{\mathbb{R}^3}((\Phi\circ\phi\circ\Phi^{-1})^* g_{\mathbb{R}^3}-g_{\mathbb{R}^3})|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C},\tag{10.21}$$

$$|\nabla^k_{\mathbb{R}^3}(r\circ(\Phi\circ\phi\circ\Phi^{-1})-r)|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C},\tag{10.22}$$

where $r(x):=|x|_{\mathbb{R}^3}$ denotes the radial distance function on $\mathbb{R}^3$.

From now on we will only work on $\mathbb{R}^3$. To simplify notation, we will write ϕ instead of $\Phi\circ\phi\circ\Phi^{-1}$. Expressing (10.21) for $k=1$ in Euclidean coordinates yields

$$\left|\sum_{s=1}^3\left(\frac{\partial^2\phi^s}{\partial x^k\partial x^i}\frac{\partial\phi^s}{\partial x^j}+\frac{\partial^2\phi^s}{\partial x^k\partial x^j}\frac{\partial\phi^s}{\partial x^i}\right)\right|\leqslant b\cdot C\alpha^{-C}D^{-E+C}.$$

Permuting the indices i, j and k cyclicly and using

$$2\frac{\partial^2\phi^s}{\partial x^i\partial x^j}\,\frac{\partial\phi^s}{\partial x^k}=\left(\frac{\partial^2\phi^s}{\partial x^i\partial x^j}\,\frac{\partial\phi^s}{\partial x^k}+\frac{\partial^2\phi^s}{\partial x^i\partial x^k}\,\frac{\partial\phi^s}{\partial x^j}\right)+\left(\frac{\partial^2\phi^s}{\partial x^j\partial x^k}\,\frac{\partial\phi^s}{\partial x^i}+\frac{\partial^2\phi^s}{\partial x^j\partial x^i}\,\frac{\partial\phi^s}{\partial x^k}\right)-\left(\frac{\partial^2\phi^s}{\partial x^k\partial x^i}\,\frac{\partial\phi^s}{\partial x^j}+\frac{\partial^2\phi^s}{\partial x^k\partial x^j}\,\frac{\partial\phi^s}{\partial x^i}\right)$$

gives us

$$\left|\sum_{s=1}^{3}\frac{\partial^2\phi^s}{\partial x^i\partial x^j}\,\frac{\partial\phi^s}{\partial x^k}\right|\leqslant b\cdot C\alpha^{-C}D^{-E+C}.$$

Combining this with (10.21) for $k=0$ implies that, under a condition of the form, (10.20)

$$|d^2\phi|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C}. \tag{10.23}$$

Let now $x_0\in A_{1/2,1}$ be a point and consider the differential $(d\phi)_{x_0}\colon\mathbb{R}^3\to\mathbb{R}^3$. By (10.21), there is a Euclidean isometry $\psi'\colon\mathbb{R}^3\to\mathbb{R}^3$ with $\psi'(x_0)=\phi(x_0)$ and

$$|(d\psi')_{x_0}-(d\phi)_{x_0}|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C}.$$

Combining this with (10.22) gives us

$$|(d\psi')_{x_0}((\nabla r)_{x_0})-(\nabla r)_{\psi'(x_0)}|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C}.$$

So, again by (10.22), we have

$$\begin{aligned}|\psi'(0)|_{\mathbb{R}^3}&=\big|\psi'(x_0)-|x_0|_{\mathbb{R}^3}(d\psi')_{x_0}((\nabla r)_{x_0})\big|_{\mathbb{R}^3}\\&\leqslant\big|\phi(x_0)-|x_0|_{\mathbb{R}^3}(\nabla r)_{\phi(x_0)}\big|_{\mathbb{R}^3}+b\cdot C\alpha^{-C}D^{-E+C}\\&\leqslant\big|\phi(x_0)-|\phi(x_0)|_{\mathbb{R}^3}(\nabla r)_{\phi(x_0)}\big|_{\mathbb{R}^3}+b\cdot C\alpha^{-C}D^{-E+C}\\&=b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned}$$

Set now $\psi:=\psi'-\psi'(0)$. Then, $\psi\in O(3)$ and, for $k=0,1$,

$$|d\psi(x_0)-d\phi(x_0)|_{\mathbb{R}^3}\leqslant b\cdot C\alpha^{-C}D^{-E+C}$$

Integrating (10.23) along paths in $A_{1/2,1}$ starting from x_0 implies that, under an assumption of the form (10.20), we have, for all $x\in A_{1/2,1}$,

$$\begin{aligned}|d\psi(x)-d\phi(x)|_{\mathbb{R}^3}&\leqslant|(d\psi)(x_0)-d\phi)(x_0)|_{\mathbb{R}^3}+10\sup_{A_{1/2,1}}|d^2\psi-d^2\phi|_{\mathbb{R}^3}\\&\leqslant b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned} \tag{10.24}$$

Integrating this bound once again along paths in $A_{1/2,1}$ yields that

$$\begin{aligned}|\psi(x)-\phi(x)|_{\mathbb{R}^3} &\leqslant |\psi(x_0)-\phi(x_0)|_{\mathbb{R}^3} + 10 \sup_{A_{1/2,1}} |d\psi(x)-d\phi(x)|_{\mathbb{R}^3} \\ &\leqslant b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned} \tag{10.25}$$

We now let $\{\zeta_1,\zeta_2\}$ be a partition of unity on $A_{1/2,1}$ such that $\zeta_1\equiv 1$ on $A_{5/6,1}$, $\zeta_2\equiv 1$ in $B_{4/6}$, and $|\nabla_{\mathbb{R}^3}\zeta_i|_{\mathbb{R}^3}\leqslant C$. Let $\hat{\phi}:=\zeta_1\phi+\zeta_2\psi$. Then, assertions (c) and (d) hold immediately and assertion (a) follows from (10.24) and (10.25). Assertion (b) follows from (10.22), the fact that $dr=d\sigma_4$ and that $r\circ\psi=r$. □

Proof of Proposition 10.2. We only need to translate the result of Lemma 10.7 back to M_{Bry}. By assertion (d) of Lemma 10.7 and the fact that M_{Bry} is rotationally symmetric, we can find an isometry $\tilde{\psi}\colon M_{\mathrm{Bry}}\to M_{\mathrm{Bry}}$ with $\tilde{\psi}(x_{\mathrm{Bry}})=x_{\mathrm{Bry}}$ and $\tilde{\psi}=\Phi^{-1}\circ\psi\circ\Phi$. Set $\tilde{\phi}:=\Phi^{-1}\circ\hat{\phi}\circ\Phi$ on $M_{\mathrm{Bry}}\big(D-\frac{1}{2}w(D),D\big)$ and $\tilde{\phi}:=\psi$ on the closure of $M_{\mathrm{Bry}}\big(D-\frac{1}{2}w(D)\big)$. By assertion (d) of Lemma 10.7, we know that $\tilde{\phi}$ is smooth. By assertion (b), the map $\tilde{\phi}$ is injective if $E\geqslant\underline{E}$ and $D\geqslant\underline{D}(\alpha)$. So, it remains to bound $\tilde{\phi}^*g_2-g_1$ on $M_{\mathrm{Bry}}\big(D-\frac{1}{2}w(D),D\big)$. To do this, we first deal with the rescaling factor λ_2 using Lemma 10.4:

$$|\tilde{\phi}^*g_2-g_1|_{g_1}\leqslant|\tilde{\phi}^*g_1-g_1|_{g_1}+|(\lambda_2^2-1)\tilde{\phi}^*g_1|_{g_1}\leqslant|\tilde{\phi}^*g_1-g_1|_{g_1}+b\cdot C\alpha^{-C}D^{-E+C}.$$

So, it remains to bound $\tilde{\phi}^*g_1-g_1$. For this purpose consider the rescaled metric

$$\bar{g}_1=w^{-2}(D)g_1=d\sigma_4^2+\bar{w}^2g_{S^2},$$

as used in the proof of Lemma 10.6, and observe that

$$\bar{g}_1=d\sigma_4^2+\bar{w}^2g_{S^2}=\frac{\bar{w}^2}{w_4^2}(d\sigma_4^2+w_4^2g_{S^2})+\left(1-\frac{\bar{w}^2}{w_4^2}\right)d\sigma_4^2.$$

Set $\chi\circ\sigma_4:=\bar{w}^2/w_4^2$ as in the proof of Lemma 10.6. As explained in this proof, we obtain using (10.19) and assertions (a) and (b) of Lemma 10.7 that, under an assumption of the form $E\geqslant\underline{E}$ and $D\geqslant\underline{D}(\alpha)$, we have

$$\begin{aligned}|\tilde{\phi}^*g_1-g_1|_{g_1} &= |\tilde{\phi}^*\bar{g}_1-\bar{g}_1|_{\bar{g}_1} \\ &\leqslant(\chi\circ\sigma_4)|\tilde{\phi}^*g_4-g_4|_{\bar{g}_1}+|\chi\circ\sigma_4\circ\tilde{\phi}-\chi\circ\sigma_4|\cdot|\tilde{\phi}^*g_4|_{\bar{g}_1} \\ &\quad+|1-(\chi\circ\sigma_4)|\cdot|\tilde{\phi}^*d\sigma_4^2-d\sigma_4^2|_{\bar{g}_1}+|\chi\circ\sigma_4\circ\tilde{\phi}-\chi\circ\sigma_4|\cdot|\tilde{\phi}^*d\sigma_4^2|_{\bar{g}_1} \\ &\leqslant b\cdot C\alpha^{-C}D^{-E+C}.\end{aligned}$$

This concludes the proof. □

Proof that Proposition 10.2 implies Proposition 10.1. Set

$$\lambda_2 := \lambda, \quad g_1 := g_{\mathrm{Bry}} \quad \text{and} \quad g_2 := \lambda_2^2 g_{\mathrm{Bry}}.$$

Assuming $\delta \leqslant \bar{\delta}(D)$, we have, using (B.4),

$$\rho_g \geqslant \tfrac{1}{2}\rho_{g_{\mathrm{Bry}}} \geqslant \tfrac{1}{4}C_B^{-1}D^{1/2} \quad \text{on } M_{\mathrm{Bry}}\big(\tfrac{1}{2}D, D\big). \tag{10.26}$$

Now, consider the map ϕ from Proposition 10.1 and note that, by the assumptions of this proposition and (10.26), that we have, for $m=0,...,4$,

$$|\nabla_g^m(\phi^* g' - g)|_g \leqslant b\rho_g^{-E} \leqslant 4^E C_B^E \cdot bD^{-E}. \tag{10.27}$$

We now claim that, for $D \geqslant \underline{D}(E,C)$ and $\delta \leqslant \bar{\delta}(E,C,D,b)$, we have

$$|\nabla_{g_1}^m(\phi^* g_2 - g_1)|_{g_1} \leqslant C_1'(E) bD^{-E} \tag{10.28}$$

for all $m=0,...,4$ and some constant

$$C_1' = C_1'(E) < \infty.$$

To see this, assume first that $D \geqslant \underline{D}(E,C)$ and $\delta \leqslant \bar{\delta}$ such that the pairs of metrics $\{g, \phi^* g'\}$, $\{g, g_1\}$, and $\{g', g_2\}$ are each 2-bilipschitz with respect to one another. So, g_1, g, $\phi^* g'$ and $\phi^* g_2$ are pairwise 8-bilipschitz. As $\lambda \in (C^{-1}, C)$, we can find a constant $C_2' = (C) < \infty$ such that by (10.1) we have, for all $m=0,...,4$,

$$|\nabla_{g_1}^m(g - g_1)|_{g_1} \leqslant C_2'\delta, \tag{10.29}$$

$$|\nabla_{\phi^* g_2}^m(\phi^* g' - \phi^* g_2)|_{g_1} \leqslant 8|\nabla_{g_2}^m(g' - g_2)|_{g_2} \circ \phi \leqslant C_2'\delta. \tag{10.30}$$

We now argue similarly as in the proof of Lemma 10.3. The tensor $\nabla_{g_1}^m(\phi^* g_2 - g_1)$ can be written as an algebraic expression in terms of the five tensors g^{-1}, $(\phi^* g_2)^{-1}$, $\nabla_g^{m'}(\phi^* g' - g)$, $\nabla_{g_1}^{m'}(g - g_1)$ and $\nabla_{\phi^* g_2}^{m'}(\phi^* g' - \phi^* g_2)$, with $m' \leqslant m$ (where we use g_1 as a background metric). So, pointwise,

$$\begin{aligned}\nabla_{g_1}^m(\phi^* g_2 - g_1) = F(&\phi^* g' - g, ..., \nabla_g^m(\phi^* g' - g),\\ &g - g_1, ..., \nabla_{g_1}^m(g - g_1), \phi^* g' - \phi^* g_2, ..., \nabla_{\phi^* g_2}^m(\phi^* g' - \phi^* g_2)),\end{aligned}$$

for some smooth, tensor-valued function F. By (10.27), (10.29) and (10.30), we therefore obtain (10.28) as long as $2^E C_B^E bD^{-E}$ and $C''\delta$ are sufficiently small.

So, the conditions of Proposition 10.2 are fulfilled for $\alpha=\alpha(E,C):=\min\{{C_1'}^{-1},C^{-1}\}$. Therefore, if

$$E\geqslant \underline{E} \quad \text{and} \quad D\geqslant \underline{D}(\alpha),$$

we obtain a diffeomorphism onto its image $\tilde{\phi}\colon M_{\mathrm{Bry}}(D)\to M_{\mathrm{Bry}}$ such that

$$\tilde{\phi}=\phi \quad \text{on } M_{\mathrm{Bry}}(D-1,D)$$

and moreover there is a universal constant $C_3'<\infty$ such that

$$|\tilde{\phi}^* g_2-g_1|_{g_1}\leqslant b\cdot C_3'\alpha^{-C_3'}D^{-E+C_3'}.$$

If $\delta\leqslant\bar{\delta}(E,b,\alpha,D)$, then we may assume that the metrics g, g_1, ϕ^*g' and ϕ^*g_2 are pairwise sufficiently bilipschitz close to another such that we still have for some universal $C_4'<\infty$

$$|\tilde{\phi}^* g'-g|_g\leqslant b\cdot C_4'\alpha^{-C_4'}D^{-E+C_4'}.$$

By (B.4), we have

$$C_4'\alpha^{-C_4'}D^{-E+C_4'}\leqslant\beta(10C_B)^{-3/2}D^{-3/2}\leqslant\beta\rho_g^{-3} \quad \text{on } M_{\mathrm{Bry}}\left(\tfrac{1}{2}D,D\right),$$

as long as

$$E\geqslant C_4'+4, \quad D\geqslant \underline{D}(E,\alpha,\beta) \quad \text{and} \quad \delta\leqslant\bar{\delta}(E,\alpha,\beta,D).$$

This implies assumption (b) of Proposition 10.1. Lastly, note that if

$$E\geqslant \underline{E} \quad \text{and} \quad D\geqslant \underline{D}(\alpha),$$

then $\tilde{\phi}$ is an immersion. So, since

$$\tilde{\phi}(M_{\mathrm{Bry}}(D-1,D))=\phi(M_{\mathrm{Bry}}(D-1,D))\subset M_{\mathrm{Bry}}(D'),$$

the image of $\tilde{\phi}$ must be contained in $M_{\mathrm{Bry}}(D')$ as well. □

11. Inductive step: extension of the comparison domain

11.1. Statement of the main result

Consider two Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$. The goal of this section is to extend a comparison domain $\mathcal{N}$ in $\mathcal{M}$ that is defined over a time-interval of the form $[0,t_J]$ by one time-step, to a comparison domain that is defined over the time-interval $[0,t_{J+1}=t_J+r_{\mathrm{comp}}^2]$. In order to carry out this construction, we will assume the existence

of a comparison from $\mathcal{M}$ to $\mathcal{M}'$ defined on $\mathcal{N}$ that together with $\mathcal{N}$ satisfies a-priori assumptions (APA1)–(APA6) for some tuple of parameters. Assuming that these parameters are chosen appropriately, we will show that the extended comparison domain and the given comparison satisfy the same a-priori assumptions for the same tuple of parameters.

The precise statement of the main result of this section is the following. We remind the reader that we are using the notation for expressing parameter bounds explained in §4.

PROPOSITION 11.1. (Extending the comparison domain) *Suppose that*

$$\begin{gathered}\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}, \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \quad \lambda \leqslant \bar{\lambda}(\delta_{\text{n}}), \quad D_{\text{cap}} \geqslant \underline{D}_{\text{cap}}(\lambda), \\ \Lambda \geqslant \underline{\Lambda}(\delta_{\text{n}}, \lambda), \quad \delta_{\text{b}} \leqslant \bar{\delta}_{\text{b}}(\lambda, \Lambda), \\ \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta_{\text{n}}, \lambda, \Lambda, \delta_{\text{b}}), \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}(\lambda, \Lambda)\end{gathered} \tag{11.1}$$

and assume that the following holds:

(i) $\mathcal{M}$ *and* $\mathcal{M}'$ *are two* $(\varepsilon_{\text{can}} r_{\text{comp}}, T)$*-complete Ricci flow spacetimes that each satisfy the* ε_{can}*-canonical neighborhood assumption at scales* $(\varepsilon_{\text{can}} r_{\text{comp}}, 1)$.

(ii) $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ *is a comparison domain in* $\mathcal{M}$ *that is defined on the time-interval* $[0, t_J]$. *We allow the case* $J{=}0$, *in which this comparison domain is empty* (*see Definition* 7.1).

(iii) $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^J)$ *is a comparison from* $\mathcal{M}$ *to* $\mathcal{M}'$ *defined on* $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ *over the (same) time-interval* $[0, t_J]$. *In the case* $J{=}0$, *this comparison is the trivial comparison* (*see the remark after Definition* 7.2).

(iv) $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$ *and* $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^J))$ *satisfy* (APA1)–(APA6) *for the parameters* $(\eta_{\text{lin}}, \delta_{\text{n}}, \lambda, D_{\text{cap}}, \Lambda, \delta_{\text{b}}, \varepsilon_{\text{can}}, r_{\text{comp}})$.

(v) $t_{J+1} := t_J + r_{\text{comp}}^2 \leqslant T$.

Then, there is a subset $\mathcal{N}^{J+1} \subset \mathcal{M}_{[t_J, t_{J+1}]}$ *such that*

$$(\mathcal{N} \cup \mathcal{N}^{J+1}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$$

is a comparison domain defined on the time-interval $[0, t_{J+1}]$ *and such that*

$$(\mathcal{N} \cup \mathcal{N}^{J+1}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$$

and $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^J)$ *satisfy the a-priori assumptions* (APA1)–(APA6) *for the same parameters* $(\eta_{\text{lin}}, \delta_{\text{n}}, \lambda, D_{\text{cap}}, \Lambda, \delta_{\text{b}}, \varepsilon_{\text{can}}, r_{\text{comp}})$.

We remind the reader that a-priori assumptions (APA1)–(APA6) allow for the possibility that the comparison $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^J)$ is defined on a shorter time-interval than the

underlying comparison domain (see Definition 7.4). In particular, (APA5) and (APA6) are only required to hold over the time-interval on which the comparison is defined, which in the context of Proposition 11.1 is $[0, t_J]$.

We briefly explain the strategy of the proof of Proposition 11.1, which will be carried out in the remainder of this section. In §11.2, we will first construct a domain $\Omega \subset \mathcal{M}_{t_{J+1}}$ such that the corresponding product domain $\Omega([t_J, t_{J+1}]) \subset \mathcal{M}_{[t_J, t_{J+1}]}$ satisfies most of the a-priori assumptions (APA1)–(APA6). The final time-slice $\mathcal{N}^{J+1}_{t_{J+1}}$ will later arise from Ω by adding certain components of its complement $\mathcal{M}_{t_{J+1}} \setminus \Omega$. This is by far the most delicate part of the proof, because we need to accommodate both a-priori assumption (APA3) (d), which forces certain components to be added to Ω, and a-priori assumption (APA5), which imposes strong restrictions whenever the addition of such components creates extension caps. The precise criterion for which components of $\mathcal{M}_{t_{J+1}} \setminus \Omega$ will be added to Ω, will be given in §11.3 and some of the less problematic a-priori assumptions will be verified in §11.4. The most important and complex step in our proof is Lemma 11.7 in §11.5, which effectively states that cap extensions only arise when a-priori assumption (APA5) is satisfied. For more details, we refer the reader to the explanations given before and after the statement of this lemma.

We make the standing assumption that hypotheses (i)–(v) of Proposition 11.1 hold for the remainder of this section. The construction of the domain $\mathcal{N}^{J+1}$ and the verification of its properties will proceed in several stages, with each stage requiring additional inequalities on the parameters. The inequalities on the parameters imposed in the assumptions of lemmas or in discussions in between lemmas will be retained for the remainder of this section. So, the assertions of these lemmas or the conclusions of these discussions continue to hold until the end of this section.

We remind the reader that, while the dependence on the parameters may seem complex, it essentially suffices to observe that the parameter order, as discussed in §7.5, is respected. We will continue our practice of introducing parameter bounds in separate displayed equations, to facilitate verification of the parameter dependences.

11.2. Choosing an almost minimal domain containing all Λr_{comp}-thick points

As a first step toward the construction of $\mathcal{N}^{J+1}$, we will construct a precursor of its final time-slice $\mathcal{N}^{J+1}_{t_{J+1}}$ — a subset $\Omega \subset \mathcal{M}_{t_{J+1}}$ bounded by central 2-spheres of δ_{n}-necks at scale r_{comp} that contains all Λr_{comp}-thick points. The final time-slice $\mathcal{N}^{J+1}_{t_{J+1}}$ of $\mathcal{N}^{J+1}$ will later emerge from Ω by the addition of certain components of its complement inside $\mathcal{M}_{t_{J+1}}$.

Consider the collection $\mathcal{S}$ of all embedded 2-spheres $\Sigma \subset \mathcal{M}_{t_{J+1}}$ that occur as central 2-spheres of δ_{n}-necks at scale r_{comp} in $\mathcal{M}_{t_{J+1}}$.

LEMMA 11.2. *We can find a subcollection $\mathcal{S}'\subset\mathcal{S}$ such that*
(a) $d_{t_{J+1}}(\Sigma_1,\Sigma_2)>10r_{\mathrm{comp}}$ *for all distinct* $\Sigma_1,\Sigma_2\in\mathcal{S}'$.
(b) *For every* $\Sigma\in\mathcal{S}$ *there is a* $\Sigma'\in\mathcal{S}'$ *such that* $d_{t_{J+1}}(\Sigma,\Sigma')<100r_{\mathrm{comp}}$.

Proof. Let $\{x_1,x_2,...\}\subset\mathcal{M}_{t_{J+1}}$ be a countable dense subset. We can successively construct a sequence of collections $\varnothing=\mathcal{S}'_0\subset\mathcal{S}'_1\subset...\subset\mathcal{S}$ by the following algorithm: If x_i is in an r_{comp}-neighborhood of some $\Sigma\in\mathcal{S}$ with the property that $d_{t_{J+1}}(\Sigma,\Sigma')>10r_{\mathrm{comp}}$ for all $\Sigma'\in\mathcal{S}'_{i-1}$, then we set $\mathcal{S}'_i:=\mathcal{S}'_{i-1}\cup\{\Sigma\}$. Otherwise, we set $\mathcal{S}'_i:=\mathcal{S}'_{i-1}$.

Set $\mathcal{S}':=\bigcup_{i=1}^{\infty}\mathcal{S}'_i$. Then, assertion (a) holds trivially and for assertion (b) observe that every $\Sigma\in\mathcal{S}$ is r_{comp}-close to some x_i. If $\mathcal{S}'_i=\mathcal{S}'_{i-1}$, then $d_{t_{J+1}}(\Sigma,\Sigma')\leqslant 10r_{\mathrm{comp}}$ for some $\Sigma'\in\mathcal{S}'_{i-1}$ and if $\mathcal{S}'_i=\mathcal{S}'_{i-1}\cup\{\Sigma'\}$, then x_i is contained in an r_{comp}-neighborhood of Σ'. In both cases, $d_{t_{J+1}}(\Sigma,\Sigma')<100r_{\mathrm{comp}}$. □

We now fix the collection $\mathcal{S}'$ for the remainder of this section.

LEMMA 11.3. *If*

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda}(\delta_{\mathrm{n}}),\quad \Lambda\geqslant\underline{\Lambda}(\delta_{\mathrm{n}}),\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta_{\mathrm{n}}),\quad r_{\mathrm{comp}}<1,$$

then the collection $\mathcal{S}'$ separates the $100\lambda r_{\mathrm{comp}}$-thin points of $\mathcal{M}_{t_{J+1}}$ from the $\Lambda r_{\mathrm{comp}}$-thick points.

Proof. Suppose that the assertion of the lemma was false. Then, there is a continuous path $\gamma\colon[0,1]\to\mathcal{M}_{t_{J+1}}\setminus\bigcup_{\Sigma\in\mathcal{S}'}\Sigma$ such that $\gamma(0)$ is $\Lambda r_{\mathrm{comp}}$-thick and $\gamma(1)$ is $100\lambda r_{\mathrm{comp}}$-thin. Without loss of generality, we may assume that γ has been chosen almost minimal in the sense that any other such path has length at least $\mathrm{length}_{t_{J+1}}(\gamma)-r_{\mathrm{comp}}$.

We first argue that we may assume in the following that

$$d_{t_{J+1}}(\gamma([0,1]),\Sigma')>1000r_{\mathrm{comp}}\quad\text{for all }\Sigma'\in\mathcal{S}',\tag{11.2}$$

Assume that $d_{t_{J+1}}(\gamma(s'),\Sigma')\leqslant 1000r_{\mathrm{comp}}$ for some $s'\in[0,1]$ and some $\Sigma'\in\mathcal{S}'$. Let $U\subset\mathcal{M}_{t_{J+1}}$ be a δ_{n}-neck at scale r_{comp} that has Σ' as a cross-sectional 2-sphere. If

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},$$

then $\gamma(s')\in U$. Moreover, if

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda}\quad\text{and}\quad\text{and}\quad \Lambda\geqslant\underline{\Lambda},$$

then no point on U is $\Lambda r_{\mathrm{comp}}$-thick or $100\lambda r_{\mathrm{comp}}$-thin and therefore $\gamma(0),\gamma(1)\notin U$. Let $\Sigma^*\subset U$ be a cross-sectional 2-sphere of U, close to its boundary such that Σ^* and Σ'

bound a domain diffeomorphic to $S^2\times[0,1]$ inside U that contains $\gamma(s')$. It follows that $\gamma|_{[0,s']}$ and $\gamma|_{[s',1]}$ intersect Σ^*. If

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},$$

then the diameter of Σ^* is less than $10r_{\mathrm{comp}}$, and Σ^* may be chosen such that the distance between Σ^* and $\gamma(s')$ is larger than $10r_{\mathrm{comp}}$. This implies that we can replace γ by a path whose length is shorter than $\mathrm{length}_{t_{J+1}}(\gamma)-r_{\mathrm{comp}}$, in contradiction to its almost minimality. Therefore, we may assume in the following that (11.2) holds.

By the intermediate value theorem, assuming

$$\lambda<\tfrac{1}{10},\quad \Lambda>1\quad\text{and}\quad r_{\mathrm{comp}}<1,$$

we may pick $s\in[0,1]$ such that $x:=\gamma(s)$ has scale $\rho(x)=r_{\mathrm{comp}}$. By the construction of $\mathcal{S}$ and (11.2), assuming

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},$$

the point x cannot be the center of a δ_{n}-neck at scale r_{comp}. So, assuming

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta_{\mathrm{n}})\quad\text{and}\quad r_{\mathrm{comp}}<1,$$

we can use Lemma 8.2 to find a compact subset $V\subset\mathcal{M}_{t_{J+1}}$ with $x\in V$ that has connected boundary and on which $C_0^{-1}r_{\mathrm{comp}}<\rho<C_0r_{\mathrm{comp}}$ holds, where $C_0=C_0(\delta_{\mathrm{n}})<\infty$. So, assuming

$$\lambda\leqslant(100C_0(\delta_{\mathrm{n}}))^{-1}\quad\text{and}\quad\Lambda\geqslant C_0(\delta_{\mathrm{n}}),$$

we can conclude that $\gamma(0),\gamma(1)\notin V$. Therefore, V must have exactly one boundary component and this component is a central 2-sphere of a δ_{n}-neck.

We claim that ∂V is disjoint from all elements of $\mathcal{S}'$. Assume by contradiction that ∂V intersects some $\Sigma'\in\mathcal{S}'$. If

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\tag{11.3}$$

then we have $\frac{1}{2}r_{\mathrm{comp}}<\rho<2r_{\mathrm{comp}}$ on $\Sigma'\cap\partial V$. Again, assuming a bound of the form (11.3), we find that ∂V is a central 2-sphere of a neck at some scale of the interval $\big(\frac{1}{4}r_{\mathrm{comp}},4r_{\mathrm{comp}}\big)$. So, the intersection of ∂V with $\gamma([0,1])$ is not further than $40r_{\mathrm{comp}}$ from the intersection with Σ', in contradiction to (11.2).

Choose now $s_1\in[0,s)$ and $s_2\in(s,1]$ such that $\gamma(s_i)\in\partial V$. By Lemma 8.2, the path $\gamma|_{[s_1,s_2]}$ can be replaced by a continuous path inside ∂V of length less than

$$\mathrm{length}_{t_{J+1}}(\gamma|_{[s_1,s_2]})-r_{\mathrm{comp}},$$

contradicting the minimality assumption of γ. □

Now let $\Omega\subset\mathcal{M}_{t_{J+1}}$ be the union of the closures of all components of

$$\mathcal{M}_{t_{J+1}}\setminus\bigcup_{\Sigma\in\mathcal{S}'}\Sigma$$

that contain $\Lambda r_{\mathrm{comp}}$-thick points. Then, by the previous lemma, Ω is weakly $100\lambda r_{\mathrm{comp}}$-thick.

LEMMA 11.4. *If*

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda),$$

then all points in Ω *survive until time* t_J.

Proof. This follows immediately from Lemma 8.5 and the fact that Ω is weakly $100\lambda r_{\mathrm{comp}}$-thick. □

LEMMA 11.5. *If* $J\geqslant 1$ *and*

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \Lambda\geqslant\underline{\Lambda},\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}\quad \textit{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}(\Lambda),$$

then for every $\Lambda r_{\mathrm{comp}}$*-thick point* $x\in\mathcal{M}_{t_{J+1}}$ *we have* $x(t_J)\in\operatorname{Int}\mathcal{N}_{t_J-}$.

Recall that $x(t_J)\in\mathcal{M}_{t_J}$ denotes the image of x under the time $-(t_{J+1}-t_J)$-flow of the time vector field ∂_t (see Definition 6.4).

Proof. Assume that $x(t_J)\notin\operatorname{Int}\mathcal{N}_{t_J-}$. By a-priori assumptions (APA3) (a) and (b), Lemma 8.5 and assuming that

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \Lambda\geqslant 2,\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}(\Lambda),$$

we have $\rho(x)\leqslant 2\Lambda r_{\mathrm{comp}}$.

Let $\delta_{\#}>0$ be a constant whose value we will determine in the course of the proof. Assuming

$$\Lambda\geqslant 1,\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta_{\#})\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}(\Lambda),$$

we can use Lemma 8.15 (for $\alpha=2\Lambda$) to argue that $(\mathcal{M}_{t_J},x(t_J))$ is $\delta_{\#}$-close to $(M_{\mathrm{Bry}},g_{\mathrm{Bry}},x_{\mathrm{Bry}})$ at scale $\rho(x)>\Lambda r_{\mathrm{comp}}$. Since ρ is uniformly bounded from below on $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$ and diverges at infinity, there is a universal constant $c>0$ such that for

$$\delta_{\#}\leqslant\bar{\delta}_{\#}$$

we can find a path $\gamma\colon[0,1]\to\mathcal{M}_{t_J}$ with

$$\gamma(0)=x(t_J),\quad \rho(\gamma(1))>\Lambda r_{\mathrm{comp}}\quad\text{and}\quad \rho(\gamma(s))>c\Lambda r_{\mathrm{comp}}$$

for all $s\in[0,1]$. So, by a-priori assumption (APA3) (b), we have $\gamma(1)\in\mathcal{N}_{t_J-}$. If

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}}\quad\text{and}\quad\Lambda\geqslant\underline{\Lambda},$$

then, by a-priori assumption (APA3) (a), the image $\gamma([0,1])$ is disjoint from $\partial\mathcal{N}_{t_J-}$. It follows that $x(t_J)=\gamma(0)\in\mathcal{N}_{t_J-}$. □

We remark that in the proof of Lemma 11.5, the use of Lemma 8.15, which is based on the rigidity theorems of Hamilton and Brendle, may be replaced by a longer but more elementary argument involving the maximum principle and the geometry of κ-solutions.

LEMMA 11.6. *If* $J \geqslant 1$ *and*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}(\delta_{\mathrm{n}}), \quad \Lambda \geqslant \underline{\Lambda}(\delta_{\mathrm{n}}), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\delta_{\mathrm{n}}) \quad \textit{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(\Lambda),$$

then $\Omega(t_J) \subset \operatorname{Int} \mathcal{N}_{t_J -}$.

Proof. Let Ω_0 be the closure of a component of $\mathcal{M}_{t_{J+1}} \setminus \bigcup_{\Sigma \in \mathcal{S}'} \Sigma$ that contains a $\Lambda r_{\mathrm{comp}}$-thick point x. Note that, by definition of Ω, we have $\Omega_0 \subset \Omega$ and the lemma follows if we can show that $\Omega_0(t_J) \subset \operatorname{Int} \mathcal{N}_{t_j -}$ for all such Ω_0.

Fix Ω_0 and a $\Lambda r_{\mathrm{comp}}$-thick point $x \in \Omega_0$ for the remainder of the proof and assume by contradiction that $\Omega_0(t_J) \not\subset \operatorname{Int} \mathcal{N}_{t_J -}$. Suppose by contradiction that there is a point Let $z \in \Omega_0$ with the property that $z(t_J) \notin \operatorname{Int} \mathcal{N}_{t_J -}$. Choose a path $\gamma_{J+1} \colon [0,1] \to \Omega_0$ within Ω_0 such that $x = \gamma_{J+1}(0)$ and $z = \gamma_{J+1}(1)$. Without loss of generality, we may assume that we have chosen z and γ_{J+1} almost minimal in the sense that for any other such choice of z' and γ'_{J+1} we have

$$\operatorname{length}_{t_{J+1}}(\gamma'_{J+1}) > \operatorname{length}_{t_{J+1}}(\gamma_{J+1}) - r_{\mathrm{comp}}. \tag{11.4}$$

By Lemma 11.5, assuming that

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}} \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(\Lambda),$$

we have $x(t_J) \in \mathcal{N}_{t_J -}$. Denote by $\gamma_J \colon [0,1] \to \mathcal{M}_{t_J}$ the curve at time t_J corresponding to γ_{J+1} under the $(-r_{\mathrm{comp}}^2)$-flow of the time vector field ∂_t, i.e. $\gamma_J(s) = (\gamma_{J+1}(s))(t_J)$. This path exists due to Lemma 11.4. Since $\gamma_J(1) = z(t_J) \notin \operatorname{Int} \mathcal{N}_{t_J -}$, we can find a parameter $s_0 \in [0,1]$ such that $\gamma_J(s_0) \in \partial \mathcal{N}_{t_J -}$ and $\gamma_J([0, s_0)) \subset \operatorname{Int} \mathcal{N}_{t_J -}$. By truncating γ_J and γ_{J+1}, we may assume without loss of generality that $s_0 = 1$ and therefore $z(t_J) = \gamma_J(1) \in \partial \mathcal{N}_{t_J -}$ and $\gamma_J([0,1)) \subset \operatorname{Int} \mathcal{N}_{t_J -}$. The almost minimality property (11.4) of z and γ_{J+1} remains preserved under this truncation process.

Let $\Sigma_J \subset \partial \mathcal{N}_{t_J -}$ be the boundary component that contains $z(t_J)$. By a-priori assumption (APA3) (a), Σ_J is a central 2-sphere of a δ_{n}-neck at scale r_{comp} in $\mathcal{M}_{t_J}$. Let $\delta_{\#} > 0$ be a constant whose value we will determine later. By Lemma 8.12, assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}(\delta_{\#}), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\delta_{\#}) \mathit{quad} \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

this implies that all points on Σ_J survive until time $t_J + \frac{1}{4} r_{\mathrm{comp}}^2$ and $\Sigma_J\big(t_J + \frac{1}{4} r_{\mathrm{comp}}^2\big)$ is a central 2-sphere of a $\delta_{\#}$-neck at scale $\frac{1}{2} r_{\mathrm{comp}}$. So, $\rho\big(z\big(t_J + \frac{1}{4} r_{\mathrm{comp}}^2\big)\big) < 0.6 r_{\mathrm{comp}}$, assuming

$$\delta_{\#} \leqslant \bar{\delta}_{\#}.$$

By Lemma 8.5, this implies that $\rho(\gamma_{J+1}(1))=\rho(z)<0.7r_{\mathrm{comp}}$, assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}} \quad \text{and} \quad r_{\mathrm{comp}} < 1.$$

Recall that at the other endpoint of γ_{J+1} we have $\rho(\gamma_{J+1}(0))=\rho(x)>\Lambda r_{\mathrm{comp}}$. So, by the intermediate value theorem, assuming

$$\Lambda > 1,$$

we can find a parameter $s\in(0,1)$ such that $y:=\gamma_{J+1}(s)$ has scale $\rho(y)=r_{\mathrm{comp}}$.

Assuming that

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}} \quad \text{and} \quad \Lambda \geqslant \underline{\Lambda},$$

we can conclude that x and z cannot lie in δ_{n}-necks at scale r_{comp}, and therefore

$$d_{t_{J+1}}(\{x,z\},\partial\Omega_0) > 2000 r_{\mathrm{comp}}.$$

So, by the almost minimal choice of γ_{J+1} we find, using the same argument as the one leading to (11.2) in the proof of Lemma 11.3, that

$$d_{t_{J+1}}(\gamma([0,1]),\partial\Omega_0) > 1000 r_{\mathrm{comp}}, \tag{11.5}$$

assuming that

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}.$$

As the interior of Ω_0 is disjoint from all elements of $\mathcal{S}'$, we can use assertion (b) of Lemma 11.2 and (11.5) to conclude that the point y cannot be a center of a δ_{n}-neck at scale r_{comp}, assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}.$$

We can hence apply Lemma 8.2 and find a smooth domain $V\subset\mathcal{M}_{t_{J+1}}$ with $y\in V$. Moreover, we have $C_0^{-1}(\delta_{\mathrm{n}})r_{\mathrm{comp}}<\rho<C_0(\delta_{\mathrm{n}})r_{\mathrm{comp}}$ on V. So, by Lemma 8.5 and assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\delta_{\mathrm{n}}),$$

all points on V survive until time t_J and

$$\rho \geqslant \tfrac{1}{2}C_0^{-1}(\delta_{\mathrm{n}})r_{\mathrm{comp}} \quad \text{on } V(t_J). \tag{11.6}$$

Also, if

$$\Lambda > C_0(\delta_{\mathrm{n}}),$$

then $x \notin V$. In particular, this implies that $\partial V \neq \varnothing$. By Lemma 8.2, the boundary ∂V is a central 2-sphere of a δ_{n}-neck. Choose $s_1 \in [0, s)$ such that $\gamma_{J+1}(s_1) \in \partial V$.

We claim that

$$z \in V. \tag{11.7}$$

If not, then we can choose $s_2 \in (s, 1]$ such that $\gamma_{J+1}(s_2) \in \partial V$. By Lemma 8.2, we can connect $\gamma_{J+1}(s_1)$ and $\gamma_{J+1}(s_2)$ by a path $\gamma' \colon [s_1, s_2] \to \partial V \subset \operatorname{Int} \Omega_0$ whose length is less than $\operatorname{length}_{t_{J+1}}(\gamma|_{[s_1,s]}) - 100 r_{\mathrm{comp}}$. The concatenation of $\gamma_{J+1}|_{[0,s_1]}$, γ' and $\gamma_{J+1}|_{[s_2,1]}$ would have length less than $\operatorname{length}_{t_{J+1}}(\gamma_{J+1}) - 100 r_{\mathrm{comp}}$, contradicting the almost minimal choice of γ_{J+1} and confirming (11.7).

Next, we argue that

$$(\partial V)(t_J) \subset \operatorname{Int} \mathcal{N}_{t_J -}. \tag{11.8}$$

Note that, by our choice of γ_{J+1}, we have $(\gamma_{J+1}(s_1))(t_J) \in \operatorname{Int} \mathcal{N}_{t_J -}$. So, if (11.8) was false, then $(\partial V)(t_J) \cap \partial \mathcal{N}_{t_J -} \neq \varnothing$. Therefore, by Lemma 8.2 we would find a continuous curve $\gamma'' \colon [s_1, 1] \to \partial V$ between $\gamma_{J+1}(s_1)$ and a point $z' \in \partial V$ with $z'(t_J) \in \partial \mathcal{N}_{t_J -}$ such that

$$\operatorname{length}_{t_{J+1}}(\gamma'') \leqslant d_{t_{J+1}}(\gamma_{J+1}(s_1), \gamma_{J+1}(s)) - 100 r_{\mathrm{comp}}.$$

The concatenation of $\gamma|_{[0,s_1]}$ with γ'' would then have length of at most

$$\operatorname{length}_{t_{J+1}}(\gamma_{J+1}) - d_{t_{J+1}}(\gamma_{J+1}(s_1), \gamma_{J+1}(s)) + \operatorname{length}_{t_{J+1}}(\gamma'') \leqslant \operatorname{length}_{t_{J+1}}(\gamma_{J+1}) - 100 r_{\mathrm{comp}}.$$

This, however, contradicts again the almost minimal choice of γ_{J+1}, confirming (11.8).

The inclusion (11.7) implies that $z(t_J) \in V(t_J)$. Let $\mathcal{C}^*$ be the component of

$$\mathcal{M}_{t_J} \setminus \operatorname{Int} \mathcal{N}_{t_J -}$$

that is adjacent to Σ_J. As $\mathcal{C}^*$ is path-connected and $z(t_J) \in \mathcal{C}^*$, we can conclude, using (11.8), that $\mathcal{C}^* \subset V$. By a-priori assumption (APA3) (d), there must be a $10\lambda r_{\mathrm{comp}}$-thin point in $\mathcal{C}^*$. So, if we choose

$$\lambda < \tfrac{1}{20} C_0^{-1}(\delta_{\mathrm{n}}),$$

then we obtain a contradiction to (11.6). □

11.3. The definition of $\mathcal{N}^{J+1}$

We will now enlarge Ω to a subset $\Omega^* \subset \mathcal{M}_{t_{J+1}}$ that will become the final time-slice $\mathcal{N}^{J+1}_{t_{J+1}}$ of the product domain $\mathcal{N}^{J+1}$. The components Z of the difference $\mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega$ fall into (at least) one of the following four types:

(I) Z has non-empty boundary and all points on Z are weakly $10\lambda r_{\mathrm{comp}}$-thick (in particular Z is not a closed component of $\mathcal{M}_{t_{J+1}}$).

(II) (a) Z is diffeomorphic to a 3-disk.
(b) $Z(t)$ is well defined and λr_{comp}-thick for all $t\in[t_J, t_{J+1}]$.
(c) $Z(t_J)\subset\mathcal{N}_{t_J-}$ if $J\geqslant 1$.

(III) (a) Z is diffeomorphic to a 3-disk.
(b) $Z(t)$ is well defined and λr_{comp}-thick for all $t\in[t_J, t_{J+1}]$.
(c) $\mathcal{C}:=Z(t_J)\setminus\operatorname{Int}\mathcal{N}_{t_J-}$ is a component of $\mathcal{M}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_J-}$, and there is a component $\mathcal{C}'\subset\mathcal{M}'_{t_J}\setminus\phi(\operatorname{Int}\mathcal{N}_{t_J-})$ such that a-priori assumptions (APA5) (a)–(e) hold, that is, the following holds:

- $\mathcal{C}$ and $\mathcal{C}'$ are 3-disks.
- $\partial\mathcal{C}'=\phi_{t_J-}(\partial\mathcal{C})$.
- There is a point $x\in\mathcal{C}$ such that $(\mathcal{M}_{t_J}, x)$ is δ_{b}-close to the pointed Bryant soliton $(M_{\text{Bry}}, g_{\text{Bry}}, x_{\text{Bry}})$ at scale $10\lambda r_{\text{comp}}$.
- There is a point $x'\in\mathcal{M}'_{t_J}$, at distance $\leqslant D_{\text{cap}} r_{\text{comp}}$ from $\mathcal{C}'$, such that $(\mathcal{M}'_{t_J}, x')$ is δ_{b}-close to the pointed Bryant soliton $(M_{\text{Bry}}, g_{\text{Bry}}, x_{\text{Bry}})$ at some scale in the interval $[D_{\text{cap}}^{-1} r_{\text{comp}}, D_{\text{cap}} r_{\text{comp}}]$.
- $\mathcal{C}$ and $\mathcal{C}'$ have diameter $\leqslant D_{\text{cap}} r_{\text{comp}}$.

(IV) None of the above.

Let Ω^* be the union of Ω with all components $Z\subset\mathcal{M}_{t_{J+1}}\setminus\operatorname{Int}\Omega$ that are of type (I), (II) or (III). Assuming $\varepsilon_{\text{can}}\leqslant\bar{\varepsilon}_{\text{can}}(\lambda)$, each component of type (I)–(III) survives until time t_J, either by definition or by Lemma 8.5. The subset Ω survives until time t_J by Lemma 11.4. Therefore, we may define $\mathcal{N}^{J+1}$ to be the product domain with final time-slice Ω^*:

$$\mathcal{N}^{J+1} := \bigcup_{t\in[t_J, t_{J+1}]} \Omega^*(t). \tag{11.9}$$

To provide some motivation for the choice of Ω^*, we point out that if $\Omega^*\subset\mathcal{M}_{t_{J+1}}$ is a manifold with boundary obtained from Ω by adding some components of its complement, and $\mathcal{N}^{J+1}$ is defined by (11.9), then one can check that $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ and $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ will only satisfy a-priori assumptions (APA1)–(APA6) if Ω^* includes all components of type (I)–(III). In this sense Ω^* is the "minimal" candidate for an extension of Ω that yields a comparison (domain) satisfying the a-priori assumptions.

In the remainder of this section we will complete the proof of Proposition 11.1 by verifying that $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ is a comparison domain, and $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ and $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ satisfy a-priori assumptions (APA1)–(APA6). Most of the verification is straightforward, using the results already established. The main difficulty will be establishing the properties of extension caps, especially (APA5). The crucial fact here, which we will prove in Lemma 11.7, is that components Z of type (I) and (II) satisfy

$Z(t_J)\subset\mathcal{N}_{t_J-}$. In other words, extension caps are only caused by components of type (III), which satisfy a-priori assumption (APA5).

The main idea of the proof of Lemma 11.7 will be to show that if $Z(t_J)\not\subset\mathcal{N}_{t_J-}$ for some component Z of type (I), then a-priori assumption (APA5) would have forced an extension cap to have occurred at some earlier time. For more details we refer to the reader to the overview preceding the proof of Lemma 11.7 in §11.5.

11.4. Verification of Proposition 11.1, except for Definition 7.1 (4) and (APA5)

We will now verify that $(\mathcal{N}\cup\mathcal{N}^{J+1},\{\mathcal{N}^j\}_{j=1}^{J+1},\{t_j\}^{J+1})$ satisfies properties (1)–(3) of the definition of a comparison domain (Definition 7.1) and that $(\mathcal{N}\cup\mathcal{N}^{J+1},\{\mathcal{N}^j\}_{j=1}^{J+1},\{t_j\}^{J+1})$ and $(\mathrm{Cut},\phi,\{\phi^j\}_{j=1}^{J})$ satisfy a-priori assumptions (APA1)–(APA4) and (APA6) (see Definition 7.4). Most of these properties and assumptions will follow fairly easily, apart from some technical points. The remaining verification of Definition 7.1 (4) and a-priori assumption (APA5) requires some deeper discussion, which we postpone to the next subsection.

We remind the reader that we assume inequalities of the form (11.1), such that the conclusions of the lemmas from the preceding subsections are valid.

Property (1) of Definition 7.1 holds by construction.

Next, let us verify property (2) of Definition 7.1. Since it is a union of Ω with connected components of its complement, Ω^* is a closed subset of $\mathcal{M}_{t_{J+1}}$, and is a domain with smooth boundary, where the boundary components are connected components of $\partial\Omega$. Since $\mathcal{N}_t^{J+1}=\Omega^*(t)$ is the image of Ω^* under the $(t-t_{J+1})$-flow of $\partial_{\mathfrak{t}}$, which is defined on a neighborhood of Ω^*, it follows that $\mathcal{N}_t^{J+1}$ is a domain with smooth boundary for all $t\in[t_J,t_{J+1}]$. Next, recall that Ω is weakly $100\lambda r_{\mathrm{comp}}$-thick by Lemma 11.3. By the definition of components of types (I)–(III) and Lemmas 8.5 and 8.6, assuming

$$\lambda\leqslant\bar{\lambda},\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda)\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

we find that, for all $t\in[t_J,t_{J+1}]$, the following statements hold:

(A) The time-slice $\mathcal{N}_t^{J+1}=\Omega^*(t)$ is $\lambda r_{\mathrm{comp}}$-thick.

(B) For every $x\in\Omega^*$ the parabolic neighborhood $P(x,r_{\mathrm{comp}})$ is unscathed and is cr_{comp}-thick, where $c=c(\lambda)>0$.

Now, suppose that $\{y_k\}\subset\mathcal{N}^{J+1}$ and $y_k\to y_\infty\in\mathcal{M}_{t_\infty}$. Then, $y_k=x_k(t_k)$ for some $x_k\in\Omega^*$, $t_k\in[t_J,t_{J+1}]$, and $t_k\to t_\infty$. Clearly, $x_k(t_\infty)\to y_\infty$. So, $\{x_k(t_\infty)\}$ is a Cauchy sequence in $\mathcal{M}_{t_\infty}$. Therefore, $\{x_k\}$ is Cauchy in $\mathcal{M}_{t_{J+1}}$ by (B) above and a distance

distortion argument. Since Ω^* is closed and λr_{comp}-thick, it is complete, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda).$$

It follows that $\{x_k\}$ converges to some $x_\infty \in \Omega^*$, and $x_\infty(t_\infty) = y_\infty$. Hence $\mathcal{N}^{J+1}$ is closed, and we have verified property (2) of Definition 7.1.

We have $\partial \mathcal{N}^{J+1}_{t_J} = (\partial \Omega^*)(t_J) \subset \Omega(t_J)$. Since $\Omega(t_J) \subset \operatorname{Int} \mathcal{N}_{t_J -}$ by Lemma 11.6, part (3) of Definition 7.1 holds.

We now turn to the a-priori assumptions.

A-priori assumption (APA1) is obvious. By (A) above, $\mathcal{N}^{J+1}$ is λr_{comp}-thick; so a-priori assumption (APA2) holds.

Note that we need only verify a-priori assumption (APA3) for $\mathcal{N}_{t_{J+1}-} = \Omega^*$. A-priori assumptions (APA3) (a)–(c) follow directly from the construction of Ω^*. To see a-priori assumption (APA3) (d), consider a component

$$Z \subset \mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \mathcal{N}_{t_{J+1}-} = \mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega^*$$

with non-empty boundary. Then, by construction, Z is a type-(IV) component of $\mathcal{M}_{t_{J+1}} \setminus \Omega$. As Z is not of type (I), it must contain a $10\lambda r_{\text{comp}}$-thin point. A-priori assumption (APA3) (e) holds, since in Proposition 11.1 the comparison is defined over the time-interval $[0, t_J]$, and does not include any cuts in $\mathcal{M}_{t_J}$.

Next, we verify a-priori assumption (APA4). Let $\mathcal{C}$ be a 3-disk component of $\mathcal{N}^J_{t_J} \setminus \operatorname{Int} \mathcal{N}^{J+1}_{t_J}$ (if $J \geqslant 1$) or $\mathcal{M}_0 \setminus \operatorname{Int} \mathcal{N}^1_0$ (if $J = 0$), such that $\partial \mathcal{C} \subset \mathcal{N}^{J+1}_{t_J}$. Assume by contradiction that all points on $\mathcal{C}$ survive until time t_{J+1} and that $\mathcal{C}(t)$ is λr_{comp}-thick for all $t \in [t_J, t_{J+1}]$. Then, $\mathcal{C}(t_{J+1})$ is contained in $\mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega^*$ by the definition of $\mathcal{N}^{J+1}$. Moreover, $\partial(\mathcal{C}(t_{J+1})) = (\partial \mathcal{C})(t_{J+1})$ is a 2-sphere contained in $\partial \Omega^*$, and hence an entire boundary component of Ω^*. It follows that $\mathcal{C}(t_{J+1})$ is a component of $\mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega^*$ that is also a component of $\mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega$ of type (IV). However, it is also of type (II), which is a contradiction.

Lastly, we point out that by the hypotheses of Proposition 11.1, we know that a-priori assumption (APA6) holds for $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ and $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ (recall that Definition 7.4 only requires the bound in a-priori assumption (APA6) to hold in the time-interval $[0, t_J]$).

11.5. Proof of Proposition 11.1, concluded

It remains to verify Definition 7.1 (4) and a-priori assumption (APA5).

We first verify the "if" direction of (APA5). To that end, suppose that $J \geqslant 1$ and that $\mathcal{C}$ is a component of $\mathcal{M}_{t_J} \setminus \operatorname{Int} \mathcal{N}_{t_J-}$ such that there is a component

$$\mathcal{C}' \subset \mathcal{M}'_{t_J} \setminus \phi_{t_J-}(\operatorname{Int} \mathcal{N}_{t_J-})$$

satisfying a-priori assumptions (APA5)(a)–(e); in other words:

(a) $\mathcal{C}$ and $\mathcal{C}'$ are 3-disks.

(b) $\partial \mathcal{C}' = \phi_{t_J-}(\partial \mathcal{C})$.

(c) There is a point $x \in \mathcal{C}$ such that $(\mathcal{M}_{t_J}, x)$ is δ_{b}-close to the pointed Bryant soliton $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at scale $10\lambda r_{\mathrm{comp}}$.

(d) There is a point $x' \in \mathcal{M}'_{t_J}$, at distance $\leqslant D_{\mathrm{cap}} r_{\mathrm{comp}}$ from $\mathcal{C}'$ such that $(\mathcal{M}'_{t_j}, x')$ is δ_{b}-close to the pointed Bryant soliton $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$ at some scale in the interval $[D_{\mathrm{cap}}^{-1} r_{\mathrm{comp}}, D_{\mathrm{cap}} r_{\mathrm{comp}}]$.

(e) $\mathcal{C}$ and $\mathcal{C}'$ have diameter $\leqslant D_{\mathrm{cap}} r_{\mathrm{comp}}$.

We now claim that, under suitable assumptions on the parameters, $\mathcal{C}$ is a component of $\mathcal{N}^{J+1}_{t_J} \setminus \operatorname{Int} \mathcal{N}^J_{t_J}$. Since $\mathcal{C}$ is a 3-disk by assumption, this will imply that $\mathcal{C}$ is an extension cap.

To see this, we will apply the Bryant slab lemma (Lemma 8.17) for

$$X_0 = \mathcal{N}^J_{t_J} \quad \text{and} \quad X_1 = \Omega.$$

Note that assumptions (i)–(iv) of the Bryant slab lemma hold due to Definition 7.1 (1), a-priori assumptions (APA3) (a)–(c) and by the construction of Ω. Assumption (v) of the Bryant slab lemma holds due to Lemma 11.6. So, the Bryant slab lemma can be applied on the time-interval $[t_J, t_{J+1}]$ if

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad 0 < \lambda < 1, \quad \Lambda \geqslant \underline{\Lambda} \quad \text{and} \quad \delta' \leqslant \bar{\delta}'(\lambda, \Lambda), \tag{11.10}$$

and if there is a map ψ with $\psi(x_{\mathrm{Bry}}, -(10\lambda)^{-2}) = x$ and a δ'-good Bryant slab $W \subset \mathcal{M}_{[t_J, t_{J+1}]}$, as required in the Bryant slab lemma. The existence of the map ψ and the δ'-good Bryant slab W follows from (c) above and Lemma 8.13, and assuming

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, \delta'), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, \delta') \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}.$$

Under assumptions of the same form as (11.10), we can also apply the Bryant slice lemma (Lemma 8.16) at time t_i, $i = J, J+1$, for $\psi = \psi_{t_i}$, $W = W_{t_i}$ and $X = X_{i-J}$.

Let $\mathcal{C}_0 := W_{t_J} \setminus \operatorname{Int} X_0$ and $\mathcal{C}_1 := W_{t_{J+1}} \setminus \operatorname{Int} X_1$ be as in the Bryant slab lemma. By the Bryant slice lemma applied at time t_{J+1}, we know that $x(t_{J+1}) = \psi(x_{\mathrm{Bry}}, 0)$ is $11\lambda r_{\mathrm{comp}}$-thin. So, by construction of Ω, we have $x(t_{J+1}) \in \mathcal{C}_1 \neq \varnothing$. By assertions (a) and (b) of the

Bryant slice lemma, we find that $\mathcal{C}_0=\mathcal{C}$ and that $\mathcal{C}_1$ is a 3-disk component of $\mathcal{M}_{t_{J+1}}\setminus\operatorname{Int}\Omega$. Assertions (a) and (b) of the Bryant slab lemma imply that $\mathcal{C}_1(t)$ is $9\lambda r_{\text{comp}}$-thick for all $t\in[t_J,t_{J+1}]$, and $\mathcal{C}_1(t_J)\supset\mathcal{C}_0$ and $\mathcal{C}=\mathcal{C}_0=\mathcal{C}_1(t_J)\setminus\operatorname{Int}\mathcal{N}^J_{t_J}$. It follows that $Z:=\mathcal{C}_1$ is a component of type (III), and so $Z(t_J)\subset\Omega^*(t_J)=\mathcal{N}^{J+1}_{t_J}$. Thus, $\mathcal{C}\subset\mathcal{N}^{J+1}_{t_J}\setminus\operatorname{Int}\mathcal{N}^J_{t_J}$, and since $\mathcal{C}$ is a component of $\mathcal{M}_{t_J}\setminus\operatorname{Int}\mathcal{N}^J_{t_J}$, it is also a component of $\mathcal{N}^{J+1}_{t_J}\setminus\operatorname{Int}\mathcal{N}^J_{t_J}$. Hence, the "if" direction of (APA5) holds.

In order to verify Definition 7.1(4) and the "only if" direction of a-priori assumption (APA5), we need the following fundamental result.

LEMMA 11.7. (Structure of extension caps) *If*

$$\eta_{\text{lin}}\leqslant\bar\eta_{\text{lin}},\quad \delta_{\text{n}}\leqslant\bar\delta_{\text{n}},\quad \lambda\leqslant\bar\lambda,\quad D_{\text{cap}}\geqslant\underline{D}_{\text{cap}}(\lambda),\quad \Lambda\geqslant\underline{\Lambda}(\lambda),$$
$$\varepsilon_{\text{can}}\leqslant\bar\varepsilon_{\text{can}}(\lambda,\Lambda,\delta_{\text{b}})\quad \textit{and}\quad r_{\text{comp}}\leqslant\bar r_{\text{comp}}(\lambda),$$

then the following holds.

If $Z\subset\mathcal{M}_{t_{J+1}}\setminus\operatorname{Int}\Omega$ is a component of type (I)*, then $Z(t_J)\subset\mathcal{N}_{t_{J-}}$.*

Before proceeding, we first explain how Lemma 11.7 completes the verification of Proposition 11.1.

For this purpose consider a component $\mathcal{C}^*\subset\mathcal{N}^{J+1}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}$. As $\Omega(t_J)\subset\operatorname{Int}\mathcal{N}_{t_{J-}}$, we have $\mathcal{C}^*\subset\mathcal{N}^{J+1}_{t_J}\setminus\Omega(t_J)$. Thus, $\mathcal{C}^*\subset\operatorname{Int}Z(t_J)$ for some component $Z\subset\mathcal{M}_{t_{J+1}}\setminus\operatorname{Int}\Omega$ of type (I), (II) or (III). By the above lemma and condition (II) (c), Z cannot be of type (I) or (II) and therefore must be of type (III). Next, observe that $\mathcal{C}^*\subset Z(t_J)\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}=:\mathcal{C}$ and

$$\mathcal{C}=Z(t_J)\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}\subset\mathcal{N}^{J+1}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}.$$

As $\mathcal{C}^*$ is a connected component of $\mathcal{N}^{J+1}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}$, it follows that $\mathcal{C}=\mathcal{C}^*$.

By (III) (c) we know that $\mathcal{C}^*=\mathcal{C}$ is a 3-disk, which proves Definition 7.1 (4). The remaining statements of (III) (c) imply that $\mathcal{C}^*=\mathcal{C}$ satisfies (APA5) (a)–(e).

Next, we provide an outline of the proof of Lemma 11.7, neglecting several technicalities.

Assume by contradiction that Z is a type-(I) component with $Z(t_J)\not\subset\mathcal{N}_{t_{J-}}$. This means that $Z(t_J)$ contains a component $\mathcal{C}$ of the complement $\mathcal{M}_{t_{J-}}\setminus\operatorname{Int}\mathcal{N}_{t_{J-}}$. As Z consists of weakly $10\lambda r_{\text{comp}}$-thick points and $\mathcal{C}$ contains a $10\lambda r_{\text{comp}}$-thin point by a-priori assumption (APA3) (d), there must be a point in $\mathcal{C}$ whose scale increases over the time-interval $[t_J,t_{J+1}]$. By Lemma 8.15, this is only possible if Z and $\mathcal{C}$ lie in a large spacetime region $W\subset\mathcal{M}$ that is very close to a Bryant soliton. More specifically, we may assume that this region is $9\lambda r_{\text{comp}}$-thick and defined over a long backward time-interval of the form $[t_{J-J_\#},t_{J+1}]$, where $J_\#\gg1$.

The existence of the component $\mathcal{C}$ and the Bryant like geometry on W will then force the existence of a sequence of components $\mathcal{C}_j \subset \mathcal{M}_{t_j} \setminus \operatorname{Int} \mathcal{N}_{t_j -}$ for $j = J, J-1, ..., J-J_\#$, where $\mathcal{C}_J = \mathcal{C}$. This will follow from a-priori assumption (APA4), which forbids the discard of components that remain λr_{comp}-thick during a time step.

Next, using the bilipschitz bound on the comparison map ϕ imposed by (APA6), and the fact that W is not too neck-like, we will find that for $t \in [t_{J-J_\#}, t_J]$, the image $\phi_t(W \cap \mathcal{N}_t)$ intersects a smoothly varying 3-disk region $\widehat{W}'_t \subset \mathcal{M}'_t$ with scale and diameter comparable to r_{comp}.

The union $\widehat{W}' \subset \mathcal{M}'_{[t_{J-J_\#}, t_J]}$ of these regions forms a "barrier region" that will help us show the existence of a point $z' \in \widehat{W}'_{t_{J-J_\#}}$ that survives until time t_J and that has the property that $z'(t) \in \widehat{W}'$ for all $t \in [t_{J-J_\#}, t_J]$. The scale of $z'(t)$ will be controlled from above and below by a constant that is independent of $J_\#$. Therefore, if we choose $J_\#$ large enough, then we can find a time-step $t_j \in [t_{J-J_\#+1}, t_{J-1}]$ such that the scale of $z(t)$ hardly decreases over the time-interval $[t_{j-1}, t_j]$. Using again Lemma 8.15 (this time in $\mathcal{M}'$), we will deduce that the geometry near $z'(t_j)$ is close to a Bryant soliton. This means that (APA5) applies and would have forced $\mathcal{C}_j$ to be an extension cap, giving a contradiction.

Proof of Lemma 11.7. Fix a type-(I) component $Z \subset \mathcal{M}_{t_{J+1}} \setminus \operatorname{Int} \Omega$ for the remainder of the proof and assume that $Z(t_J) \not\subset \mathcal{N}_{t_J}$. So, $Z(t_J)$ intersects a component $\mathcal{C}$ of $\mathcal{M}_{t_J} \setminus \operatorname{Int} \mathcal{N}_{t_J -}$. Because $Z(t_J)$ is a closed subset, its topological boundary in $\mathcal{M}_{t_J}$ is $\partial Z(t_J)$, and since $\partial Z(t_J) \subset \partial \Omega(t_J) \subset \operatorname{Int} \mathcal{N}_{t_J -}$, it is disjoint from $\mathcal{M}_{t_J} \setminus \operatorname{Int} \mathcal{N}_{t_J -}$. The connectedness of $\mathcal{C}$ now implies that $\mathcal{C} \subset Z(t_J)$.

By a-priori assumption (APA3) (d) there is a $10\lambda r_{\text{comp}}$-thin point $x \in \mathcal{C} \subset Z(t_J)$. By the type-(I) property and the discussion in §11.3, we know that x survives until time t_{J+1} and that $x(t_{J+1})$ is weakly $10\lambda r_{\text{comp}}$-thick. Moreover, by Lemma 8.5, we find that $x(t_{J+1})$ is $11\lambda r_{\text{comp}}$-thin, assuming

$$\lambda \leqslant \tfrac{1}{10}, \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda) \quad \text{and} \quad r_{\text{comp}} \leqslant \tfrac{1}{10}.$$

We can therefore apply 8.15 to x and obtain that a large spacetime neighborhood of $x(t_{J+1})$ is close to a Bryant soliton. More specifically, Lemma 8.15 implies the following. Let $\delta_\# > 0$ and $J_\# < \infty$ be constants whose values will be determined in the course of the proof. Then, under a condition of the form

$$\lambda \leqslant \tfrac{1}{20}, \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda, J_\#, \delta_\#) \quad \text{and} \quad r_{\text{comp}} \leqslant 1,$$

we can find a $(10\lambda r_{\text{comp}})^2$-time equivariant and ∂_t-preserving diffeomorphism

$$\psi \colon W^* := \overline{M_{\text{Bry}}(\delta_\#^{-1})} \times [-\min\{J, J_\# + 1\} \cdot (10\lambda)^{-2}, 0] \longrightarrow \mathcal{M}$$

onto its image such that $\psi(x_{\mathrm{Bry}},0)=x(t_{J+1})$ and

$$\|(10\lambda r_{\mathrm{comp}})^{-2}\psi^*g-g_{\mathrm{Bry}}\|_{C^{[\delta_\#^{-1}]}(W^*)}<\delta_\#. \tag{11.11}$$

Let $W=\psi(W^*)$. Note that W^* has been chosen in such a way that its image W has initial time-slice $t_{J-J_\#}$ if $J_\#\leqslant J-1$, and t_1 otherwise.

Next, we show that the existence of the component $\mathcal{C}$ of $\mathcal{M}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_J-}$ forces the existence of components $\mathcal{C}_j\subset\mathcal{M}_{t_j}\setminus\operatorname{Int}\mathcal{N}_{t_j-}$ at a large number of earlier times $t_j\leqslant t_J$. The existence of these components will be deduced using a-priori assumption (APA4) and the Bryant-like geometry on W.

Claim 11.8. (Cap hierarchy) *If, in addition,*

$$\delta_{\mathrm{n}}\leqslant\bar\delta_{\mathrm{n}},\quad \lambda\leqslant\bar\lambda,\quad \Lambda\geqslant\underline{\Lambda}\quad \textit{and}\quad \delta_\#\leqslant\bar\delta_\#(\lambda,\Lambda,J_\#),$$

then $J\geqslant J_\#+1$ *and the following statements hold*:

(a) *For all* $J-J_\#\leqslant j\leqslant J$ *the subset* $\mathcal{C}_j:=W_{t_j}\setminus\operatorname{Int}\mathcal{N}_{t_j-}$ *is a 3-disk.*

(b) *For all* $J-J_\#+1\leqslant j\leqslant J$ *all points on* $\mathcal{C}_j$ *survive until time* t_{j-1} *and* $\mathcal{C}_{j-1}\subset\mathcal{C}_j(t_{j-1})$.

(c) $\mathcal{C}=\mathcal{C}_J$.

Proof. In the following we will apply the Bryant slice lemmma (Lemma 8.16) at time t_j for $X=\mathcal{N}_{t_j-}$, where $J-J_\#\leqslant j\leqslant J$. We will also apply the Bryant slab lemma (Lemma 8.17) for $X_0=\mathcal{N}_{t_{j-1}-}$ and $X_1=\mathcal{N}_{t_j-}$, where $J-J_\#+1\leqslant j\leqslant J$. Note that assumptions (i)–(iv) of the Bryant slice lemma hold due to a-priori assumptions (APA3) (a)–(c) and assumption (v) of the Bryant slab lemma holds due to Definition 7.1(3). If

$$\delta_{\mathrm{n}}\leqslant\bar\delta_{\mathrm{n}},\quad 0<\lambda<1,\quad \Lambda\geqslant\underline{\Lambda}\quad\text{and}\quad \delta_\#\leqslant\bar\delta_\#(J_\#,\lambda,\Lambda),$$

then the remaining assumptions of both the Bryant slice and the Bryant slab lemma are satisfied. This means, in particular, that the time-slice W_{t_j} and the slab $W_{[t_{j-1},t_j]}$ satisfy the assumptions of the Bryant slice and slab lemmas, for all $J-J_\#\leqslant j\leqslant J$ and $J-J_\#+1\leqslant j\leqslant J$, respectively.

Since $x\in\mathcal{C}\cap(W_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_J-})$, we know by the Bryant slice lemma at time t_J that $\mathcal{C}_J:=W_{t_J-}\setminus\operatorname{Int}\mathcal{N}_{t_J}$ is a 3-disk and is a component of $\mathcal{M}_{t_J}\setminus\operatorname{Int}\mathcal{N}_{t_J-}$. Hence, it coincides with $\mathcal{C}$, which proves assertion (c).

Fix some j with $J-J_\#\leqslant j\leqslant J$. Assume inductively that $j\geqslant 1$, and that assertion (a) holds for all $j\leqslant j'\leqslant J$ and assertion (b) holds for all $j+1\leqslant j'\leqslant J$. If $j=J-J_\#$, then $J\geqslant J_\#+1$, as claimed, and assertions (a) and (b) hold. So, assume in the following that $j>J-J_\#$.

By assertion (a) of the Bryant slab lemma, $\mathcal{C}_j(t)$ is defined and $9\lambda r_{\mathrm{comp}}$-thick for all $t\in[t_{j-1},t_j]$. Moreover, the subset $\mathcal{C}_j(t_{j-1})$ is a 3-disk component of $\mathcal{M}_{t_{j-1}}\setminus\operatorname{Int}\mathcal{N}_{t_{j-1}+}$ and $\partial\mathcal{C}_j(t_{j-1})\subset\partial\mathcal{N}_{t_{j-1}+}$. It follows from a-priori assumption (APA4) that $j-1\geqslant 1$. Now suppose that $\mathcal{C}_{j-1}=W_{t_{j-1}}\setminus\operatorname{Int}\mathcal{N}_{t_{j-1}-}=\varnothing$. It follows that $\mathcal{C}_j(t_{j-1})\subset W_{t_{j-1}}\subset\operatorname{Int}\mathcal{N}_{t_{j-1}-}$. Therefore, $\mathcal{C}_j(t_{j-1})\subset\mathcal{N}_{t_{j-1}-}\setminus\operatorname{Int}\mathcal{N}_{t_{j-1}+}$, and since it is a 3-disk with boundary contained in $\partial\mathcal{N}_{t_{j-1}+}$, it is a component of $\mathcal{N}_{t_{j-1}-}\setminus\operatorname{Int}\mathcal{N}_{t_{j-1}+}$. This contradicts a-priori assumption (APA4). Thus, $\mathcal{C}_{j-1}\neq\varnothing$ and by the Bryant slice lemma at time t_{j-1} it must be a 3-disk. So, assertion (a) holds for $j-1$ and assertion (b) holds for j by the Bryant slab lemma.

By induction we conclude that $J\geqslant J_{\#}+1$, and (a) and (b) hold. □

Next, we will construct the "barrier" region $\widehat{W}'$ mentioned in the outline given above. We remark that in the following construction, we have to choose $\widehat{W}$ larger than the reader may anticipate. The reason is purely technical: Due to the fact that a-priori assumption (APA6) only gives us C^0 bounds on the metric distortion of ϕ, the weakness of the resulting scale distortion control (see Lemma 8.9) forces us to work in a region whose boundary has scale a large multiple of r_{comp}.

We will now construct the subset $\widehat{W}\subset W$. For this purpose fix the (universal) constant C_{SD} from Lemma 8.9 and assume without loss of generality that $C_{\mathrm{SD}}>100$. Define

$$\widehat{W}^*\subset M_{\mathrm{Bry}}\times[-(J_{\#}+1)(10\lambda)^{-2},-(10\lambda)^{-2}]$$

to be the subset of points on which $\rho\leqslant 20C_{\mathrm{SD}}^2\cdot(10\lambda)^{-1}$. Then, $\widehat{W}^*$ is closed and connected, and its time-slices $\widehat{W}^*_t$ are pairwise isometric 3-disks for all

$$t\in[-(J_{\#}+1)(10\lambda)^{-2},-(10\lambda)^{-2}].$$

If

$$\delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,J_{\#}),$$

then

$$\widehat{W}^*\subset W^*=\overline{M_{\mathrm{Bry}}(\delta_{\#}^{-1})}\times[-(J_{\#}+1)(10\lambda)^{-2},-(10\lambda)^{-2}].$$

So, we may define

$$\widehat{W}:=\psi(\widehat{W}^*)\subset\mathcal{M}_{[t_{J-J_{\#}},t_J]}.$$

Then, assuming

$$\delta_{\#}\leqslant\bar{\delta}_{\#}(\lambda,J_{\#})\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

we obtain that, for all $t\in[t_{J-J_{\#}},t_J]$, the time-slice $\widehat{W}_t$ is a 3-disk and

$$\begin{gathered}10C_{\mathrm{SD}}^2r_{\mathrm{comp}}<\rho=\rho_1<40C_{\mathrm{SD}}^2r_{\mathrm{comp}}\quad\text{on }\partial\widehat{W}_t,\\ W_t\setminus\operatorname{Int}\widehat{W}_t\quad\text{is }10C_{\mathrm{SD}}^2r_{\mathrm{comp}}\text{-thick},\\ \widehat{W}_t\quad\text{is }40C_{\mathrm{SD}}^2r_{\mathrm{comp}}\text{-thin}\end{gathered}\tag{11.12}$$

CLAIM 11.9. *If*

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},\quad \delta_{\#}\leqslant\bar\delta_{\#}(\lambda,\Lambda,J_{\#}),\quad \varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}\quad \textit{and}\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

then, the following statements hold:

(a) $W\subset\mathcal{M}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}}\mathcal{D}$, *and hence by Definition* 7.2 (5) *the map* ϕ *is well defined on* $W_{[t_{J-J_{\#}},t_J]}\cap\mathcal{N}$.

(b) *For all* $J-J_{\#}<j\leqslant J$ *and* $t\in[t_{j-1},t_j]$,

$$W_t\setminus\operatorname{Int}\widehat{W}_t\subset\operatorname{Int}\mathcal{N}_t^j.$$

(c) $\mathcal{C}_j\subset\operatorname{Int}\widehat{W}_{t_j}$ *for all* $J-J_{\#}\leqslant j\leqslant J$.

Proof. If

$$\delta_{\#}\leqslant\bar\delta_{\#}(\lambda,\Lambda,J_{\#}),$$

then for every $J-J_{\#}\leqslant j\leqslant J+1$ we get that ∂W_{t_j} is $\Lambda r_{\rm comp}$-thick.

Suppose that $\mathcal{D}\cap W_{t_j}\neq\varnothing$ for some $\mathcal{D}\in\mathrm{Cut}$. Note that this implies that $J-J_{\#}\leqslant j<J$. Since $\mathcal{D}$ is $\Lambda r_{\rm comp}$-thin by (APA3) (e), it is disjoint from ∂W_{t_j}. So, since $\mathcal{D}$ is connected by Definition 7.2 (2), we have $\mathcal{D}\subset W_{t_j}$. By Definition 7.2 (3) the cut $\mathcal{D}$ contains an extension cap. However, this contradicts assertion (b) of Claim 11.8. So, we have shown assertion (a) of this claim.

Now suppose that $J-J_{\#}<j\leqslant J$ and $t\in[t_{j-1},t_j]$ or $j=J-J_{\#}$ and $t=t_j$. As ∂W_{t_j} is $\Lambda r_{\rm comp}$-thick, it is contained in $\mathcal{N}_{t_j-}$ by (APA3) (b). Thus $\partial W_t\subset\mathcal{N}^j$. Moreover, if

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},\quad \varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}\quad\text{and}\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

then we obtain from Lemma 8.12 that $\partial\mathcal{N}_t^j$ is $2.1r_{\rm comp}$-thin. In view of the fact that $W_t\setminus\operatorname{Int}\widehat{W}_t$ is connected and $10C_{\rm SD}^2r_{\rm comp}$-thick by (11.12), it is disjoint from $\partial\mathcal{N}_t^j$, and hence contained in $\mathcal{N}_t$. This proves assertion (b) and assertion (c) follows in the case $t=t_j$. □

Next, we consider the image of $W_t\setminus\operatorname{Int}\widehat{W}_t$ under ϕ, and show that the boundary component $\phi_t(\partial\widehat{W}_t)$ is adjacent to a region with controlled geometry.

CLAIM 11.10. *Assuming*

$$\eta_{\rm lin}\leqslant\bar\eta_{\rm lin},\quad \delta_{\rm n}\leqslant\bar\delta_{\rm n},\quad \lambda\leqslant\bar\lambda,\quad \Lambda\geqslant\underline{\Lambda}(\lambda),\quad \delta_{\#}\leqslant\bar\delta_{\#}(\lambda,J_{\#}),$$
$$\varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}(\lambda),\quad r_{\rm comp}\leqslant\bar r_{\rm comp}(\lambda),$$

there is a constant $C_1=C_1(\lambda)<\infty$ *with the following property.*

There is a subset $\widehat{W}' \subset \mathcal{M}'_{[t_{J-J_\#}, t_J]}$ *such that, for every* $t \in [t_{J-J_\#}, t_J]$,

(a) $\widehat{W}'_t$ *is a 3-disk.*

(b) $\widehat{W}'_t \cap \phi_t(W_t \setminus \operatorname{Int} \widehat{W}_t) = \phi_t(\partial \widehat{W}_t) = \partial \widehat{W}'_t$.

(c) $\widehat{W}'$ *is compact and its relative topological boundary inside the time-slab*

$$\mathcal{M}'_{[t_{J-J_\#}, t_J]}$$

is equal to

$$\bigcup_{t \in [t_{J-J_\#}, t_J]} \partial \widehat{W}'_t.$$

(d) $\widehat{W}'_t$ *is* $C_1 r_{\text{comp}}$*-thin and* $C_1^{-1} r_{\text{comp}}$*-thick and* $\operatorname{diam}_t \widehat{W}'_t \leqslant C_1 r_{\text{comp}}$.

(e) $\partial \widehat{W}'_t$ *is* $10 C_{\text{SD}} r_{\text{comp}}$*-thick.*

(f) *For any* $J - J_\# \leqslant j \leqslant J$ *the difference* $\mathcal{C}'_j := \widehat{W}'_{t_j} \setminus \phi_{t_j -}(\operatorname{Int} \mathcal{N}_{t_j -})$ *is a 3-disk component of* $\mathcal{M}'_{t_j} \setminus \phi_{t_j -}(\operatorname{Int} \mathcal{N}_{t_j -})$, *and we have* $\partial \mathcal{C}'_j = \phi_{t_j -}(\partial \mathcal{C}_j)$.

Proof. Fix $t \in [t_{J-J_\#}, t_J]$. By (11.12), a-priori assumptions (APA2), (APA3) (a), (c), (APA6) and Lemma 8.9, assuming

$$\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}, \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \quad \Lambda \geqslant 2, \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda) \quad \text{and} \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}, \tag{11.13}$$

we have

$$10 C_{\text{SD}} r_{\text{comp}} < \rho_1 = \rho < 40 C_{\text{SD}}^3 r_{\text{comp}} \quad \text{on} \quad \phi_t(\partial \widehat{W}_t), \tag{11.14}$$

and

$$\operatorname{diam} \phi_t(\partial \widehat{W}_t) < 10 \operatorname{diam} \partial \widehat{W}_t < C'_1 r_{\text{comp}},$$

where $C'_1 < \infty$ is a universal constant that can be determined in terms of C_{SD}.

Choose $x \in \partial \widehat{W}_t$. Using (11.12) and assuming

$$\delta_\# \leqslant \bar{\delta}_\#(\lambda, J_\#) \quad \text{and} \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}},$$

we can find a point $y \in W_t \setminus \operatorname{Int} \widehat{W}_t$ with $\rho_1(y) = \rho(y) = 80 C_{\text{SD}}^4 r_{\text{comp}}$ that can be connected to x by a path of length at most $C'_2 r_{\text{comp}}$ inside $W_t \setminus \operatorname{Int} \widehat{W}_t$, for some $C'_2 = C'_2(\lambda) < \infty$. Let $x' = \phi_t(x)$ and $y' = \phi_t(y)$. Again, by the scale distortion Lemma 8.9, a-priori assumptions (APA2), (APA3) (a), (c), (APA6), and assuming a bound of the form (11.13), we conclude using (11.14) that

$$\rho(y') > 80 C_{\text{SD}}^3 r_{\text{comp}} > 2 \rho(x') \tag{11.15}$$

and $d_{\mathcal{M}'_t}(x', y') < 2 C'_2 r_{\text{comp}}$. So, there is a constant $\delta^\circ = \delta^\circ(\lambda) > 0$ such that x' cannot be a center of a δ°-neck in $\mathcal{M}'_t$.

Let us now apply Lemma 8.2 to x' for $\delta=\delta^\circ(\lambda)$. In order to satisfy the assumptions of this lemma, we need to assume that

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda) \quad \text{and} \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}.$$

We obtain a constant $C_0=C_0(\delta^\circ(\lambda))<\infty$ and a compact subset $V'\subset\mathcal{M}'_t$, containing x' such that (compare with (11.14))

$$10C_0^{-1}C_{\text{SD}}r_{\text{comp}} < \rho_1 = \rho < 40C_0C_{\text{SD}}^3 r_{\text{comp}} \quad \text{on } V', \tag{11.16}$$

and such that $\operatorname{diam}_t V'<40C_0C_{\text{SD}}^3r_{\text{comp}}$. Moreover, we may assume that $\delta^\circ(\lambda)$ is chosen small enough such that $B(x',\max\{C'_1,2C'_2\}r_{\text{comp}})\subset V'$. This implies that $y'\in\operatorname{Int} V'$ and

$$\phi_t(\partial\widehat{W}_t)\subset \operatorname{Int} V'. \tag{11.17}$$

We claim that V' is a 3-disk. To see this, we assume

$$\delta_{\#} \leqslant \bar{\delta}_{\#}(\lambda, J_{\#}) \quad \text{and} \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}(\lambda),$$

such that ∂W_t is $40C_0C_{\text{SD}}^4r_{\text{comp}}$-thick and that $40C_0C_{\text{SD}}^4r_{\text{comp}}\leqslant 1$. So, again, by the scale distortion Lemma 8.9, a-priori assumptions (APA2), (APA3) (a), (c), (APA6), and assuming (11.13), we obtain that $\phi_t(\partial W_t)$ is $40C_0C_{\text{SD}}^3r_{\text{comp}}$-thick. Thus, by (11.16), we have

$$\phi_t(\partial W_t)\cap V' = \varnothing. \tag{11.18}$$

As x' and $\phi_t(\partial W_t)$ lie in the same connected component of $\mathcal{M}'_t$, we must have $\partial V'\neq\varnothing$, and due to (11.15), Lemma 8.2 implies that V' is a 3-disk.

By (11.17) the 2-sphere $\phi_t(\partial\widehat{W}_t)$ bounds a 3-disk $\widehat{W}'_t\subset V'$. We now repeat the construction above for all $t\in[t_{J-J_{\#}},t_J]$ and set

$$\widehat{W}' := \bigcup_{t\in[t_{J-J_{\#}},t_J]} \widehat{W}'_t.$$

Then, assertion (a) holds automatically.

Next, observe that $\widehat{W}'_t$ and $\phi_t(W_t\setminus\operatorname{Int}(\widehat{W}_t))$ are compact connected domains with smooth boundary that share a single boundary component $\phi(\partial\widehat{W}_t)$. Therefore, assertion (b) of this claim can fail only if $\phi_t(W_t\setminus\operatorname{Int}\widehat{W}_t)\subset\widehat{W}'_t\subset V'$, which would contradict (11.18). Thus assertion (b) holds.

In order to show assertion (c), it suffices to show that for all $t\in[t_{J-J_{\#}},t_J]$, every point $q\in\operatorname{Int}\widehat{W}'_t$ is not contained in the relative boundary of $\widehat{W}'$ inside $\mathcal{M}'_{[t_{J-J_{\#}},t_J]}$. To see this,

let $U \subset \mathcal{M}'_{[t_{J-J_\#},t_J]}$ be a product domain containing $\widehat{W}'_t$ that is open in $\mathcal{M}'_{[t_{J-J_\#},t_J]}$. By Claim 11.9 and Definition 7.2, the map ϕ is well defined and smooth on a neighborhood of

$$\bigcup_{\bar{t}\in[t_{J-J_\#},t_J]} W_{\bar{t}} \setminus \operatorname{Int} \widehat{W}_{\bar{t}}.$$

Therefore, after shrinking U if necessary, we may assume that if $\bar{t}\in[t_{J-J_\#},t_J]$ is close to t, then $(\phi(W_{\bar{t}}\setminus\operatorname{Int}\widehat{W}_{\bar{t}})\cap U_{\bar{t}})(t)$ is defined and moves by smooth isotopy as $\bar{t}$ varies. So, by assertion (b) the 3-disk $(\widehat{W}'_{\bar{t}}\cap U_{\bar{t}})(t)$ varies by smooth isotopy as well and therefore it contains a small neighborhood of q inside $\mathcal{M}'_t$ for $\bar{t}$ close to t. This implies that a small neighborhood of q is contained in $\bigcup_{\bar{t}\in[t_{J-J_\#},t_J]} \widehat{W}'_{\bar{t}}$, which finishes the proof of assertion (c). The same argument implies that $\widehat{W}'$ is a finite union of compact subsets and must therefore be compact.

Assertions (d) and (e) follow by construction of $\widehat{W}'_t$ and (11.16) and (11.14), as long as $C_1 > 40 C_0 C_{\mathrm{SD}}^3$.

Lastly, consider assertion (f). Suppose that $J-J_\# \leqslant j \leqslant J$.

Recall that $\mathcal{C}_j = W_{t_j}\setminus\operatorname{Int}\mathcal{N}_{t_j-}$ is a 3-disk by assertion (a) of Claim 11.8. By assertion (b) of Claim 11.9, we have $W_{t_j}\setminus\operatorname{Int}\widehat{W}_{t_j}\subset\operatorname{Int}\mathcal{N}_{t_j-}$. So, $\mathcal{C}_j\subset\operatorname{Int}\widehat{W}_{t_j}$. Therefore,

$$\widehat{W}_{t_j}\setminus\operatorname{Int}\mathcal{C}_j = \mathcal{N}_{t_j-}\cap\widehat{W}_{t_j}$$

is a compact connected manifold with boundary.

As $\phi_{t_j-}(\partial\widehat{W}_{t_j}) = \partial\widehat{W}'_{t_j}$ and $\phi_{t_j-}:\mathcal{N}_{t_j-}\to\mathcal{M}'_{t_j}$ is injective, the image

$$\phi_{t_j-}(\operatorname{Int}\widehat{W}_{t_j}\cap\mathcal{N}_{t_j-})$$

must either be contained in $\widehat{W}'_{t_j}$ or in its complement. Since ϕ_{t_j-} maps a neighborhood of $\partial\widehat{W}_{t_j}$ in $\widehat{W}_{t_j}$ into $\widehat{W}'_{t_j}$, we obtain by connectedness that

$$\phi_{t_j-}(\mathcal{N}_{t_j-}\cap\widehat{W}_{t_j})\subset\widehat{W}'_{t_j}.$$

By the same argument, if N_0 is the component of $\mathcal{N}_{t_j-}$ that contains $\partial\widehat{W}_{t_j}$, then

$$\phi_{t_j-}(N_0\setminus\widehat{W}_{t_j})$$

is disjoint from $\widehat{W}'_{t_j}$.

Assume now that there is a component $N_1\neq N_0$ of $\mathcal{N}_{t_j-}$ with the property that $\phi_{t_j-}(N_1)$ intersects $\widehat{W}'_{t_j}$. Then, again, since $\phi_{t_j-}(\partial\widehat{W}_{t_j})=\partial\widehat{W}'_{t_j}$ and ϕ_{t_j-} is injective, we must have $\phi_{t_j-}(N_1)\subset\widehat{W}'_{t_j}$. By a-priori assumption (APA3) (c), we know that N_1 must contain a $\Lambda r_{\mathrm{comp}}$-thick point. So, by Lemma 8.9, a-priori assumptions (APA2),

(APA3) (a), (c), (APA6) and assuming (11.13), these points must be mapped by ϕ_{t_j-} to $C_{\mathrm{SD}}^{-1}\Lambda r_{\mathrm{comp}}$-thick points in $\widehat{W}'_{t_j}$. Assuming

$$\Lambda \geqslant \underline{\Lambda}(\lambda) \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(\lambda),$$

this, however, contradicts assertion (d) of this claim.

Combining the conclusions of the last two paragraphs, we obtain that

$$\phi_{t_j-}(\mathcal{N}_{t_j-}) \cap \widehat{W}'_{t_j} = \phi_{t_j-}(N_0) \cap \widehat{W}'_{t_j} = \phi_{t_j-}(N_0 \cap \widehat{W}_{t_j}) = \phi_{t_j-}(\mathcal{N}_{t_j-} \cap \widehat{W}_{t_j}).$$

Since $\partial \widehat{W}'_{t_j} = \phi_{t_j-}(\partial \widehat{W}_{t_j}) \subset \phi_{t_j-}(\operatorname{Int} \mathcal{N}_{t_j-})$ we obtain from Alexander's theorem that

$$\mathcal{C}'_{t_j} = \widehat{W}'_{t_j} \setminus \phi_{t_j-}(\operatorname{Int} \mathcal{N}_{t_j-}) = \widehat{W}'_{t_j} \setminus \phi_{t_j-}(\widehat{W}_{t_j} \cap \operatorname{Int} \mathcal{N}_{t_j-})$$

is a 3-disk. As $\mathcal{C}'_{t_j} \subset \operatorname{Int} \widehat{W}'_{t_j}$, it is also a component of $\mathcal{M}'_{t_j} \setminus \phi_{t_j-}(\operatorname{Int} \mathcal{N}_{t_j-})$. This establishes assertion (f). □

Choose $z' = \phi_{t_{J-J_\#}-}(z) \in \phi_{t_{J-J_\#}-}(\partial \mathcal{C}_{J-J_\#})$. We now show that z' survives until time t_J and at some time lies at the tip of an approximate Bryant soliton.

Claim 11.11. *If*

$$\eta_{\mathrm{lin}} \leqslant \bar{\eta}_{\mathrm{lin}}, \quad \delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad J_\# \geqslant \underline{J}_\#(\lambda, \delta_{\mathrm{b}}), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, \delta_{\mathrm{b}}, J_\#), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(\lambda),$$

then, the following statements hold:

(a) *$z'(t)$ is defined and contained in $\widehat{W}'_t$ for all $t \in [t_{J-J_\#}, t_J]$.*

(b) *There is a $j_0 \in \{J-J_\#, ..., J-1\}$ such that $(\mathcal{M}'_{t_{j_0}}, z'(t_{j_0}))$ is δ_{b}-close to*

$$(M_{\mathrm{Bry}}, g_{\mathrm{Bry}}, x_{\mathrm{Bry}})$$

at scale $\rho(z'(t_{j_0})) < C_1(\lambda) r_{\mathrm{comp}}$.

Proof. By the scale distortion lemma (Lemma 8.9) and a-priori assumptions (APA2), (APA3) (a), (c), (APA6) we obtain, assuming

$$\eta_{\mathrm{lin}} \leqslant \eta_{\mathrm{lin}}, \quad \delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \Lambda \geqslant 2, \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda) \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

that

$$\rho(z') = \rho_1(z') < C_{\mathrm{SD}} \rho_1(z) < 2 C_{\mathrm{SD}} r_{\mathrm{comp}}.$$

Next, recall that the scalar curvature on any κ-solution is pointwise non-decreasing in time (see assertion (e) of Lemma C.1). So, assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(J_\#) \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

we obtain from Lemma 6.2 that at any $y \in \mathcal{M}'$, with $\varepsilon_{\text{can}} r_{\text{comp}} < \rho(y) < 4C_{\text{SD}} r_{\text{comp}}$, we have

$$\partial_t \rho^2(y) = \partial_t\big(\tfrac{1}{3}R\big) < C_{\text{SD}}^2 J_\#^{-1}. \tag{11.19}$$

Choose $t^* \in [t_{J-J_\#}, t_J]$ maximal such that $z'(t)$ is well defined for all $t \in [t_{J-J_\#}, t^*)$. By (11.19) we have $\rho^2(z'(t)) < 5C_{\text{SD}}^2 r_{\text{comp}}^2$ for all $t \in [t_{J-J_\#+1}, t^*)$, and therefore

$$\rho(z'(t)) < 10 C_{\text{SD}} r_{\text{comp}}$$

for all such t. Suppose that $t^* < t_J$. As $\widehat{W}'$ is compact, we must have $z'(t) \notin \widehat{W}'$ for t close to t^*. So, there is a $t' \in [t_{J-J_\#}, t^*)$ such that $z(t')$ lies on the relative topological boundary of $\widehat{W}'$ inside $\mathcal{M}'_{[t_{J-J_\#}, t_J]}$. By assertions (c) and (e) of Claim 11.10 this, however, implies that $\rho(z'(t')) > 10 C_{\text{SD}} r_{\text{comp}}$, contradicting our previous conclusion. Therefore, $t^* = t_J$ and $z'(t) \in \widehat{W}'$ for all $t \in [t_{J-J_\#}, t_J]$.

Since

$$\sum_{j=J-J_\#+1}^{J-1} (\rho(z'(t_j)) - \rho(z'(t_{j-1}))) = \rho(z'(t_{J-1})) - \rho(z') > -\rho(z') > -2C_{\text{SD}} r_{\text{comp}},$$

we can find a $j_0 \in \{J - J_\# + 1, \dots J-1\}$ such that

$$\rho(z'(t_{j_0})) - \rho(z'(t_{j_0-1})) > -\frac{2C_{\text{SD}}}{J_\# - 1} r_{\text{comp}}.$$

Next, observe that

$$C_1^{-1}(\lambda) r_{\text{comp}} < \rho(z'(t_{j_0})) < C_1(\lambda) r_{\text{comp}},$$

by assertion (d) of Claim 11.10. So, by Lemma 8.15, assuming

$$J_\# \geqslant \underline{J}_\#(\lambda, \delta_{\text{b}}), \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda, \delta_{\text{b}}), \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}(\lambda),$$

we find that $(\mathcal{M}'_{t_{j_0}}, z'(t_{j_0}))$ is δ_{b}-close to $(M_{\text{Bry}}, g_{\text{Bry}}, x_{\text{Bry}})$ at scale $\rho(z'(t_{j_0}))$. □

Consider the component $\mathcal{C}'_{j_0}$ from assertion (f) of Claim 11.10. We will now verify that (APA5) (a)–(e) hold for $\mathcal{C} = \mathcal{C}_{j_0}$ and $\mathcal{C}' = \mathcal{C}'_{j_0}$, forcing the existence of an extension cap at time t_{j_0}. A-priori assumptions (APA5) (a) and (b) hold by Claim 11.8 (a) and Claim 11.10 (f). A-priori assumption (APA5) (c) is implied by (11.11), assuming

$$\delta_\# \leqslant \bar{\delta}_\#(\delta_{\text{b}}, J_\#).$$

A-priori assumption (APA5) (d) and the diameter bound on $\mathcal{C}'_{j_0}$ in a-priori assumption (APA5) (e) hold by Claim 11.10 (d) and Claim 11.11 (b), as long as

$$D_{\text{cap}} \geqslant C_1(\lambda).$$

Lastly, the diameter bound on $\mathcal{C}_{j_0}$ in a-priori assumption (APA5) (e) follows from the fact that $\mathcal{C}_{j_0} \subset \widehat{W}_{t_{j_0}}$ and, by construction, assuming that

$$D_{\text{cap}} \geqslant \underline{D}_{\text{cap}}(\lambda).$$

The conclusions of the previous paragraph combined with (APA5) imply that $\mathcal{C}_{j_0}$ must be an extension cap, which implies that $\mathcal{C}_{j_0} \subset \mathcal{N}_{t_{j_0}+}$. This, however, contradicts assertion (b) of Claim 11.8 for $j = j_0+1$, which finishes the proof. □

12. Inductive step: extension of the comparison map

12.1. Statement of the main result

In this section we consider a comparison domain defined on the time-interval $[0, t_{J+1}]$, as constructed in §11, and a comparison defined on the time-interval $[0, t_J]$. Our goal will be to extend the comparison to the time-interval $[0, t_{J+1}]$. The following proposition will be the main result of this section.

PROPOSITION 12.1. (Extending the comparison map by one step) *Suppose that*

$$\begin{gathered}
T > 0, \quad E \geqslant \underline{E}, \quad H \geqslant \underline{H}(E), \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}(E), \\
\nu \leqslant \bar{\nu}(T, E, H, \eta_{\text{lin}}), \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}(T, E, H, \eta_{\text{lin}}), \quad \lambda \leqslant \bar{\lambda}, \\
D_{\text{cap}} > 0, \quad \eta_{\text{cut}} \leqslant \bar{\eta}_{\text{cut}}, \quad D_{\text{cut}} \geqslant \underline{D}_{\text{cut}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}), \\
W \geqslant \underline{W}(E, \lambda, D_{\text{cut}}), \quad A \geqslant \underline{A}(E, \lambda, W), \quad \Lambda \geqslant \underline{\Lambda}(\lambda, A), \\
\delta_{\text{b}} \leqslant \bar{\delta}_{\text{b}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}, D_{\text{cut}}, A, \Lambda), \\
\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}, D_{\text{cut}}, W, A, \Lambda), \\
r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}(T, H, \lambda, D_{\text{cut}}),
\end{gathered} \tag{12.1}$$

and assume that

(i) $\mathcal{M}$ *and* M' *are two* $(\varepsilon_{\text{can}} r_{\text{comp}}, T)$*-complete Ricci flow spacetimes that each satisfy the* ε_{can}*-canonical neighborhood assumption at scales* $(\varepsilon_{\text{can}} r_{\text{comp}}, 1)$.

(ii) $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ *is a comparison domain in* $\mathcal{M}$*, which is defined over the time-interval* $[0, t_{J+1}]$. *We allow the case* $J = 0$.

(iii) *The tuple* $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ *is a comparison from* $\mathcal{M}$ *to* $\mathcal{M}'$ *defined on*

$$(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$$

over the time-interval $[0, t_J]$. *If* $J = 0$*, then this comparison is trivial, as explained in Definition* 7.2.

(iv) $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ *and* $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ *satisfy* (APA1)–(APA6) *for the parameters*

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}})$$

and a-priori assumptions (APA7)–(APA13) *for the parameters*

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}}).$$

(v) $t_{J+1} \leqslant T$.

(vi) *If* $J=0$, *then we assume in addition the existence of a map* $\zeta: X \to \mathcal{M}_0'$ *with the following properties. First,* $X \subset \mathcal{M}_0$ *is an open set that contains the* $\delta_{\mathrm{n}}^{-1} r_{\mathrm{comp}}$*-tubular neighborhood around* $\mathcal{N}_0$. *Second,* $\zeta: X \to \mathcal{M}_0'$ *is a diffeomorphism onto its image that satisfies the following bounds on* X:

$$\begin{aligned} |\zeta^* g_0' - g_0| &\leqslant \eta_{\mathrm{lin}}, \\ e^{HT} \rho_1^E |\zeta^* g_0' - g_0| &\leqslant \nu \overline{Q} = \nu \cdot 10^{-E-1} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E, \\ e^{HT} \rho_1^3 |\zeta^* g_0' - g_0| &\leqslant \nu \overline{Q}^* = \nu \cdot 10^{-1} \eta_{\mathrm{lin}} (\lambda r_{\mathrm{comp}})^3. \end{aligned}$$

Assume, moreover, that the $\varepsilon_{\mathrm{can}}$*-canonical neighborhood assumption holds at scales* $(0,1)$ *on the image* $\zeta(X)$.

Then, under the above assumptions, there are a set Cut^J *of pairwise disjoint disks in* $\mathcal{M}_{t_J}$, *a time-preserving diffeomorphism onto its image* $\phi^{J+1}: \mathcal{N}^{J+1} \to \mathcal{M}'$ *and a continuous map*

$$\bar{\phi}: \mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D} \longrightarrow \mathcal{M}'$$

such that the following holds.

The tuple

$$(\mathrm{Cut} \cup \mathrm{Cut}^J, \bar{\phi}, \{\phi^j\}_{j=1}^{J+1})$$

is a comparison from $\mathcal{M}$ *to* $\mathcal{M}'$ *defined on* $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ *over the time-interval* $[0, t_{J+1}]$. *This comparison and the corresponding domain still satisfy a-priori assumptions* (APA1)–(APA6) *for the parameters*

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}})$$

and a-priori assumptions (APA7)–(APA13) *for the parameters*

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}}).$$

Lastly, in the case $J=0$ *we have* $\phi_0^1 = \zeta|_{\mathcal{N}_0^1}$.

The proof of Proposition 12.1 is divided into three steps, which are of rather different character. These are presented in §§12.2– 12.4, respectively.

In the first step, we identify the set of disks Cut^J, and construct the initial map $\phi_{t_J}^{J+1}$, at time t_J, so that it is defined on the union of $\mathcal{N}_{t_J-}$ with the extension caps, and agrees with ϕ_{t_J-} away from the cuts in Cut^J. Here we use the Bryant extension proposition (Proposition 10.1).

In the second step, we promote this extended map to a map ϕ^{J+1} that is defined on a time-interval of the form $[t_J, t^*]$, for some $t^* \in (t_J, t_{J+1}]$, by solving the harmonic map heat flow equation. Unfortunately, at this point we cannot guarantee a priori that the harmonic map heat flow equation admits a solution on the *entire* time-interval $[t_J, t_{J+1}]$, as it may develop a singularity at an earlier time. However, we can rule out such a singularity as long as the solution satisfies certain uniform bounds. In such a case we can indeed choose $t^* = t_{J+1}$.

In the third step, we verify that the map ϕ^{J+1}, as constructed in the second step, satisfies a-priori assumptions (APA1)–(APA12). Our main focus will be on a-priori assumptions (APA6)–(APA9), as the remaining a-priori assumptions follow relatively easily from our construction. Once this is done, a-priori assumption (APA6) provides sufficient control on the map ϕ^{J+1} to rule out the development of a singularity up to time t^* and slightly after. It thus follows a posteriori that $t^* = t_{J+1}$, which finishes the proof.

Readers interested in a more detailed description of the steps above will find further explanations embedded in §§12.2–12.4.

This section is organized as follows. The intermediate results, Propositions 12.2 and 12.8, are presented in the next two subsections. In order to reduce complexity, we have organized the discussion in each of these subsections to be independent from the remaining subsections; no assumptions are implicitly carried over to from one subsection to the next. The last subsection (§12.4) contains the proof of the main proposition (Proposition 12.1). This proof is linked to §12.2 and §12.3 only via the intermediate results, Propositions 12.2 and 12.8, and does not depend on the details of their proofs.

As in §11, we introduce parameter bounds in displayed equations.

12.2. Extending the comparison over the extension caps

In this subsection, we consider a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$, which is defined on the time-interval $[0, t_{J+1}]$, and a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$, which is defined on the time-interval $[0, t_J]$. Based on this data, we will construct a collection of cuts Cut^J at time t_J and a map $\hat{\phi}: \mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+} \to \mathcal{M}'_{t_J}$, which can be seen as an extension of ϕ_{t_J-} away from the cuts. In Proposition 12.8, which is the main result of the next subsection,

the initial value ϕ_{t_J+} of the map ϕ^{J+1} will be taken to be the restriction of $\hat{\phi}$ to $\mathcal{N}_{t_J+}$. In the proof of this proposition, it will turn out to be necessary that $\hat{\phi}$ is defined on a slightly larger domain than ϕ_{t_J+} due to technical reasons having to do with our process for promoting ϕ_{t_J+} to later times $t>t_J$.

PROPOSITION 12.2. (Extending the comparison over the extension caps) *Suppose that*

$$\begin{gathered} E \geqslant \underline{E}, \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}, \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \quad \lambda \leqslant \bar{\lambda}, \\ D_{\text{cut}} \geqslant \underline{D}_{\text{cut}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}), \quad \Lambda \geqslant \underline{\Lambda}, \\ \delta_{\text{b}} \leqslant \bar{\delta}_{\text{b}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}, D_{\text{cut}}, A, \Lambda), \\ \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}, D_{\text{cut}}, A, \Lambda), \\ r_{\text{comp}} \leqslant \bar{r}_{\text{comp}}(T, H, \lambda, D_{\text{cut}}) \end{gathered} \tag{12.2}$$

and assume that assumptions (i)–(v) *of Proposition* 12.1 *hold and that* $J \geqslant 1$.

Then, there is a set of cuts Cut^J *at time* t_J, *i.e. a family of pairwise disjoint* 3*-disks in* $\text{Int}\,\mathcal{N}_{t_J+}$, *and a diffeomorphism onto its image* $\hat{\phi}: \mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+} \to \mathcal{M}'_{t_J}$ *such that the following hold*:

(a) *Each* $\mathcal{D} \in \text{Cut}^J$ *contains exactly one extension cap of the comparison domain*

$$(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$$

and each extension cap of this comparison domain that is in $\mathcal{M}_{t_J}$ *is contained in one* $\mathcal{D} \in \text{Cut}^J$.

(b) $\hat{\phi} = \phi_{t_J-}$ *on* $\mathcal{N}_{t_J-} \setminus \bigcup_{\mathcal{D} \in \text{Cut}^J} \mathcal{D}$.

(c) *Every cut* $\mathcal{D} \in \text{Cut}^J$ *has diameter* $< D_{\text{cut}} r_{\text{comp}}$ *and contains a* $\frac{1}{10} D_{\text{cut}} r_{\text{comp}}$*-neighborhood of the corresponding extension cap in* $\mathcal{D}$.

(d) *The associated perturbation* $\widehat{h} := \hat{\phi}^* g'_{t_J} - g_{t_J}$ *satisfies* $|\widehat{h}| \leqslant \eta_{\text{lin}}$ *on* $\mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+}$ *and*

$$e^{H(T-t_J)} \rho_1^3 |\widehat{h}| \leqslant \eta_{\text{cut}} \overline{Q}^*$$

on each $\mathcal{D} \in \text{Cut}^J$.

(e) *The* ε_{can}*-canonical neighborhood assumption holds at scales* $(0, 1)$ *on the image* $\hat{\phi}(\mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+})$.

The main idea of the proof of this proposition is to use the Bryant extension proposition (Proposition 10.1) in order to construct the cuts $\mathcal{D} \in \text{Cut}^J$ and the map $\hat{\phi}$ on each $\mathcal{D}$. The assumptions of that proposition hold due to a-priori assumptions (APA5) and (APA7): the former implies that regions in $\mathcal{M}$ that are close to extension caps, as well as the corresponding regions in $\mathcal{M}'$, are geometrically close to Bryant solitons; the latter gives the bound $Q \leqslant \overline{Q}$ near each extension cap.

While the strategy of proof can be summarized in a relatively straightforward way, there are several technical issues that we need to address. First, we need to argue that extension caps at time t_J are positioned close enough to a tip of an almost Bryant soliton region and that those regions are far enough away from one another to allow a separate construction of $\hat{\phi}$ in a large neighborhood of each extension cap. Second, we need to verify the condition under which a-priori assumption (APA7) guarantees the bound $Q \leqslant \overline{Q}$. Lastly, once the cuts $\mathcal{D}$ and the extensions have been constructed on each $\mathcal{D}$, we need to verify that the resulting map $\hat{\phi}$ satisfies all the desired properties, for example that it is a diffeomorphism onto its image.

Proof. In the following proof we will always assume, without further mention, that

$$\eta_{\text{lin}}, \lambda, r_{\text{comp}} < 10^{-2}, \tag{12.3}$$

and that

$$\delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}$$

is chosen small enough such that by a-priori assumption (APA3) (a) we have

$$0.9 r_{\text{comp}} < \rho_1 = \rho < 1.1 r_{\text{comp}} \quad \text{on } \partial \mathcal{N}_{t_J -}. \tag{12.4}$$

By definition of the comparison domain

$$(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1}),$$

we know that $\mathcal{N}_{t_J +} \setminus \operatorname{Int} \mathcal{N}_{t_J -}$ is a disjoint union of (possibly infinitely many) extension caps $\mathcal{C}_i$, $i \in I$, which are 3-disks. A-priori assumption (APA5) implies the existence of components $\mathcal{C}'_i$, $i \in I$, of $\mathcal{M}'_{t_J} \setminus \phi_{t_J -}(\operatorname{Int} \mathcal{N}_{t_J -})$ such that the following holds for all $i \in I$:

(1) $\mathcal{C}'_i$ is a 3-disk.

(2) $\phi_{t_J -}(\partial \mathcal{C}_i) = \partial \mathcal{C}'_i$.

(3) There is a diffeomorphism $\psi_i \colon M_{\text{Bry}}(\delta_{\text{b}}^{-1}) \to W_i \subset \mathcal{M}_{t_J}$ such that $\psi_i(x_{\text{Bry}}) \in \mathcal{C}_i$ and

$$\| (10 \lambda r_{\text{comp}})^{-2} \psi_i^* g_{t_J} - g_{\text{Bry}} \|_{C^{[\delta_{\text{b}}^{-1}]}(M_{\text{Bry}}(\delta_{\text{b}}^{-1}))} < \delta_{\text{b}}.$$

(4) There is a diffeomorphism $\psi'_i \colon M_{\text{Bry}}(\delta_{\text{b}}^{-1}) \to W'_i \subset \mathcal{M}'_{t_J}$ such that

$$d_{t_J}(\psi'_i(x_{\text{Bry}}), \mathcal{C}'_i) \leqslant D_{\text{cap}} r_{\text{comp}}$$

and

$$\| a_i^{-2} (\psi'_i)^* g'_{t_J} - g_{\text{Bry}} \|_{C^{[\delta_{\text{b}}^{-1}]}(M_{\text{Bry}}(\delta_{\text{b}}^{-1}))} < \delta_{\text{b}}$$

for some scale $a_i \in [D_{\text{cap}}^{-1} r_{\text{comp}}, D_{\text{cap}} r_{\text{comp}}]$.

(5) $\mathcal{C}_i$ and $\mathcal{C}'_i$ have diameter $\leqslant D_{\text{cap}} r_{\text{comp}}$.

Since $\phi_{t_J-}\colon \mathcal{N}_{t_J-} \to \phi(\mathcal{N}_{t_J-}) \subset \mathcal{M}'_{t_J}$ is a diffeomorphism onto its image, we obtain from items (1) and (2) that the components $\mathcal{C}'_i$, $i \in I$, are pairwise distinct.

We will assume in the following that

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}$$

is chosen sufficiently small such that for all $i \in I$.

(6) lengths of curves in $M_{\mathrm{Bry}}(\delta_{\mathrm{b}}^{-1})$ are distorted by ψ_i by a factor of at least $9\lambda r_{\mathrm{comp}}$ and at most $11\lambda r_{\mathrm{comp}}$.

We now fix a constant $D_{\#} < \infty$ whose value we will determine in the course of the proof. This constant controls the size of the neighborhood around each extension cap $\mathcal{C}_i$ in which we will carry out our construction of $\hat{\phi}$. More specifically, each such neighborhood will be of the form $\psi_i(M_{\mathrm{Bry}}(D_{\#})) \supset \mathcal{C}_i$; in particular, its diameter will be approximately $D_{\#} \cdot 10\lambda r_{\mathrm{comp}}$. Outside these neighborhoods, we will set $\hat{\phi} := \phi_{t_J-}$ and we will choose the cuts Cut^J to be disks that are contained in the corresponding $\psi_i(M_{\mathrm{Bry}}(D_{\#}))$.

As we proceed with the proof of Proposition 12.2, we will establish several claims, which hold under certain bounds on the parameters. At any point in the proof we will assume that the parameter bounds of the preceding claims hold, so that we can apply the assertions of these claims without restating the parameter bounds.

We first show that, under certain assumptions on our parameters, the neighborhoods $\psi_i(M_{\mathrm{Bry}}(D_{\#}))$ are pairwise disjoint and the extension caps $\mathcal{C}_i$ lie in bounded domains of the form $\psi_i(M_{\mathrm{Bry}}(D_0(\lambda)))$.

Claim 12.3. *There is a constant $D_0 = D_0(\lambda) < \infty$ such that, if*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \Lambda \geqslant \underline{\Lambda}, \quad D_{\#} \geqslant D_0(\lambda), \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, \Lambda, D_{\#}) \quad \textit{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(D_{\#}),$$

then $D_{\#} \leqslant \delta_{\mathrm{b}}^{-1}$ and the images $\psi_i(M_{\mathrm{Bry}}(D_{\#}))$, $i \in I$, are pairwise disjoint. Moreover, for all $i \in I$, we have

$$\mathcal{C}_i \subset \psi_i(M_{\mathrm{Bry}}(D_0)) \quad \textit{and} \quad \psi_i(M_{\mathrm{Bry}}(D_{\#})) \subset W_i \subset \mathcal{N}_{t_J-} \cup \mathcal{C}_i \tag{12.5}$$

and

$$9\lambda r_{\mathrm{comp}} \rho(x) < \rho(\psi_i(x)) = \rho_1(\psi_i(x)) < 11\lambda r_{\mathrm{comp}} \rho(x) \quad \textit{for all } x \in M_{\mathrm{Bry}}(D_{\#}). \tag{12.6}$$

Proof. Fix some index $i \in I$. The bound (12.6) follows immediately from (3), provided that

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(D_{\#}) \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(D_{\#}).$$

Next, we invoke the Bryant slice lemma (Lemma 8.16), for $X=\mathcal{N}_{t_J+}$, assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \Lambda \geqslant \underline{\Lambda} \quad \text{and} \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, \Lambda).$$

Assumptions (i)–(iv) of this lemma hold due to Definition 7.1 and a-priori assumptions (APA3) (a)–(c). The first inclusion in (12.5) is a restatement of assertion (d) of the Bryant slice lemma and the second string of inclusions is a consequence of assertion (a).

Finally, assume that

$$\psi_{i_1}(M_{\mathrm{Bry}}(D_{\#})) \cap \psi_{i_2}(M_{\mathrm{Bry}}(D_{\#})) \neq \varnothing$$

for some $i_1 \neq i_2$. Then, assuming

$$\delta_{\mathrm{b}} \leqslant 10^{-2} D_{\#}^{-1},$$

we must have $\mathcal{C}_{i_2} \subset \psi_{i_2}(M_{\mathrm{Bry}}(D_{\#})) \subset W_{i_1}$, contradicting the second string of inclusions of (12.5). This finishes the proof of the claim. □

In the second claim we show that the neighborhoods $\psi_i(M_{\mathrm{Bry}}(D_{\#}))$ around the extension caps $\mathcal{C}_i$ are mapped by ϕ_{t_J-} into the regions W_i', which are geometrically close to Bryant solitons.

Claim 12.4. *If*

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cap}}, D_{\#}),$$

then $D_{\#} < \delta_{\mathrm{b}}^{-1}$ *and* $\phi_{t_J-}(\psi_i(M_{\mathrm{Bry}}(D_{\#})) \setminus \operatorname{Int} \mathcal{C}_i) \subset W_i'$ *for all* $i \in I$.

Proof. Fix an index $i \in I$ and a point $y \in \psi_i(M_{\mathrm{Bry}}(D_{\#})) \setminus \operatorname{Int} \mathcal{C}_i$. By property (6) above, we can find a continuous path $\gamma\colon [0,1] \to \mathcal{N}_{t_J-}$ between y and a point $z \in \partial \mathcal{C}_i$ whose length is at most $11\lambda r_{\mathrm{comp}} \cdot 2D_{\#}$. Assuming (12.3), and using a-priori assumption (APA6), we find that the length of its image $\phi_{t_J-} \circ \gamma$ is bounded by $100 D_{\#} \lambda r_{\mathrm{comp}}$. So, since $\phi_{t_J-}(\partial \mathcal{C}_i) = \partial \mathcal{C}_i'$, we have

$$d_{t_J}(\phi_{t_J-}(y), \partial \mathcal{C}_i') \leqslant 100 D_{\#} \lambda r_{\mathrm{comp}}.$$

On the other hand, by properties (4) and (5) above, we have

$$\partial \mathcal{C}_i' \subset B(\psi_i'(x_{\mathrm{Bry}}), 2D_{\mathrm{cap}} r_{\mathrm{comp}}).$$

So, if

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cap}}, D_{\#}),$$

then we obtain that $y \in W_i'$, as desired. □

This concludes our discussion on the relative positions of the components $\mathcal{C}_i$ and $\mathcal{C}'_i$, and the images of the maps ψ_i and ψ'_i. We will now focus on the associated perturbation $h_{t_J-}=\phi^*_{t_J-}g'_{t_J}-g_{t_J}$. In the next claim, and its proof, we use the bound $Q\leqslant\overline{Q}$, as asserted by a-priori assumption (APA7), to deduce a bound on the weighted norm $\rho^E|h_{t_J-}|$ on $\psi_i(M_{\mathrm{Bry}}(D_\#))\backslash\mathcal{C}_i$. Using a standard local derivative estimate, we will also deduce similar weighted bounds on covariant derivatives of the form $\nabla^m h_{t_J-}$.

Claim 12.5. *There is a constant $C=C(E)<\infty$ such that, if*

$$\eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}},\quad \delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda},\quad \Lambda\geqslant\underline{\Lambda}\quad \delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda,D_{\mathrm{cut}},A,\Lambda,D_\#),$$
$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,D_{\mathrm{cut}},A,\Lambda,D_\#),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

then, for the associated perturbation $h_{t_J-}=\phi^*_{t_J-}g'_{t_J}-g_{t_J}$, the following holds for all $i\in I$ and all $m=0,1,...,4$:

$$e^{H(T-t_J)}\rho^E|\nabla^m h_{t_J-}|\leqslant C\lambda^{-m}r_{\mathrm{comp}}^{-m+E}\quad \text{on } \psi_i(M_{\mathrm{Bry}}(D_0+1,D_\#-1)). \tag{12.7}$$

As mentioned earlier, the main idea of the proof of this claim is to invoke the bound $Q\leqslant\overline{Q}$ from a-priori assumption (APA7). However, this bound is predicated on the remoteness of cuts. In order to verify this remoteness, we will invoke Lemma 8.19.

Proof. Fix an index $i\in I$ and a point $x\in\psi_i(M_{\mathrm{Bry}}(D_0+1,D_\#-1))$ for the remainder of this proof. Then, by Claim 12.3, (12.3) and property (6) above, we have

$$B(x,\lambda r_{\mathrm{comp}})\subset\psi_i(M_{\mathrm{Bry}}(D_0,D_\#))\subset\mathcal{N}_{t_J-}. \tag{12.8}$$

So, for the corresponding parabolic neighborhood we have

$$P(x,\lambda r_{\mathrm{comp}})\subset\mathcal{N}^J_{(t_{J-1},t_J]}\subset\mathcal{N}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}}\mathcal{D},$$

since $\mathcal{D}\subset\mathcal{M}_{[0,t_{J-1}]}$ for all $\mathcal{D}\in\mathrm{Cut}$.

Our goal will be to use a-priori assumption (APA7) to deduce the bound $Q\leqslant\overline{Q}$ on $P(x,\lambda r_{\mathrm{comp}})$. So, consider a point $y\in P(x,\lambda r_{\mathrm{comp}})$ and set $t':=\mathfrak{t}(y)$. We now claim that, for an appropriate choice of constants, we have

$$P(y,A\rho_1(y))\cap\mathcal{D}=\varnothing\quad\text{for all } \mathcal{D}\in\mathrm{Cut}. \tag{12.9}$$

To see this, choose a point $z\in\partial\mathcal{C}_i\subset\partial\mathcal{N}_{t_J-}$ nearest to $y(t_J)$. Then, by (12.8) and properties (3) and (6) above,

$$d_{t_J}(y(t_J),z)<11D_\#\lambda r_{\mathrm{comp}}. \tag{12.10}$$

Let $z' := z(t')$. Since $t' \in (t_{J-1}, t_J]$, we can use the curvature bound on the product domain $\mathcal{N}^J$ from a-priori assumption (APA2) to derive a distortion estimate of the minimizing geodesic between $y(t_J)$ and z over the time-interval $[t', t_J]$. Since $t_J - t' \leqslant (\lambda r_{\mathrm{comp}})^2$, we obtain that, for some universal constant $C_1' < \infty$,

$$d_{t'}(y, z') < 11 e^{C_1'} D_{\#} \lambda r_{\mathrm{comp}}. \tag{12.11}$$

Next, let us apply bounded curvature at bounded distance, Lemma 8.6, at z, along with (12.10), while assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(D_{\#}).$$

We obtain a constant $C_2' = C_2'(D_{\#}) < \infty$ such that, by (12.4),

$$C_2' \rho_1(y) \geqslant \rho_1(z) \geqslant 0.9 r_{\mathrm{comp}}.$$

Combining this with (12.11), yields that

$$d_{t'}(y, z') < D' \rho_1(y)$$

for some $D' = D'(D_{\#}) < \infty$. So, if $t' < t_J$, then $B(y, D'\rho_1(y)) \not\subset \mathcal{N}_{t'-}$. We can now apply Lemma 8.19 (boundaries and cuts are far apart), along with a-priori assumption (APA11), assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A, \Lambda, D'(\lambda, D_{\#})),$$
$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A, \Lambda, D'(\lambda, D_{\#})), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

and obtain (12.9). The case $t' = t_J$ follows from the case $t' < t_J$ by continuity.

Using (12.9) and a-priori assumption (APA7), we can now deduce that

$$e^{H(T-t_J)} \rho_1^E(y) |h_{t'-}(y)| \leqslant Q(y) \leqslant \overline{Q} = 10^{-E-1} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E. \tag{12.12}$$

Next, we apply bounded curvature at bounded distance, Lemma 8.6, at x, along with a-priori assumption (APA2), while assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda).$$

We obtain that there is a universal constant $C_3' < \infty$ such that

$$\rho(x) = \rho_1(x) \leqslant C_3' \rho_1(y). \tag{12.13}$$

The equality statement follows from (12.6). Combining (12.12) with (12.13) yields

$$e^{H(T-t_J)} \rho^E(x) |h_{t_J-}(y)| \leqslant 10^{-E-1} {C_3'}^E \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E. \tag{12.14}$$

If $y=x$, then this bound implies (12.7) for $m=0$. The bounds on the higher derivatives follow from (12.14) using (12.13), a-priori assumption (APA6), Shi's estimates and standard local gradient estimates for the Ricci–DeTurck flow (see also Lemma A.2), assuming

$$\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}.$$

This finishes the proof. □

We will now apply the Bryant extension proposition (Proposition 10.1) to the restrictions of the map ϕ_{t_J-} to each W_i, for suitably chosen $D_\#$. The resulting maps, which will be denoted by $\tilde{\phi}_i$, will be only defined on the domains $\psi_i(M_{\text{Bry}}(D_\#))$, but will be equal to ϕ_{t_J-} near the boundaries of these domains.

Claim 12.6. *If*

$$E \geqslant \underline{E}, \quad \eta_{\text{cut}} \leqslant \bar{\eta}_{\text{cut}}, \quad D_\# \geqslant \underline{D}_\#(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}),$$
$$\delta_{\text{b}} \leqslant \bar{\delta}_{\text{b}}(T, E, H, \eta_{\text{lin}}, \lambda, D_{\text{cap}}, \eta_{\text{cut}}, D_\#),$$

then, for each $i\in I$, there is a diffeomorphism onto its image

$$\tilde{\phi}_i \colon \psi_i(M_{\text{Bry}}(D_\#-1)) \longrightarrow W_i'$$

and a 3-disk

$$\mathcal{D}_i := \psi_i(\overline{M_{\text{Bry}}(D_\#-2)}) \subset \mathcal{M}_{t_J}$$

such that the following holds:

(a) $\mathcal{C}_i \subset \operatorname{Int} \mathcal{D}_i$.

(b) $\tilde{\phi}_i = \phi_{t_J-}$ *on* $\psi_i(M_{\text{Bry}}(D_\#-1)) \setminus \operatorname{Int} \mathcal{D}_i$.

(c) *The perturbation* $\tilde{h}_i := \tilde{\phi}_i^* g'_{t_J} - g_{t_J}$ *satisfies the following bounds on* $\mathcal{D}_i$:

$$|\tilde{h}_i| \leqslant \eta_{\text{lin}} \quad \textit{and} \quad e^{H(T-t_J)} \rho_1^3 |\tilde{h}_i| \leqslant \eta_{\text{cut}} \overline{Q}^* = \eta_{\text{cut}} \cdot 10^{-1} \eta_{\text{lin}} (\lambda r_{\text{comp}})^3.$$

(d) $\tilde{\phi}_i(\mathcal{D}_i) = \phi_{t_J-}(\mathcal{D}_i \cap \mathcal{N}_{t_J-}) \cup \operatorname{Int} \mathcal{C}_i'$.

(e) $\mathcal{D}_i$ *contains the* $8\lambda D_\# r_{\text{comp}}$*-tubular neighborhood around* $\mathcal{C}_i$.

Proof. Fix some $i\in I$. Set $b_i := a_i (10\lambda r_{\text{comp}})^{-1}$ and notice that property (4) from above gives

$$b_i \in [(10\lambda)^{-1} D_{\text{cap}}^{-1}, (10\lambda)^{-1} D_{\text{cap}}].$$

Assume that

$$D_\# \geqslant 2(D_0(\lambda)+1)+1 \tag{12.15}$$

and consider the map

$$\phi_i^\circ := {\psi_i'}^{-1} \circ \phi_{t_J-} \circ \psi_i \colon M_{\mathrm{Bry}}\big(\tfrac{1}{2}(D_\#-1), D_\#-1\big) \longrightarrow M_{\mathrm{Bry}}(\delta_{\mathrm{b}}^{-1}),$$

which is well defined by Claims 12.3 and 12.4. Let

$$g_i^\circ := (10\lambda r_{\mathrm{comp}})^{-2}\psi_i^* g_{t_J} \quad \text{and} \quad g_i'^\circ := (10\lambda r_{\mathrm{comp}})^{-2}\psi_i'^* g'_{t_J}$$

be the pull-back metrics on W_i and W_i' to $M_{\mathrm{Bry}}(D_\#)$. Notice that these pull-backs are close to g_{Bry} and $b_i^2 g_{\mathrm{Bry}}$, respectively, by properties (3) and (4) above. Rescaling (12.7) from Claim 12.5 by $(10\lambda r_{\mathrm{comp}})^{-1}$ yields, for $h_i^\circ := (\phi_i^\circ)^* g_i'^\circ - g_i^\circ$,

$$\rho^E |\nabla^m h_i^\circ|_{g_i^\circ} \leqslant C(E) e^{-H(T-t_J)} (10\lambda)^{-E} \cdot 10^m \leqslant C(E)(10\lambda)^{-E} \cdot 10^4,$$

for all $m = 0, 1, \ldots, 4$. Here, we have used $t_J \leqslant T$. Note that ρ is taken with respect to g_i°.

We now apply the Bryant extension proposition (Proposition 10.1) with $D = D_\# - 1$, $b = C(E)(10\lambda)^{-E}\cdot 10^4$, $\beta = e^{-H(T-t_J)}\eta_{\mathrm{cut}} 10^{-4} \cdot \eta_{\mathrm{lin}} \cdot b^{-1}$, $C = \max\{(10\lambda)^{-1} D_{\mathrm{cap}}, b\}$, $\phi = \phi_i^\circ$, $g = g_i^\circ$, $g' = g_i'^\circ$. We obtain that, if

$$E \geqslant \underline{E}, \quad D_\# \geqslant \underline{D}_\#(T, E, H, \eta_{\mathrm{lin}}, \lambda, D_{\mathrm{cap}}, \eta_{\mathrm{cut}}),$$
$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(T, E, H, \eta_{\mathrm{lin}}, \lambda, D_{\mathrm{cap}}, \eta_{\mathrm{cut}}, D_\#),$$

then there is a smooth map $\tilde{\phi}_i^\circ \colon M_{\mathrm{Bry}}(D_\# - 1) \to M_{\mathrm{Bry}}(\delta_{\mathrm{b}}^{-1})$, with

$$\tilde{\phi}_i^\circ = \phi_i^\circ \quad \text{on } M_{\mathrm{Bry}}(D_\# - 2, D_\# - 1), \tag{12.16}$$

such that, for $\tilde{h}_i^\circ := (\tilde{\phi}_i^\circ)^* g_i'^\circ - \psi'^* g_i^\circ$, we have

$$\rho^3 |\tilde{h}_i^\circ|_{g_i^\circ} \leqslant \eta_{\mathrm{cut}} \cdot e^{-H(T-t_J)} \cdot 10^{-4} \eta_{\mathrm{lin}}. \tag{12.17}$$

Now, set $\tilde{\phi}_i := \psi_i' \circ \tilde{\phi}_i^\circ \circ \psi_i^{-1}$. Then, assertion (b) holds due to (12.16). Rescaling (12.17) by $10\lambda r_{\mathrm{comp}}$ implies the second bound in assertion (c). The first bound in assertion (c) follows from the second assuming

$$\eta_{\mathrm{cut}} \leqslant \bar{\eta}_{\mathrm{cut}} \quad \text{and} \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}.$$

Assertion (a) follows from Claim 12.3 and (12.15).

To see assertion (d) observe first that, by assertion (b) and (12.5), from Claim 12.3 we have

$$\partial(\tilde{\phi}_i(\mathcal{D}_i)) = \tilde{\phi}_i(\partial\mathcal{D}_i) = \phi_{t_J-}(\partial\mathcal{D}_i) = \partial(\phi_{t_J-}(\mathcal{D}_i \cap \mathcal{N}_{t_J-}) \cup \operatorname{Int} \mathcal{C}_i').$$

So, the smooth domains on both sides of the equation in assertion (d) share the same boundary and by assertion (b) these domains lie on the same side of this boundary. So, they have to agree.

Assertion (e) follows for

$$D_{\#} \geqslant \underline{D}_{\#}(\lambda) \quad \text{and} \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda)$$

from (12.5) in Claim 12.3 and property (6) from above. □

Next, we combine the maps $\tilde{\phi}_i$ and $\phi_{t_{J^-}}$ to a map $\hat{\phi}\colon \mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}} \to \mathcal{M}'_{t_J}$. To do this, recall that by Claim 12.3, the subsets $\psi_i(M_{\mathrm{Bry}}(D_{\#}))$, $i \in I$, are pairwise disjoint. So, the 3-disks $\mathcal{D}_i$, $i \in I$, are pairwise disjoint as well. Moreover, recall that, by Claim 12.3 and Claim 12.6 (a), we have

$$\mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}} = \mathcal{N}_{t_{J^-}} \bigcup_{i \in I} \operatorname{Int} \mathcal{C}_i = \mathcal{N}_{t_{J^-}} \bigcup_{i \in I} \mathcal{D}_i.$$

Therefore, we can define $\hat{\phi}\colon \mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}} \to \mathcal{M}'_{t_J}$ as follows:

$$\hat{\phi} := \begin{cases} \tilde{\phi}_i, & \text{on each } \mathcal{D}_i,\ i \in I, \\ \phi_{t_{J^-}}, & \text{on } \mathcal{N}_{t_{J^-}} \setminus \bigcup_{i \in I} \mathcal{D}_i. \end{cases}$$

Claim 12.7. *$\hat{\phi}$ is a diffeomorphism onto its image.*

Proof. By assertions (a) and (b) of Claim 12.6, we know that $\hat{\phi}$ is smooth and has non-degenerate differential. Next, we argue that $\hat{\phi}$ is injective. To see this, observe that the maps $\tilde{\phi}_i$, $i \in I$, and $\phi_{t_{J^-}}$ are each injective. So, it suffices to show that the images $\tilde{\phi}_i(\mathcal{D}_i)$, $i \in I$, and $\phi_{t_{J^-}}(\mathcal{N}_{t_{J^-}} \setminus \bigcup_{i \in I} \mathcal{D}_i)$ are pairwise disjoint. Using Claim 12.6 (d) and the fact that the 3-disks $\mathcal{D}_i$, as well as the 3-disks $\mathcal{C}'_i$, $i \in I$, are pairwise disjoint, it follows immediately that the images $\tilde{\phi}_i(\mathcal{D}_i)$, $i \in I$, are pairwise disjoint. Similarly, using Claim 12.6 (d), we have, for all $i \in I$,

$$\tilde{\phi}_i(\mathcal{D}_i) \cap \phi_{t_{J^-}}\Big(\mathcal{N}_{t_{J^-}} \setminus \bigcup_{i \in I} \mathcal{D}_i\Big) = (\phi_{t_{J^-}}(\mathcal{D}_i \cap \mathcal{N}_{t_{J^-}}) \cup \operatorname{Int} \mathcal{C}'_i) \cap \phi_{t_{J^-}}\Big(\mathcal{N}_{t_{j^-}} \setminus \bigcup_{i \in I} \mathcal{D}_i\Big) = \varnothing,$$

as desired.

So, $\hat{\phi}$ is an injective smooth map with non-degenerate differential. In order to see that $\hat{\phi}$ is even a diffeomorphism onto its image, it suffices to show that

$$\hat{\phi}^{-1}\colon \operatorname{Im} \hat{\phi} \longrightarrow \mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}}$$

is continuous, i.e. for any sequence $x_k \in \mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}}$ and any point $x_\infty \in \mathcal{N}_{t_{J^-}} \cup \mathcal{N}_{t_{J^+}}$, if

$$\lim_{k \to \infty} \hat{\phi}(x_k) = \hat{\phi}(x_\infty),$$

then $\lim_{k\to\infty} x_k = x_\infty$ itself. This can be seen as follows: If x_∞ lies in the interior of $\mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+}$, then we are done by the inverse function theorem and the fact that $\hat{\phi}$ is injective and has non-degenerate differential. So, assume that

$$x_\infty \in \partial(\mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+}) = \partial\mathcal{N}_{t_J-} \setminus \bigcup_{i\in I} \partial\mathcal{C}_i = \partial\mathcal{N}_{t_J-} \setminus \bigcup_{i\in I} W_i. \tag{12.18}$$

The first equality follows from Definition 7.1 (3) and the last equality follows from (12.5) in Claim 12.3. If for some k we have $x_k \in \mathcal{D}_{i_k}$ for some $i_k \in I$, then, by Claim 12.6 (d), by the construction of $\mathcal{D}_{i_k}$ and by a-priori assumption (APA6) and (12.3), a ball of uniform radius around $\hat{\phi}(x_k)$ must still be contained in

$$\hat{\phi}(\psi_{i_k}(M_{\mathrm{Bry}}(D_\#))) \subset \hat{\phi}(W_{i_k}).$$

Therefore, by (12.18), the distance $d_{t_J}(\hat{\phi}(x_\infty), \hat{\phi}(x_k))$ must be bounded from below by a uniform constant. It follows that, for large k, we have $x_k \in \mathcal{N}_{t_J-} \setminus \bigcup_{i\in I} \mathcal{D}_i$, and thus $\hat{\phi}(x_k) = \phi_{t_J-}(x_k)$ by Claim 12.6 (b). Since ϕ_{t_J-} is a diffeomorphism onto its image, we must have $\lim_{k\to\infty} x_k = x_\infty$, which proves our claim. □

Now let

$$\mathrm{Cut}^J := \{\mathcal{D}_i : i \in I\}.$$

Then, assertion (a) of this proposition holds due to Claim 12.6 (a). Assertion (b) holds by Claim 12.6 (b) and by the construction of $\hat{\phi}$. Assertion (d) of the proposition follows from Claim 12.6 (c) and a-priori assumption (APA6). For assertion (e) recall that by Claim 12.6 (d) we have

$$\hat{\phi}(\mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+}) = \phi_{t_J-}(\mathcal{N}_{t_J-}) \bigcup_{i\in I} \mathcal{C}_i'$$

and $\mathcal{C}_i' \subset W_i'$ for all $i \in I$. By a-priori assumption (APA6) we know that the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds at scales $(0,1)$ on $\phi_{t_J-}(\mathcal{N}_{t_J-})$ and by property (4) above we have $\rho > \frac{1}{2} D_{\mathrm{cap}}^{-1} r_{\mathrm{comp}} > \varepsilon_{\mathrm{can}} r_{\mathrm{comp}}$ on W_i' for all $i \in I$, assuming

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}} \quad \text{and} \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(D_{\mathrm{cap}}).$$

Therefore, the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds at scales $(0,1)$ on W_i' as well, which implies assertion (e).

Lastly, we argue that assertion (c) holds if we choose

$$D_{\mathrm{cut}} = 22\lambda D_\#. \tag{12.19}$$

Fix some $i \in I$. By property (6) from the beginning of this proof, we have

$$\operatorname{diam} \mathcal{D}_i < 2 \cdot 11 \lambda D_{\#} r_{\mathrm{comp}} = D_{\mathrm{cut}} r_{\mathrm{comp}}.$$

On the other hand, Claim 12.6 (e) states that $\mathcal{D}_i$ contains a $8\lambda D_{\#} r_{\mathrm{comp}}$-tubular neighborhood around $\mathcal{C}_i$ and

$$8\lambda D_{\#} \geqslant \tfrac{1}{10} 22 \lambda D_{\#} = \tfrac{1}{10} D_{\mathrm{cut}}.$$

Lastly, let us review the choice of parameters. In the course of the proof, we have introduced the auxiliary parameter $D_{\#}$, which is related to λ and D_{cut} via (12.19). Once λ has been fixed, any lower bound on $D_{\#}$ implies a lower bound on D_{cut}, as indicated in (12.2). After fixing D_{cut}, the auxiliary parameter $D_{\#}$ can be viewed as a constant of the form $D_{\#}(\lambda, D_{\mathrm{cut}})$. This constant influences the choices of δ_{b}, $\varepsilon_{\mathrm{can}}$ and r_{comp}. So, these parameters are bounded in terms of λ and D_{cut}, as shown in (12.2).

This completes the proof of Proposition 12.2. □

12.3. Extending the comparison map past time t_J

The goal of this subsection is to evolve the map $\hat{\phi}$, as constructed in Proposition 12.2, forward in time by the harmonic map heat flow. More specifically, we consider again a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=1}^{J+1})$, defined over the time-interval $[0, t_{J+1}]$, and a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ from $\mathcal{M}$ to $\mathcal{M}'$ defined over the time-interval $[0, t_J]$. We moreover consider the map $\hat{\phi}: \mathcal{N}_{t_J -} \cup \mathcal{N}_{t_J +} \to \mathcal{M}'$ from Proposition 12.2. We will then promote the map $\hat{\phi}|_{\mathcal{N}_{t_J +}}$ to a map $\phi^{J+1}: \mathcal{N}^{J+1}_{[t_J, t^*]} \to \mathcal{M}'$, which is defined on a time-interval of the form $[t_J, t^*]$, where $t^* \in (t_J, t_{J+1}]$. In this subsection we will not be able to guarantee that $t^* = t_{J+1}$ — in fact t^* may be quite close to t_J — since we will only solve the harmonic map heat flow until $|h|$ reaches a certain threshold. However, we will find that if $|h|$ does not reach this threshold on the time-interval $[t_J, t^*]$, then in fact $t^* = t_{J+1}$. In the next subsection, we will then deduce various bounds on $|h|$, which will imply that $|h|$ stays below this threshold. Hence, it will follow that $t^* = t_{J+1}$ and so ϕ^{J+1} can indeed be used to extend the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ to the time-interval $[0, t_{J+1}]$.

In the course of our construction, we will also discuss the case $J = 0$, i.e. the case in which ϕ^{J+1} is the comparison map in the first time-step. In this case, the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ is empty to start with and Proposition 12.2 does not apply. Instead, we will assume in this case that $\hat{\phi}$ is the initial map ζ, as introduced in the assumptions of Proposition 12.1.

Let us now state our main result of this subsection.

PROPOSITION 12.8. (Extending the comparison map until we lose control) *If*

$$\begin{gathered} E>2, \quad F>0, \quad H\geqslant \underline{H}(E), \quad \eta_{\mathrm{lin}}\leqslant \bar{\eta}_{\mathrm{lin}}(E), \\ \nu\leqslant \bar{\nu}(T,E,F,H,\eta_{\mathrm{lin}}), \quad \delta_{\mathrm{n}}\leqslant \bar{\delta}_{\mathrm{n}}(T,E,F,H,\eta_{\mathrm{lin}}), \quad \lambda\leqslant\bar{\lambda}, \\ \delta_{\mathrm{b}}\leqslant \bar{\delta}_{\mathrm{b}}(T,E,F,H,\eta_{\mathrm{lin}},\lambda,D_{\mathrm{cut}},A,\Lambda), \\ \varepsilon_{\mathrm{can}}\leqslant \bar{\varepsilon}_{\mathrm{can}}(T,E,F,H,\eta_{\mathrm{lin}},\lambda,D_{\mathrm{cut}},A,\Lambda), \quad r_{\mathrm{comp}}\leqslant \bar{r}_{\mathrm{comp}}, \end{gathered} \tag{12.20}$$

then the following holds.

Assume that assumptions (i)–(vi) *of Proposition* 12.1 *hold.*

Recall that, in the case $J=0$, *assumption* (vi) *imposes the existence of a domain* $X\subset\mathcal{M}_{t_J}$ *and map* $\zeta\colon X\to\mathcal{M}'_{t_J}$ *with certain properties. In this case, we set* $\hat{\phi}:=\zeta$.

In the case $J\geqslant 1$, *we set* $X:=\mathcal{N}_{t_J-}\cup\mathcal{N}_{t_J+}$ *and consider the set* Cut^J *and the map* $\hat{\phi}\colon X\to\mathcal{M}'_{t_J}$ *satisfying all assertions of Proposition* 12.2.

Then, there is some time $t^*\in(t_J,t_{J+1}]$ *and a smooth, time-preserving diffeomorphism onto its image*

$$\phi^{J+1}\colon \mathcal{N}^{J+1}_{[t_J,t^*]}\longrightarrow\mathcal{M}',$$

with $\phi^{J+1}_{t_J}=\widehat{\phi}|_{\mathcal{N}_{t_J+}}$, *whose inverse*

$$(\phi^{J+1})^{-1}\colon \phi^{J+1}(\mathcal{N}^{J+1}_{[t_J,t^*]})\longrightarrow\mathcal{N}^{J+1}_{[t_J,t^*]}$$

evolves by harmonic map heat flow (*see Definition* 6.13) *and such that the following holds for the associated perturbation* $h^{J+1}:=(\phi^{J+1})^*g'-g$ (*which is a Ricci–DeTurck flow*):

(a) $|h^{J+1}|\leqslant 10\eta_{\mathrm{lin}}$ *on* $\mathcal{N}^{J+1}_{[t_J,t^*]}$.

(b) *For any* $t\in[t_J,t^*]$ *and* $x\in\mathcal{N}^{J+1}_t$ *whose time-*t *distance to* $\partial\mathcal{N}^{J+1}_t$ *is smaller than* Fr_{comp}, *we have*

$$Q_+(x)=e^{H(T-\mathfrak{t}(x))}\rho_1^E(x)|h^{J+1}(x)|<\overline{Q}=10^{-E-1}\eta_{\mathrm{lin}}r^E_{\mathrm{comp}}.$$

(c) *If even* $|h^{J+1}|\leqslant\eta_{\mathrm{lin}}$ *on* $\mathcal{N}^{J+1}_{t^*}$, *then* $t^*=t_{J+1}$.

(d) $\phi^{J+1}(\mathcal{N}^{J+1}_{[t_J,t^*]})$ *is* $\varepsilon_{\mathrm{can}}r_{\mathrm{comp}}$*-thick.*

We emphasize that we have introduced another auxiliary parameter, F, which we will choose in §12.4, depending only on E. The bound in assertion (b), which holds Fr_{comp}-close to the boundary of $\partial\mathcal{N}^{J+1}$, will be helpful later, as we are not able to apply the semi-local maximum principle, Proposition 9.1, too close to the boundary. For this purpose, we will later choose $F\geqslant L(E)$, where the latter is the constant from Proposition 9.1.

Let us now explain the main strategy of the proof of Proposition 12.8. Observe first that the parabolic domain $\mathcal{N}^{J+1}\subset\mathcal{M}$ is a product domain and the Ricci flow on it can

be viewed as a conventional, non-singular Ricci flow. A similar domain, which contains the image $\hat{\phi}(\mathcal{N}_{t_J+})$, can be found in $\mathcal{M}'$. So, the proof of Proposition 12.8 can be reduced to a relatively standard short-time and long-time existence statement for the harmonic map heat flow between conventional Ricci flows on manifolds with boundary. Rather than solving the harmonic map heat flow equation with a boundary condition, we found it technically simpler to use a "grafting" construction to eliminate the boundary.

A large part of the following proof will be devoted to the characterization of the geometry near the boundary of $\mathcal{N}_{t_J+}$ and the boundary of its image $\hat{\phi}(\mathcal{N}_{t_J+})$, which will serve as a setup for the subsequent grafting construction. More specifically, our goal will be to show that the boundary of $\mathcal{N}_{t_J+}$ and its image $\hat{\phi}(\mathcal{N}_{t_J+})$ are contained in regions that look sufficiently neck-like on the time-interval $[t_J, t_{J+1}]$. To achieve this, we will employ the following strategy. A-priori assumption (APA3) (a) provides neck structures near $\partial\mathcal{N}_{t_{J+1}-}$ at time t_{J+1}. Using Lemma 8.12, these neck structures can be promoted backwards onto the time-interval $[t_J, t_{J+1}]$. The newly constructed neck structure at time t_J, near $\partial\mathcal{N}_{t_J+}$, a-priori assumption (APA7) and the interior decay estimate, Proposition 9.2, can then be used to identify C^0-neck structures near the boundary of $\hat{\phi}(\mathcal{N}_{t_J+}) \subset \mathcal{M}'_{t_J}$. Using the canonical neighborhood assumption and the self-improving property of necks in κ-solutions, Lemma 8.10, these C^0-neck structures imply the existence of neck structures of higher regularity in $\mathcal{M}'_{t_J}$. Lastly, we use Lemma 8.12, to promote these neck structures forward in $\mathcal{M}'$, onto the time-interval $[t_J, t_{J+1}]$.

Based on this characterization of the boundary of $\mathcal{N}_{t_J+}$ and its image, we perform a grafting construction in the last phase of the proof. This grafting construction involves cutting $\mathcal{M}_{[t_J,t_{J+1}]}$ and $\mathcal{M}'_{[t_J,t_{J+1}]}$ inside the previously identified neck regions, gluing on shrinking round half-cylinders, and passing to a map between the grafted spacetimes. We have thus reduced our discussion to standard existence results for the harmonic map heat flow between complete manifolds. We remark that our approach is facilitated by the fact that $\hat{\phi}$ is already defined on a larger neighborhood of $\mathcal{N}_{t_J+}$, therefore providing enough space for an interpolation between the metric on $\mathcal{M}_{[t_J,t_{J+1}]}$ and the cylindrical metric.

Proof of Proposition 12.8. Let $\delta_\# > 0$ be a constant whose value we will determine at the end of the proof. It will only depend on T, E, H and η_{lin}, and influence only the parameters ν, δ_{n}, δ_{b} and ε_{can}. So, it lies between η_{lin} and ν in the parameter order introduced in §7.5. To avoid an accumulation of a large number of different constants, in what follows we will be using the standard practice of making a series of adjustments to the constant $\delta_\#$. This means, strictly speaking, that $\delta_\#$ is not really a single constant, but takes on different values at different places in the proof, and the earlier values are adjusted as functions of the later values.

By a-priori assumption (APA3) (a), each boundary component $\Sigma\subset\partial\mathcal{N}^{J+1}_{t_{J+1}}$ is the central 2-sphere of a $\delta_{\rm n}$-neck $U_\Sigma\subset\mathcal{M}_{t_{J+1}}$ at scale $r_{\rm comp}$. Lemma 8.12 implies that, if

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n}(\delta_\#),\quad \varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}(\delta_\#)\quad\text{and}\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

then, for each such Σ, there is a product domain $U^*_\Sigma\subset\mathcal{M}_{[t_J,t_{J+1}]}$ that contains Σ and on which the flow is $\delta_\#$-close at scale $r_{\rm comp}$ to the round shrinking cylinder on the time-interval $[-1,0]$. By this we mean the following: we can find an $r^2_{\rm comp}$-time-equivariant and $\partial_{\rm t}$-preserving diffeomorphism

$$\psi_\Sigma\colon S^2\times(-\delta_\#^{-1},\delta_\#^{-1})\times[-1,0]\longrightarrow U^*_\Sigma$$

such that $\Sigma=\psi_\Sigma(S^2\times\{0\}\times\{0\})$ and

$$\|r^{-2}_{\rm comp}\psi^*_\Sigma g-g^{S^2\times\mathbb{R}}\|_{C^{[\delta_\#^{-1}]}}<\delta_\#. \tag{12.21}$$

Here, $g^{S^2\times\mathbb{R}}$ denotes the metric of the standard round shrinking cylinder spacetime and the norm is taken over the domain of ψ_Σ.

By (12.21) and assuming

$$\delta_\#\leqslant\bar\delta_\#\quad\text{and}\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

we have

$$1.9r_{\rm comp}<\rho_1=\rho<2.1r_{\rm comp}\quad\text{on } U^*_{\Sigma,t_J}. \tag{12.22}$$

So, by a-priori assumption (APA3) (a), applied at time t_J, and assuming

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},$$

we find that U^*_{Σ,t_J} is disjoint from $\partial\mathcal{N}^J_{t_J}$ if $J\geqslant1$. So, if $J\geqslant1$, since

$$\Sigma(t_J)\subset\partial\mathcal{N}^{J+1}_{t_J}\subset\mathcal{N}_{t_J-},$$

it follows that $U^*_{\Sigma,t_J}\subset\mathcal{N}_{t_J-}\subset X$.

On the other hand, if $J=0$, and

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},$$

then U^*_{Σ,t_J} has diameter $<10\delta_\#^{-1}r_{\rm comp}$. So, assuming

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n}(\delta_\#),$$

we have $U^*_{\Sigma,t_J} \subset X$. So, in summary,

$$\begin{cases} U^*_{\Sigma,t_J} \subset \mathcal{N}_{t_J-} \subset X, & \text{if } J \geqslant 1, \\ U^*_{\Sigma,t_J} \subset X, & \text{if } J = 0. \end{cases} \tag{12.23}$$

Consider the Ricci–DeTurck perturbation $(h, \{h^j\}_{j=1}^J)$ associated with the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^J)$, and let Q be defined as in Definition 7.5 of the a-priori assumptions (APA7)–(APA13). We will now use a-priori assumption (APA7) to show that we have a bound on Q in large parabolic neighborhoods near the boundary of $\mathcal{N}_{t_J+}$. In Claim 12.10 this bound will later be used to obtain an improved bound on Q, and therefore on h, via the interior decay estimate, Proposition 9.2. For this purpose, let $A_\# < \infty$ be a constant whose value will be determined in the proof of Claim 12.10 (depending only on E and $\delta_\#$).

Claim 12.9. *If $J \geqslant 1$ and*

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}(A_\#), \quad \lambda \leqslant \bar{\lambda}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(A_\#, \lambda, D_{\mathrm{cut}}, \Lambda, A),$$
$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(A_\#, \lambda, D_{\mathrm{cut}}, \Lambda, A), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

then, for any $x \in \partial\mathcal{N}^{J+1}_{t_J}$, the parabolic neighborhood $P(x, A_\# r_{\mathrm{comp}})$ is unscathed,

$$P(x, A_\# r_{\mathrm{comp}}) \subset \mathcal{N} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}, \tag{12.24}$$

and we have the bound

$$Q \leqslant \overline{Q} \quad \text{on } P(x, A_\# r_{\mathrm{comp}}). \tag{12.25}$$

Proof. Choose a boundary component $\Sigma \subset \partial\mathcal{N}^{J+1}_{t_{J+1}}$ such that $x \in \Sigma(t_J) \subset \partial\mathcal{N}^{J+1}_{t_J}$. So, $x \in U^*_{\Sigma,t_J}$. Assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}(A_\#),$$

we obtain, by similar arguments as those that led to (12.23), that

$$B(x, A_\# r_{\mathrm{comp}}) \subset \mathcal{N}_{t_J-}. \tag{12.26}$$

Next, using Lemma 8.7 along with (12.22), and assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(A_\#, A),$$

we can find a constant $A' = A'(A_\#, A) < \infty$, with $A' \geqslant A$, such that $P(x, A'\rho_1(x))$ is unscathed and

$$P(y, A\rho_1(y)) \subset P(x, A'\rho_1(x)) \quad \text{for all } y \in P(x, A_\# r_{\mathrm{comp}}). \tag{12.27}$$

We now show that $P(x, A'\rho_1(x))$ is disjoint from the cuts. To do this, observe that, for any $t' \in (t_J, t_{J+1}]$, we have $B(x(t'), A'\rho_1(x)) \not\subset \mathcal{N}$. So, by Lemma 8.19, along with (12.22) and a-priori assumption (APA11), assuming

$$\delta_{\rm n} \leqslant \bar{\delta}_{\rm n}, \quad \lambda \leqslant \bar{\lambda}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\rm b} \leqslant \bar{\delta}_{\rm b}(\lambda, D_{\rm cut}, A'(A_\#, A), \Lambda),$$
$$\varepsilon_{\rm can} \leqslant \bar{\varepsilon}_{\rm can}(\lambda, D_{\rm cut}, A'(A_\#, A), \Lambda), \quad r_{\rm comp} \leqslant \bar{r}_{\rm comp},$$

we find that $P(x, A'\rho_1(x)) \cap \mathcal{D} = \varnothing$ for all $\mathcal{D} \in {\rm Cut} \cup {\rm Cut}^J$. Combining this with (12.26) gives us (12.24) via Lemma 8.18. Combining it further with (12.27) and a-priori assumption (APA7) yields (12.25). □

Next we improve the estimate from Claim 12.9 and use it to identify more precise necks in $\mathcal{M}'$.

Claim 12.10. *If*

$$\begin{gathered} E > 2, \quad H \geqslant \underline{H}(E). \quad \eta_{\rm lin} \leqslant \bar{\eta}_{\rm lin}(E), \quad A_\# \geqslant \underline{A}_\#(E, \delta_\#), \\ \nu \leqslant \bar{\nu}(E, \delta_\#), \quad \varepsilon_{\rm can} \leqslant \bar{\varepsilon}_{\rm can}(E, \delta_\#), \quad r_{\rm comp} \leqslant \bar{r}_{\rm comp}, \end{gathered} \tag{12.28}$$

then, for any component $\Sigma \subset \partial \mathcal{N}^{J+1}_{t_{J+1}}$, *we have*

$$|\hat{\phi}^* g'_{t_J} - g_{t_J}| < \delta_\# \quad \text{on } U^*_{\Sigma, t_J}, \tag{12.29}$$

and

$$\|r_{\rm comp}^{-2} (\hat{\phi} \circ \psi_{\Sigma, t_J})^* g'_{t_J} - g_{-1}^{S^2 \times \mathbb{R}}\|_{C^0} < \delta_\#. \tag{12.30}$$

Proof. Consider first the case $J = 0$. In this case, by (12.23) and assumption (vi) of this proposition we have, on U^*_{Σ, t_0},

$$e^{HT} \rho_1^E |\widehat{h}| \leqslant \nu \overline{Q} = \nu \cdot 10^{-E-1} \eta_{\rm lin} r_{\rm comp}^E ;$$

recall that $\widehat{h} = \tilde{\phi}^* g'_{t_0} - g_{t_0}$. So, (12.29) follows from (12.22), assuming

$$\eta_{\rm lin} \leqslant 1 \quad \text{and} \quad \nu \leqslant \bar{\nu}(\delta_\#).$$

Second, consider the case $J \geqslant 1$. By (12.24) and assuming

$$A_\# \geqslant \underline{A}_\#(\delta_\#),$$

we have $U^*_{\Sigma, t_J} \cap \mathcal{D} = \varnothing$ for all $\mathcal{D} \in {\rm Cut}^J$. So, therefore on U^*_{Σ, t_J} we have $\hat{\phi} = \phi_{t_J -}$, and hence

$$\hat{\phi}^* g'_{t_J} - g_{t_J} = h_{t_J -}.$$

We will now apply Proposition 9.2 at every point of U^*_{Σ,t_J}. To do this, note that by (12.24) the perturbation h is defined and smooth on all of $P(x, A_\# r_{\text{comp}})$ and by a-priori assumption (APA6) and (12.25) we have $|h| \leqslant \eta_{\text{lin}}$ and $Q \leqslant \overline{Q}$ everywhere on this parabolic neighborhood. Moreover, if $P(x, A_\# r_{\text{comp}})$ intersects the initial time-slice $\mathcal{M}_0$, then by a-priori assumption (APA12) we have $Q \leqslant \nu \overline{Q}$ on the intersection. Lastly, note that the diameter of U^*_{Σ,t_J} is bounded by $10\delta_\#^{-1} r_{\text{comp}}$ for sufficiently small $\delta_\#$. So, assuming

$$E > 2, \quad H \geqslant \underline{H}(E), \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}(E), \quad A_\# \geqslant \underline{A}_\#(E, \delta_\#),$$
$$\nu \leqslant \bar{\nu}(E, \delta_\#), \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(E, \delta_\#), \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}},$$

we conclude by Proposition 9.2 that $Q < \delta_\# \overline{Q}$ on U^*_{Σ,t_J}. Note that here we have used (12.22) and we applied Proposition 9.2 centered at all points in U^*_{Σ,t_J}, with an appropriate choice for the radius A.

So, on U^*_{Σ,t_J},

$$e^{H(T-t_J)} \rho_1^E |h_{t_J -}| = Q < \delta_\# \overline{Q} = \delta_\# \cdot 10^{-E-1} \eta_{\text{lin}} r_{\text{comp}}^E.$$

Using (12.22) and the fact that $t_J \leqslant T$, due to assumption (v) of this proposition, we obtain (12.29) assuming

$$\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}.$$

Finally, the bound (12.30) follows by combining (12.29) with (12.21) and adjusting (the earlier instance of) $\delta_\#$. □

Next, we use (12.30) to establish the existence of a $\delta_\#$-neck in $\mathcal{M}'$.

Claim 12.11. *If*

$$\delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \tag{12.31}$$

then following holds. For any component $\Sigma \subset \partial \mathcal{N}_{t_{J+1}-}$ there is a $\delta_\#$-neck $U'_\Sigma \subset \mathcal{M}'_{t_J}$ at scale $2r_{\text{comp}}$ that has a central 2-sphere which intersects $\hat{\phi}(\Sigma(t_J)) \subset \mathcal{M}'_{t_J}$.

Proof. Note that $\hat{\phi}(\Sigma(t_J)) = \phi^J_{t_J}(\Sigma(t_J))$, as $\mathcal{D} \subset \operatorname{Int} \mathcal{N}_{t_J+}$ for all $\mathcal{D} \in \operatorname{Cut}^J$ (see Definition 7.2). The ε_{can}-canonical neighborhood assumption holds on $\hat{\phi}(\Sigma(t_J))$ by assumption (vi) of this proposition (if $J = 0$) and by assertion (e) of Proposition 12.2 (if $J \geqslant 1$). The statement now follows from Lemma 8.10, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta_\#),$$

after possibly adjusting $\delta_\#$. □

By Lemma 8.12 and assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\delta_{\#}) \quad \text{and} \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}},$$

we obtain furthermore, after adjusting $\delta_{\#}$, the following.

CLAIM 12.12. *Assuming parameter bounds of the same form as in* (12.28) *and* (12.31), *the following holds.*

For any component $\Sigma \subset \partial \mathcal{N}_{t_{J+1}-}$, *there is a product domain* $U'^{*}_{\Sigma} \subset \mathcal{M}'_{[t_J,t_{J+1}]}$ *on the time-interval* $[t_J, t_{J+1}]$, *with* $\hat{\phi}(\Sigma(t_J)) \subset U'^{*}_{\Sigma,t_J}$, *on which the metric is* $\delta_{\#}$*-close at scale* r_{comp} *to the standard round shrinking cylinder. More specifically, there is an* r^2_{comp}*-time-equivariant and* ∂_t*-preserving diffeomorphism*

$$\psi'_{\Sigma} \colon S^2 \times (-\delta_{\#}^{-1}, \delta_{\#}^{-1}) \times [-1, 0] \longrightarrow U'^{*}_{\Sigma}$$

such that

$$\| r_{\text{comp}}^{-2} \psi'^{*}_{\Sigma} g' - g^{S^2 \times \mathbb{R}} \|_{C^{[\delta_{\#}^{-1}]}} < \delta_{\#}. \tag{12.32}$$

We furthermore have

$$\psi'_{\Sigma}(S^2 \times \{0\} \times \{-1\}) \cap \hat{\phi}(\Sigma(t_J)) \neq \varnothing.$$

We now carry out the grafting construction. We begin by identifying product domains in the time slabs $\mathcal{M}_{[t_J,t_{J+1}]}$ and $\mathcal{M}'_{[t_J,t_{J+1}]}$ that will be used in the construction.

For $k = 0, ..., 5$, let N_k be the (open) $100kr_{\text{comp}}$-tubular neighborhood around $\mathcal{N}_{t_J+}$ in $\mathcal{M}_{t_J}$, and set $N'_k := \hat{\phi}(N_k)$. Assuming

$$\delta_{\#} \leqslant \bar{\delta}_{\#} \quad \text{and} \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}},$$

we obtain from (12.23) and assumption (vi) of this proposition that

$$\begin{aligned} &N_0 \subset N_1 \subset ... \subset N_5 \subset X, \\ &N'_0 \subset N'_1 \subset ... \subset N'_5 \subset \hat{\phi}(X). \end{aligned} \tag{12.33}$$

Moreover, assuming

$$\eta_{\text{lin}} \leqslant 10^{-2} \quad \text{and} \quad \delta_{\#} \leqslant \bar{\delta}_{\#},$$

Claim 12.12 and a-priori assumption (APA6) (if $J \geqslant 1$) or the assumptions from the proposition (if $J = 0$) yield

$$N_5 \setminus N_0 \subset \bigcup_{\Sigma \subset \partial \mathcal{N}^{J+1}_{t_{J+1}}} U^{*}_{\Sigma,t_J} \quad \text{and} \quad N'_5 \setminus N'_0 \subset \bigcup_{\Sigma \subset \partial \mathcal{N}^{J+1}_{t_{J+1}}} U'^{*}_{\Sigma,t_J}. \tag{12.34}$$

By construction and by (12.34), all points on $N_5 = \mathcal{N}_{t_J+} \cup (N_5 \setminus N_0)$ and $N_5' \setminus N_0'$ survive until time t_{J+1}. A-priori assumptions (APA2), (APA3) (a) and (c), (APA6), assertions (d) and (e) of Proposition 12.2, (12.33), assumption (vi) of this proposition and Lemma 8.9, as well as (12.32) and (12.34), imply, assuming

$$\delta_\# \leqslant \bar{\delta}_\#, \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}, \quad \delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \quad \lambda \leqslant 1, \quad \Lambda \geqslant 2, \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda), \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}},$$

that

$$\rho > C_{\text{SD}}^{-1} \lambda r_{\text{comp}} > \varepsilon_{\text{can}} r_{\text{comp}} \quad \text{on } N_5'.$$

Let $t_1^* \in [t_J, t_{J+1}]$ be maximal with the property that $N_5'(t)$ is defined and weakly $\frac{1}{2} C_{\text{SD}}^{-1} \lambda r_{\text{comp}}$-thick for all $t \in [t_J, t_1^*]$, where C_{SD} is the constant from Lemma 8.9. Note here that t_1^* is well defined by the $(\varepsilon_{\text{can}} r_{\text{comp}}, T)$-completeness of $\mathcal{M}'$ and Lemma 8.4, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(\lambda).$$

We can now express the flows g and g' restricted to the product domains $N_5([t_J, t_1^*])$ and $N_5'([t_J, t_1^*])$ by conventional Ricci flows $(N_5, (g_t)_{t \in [t_J, t_1^*]})$ and $(N_5', (g_t')_{t \in [t_J, t_1^*]})$.

Claim 12.13. (Grafting on round half-cylinders) *After adjusting $\delta_\#$, there are smoothly varying Riemannian metrics $(g_t^+)_{t \in [t_J, t_1^*]}$ and $(g_t'^+)_{t \in [t_J, t_1^*]}$ on smooth manifolds N^+ and N'^+, respectively, and a diffeomorphism $\phi^+ \colon N^+ \to N'^+$ such that the following holds*:

(a) *N_5 and N_5' can be viewed as open subsets of N^+ and N'^+, respectively.*

(b) *For all $t \in [t_J, t_1^*]$, we have $g_t^+ = g_t$ on $N_1 \subset N^+$ and $g_t'^+ = g_t'$ on $N_1' \subset N'^+$.*

(c) *g_t^+ and $g_t'^+$ are complete for all $t \in [t_J, t_1^*]$.*

(d) *For some constant $C = C(\lambda) < \infty$, we have*

$$|\text{Rm}_{g_t^+}|, |\text{Rm}_{g_t'^+}| \leqslant C r_{\text{comp}}^{-2}.$$

(e) *$(g_t^+)_{t \in [t_J, t_1^*]}$ and $(g_t'^+)_{t \in [t_J, t_1^*]}$ are "$\delta_\#$-approximate Ricci flows"*:

$$-\delta_\# r_{\text{comp}}^{-2} g_t^+ < \partial_t g_t^+ + 2 \,\text{Ric}_{g_t^+} < \delta_\# r_{\text{comp}}^{-2} g_t^+,$$
$$-\delta_\# r_{\text{comp}}^{-2} g_t'^+ < \partial_t g_t'^+ + 2 \,\text{Ric}_{g_t'^+} < \delta_\# r_{\text{comp}}^{-2} g_t'^+.$$

(f) *For some $C^* = C^*(\lambda) < \infty$,*

$$|\nabla_{g_t^+}^m \text{Rm}(g_t^+)|_{g_t^+}, |\nabla_{g_t'^+}^m \text{Rm}(g_t'^+)|_{g_t'^+} < C^* r_{\text{comp}}^{-2} (t - t_J)^{-m/2}$$
$$|\nabla_{g_t^+}^{m_1} \partial_t^{m_2} g_t^+|_{g_t^+}, |\nabla_{g_t'^+}^{m_1} \partial_t^{m_2} g_t'^+|_{g_t'^+} < C^* r_{\text{comp}}^{-2} (t - t_J)^{-(m_2 - 1 + m_1/2)}$$

for all $t\in(t_J,t_1^*]$ *and* $m,m_1,m_2=0,...,100$.

(g) *There is a universal constant* $C^{**}<\infty$ *such that, at every* $x\in N^+$ *with*

$$d_{g^+_{t_J}}(x,N^+\setminus N_0)<\delta_\#^{-1}r_{\mathrm{comp}},$$

we have

$$|\nabla^m_{g^+_t}\mathrm{Rm}(g^+_t)|_{g^+_t}(x)<C^{**}r_{\mathrm{comp}}^{-2-m}$$
$$|\nabla^{m_1}_{g^+_t}\partial_t^{m_2}g^+_t|_{g^+_t}(x)<C^{**}r_{\mathrm{comp}}^{-m_1-2m_2}$$

for all $t\in(t_J,t_1^*]$ *and* $m,m_1,m_2=0,...,100$.

(h) $\phi^+=\hat\phi$ *on* N_2.

(i) *We have* $|(\phi^+)^*g'^+_{t_J}-g_{t_J}|_{g^+_t}<\delta_\#$ *at every point* $x\in N^+$, *with*

$$d_{g^+_{t_J}}(x,N^+\setminus N_0)<\delta_\#^{-1}r_{\mathrm{comp}}.$$

(j) $t_1^*>t_J$ *and if* $t_1^*<t_{J+1}$, *then*

$$\bigcup_{t\in[t_J,t_1^*]}N'_5(t)$$

must contain a $C_{\mathrm{SD}}^{-1}\lambda r_{\mathrm{comp}}$*-thin point.*

Proof. Using Lemma 8.4, we find that $t_1^*>t_J$ and that if $t_1^*<t_{J+1}$, then $N'_5([t_J,t_1^*])$ must contain a $C_{\mathrm{SD}}^{-1}\lambda r_{\mathrm{comp}}$-thin point. This proves assertion (j).

For each component $\Sigma\subset\partial\mathcal{N}^{J+1}_{t_{J+1}}$, we may push forward $r_{\mathrm{comp}}^{-2}g^{S^2\times\mathbb{R}}$ under ψ_Σ and ψ'_Σ to obtain spacetime metrics on U^*_Σ and U'^*_Σ. Using the product structure on U^*_Σ and U'^*_Σ, these yield evolving metrics $(g^\Sigma_t)_{t\in[t_J,t_{J+1}]}$ and $(g'^\Sigma_t)_{t\in[t_J,t_{J+1}]}$ on the initial time-slices U^*_{Σ,t_J}, U'^*_{Σ,t_J}, and hence also on $N_5\setminus N_0$ and $N'_5\setminus N'_0$, by (12.34).

By a standard interpolation argument, we can construct smooth families of metrics $(\tilde g_t)_{t\in[t_J,t_1^*]}$ and $(\tilde g')_{t\in[t_J,t_1^*]}$ on N_5 and N'_5 such that $\tilde g_t=g_t$ and $\tilde g'_t=g'_t$ on N_1 and N'_1, respectively, and such that, for every component $\Sigma\subset\partial\mathcal{N}^{J+1}_{t_{J+1}}$, we have $\tilde g_t=g^\Sigma_t$ and $\tilde g'=g'^\Sigma_t$ on $(N_5\setminus N_2)\cap U^*_{\Sigma,t_J}$ and $(N'_5\setminus N'_2)\cap U'^*_{\Sigma,t_J}$, respectively. Moreover, using (12.21) and (12.32), and after possibly adjusting $\delta_\#$, we may assume that, for every component $\Sigma\subset\partial\mathcal{N}^{J+1}_{t_{J+1}}$, $m_1,m_2<\delta_\#^{-1}$, and $t\in[t_J,t_1^*]$, we have

$$\begin{aligned}|\nabla^{m_1}_{g^\Sigma_{t_J}}\partial_t^{m_2}(\tilde g_t-g^\Sigma_t)|_{g^\Sigma_t}<\delta_\#r_{\mathrm{comp}}^{-m_1-2m_2},\\ |\nabla^{m_1}_{g'^\Sigma_{t_J}}\partial_t^{m_2}(\tilde g'_t-g'^\Sigma_t)|_{g'^\Sigma_t}<\delta_\#r_{\mathrm{comp}}^{-m_1-2m_2}\end{aligned}\tag{12.35}$$

on $(N_5\setminus N_0)\cap U^*_{\Sigma,t_J}$ and $(N'_5\setminus N'_0)\cap U'^*_{\Sigma,t_J}$. So, after possibly adjusting $\delta_\#$ once again, we may assume that, on N_5 and N'_5, we have

$$-\delta_\# r_{\mathrm{comp}}^{-2}\tilde{g}_t < \partial_t\tilde{g}_t + 2\,\mathrm{Ric}_{\tilde{g}_t} < \delta_\# r_{\mathrm{comp}}^{-2}\tilde{g}_t,$$
$$-\delta_\# r_{\mathrm{comp}}^{-2}\tilde{g}'_t < \partial_t\tilde{g}'_t + 2\,\mathrm{Ric}_{\tilde{g}'_t} < \delta_\# r_{\mathrm{comp}}^{-2}\tilde{g}'_t.$$

Since these flows are isometric to round shrinking cylindrical flows near the ends of N_5 and N'_5, we can attach round shrinking half-cylinders to these flows at each end. This produces flows $(g^+_t)_{t\in[t_J,t^*_1]}$ and $(g'^+_t)_{t\in[t_J,t^*_1]}$ on $N^+\supset N_4$ and $N'^+\supset N'_4$ satisfying assertions (a)–(e) of this claim. Assertion (g) also follows from (12.21), (12.32) and (12.35), after adjusting $\delta_\#$. Assertion (f) follows from Shi's estimates in N_1, N'_1, $N^+\setminus N_2$, $N'^+\setminus N'_2$ and assertion (g).

Since $\hat{\phi}$ is a $(1+\delta_\#)$-bilipschitz map on $N_5\setminus N_2\subset\bigcup_\Sigma U^*_{\Sigma,t_J}$ (see (12.29) and (12.34)), and $\tilde{g}_{t_J}$ and $\tilde{g}'_{t_J}$ are isometric to subsets of round cylinders on the interior of $N_5\setminus N_2$ and $N'_5\setminus N'_2$, respectively, we can use a smoothing procedure (see also Lemma D.1) to construct a diffeomorphism onto its image $\tilde{\phi}\colon N_4\to\tilde{\phi}(N_4)\subset N'_5$ such that $\tilde{\phi}=\hat{\phi}$ on N_2 and $\tilde{\phi}^*g'^+_{t_J}=\tilde{\phi}^*\tilde{g}'_{t_J}=\tilde{g}_{t_J}=g^+_{t_J}$ on $N_4\setminus N_3$, and such that, after adjusting $\delta_\#$,

$$|\tilde{\phi}^*g'^+_{t_J}-g^+_{t_J}|<\delta_\# \quad \text{on } U^*_{\Sigma,t_J}\cap N_4, \tag{12.36}$$

for every component $\Sigma\subset\partial\mathcal{N}^{J+1}_{t_{J+1}}$. We can now extend the diffeomorphism $\tilde{\phi}\colon N_4\to\tilde{\phi}(N_4)$ to a diffeomorphism $\phi^+\colon N^+\to N'^+$ such that it remains an isometry on $N^+\setminus N_4$. Adjusting $\delta_\#$ again, the map ϕ^+ will satisfy assertions (h) and (i) of this claim, by (12.36) and the fact that $(\phi^+)^*g'^+_{t_J}=g^+_{t_J}$ on $N^+\setminus N_3$. □

We now construct the map ϕ^{J+1} by solving the harmonic map heat flow equation starting from $(\phi^+)^{-1}$, where $\phi^+\colon N^+\to N'^+$ is the map constructed in Claim 12.13.

Using Claim 12.13 and Proposition A.9 from the appendix, we obtain that, if

$$\eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}} \quad\text{and}\quad \delta_\#\leqslant\bar{\delta}_\#(\eta_{\mathrm{lin}}),$$

then we can find a time $t^*\in(t_J,t^*_1]$ and a solution $(\chi_t)_{t\in[t_J,t^*]}$, with $\chi_t\colon N'^+\to N^+$, to the harmonic map heat flow equation with respect to $(g'^+_t)_{t\in[t_J,t^*_1]}$ and $(g^+_t)_{t\in[t_J,t^*_1]}$ with the following properties:

(1) $\chi_{t_J}=(\phi^+)^{-1}$.
(2) χ_t is a diffeomorphism for all $t\in[t_J,t^*]$.
(3) $|(\chi_t^{-1})^*g'^+_t-g^+_t|_{g^+_t}<2\eta_{\mathrm{lin}}$ for all $t\in[t_J,t^*]$.
(4) If
$$|(\chi_{t^*}^{-1})^*g'^+_{t^*}-g^+_{t^*}|_{g^+_{t^*}}<1.9\eta_{\mathrm{lin}}$$
holds on N'^+, then $t^*=t^*_1$.

We first show that

$$\chi_t^{-1}(N_0) \subset N_1' \quad \text{for all } t \in [t_J, t^*]. \tag{12.37}$$

After rescaling by r_{comp}^{-2}, assuming

$$\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}} \quad \text{and} \quad \delta_{\#} \leqslant \bar{\delta}_{\#},$$

we may apply assertion (g) of Claim 12.13 and Proposition A.10, taking the constants δ and A in the hypotheses to be $\delta=1$ and $A=C^{**}$, to conclude that $\chi_t^{-1}(\partial N_0)\subset N_1'$ for all $t\in[t_J,t^*]$. Here, we have used a continuity argument, the fact that

$$\chi_{t_J}^{-1}(\partial N_0) = \phi^+(\partial N_0) = \hat{\phi}(\partial N_0) = \partial N_0'$$

by assertion (h) of Claim 12.13, and property (3) above, to retain the hypotheses of Proposition A.10. Therefore, since χ_t^{-1} is a smoothly varying diffeomorphism, (12.37) follows.

Now, set $\bar{\phi}_t := \chi_t^{-1}|_{N_0}$ for $t\in[t_J,t^*]$. As $g_t^+ = g_t$ on N_0 and $g_t'^+ = g_t'$ on N_1' by assertion (b) of Claim 12.13, we can view $(\bar{\phi}_t)_{t\in[t_J,t^*]}$ as a smooth, time-preserving diffeomorphism onto its image of the form

$$\phi^{J+1}: \mathcal{N}^{J+1}_{[t_J,t^*]} \longrightarrow N_1'([t_J,t^*]) \subset \mathcal{M}',$$

whose inverse evolves by harmonic map heat flow equation with respect to g' and g.

Consider the perturbation $h^{J+1} = (\phi^{J+1})^* g' - g$. Assertion (a) of this proposition follows from property (3) above.

If

$$\eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}} \quad \text{and} \quad \delta_{\#} \leqslant \bar{\delta}_{\#}(T, E, H, \eta_{\text{lin}}),$$

then for every $x\in N^+$ with $d_{g_{t_J}^+}(x, N^+\setminus N_0) < \delta_{\#}^{-1} r_{\text{comp}}$, after adjusting $\delta_{\#}$, and by assertions (e), (g), (i) of Claim 12.13 and property (3) above, we may apply Proposition A.13 to conclude that

$$|(\chi_t^{-1})^* g_t'^+ - g_t^+|_{g_t^+}(x) < e^{-HT} 10^{-E-1} r_{\text{comp}}^E \rho_1^{-E}(x) \eta_{\text{lin}} < \eta_{\text{lin}} \tag{12.38}$$

for all $t\in[t_J,t^*]$. If

$$\delta_{\#} \leqslant \bar{\delta}_{\#}(F),$$

then (12.21) implies that, for every $t\in[t_J,t^*]$ and every $x\in\mathcal{N}_t^{J+1}$ such that

$$d_t(x, \partial\mathcal{N}_t^{J+1}) < F r_{\text{comp}},$$

we have

$$d_{g^+_{t_J}}(x(t_J), N^+\setminus N_0) < \delta_\#^{-1} r_{\mathrm{comp}}.$$

Hence,

$$|h^{J+1}(x)| = |(\chi_t^{-1})^* g'^+_t - g^+_t|_{g^+_t}(x(t_J)) < e^{-HT} 10^{-E} r_{\mathrm{comp}}^E \rho_1^{-E-1}(x)\eta_{\mathrm{lin}}$$

by (12.38). This yields assertion (b) of this proposition.

Finally, we verify assertions (c) and (d) of this proposition. We first apply Lemma 8.9 and a-priori assumptions (APA2), (APA3) (a) and (c), (APA6), assuming

$$\eta_{\mathrm{lin}} \leqslant \bar\eta_{\mathrm{lin}}, \quad \delta_{\mathrm{n}} \leqslant \bar\delta_{\mathrm{n}}, \quad \lambda \leqslant \bar\lambda, \quad \Lambda \geqslant 2, \quad \varepsilon_{\mathrm{can}} \leqslant \bar\varepsilon_{\mathrm{can}}(\lambda), \quad r_{\mathrm{comp}} \leqslant \bar r_{\mathrm{comp}},$$

to find that, for all $t\in[t_J, t^*]$, the following holds: If the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds at scales $(0,1)$ on $\phi_t^{J+1}(\mathcal{N}_t^{J+1})$, then $\phi^{J+1}(\mathcal{N}_t^{J+1})$ is $C_{\mathrm{SD}}\lambda r_{\mathrm{comp}}$-thick. By assertion (e) of Proposition 12.2, this condition holds for $t=t_J$. Therefore, assuming

$$\varepsilon_{\mathrm{can}} < C_{\mathrm{SD}}^{-1}\lambda,$$

it holds for all $t\in[t_J, t^*]$ by continuity. This shows assertion (d) of this proposition and the fact that $\phi_{t^*}^{J+1}(\mathcal{N}_{t^*}^{J+1})$ is $C_{\mathrm{SD}}\lambda r_{\mathrm{comp}}$-thick.

To see assertion (c) of this proposition, assume that $|h^{J+1}|\leqslant\eta_{\mathrm{lin}}$ on $\mathcal{N}_{t^*}^{J+1}$. Then, by the definition of ϕ^{J+1} we have the bound $|(\chi_{t^*}^{-1})^* g'^+_{t^*} - g^+_{t^*}|_{g^+_{t^*}} \leqslant \eta_{\mathrm{lin}}$ on N_0. By (12.38) it follows that $|(\chi_{t^*}^{-1})^* g'^+_{t^*} - g^+_{t^*}|_{g^+_{t^*}} \leqslant \eta_{\mathrm{lin}}$ holds everywhere on N^+. So, property (4) above implies $t^*=t_1^*$. Combining this with the conclusion from the previous paragraph and applying assertion (j) of Claim 12.13, yields $t^*=t_1^*=t_{J+1}$, as desired. □

12.4. Proof of Proposition 12.1, concluded

In this subsection we use the results of the previous subsections to prove our main Proposition 12.1. More specifically, we will analyze the map

$$\phi^{J+1}: \mathcal{N}^{J+1}_{[t_J,t^*]} \longrightarrow \mathcal{M}'$$

that was constructed in Proposition 12.8. We will verify that this map satisfies a-priori assumptions (APA1)–(APA12), and show that $t^*=t_{J+1}$. Therefore, ϕ^{J+1} can be used to extend the comparison $(\phi, \{\phi^j\}_{j=1}^J, \mathrm{Cut})$ to the time-interval $[0, t_{J+1}]$. This will finish the proof of Proposition 12.1.

Our proof can be roughly summarized as follows: By the assumptions of Proposition 12.1, we may assume that a-priori assumptions (APA1)–(APA5) already hold until

time t_{J+1}, and a-priori assumptions (APA6) and (APA7)–(APA12) already hold until time t_J. We will refer to this assumption as the "induction hypothesis" henceforth. Using the induction hypothesis and the conclusions of Propositions 12.2 and 12.8, we will then establish that a-priori assumptions (APA1)–(APA12) hold up to time t^*. The only non-trivial assumptions in this step will be a-priori assumptions (APA6)–(APA9). We will prove these assumptions using another continuity argument: We will assume that relaxed versions of a-priori assumptions (APA7)–(APA9) hold up to some almost maximal time $t^{**} \leqslant t^*$ and, based on these extra assumptions, we prove a-priori assumptions (APA6)–(APA9) up to time t^{**}. By a straightforward openness argument, it therefore follows that $t^{**} = t^*$, and therefore that a-priori assumptions (APA6)–(APA9) hold up to time t^*. Eventually, the fact that a-priori assumption (APA6) holds up to time t^* and assertion (c) of Proposition 12.8 imply that we indeed have $t^* = t_{J+1}$. This will finish our proof.

We remark that throughout this entire subsection, we will introduce global terminology and assumptions on the parameters, which will be understood to remain valid for the remainder of the subsection. In particular, conditions on the parameters that can be found in the following lemmas will be assumed to hold for the remainder of the subsection, so that the conclusions of these lemmas can be applied immediately.

This subsection is structured as follows: We first set up our argument by recalling the important assumptions from Proposition 12.1. In Lemma 12.14, we will then summarize and put into context the results of the constructions from Propositions 12.2 and 12.8. Next, we introduce the relaxed versions of a-priori assumptions (APA7)–(APA9) in equations (12.40)–(12.42), which hold up to some time $t^{**} \leqslant t^*$. In Lemma 12.15, we show that $t^{**} > t_J$ and that if the strong versions of a-priori assumptions (APA7)–(APA9) hold up to time t^{**}, then we must in fact have $t^{**} = t^*$. Based on these relaxed versions of a-priori assumptions (APA7)–(APA9), we will establish a-priori assumptions (APA6)–(APA9) in Lemmas 12.16–12.19 — one lemma per a-priori assumption and in this order. Lastly, we wrap up our discussion, argue that $t^{**} = t^* = t_{J+1}$ and verify the assertions of Proposition 12.1.

Further explanations of the arguments may be found after the statements of the lemmas below.

In what follows, we will be considering the setup as described in assumptions (i)–(vi) of Proposition 12.1. So, among other things, we assume that $\mathcal{M}$ and $\mathcal{M}'$ are $(\varepsilon_{\text{can}} r_{\text{comp}}, T)$-complete and satisfy the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon_{\text{can}} r_{\text{comp}}, 1)$. We consider a comparison domain $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ over the time-interval $[0, t_{J+1}]$, for $J \geqslant 0$, and a comparison $(\text{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ from $\mathcal{M}$ to $\mathcal{M}'$ defined on this comparison domain over the time-interval $[0, t_J]$. If $J = 0$, then this comparison

is trivial, as explained after Definition 7.2. We also assume $(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$ and $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ satisfy a-priori assumptions (APA1)–(APA6) for the parameters

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}})$$

and a-priori assumptions (APA7)–(APA13) for the parameters

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}}).$$

Moreover, we assume in the following that

$$t_{J+1} \leqslant T. \tag{12.39}$$

If $J \geqslant 1$, then assumptions (i)–(v) of Proposition 12.1 allow us to apply Proposition 12.2. Doing so yields the map $\hat{\phi}: \mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+} \to \mathcal{M}'_{t_J}$ and the set of cuts Cut^J with the properties as explained in assertions (a)–(e) of this proposition. If $J=0$, then we skip this step.

Next, we fix an auxiliary constant $F < \infty$, whose value will be determined in the course of this subsection depending only on E. We can then apply Proposition 12.8 for $\hat{\phi}: X := \mathcal{N}_{t_J-} \cup \mathcal{N}_{t_J+} \to \mathcal{M}'_{t_J}$ from Proposition 12.2 (if $J \geqslant 1$) or $\hat{\phi} = \zeta: X \to \mathcal{M}'_0$ from assumption (vi) of Proposition 12.1 (if $J=0$). Then, Proposition 12.8 yields a time $t^* \in (t_J, t_{J+1}]$ and a map $\phi^{J+1}: \mathcal{N}^{J+1}_{[t_J, t^*]} \to \mathcal{M}'$ satisfying assertions (a)–(d) of this proposition. Note that Propositions 12.2 and 12.8 are only applicable if our parameters satisfy the bounds (12.2) and (12.20). These bounds are implied by bounds of the form (12.1) and

$$\nu \leqslant \bar{\nu}(T, E, F, H, \eta_{\mathrm{lin}}), \quad \delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}(T, E, F, H, \eta_{\mathrm{lin}}),$$
$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(T, E, F, H, \eta_{\mathrm{lin}}, \lambda, D_{\mathrm{cut}}, A, \Lambda), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(T, E, F, H, \eta_{\mathrm{lin}}, \lambda, D_{\mathrm{cut}}, A, \Lambda).$$

(Note that assuming $F = F(E)$, these bounds also follow from bounds of the form (12.1).)

In the following lemma we summarize the important properties of Cut^J and ϕ^{J+1}, and we show how these objects can be used to extend the comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ to a comparison that is defined over the time-interval $[0, t^*]$.

Lemma 12.14. *There is a unique map*

$$\bar{\phi}: \mathcal{N}_{[0,t^*]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D} \longrightarrow \mathcal{M}'$$

such that $(\mathrm{Cut} \cup \mathrm{Cut}^J, \bar{\phi}, \{\phi^j\}_{j=0}^{J+1})$ *is a comparison defined on the comparison domain* $(\mathcal{N}_{[0,t^*]}, \{\mathcal{N}^j, \mathcal{N}^{J+1}_{[t_J,t^*]}\}_{j=1}^{J}, \{t_j, t^*\}_{j=0}^{J})$. *This comparison is an extension of the comparison* $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^{J})$ *in the sense that*

$$\bar{\phi} = \phi \quad \text{on} \bigcup_{j=1}^{J} \mathcal{N}^j \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}.$$

Furthermore, this extended comparison and the comparison domain

$$(\mathcal{N}_{[0,t^*]}, \{\mathcal{N}^j, \mathcal{N}^{J+1}_{[t_J,t^*]}\}_{j=1}^{J}, \{t_j, t^*\}_{j=0}^{J})$$

satisfy the following properties:

(a) *They satisfy a-priori assumptions* (APA10)–(APA13) *for the parameters*

$$(T, E, H, \eta_{\text{lin}}, \nu, \lambda, \eta_{\text{cut}}, D_{\text{cut}}, W, A, r_{\text{comp}}).$$

(b) *Let* $(h, \{h^j\}_{j=1}^{J+1})$ *be the associated Ricci–DeTurck perturbation (note that* h^{J+1} *is only defined over the time-interval* $[t_J, t^*]$*). Then,* $|h| \leqslant \eta_{\text{lin}}$ *on*

$$\bigcup_{j=1}^{J} \mathcal{N}^j \setminus \bigcup_{\mathcal{D} \in \text{Cut} \cup \text{Cut}^J} \mathcal{D}$$

and $|h^{J+1}| \leqslant 10\eta_{\text{lin}}$ *on* $\mathcal{N}^{J+1}_{[t_J,t^*]}$*. Moreover,* $\phi^{J+1}(\mathcal{N}^{J+1}_{[t_J,t^*]})$ *is* $\varepsilon_{\text{can}} r_{\text{comp}}$*-thick.*

(c) *For any* $t \in [t_J, t^*]$ *and* $x \in \mathcal{N}^{J+1}_t$ *with* $d_t(x, \partial\mathcal{N}^{J+1}_t) < Fr_{\text{comp}}$*, we have*

$$Q_+(x) = e^{H(T-t)} \rho_1^E(x) |h^{J+1}(x)| \leqslant \overline{Q} = 10^{-E-1} \eta_{\text{lin}} r_{\text{comp}}^E.$$

(d) *If even* $|h| \leqslant \eta_{\text{lin}}$ *on* $\mathcal{N}^{J+1}_{t^*}$*, then* $t^* = t_{J+1}$.

Proof. The construction of the map $\bar{\phi}$ and the verification of the properties of Definitions 7.1 and 7.2 are straightforward. A-priori assumptions (APA10) and (APA11) follow directly from the corresponding a-priori assumptions of the induction hypothesis and Proposition 12.2 (d) and (c). A-priori assumption (APA12) follows directly from a-priori assumption (APA12) of the induction hypothesis (if $J \geqslant 1$) or from assumption (vi) in Proposition 12.1 (if $J = 0$). Assertions (b)–(d) are just restatements of assertions (a)–(d) of Proposition 12.8 combined with the induction hypthesis. □

Note that by the assumptions of Proposition 12.1, the comparison domain

$$(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^{J+1}, \{t_j\}_{j=0}^{J+1})$$

satisfies a-priori assumptions (APA1)–(APA5) for the parameters

$$(\eta_{\text{lin}}, \delta_{\text{n}}, \lambda, D_{\text{cap}}, \Lambda, \delta_{\text{b}}, \varepsilon_{\text{can}}, r_{\text{comp}}).$$

For the remainder of this section, references to (APA1)–(APA5) will implicitly refer to this larger comparison domain, rather than the comparison domain defined on the shorter interval $[0, t^*]$.

It remains to verify a-priori assumptions (APA6)–(APA9). Once this has been accomplished, assertion (d) of Lemma 12.14 will immediately imply that $t^*=t_{J+1}$. So, we have reduced our discussion to an analysis of the associated Ricci–DeTurck perturbation and its derived quantities Q and Q^*.

We will verify a-priori assumptions (APA6)–(APA9) via another continuity argument, which we will set up now. Consider the comparison $(\mathrm{Cut}\cup\mathrm{Cut}^J, \bar\phi, \{\phi^j\}_{j=0}^{J+1})$ from Lemma 12.14 and let $(h, \{h^j\}_{j=1}^{J+1})$ be the associated Ricci–DeTurck perturbation, as mentioned in assertion (b) of this lemma. As in Definition 7.5 we define the quantities

$$Q = e^{H(T-\mathfrak{t})}\rho_1^E|h| \quad \text{and} \quad Q^* = e^{H(T-\mathfrak{t})}\rho_1^3|h|,$$

and the extensions $Q_\pm$ and $Q^*_\pm$ to $\mathcal{N}_{t_j\pm}$. Moreover, again as in Definition 7.5, we set

$$\overline{Q} := 10^{-E-1}\eta_{\mathrm{lin}} r_{\mathrm{comp}}^E \quad \text{and} \quad \overline{Q}^* := 10^{-1}\eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3.$$

Choose a time $t^{**}\in[t_J, t^*]$ such that, for all

$$x \in \mathcal{N}^{J+1}_{[t_J,t^{**}]} \setminus \bigcup_{\mathcal{D}\in\mathrm{Cut}^J} \mathcal{D},$$

the following conditions hold:

$$Q(x) \leqslant 10\overline{Q} \qquad \text{whenever } P(x, 10A\rho_1(x))\cap\mathcal{D} = \varnothing \text{ for all } \mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J, \tag{12.40}$$

$$Q(x) \leqslant 10W\overline{Q}, \tag{12.41}$$

$$Q^*(x) \leqslant 10\overline{Q}^* \qquad \text{whenever } B(x, 10A\rho_1(x)) \subset \mathcal{N}_{\mathfrak{t}(x)-}. \tag{12.42}$$

Note that these conditions are relaxed versions of a-priori assumptions (APA7)–(APA9). The main objective of this subsection will be to show — under certain bounds on our parameters — that assumptions (12.40)–(12.42) imply a-priori assumptions (APA6) and (APA7)–(APA9) up to time t^{**}. The following lemma will help us conclude that it is possible to choose $t^{**}=t^*$ if a-priori assumptions (APA7)–(APA9) have been established.

Lemma 12.15. *If*

$$E \geqslant \underline{E}, \quad F \geqslant \underline{F}, \quad \eta_{\mathrm{lin}} \leqslant \bar\eta_{\mathrm{lin}}, \quad \delta_{\mathrm{n}} \leqslant \bar\delta_{\mathrm{n}}, \quad \lambda \leqslant \bar\lambda, \quad \eta_{\mathrm{cut}} \leqslant \bar\eta_{\mathrm{cut}},$$
$$D_{\mathrm{cut}} \geqslant \underline{D}_{\mathrm{cut}}(\lambda), \quad W \geqslant \underline{W}(E, \lambda, D_{\mathrm{cut}}), \quad A \geqslant \underline{A}, \quad \Lambda \geqslant \underline{\Lambda},$$
$$\delta_{\mathrm{b}} \leqslant \bar\delta_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A, \Lambda), \quad \varepsilon_{\mathrm{can}} \leqslant \bar\varepsilon_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A, \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar r_{\mathrm{comp}}(\lambda),$$

then we can choose $t^{**} > t_J$.

Furthermore, there is a constant $\tau=\tau(T,E,H,\eta_{\text{lin}},\lambda,A,r_{\text{comp}})>0$ *with the following property. If a-priori assumptions* (APA7)–(APA9) *hold up to time* t^{**}*, meaning that, for all* $x\in\mathcal{N}^{J+1}_{[t_J,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\text{Cut}^J}\mathcal{D}$*,*

$$Q(x)\leqslant\overline{Q}\qquad \text{whenever } P(x,A\rho_1(x))\cap\mathcal{D}=\varnothing \text{ for all } \mathcal{D}\in\text{Cut}\cup\text{Cut}^J, \tag{12.43}$$

$$Q(x)\leqslant W\overline{Q}, \tag{12.44}$$

$$Q^*(x)\leqslant\overline{Q}^*\qquad \text{whenever } B(x,A\rho_1(x))\subset\mathcal{N}_{\mathrm{t}(x)-}, \tag{12.45}$$

then (12.40)–(12.42) *even hold for all*

$$x\in\mathcal{N}^{J+1}_{[t_J,\min\{t^{**}+\tau,t^*\}]}\setminus\bigcup_{\mathcal{D}\in\text{Cut}^J}\mathcal{D}.$$

In other words, if (APA7)–(APA9) hold up to time t^{**}, then we may replace t^{**} by $\min\{t^{**}+\tau,t^*\}$. The important point here is that τ can be chosen independently of t^{**}. In Lemmas 12.16–12.19 below, we will show that a-priori assumptions (APA7)–(APA9) indeed hold up to time t^{**}, regardless of the choice of t^{**}. It will then follow by iterating Lemma 12.15 that we can choose $t^{**}=t^*$ and that a-priori assumptions (APA7)–(APA9) hold up to time t^*.

The main idea of the proof of Lemma 12.15 is that the relaxed conditions (12.40)–(12.42) hold in the neighborhood of any point at which the stricter conditions (12.43)–(12.45) are satisfied. Using the canonical neighborhood assumption, we will find a uniform lower bound on the size of such a neighborhood. Extra care has to be taken near the cuts at time t_J. Here we will use the a-priori assumptions from our induction hypothesis along with the geometry of the cuts to deduce that (12.44) and (12.45) even hold on and near the cuts.

Proof. We first show that for all $x\in\mathcal{N}_{t_J+}$ (which may possibly lie on a cut)

$$Q_+(x)\leqslant W\overline{Q}, \tag{12.46}$$

$$Q^*_+(x)\leqslant\overline{Q}^*\qquad \text{whenever } B(x,A\rho_1(x))\subset\mathcal{N}_{t_J+}. \tag{12.47}$$

Note that the condition in (12.47) refers to the *forward* time-slice, in contrast to (12.45).

Let us first prove (12.47). If $x\in\mathcal{D}\in\text{Cut}^J$, then (12.47) follows from a-priori assumption (APA10) (see assertion (a) of Lemma 12.14), assuming

$$\eta_{\text{cut}}\leqslant 1.$$

So, assume that

$$x \in \mathcal{N}_{t_{J^+}} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}^J} \mathcal{D}.$$

If $B(x, A\rho_1(x)) \subset \mathcal{N}_{t_{J^-}}$, then (12.47) follows from a-priori assumption (APA9) of the induction hypothesis. So, assume that $B(x, A\rho_1(x)) \not\subset \mathcal{N}_{t_{J^-}}$, but $B(x, A\rho_1(x)) \subset \mathcal{N}_{t_{J^+}}$. In other words, $B(x, A\rho_1(x))$ intersects an extension cap $\mathcal{C}_0 \subset \mathcal{N}_{t_{J^+}} \setminus \mathrm{Int}\, \mathcal{N}_{t_{J^-}}$. Choose $\mathcal{D}_0 \in \mathrm{Cut}^J$ with $\mathcal{C}_0 \subset \mathcal{D}_0$ and let $C_\# < \infty$ be a constant whose value we will determine in the course of the proof. We now apply Lemma 8.20 for $A_0 = A$ and $T_0 = 0$, assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad D_{\mathrm{cut}} \geqslant \underline{D}_{\mathrm{cut}}(\lambda, C_\#), \quad \Lambda \geqslant \underline{\Lambda},$$
$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, C_\#, D_{\mathrm{cut}}, A, \Lambda), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A, \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(C_\#)$$

Note that the assumptions of this lemma on the set $\mathrm{Cut} \cup \mathrm{Cut}^J$ hold due to a-priori assumption (APA11), which holds due to assertion (a) of Lemma 12.14. We find that

$$\rho_1(x) \geqslant C_\# r_{\mathrm{comp}}, \tag{12.48}$$

and that $P(x, A\rho_1(x)) \cap \mathcal{D} = \varnothing$ for all $\mathcal{D} \in \mathrm{Cut}$. So, by a-priori assumption (APA7) of the induction hypothesis we have

$$e^{H(T-t_J)} \rho_1^E(x) |h(x)| = Q(x) \leqslant \overline{Q} = 10^{-E-1} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E.$$

Combining this with (12.48) yields

$$Q^*(x) = e^{H(T-t_J)} \rho_1^3(x) |h(x)| \leqslant \rho_1^{3-E}(x) \cdot 10^{-E-1} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E \leqslant C_\#^{3-E} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^3.$$

It follows that $Q^*(x) \leqslant \overline{Q}^*$ if $C_\#^{3-E} \leqslant 10^{-1} \lambda^3$, which holds assuming

$$E \geqslant 4 \quad \text{and} \quad C_\# \geqslant \underline{C}_\#(\lambda).$$

This finishes the proof of (12.47).

To see the bound (12.46), we only need to consider the case $x \in \mathcal{D} \in \mathrm{Cut}^J$, due to a-priori assumption (APA8) of the induction hypothesis. Then, again by a-priori assumption (APA10) (see assertion (a) of Lemma 12.14) we have $Q_+^*(x) \leqslant \eta_{\mathrm{cut}} \overline{Q}^*$. By a-priori assumptions (APA5) and (APA11), and assuming

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}),$$

we conclude that there is a constant $C' = C'(\lambda, D_{\mathrm{cut}}) < \infty$ such that $\rho_1 < C' r_{\mathrm{comp}}$ on $\mathcal{D}$. So

$$\begin{aligned} Q_+(x) &= e^{H(T-t_J)} \rho_1^E(x) |h_{t_{J^+}}(x)| = \rho_1^{E-3}(x) Q_+^*(x) \\ &\leqslant (C')^{E-3} r_{\mathrm{comp}}^{E-3} \cdot \eta_{\mathrm{cut}} \overline{Q}^* = (C')^{E-3} r_{\mathrm{comp}}^{E-3} \cdot \eta_{\mathrm{cut}} \cdot 10^{-1} \eta_{\mathrm{lin}} (\lambda r_{\mathrm{comp}})^3 \\ &= (C')^{E-3} \lambda^3 10^E \eta_{\mathrm{cut}} \cdot \overline{Q}. \end{aligned}$$

It follows that $Q_+(x)\leqslant W\overline{Q}$, assuming

$$\eta_{\mathrm{cut}}\leqslant 1 \quad \text{and} \quad W\geqslant \underline{W}(E,\lambda,D_{\mathrm{cut}}).$$

This finishes the proof of (12.46).

We will show that (12.40)–(12.42) hold slightly beyond time t^{**} if $t^{**}<t^*$. The fact that we can choose $t^{**}>t_J$ will follow along the lines of the proof.

Let $\tau>0$ be a constant to be determined in the course of the proof. It suffices to argue that, if $t^{**}<t^*$ and if (12.43)–(12.45) hold on $\mathcal{N}^{J+1}_{[t_J,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}$ and (12.46), (12.47) hold on $\mathcal{N}_{t_J+}$, then (12.40)–(12.42) hold in $\mathcal{N}^{J+1}_{t'}$ whenever $t'\in(t^{**},t^*]$ and $t'-t^{**}<\tau$, where $\tau\leqslant\bar\tau(T,E,H,\eta_{\mathrm{lin}},\lambda,A,r_{\mathrm{comp}})$. To that end, choose $t'\in(t^{**},t^*]$ with $t'-t^{**}<\tau$, and a point $x'\in\mathcal{N}^{J+1}_{t'}$. Since $\mathcal{N}^{J+1}_{[t_J,t*]}$ is a product domain, we have $x'=x(t')$ for some $x\in\mathcal{N}^{J+1}_{t^{**}}$.

First suppose that $d_{t^{**}}(x,\partial\mathcal{N}^{J+1}_{t^{**}})\leqslant r_{\mathrm{comp}}$. Assuming

$$\tau\leqslant\bar\tau(\lambda,r_{\mathrm{comp}}),$$

then by a distance distortion estimate based on a-priori assumption (APA2) and the fact that $\mathcal{N}^{J+1}_{[t_J,t^*]}$ is a product domain, we obtain $d_{t'}(x(t'),\partial\mathcal{N}^{J+1}_{t'})<10r_{\mathrm{comp}}$. So, assuming

$$F\geqslant 10,$$

assertion (c) of Lemma 12.14 implies that (12.40) holds for $x(t')$. Thus, if

$$W\geqslant 1,$$

then (12.41) holds as well. Next, by a-priori assumption (APA3) (a), Lemma 8.12, and assuming

$$\delta_{\mathrm{n}}\leqslant\bar\delta_{\mathrm{n}},\quad \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}} \quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar r_{\mathrm{comp}},$$

we obtain that $\rho(x(t'))>\frac{1}{2}r_{\mathrm{comp}}$. So, $B(x(t'),A\rho_1(x(t')))\not\subset\mathcal{N}$, assuming

$$A>20,$$

and thus (12.42) holds.

Now, suppose $d_{t^{**}}(x,\partial\mathcal{N}^{J+1}_{t^{**}})>r_{\mathrm{comp}}$. By a-priori assumption (APA2), Lemma 8.4, and assuming

$$\tau\leqslant\bar\tau(E,\lambda,r_{\mathrm{comp}}) \quad\text{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}}(\lambda),$$

we obtain that

$$(0.9)^{1/E}\rho_1(x)\leqslant\rho_1\leqslant(1.1)^{1/E}\rho_1(x) \quad\text{on } P(x,\tau,\tau). \tag{12.49}$$

Thus, on $P(x,\tau,\tau)$,

$$0.9\cdot e^{H(T-\mathrm{t})}\rho_1^E(x)\cdot|h|\leqslant Q\leqslant 1.1\cdot e^{H(T-\mathrm{t})}\rho_1^E(x)\cdot|h|. \tag{12.50}$$

Assume now that $P(x(t'),10A\rho_1(x(t')))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$. By a-priori assumption (APA2) and bounded curvature at bounded distance, Lemma 8.6, and assuming

$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,A),$$

we conclude that $P(x(t'),10A\rho_1(x(t')))$ is unscathed. We also obtain a curvature bound on this parabolic neighborhood, which implies, via a distance distortion estimate, that

$$P(x,9A\rho_1(x(t')))\subset P(x(t'),10A\rho_1(x(t'))),$$

assuming

$$\tau\leqslant\bar{\tau}(\lambda,A,r_{\mathrm{comp}}).$$

Combining this with (12.49) and a-priori assumption (APA2), and assuming

$$E\geqslant 1,\quad A\geqslant\underline{A}\quad\text{and}\quad\tau\leqslant\bar{\tau}(\lambda,A,r_{\mathrm{comp}}),$$

we obtain that $P(y,A\rho_1(y))\subset P(x(t'),10A\rho_1(x(t')))$ for all $y\in B(x,\tau)$. This implies that, for all such y, we have $P(y,A\rho_1(y))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$. Therefore, by (12.43), we have $Q\leqslant\overline{Q}$ on $B(x,\tau)$. So, if

$$\eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}},$$

then we can use Proposition A.12 together with a-priori assumption (APA2), (12.49), (12.50) and assertion (b) of Lemma 12.14, and assuming

$$\tau\leqslant\bar{\tau}(T,E,H,\eta_{\mathrm{lin}},\lambda,r_{\mathrm{comp}}), \tag{12.51}$$

to get that (12.40) holds.

Using similar arguments, properties (12.41) and (12.42) can be verified at $x(t')$ as well, assuming a bound of the form (12.51). Note that if $t^{**}=t_J$, then we need to use the bounds (12.46) and (12.47). □

Assume, for the remainder of this subsection, that the parameter bounds that appear in Lemma 12.15 hold and that $t^{**}>t_J$.

In the following we will verify a-priori assumptions (APA6)–(APA9) up to time t^{**}. Whenever we say that "a-priori assumption (APAx) holds", then we mean that

$$(\mathcal{N}_{[0,t^{**}]},\{\mathcal{N}^j,\mathcal{N}^{J+1}_{[t_J,t^{**}]}\}_{j=1}^J,\{t_j,t^{**}\}_{j=1}^J)$$

and

$$(\mathrm{Cut}\cup\mathrm{Cut}^J, \bar\phi|_{\mathcal{N}_{[0,t^{**}]}}, \{\phi^j, \phi^{J+1}|_{\mathcal{N}^{J+1}_{[t_J,t^{**}]}}\}_{j=1}^J)$$

satisfy a-priori assumption (APAx) for the set of parameters

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}}).$$

Note that it follows from Lemma 12.14 (a), that a-priori assumptions (APA10)–(APA13) hold.

Let us first verify a-priori assumption (APA6).

LEMMA 12.16. (Verification of (APA6)) *If*

$$\delta_{\mathrm{n}} \leqslant \bar\delta_{\mathrm{n}}, \quad \lambda \leqslant \bar\lambda, \quad \Lambda \geqslant \underline{\Lambda}(\lambda, A), \quad \delta_{\mathrm{b}} \leqslant \bar\delta_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A, \Lambda),$$
$$\varepsilon_{\mathrm{can}} \leqslant \bar\varepsilon_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A, \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar r_{\mathrm{comp}},$$

then a-priori assumption (APA6) *holds. In other words, we have* $|h|\leqslant\eta_{\mathrm{lin}}$ *on*

$$\mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J} \mathcal{D}$$

and the $\varepsilon_{\mathrm{can}}$*-canonical neighborhood assumption holds at scales* $(0,1)$ *on*

$$\bigcup_{j=1}^J \phi^j(\mathcal{N}^j) \cup \phi^{J+1}(\mathcal{N}^{J+1}_{[t_J,t^{**}]}).$$

We summarize the idea of the proof. It only remains to establish the bound $|h|\leqslant\eta_{\mathrm{lin}}$. For points that are far enough away (compared to A) from the neck-like boundary of $\mathcal{N}\cup\mathcal{N}^{J+1}$, we have $Q^*\leqslant 10\overline{Q}^*$ from (12.42), and together with the lower bound $\rho_1 > \lambda r_{\mathrm{comp}}$ on $\mathcal{N}\cup\mathcal{N}^{J+1}$ from a-priori assumption (APA2), this implies $|h|\leqslant\eta_{\mathrm{lin}}$. On the other hand, points that are close to this boundary are far from the cuts, by Lemma 8.19. So, at these points we may rely instead on the bound $Q\leqslant 10\overline{Q}$ from (12.40). This bound implies $|h|\leqslant\eta_{\mathrm{lin}}$ as long as $\rho_1 \geqslant \frac{1}{10} r_{\mathrm{comp}}$, a fact which follows from the neck-like structure of the boundary of $\mathcal{N}^{J+1}$ and almost non-negative curvature (see Lemma 8.11).

Proof. The second part of a-priori assumption (APA6) follows from assertion (b) of Lemma 12.14 and the induction hypothesis. So, it remains to prove the bound $|h|\leqslant\eta_{\mathrm{lin}}$. To this end, consider a point $x\in\mathcal{N}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}$ and set $t:=\mathfrak{t}(x)$. Our goal will be to show $|h(x)|\leqslant\eta_{\mathrm{lin}}$. In the case $t\in[0,t_J]$, we are done by a-priori assumption (APA6) from our induction hypothesis, and the fact that $\mathcal{N}_{t_J+}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}\subset\mathcal{N}_{t_J-}$. So, assume that $t\in(t_J,t^{**}]$ and therefore $x\in\mathcal{N}^{J+1}_{(t_J,t^{**}]}$.

We now distinguish the following two cases.

Case 1. $B(x,10A\rho_1(x))\subset\mathcal{N}_t^{J+1}=\mathcal{N}_{t-}$.

In this case we can apply (12.42) and obtain that

$$e^{H(T-t)}\rho_1^3(x)|h(x)| = Q^*(x) \leqslant 10\overline{Q}^* = \eta_{\rm lin}(\lambda r_{\rm comp})^3.$$

Since by a-priori assumption (APA2) and assumption (12.39) we have $\rho_1(x)>\lambda r_{\rm comp}$ and $t\leqslant t_{J+1}\leqslant T$, this implies $|h(x)|\leqslant\eta_{\rm lin}$.

Case 2. $B(x,10A\rho_1(x))\not\subset\mathcal{N}_t^{J+1}$.

Let us first apply Lemma 8.19 along with a-priori assumption (APA11). We obtain that, if

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},\quad \lambda\leqslant\bar\lambda,\quad \Lambda\geqslant\underline{\Lambda},\quad \delta_{\rm b}\leqslant\bar\delta_{\rm b}(\lambda,D_{\rm cut},A,\Lambda),$$
$$\varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}(\lambda,D_{\rm cut},A,\Lambda),\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

then $P(x,10A\rho_1(x))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in{\rm Cut}\cup{\rm Cut}^J$. So, by (12.40), we have

$$e^{H(T-t)}\rho_1^E(x)|h(x)| = Q(x) \leqslant 10\overline{Q} = 10^{-E}\eta_{\rm lin}r_{\rm comp}^E.$$

By assumption (12.39), we have $t\leqslant t_{J+1}\leqslant T$. So, in order to show that $|h(x)|\leqslant\eta_{\rm lin}$, it suffices to verify the bound

$$\rho_1(x)\geqslant\tfrac{1}{10}r_{\rm comp}. \tag{12.52}$$

To see that (12.52) holds, choose first some point $y\in\partial\mathcal{N}_t^{J+1}$ with $d_t(x,y)<10A\rho_1(x)$. Let $\Sigma\subset\partial\mathcal{N}_t^{J+1}$ be the (spherical) boundary component of $\mathcal{N}_t^{J+1}$ that contains y. Consider the constant $\delta_0>0$ from Lemma 8.11. If

$$\delta_{\rm n}\leqslant\bar\delta_{\rm n},\quad \varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}\quad\text{and}\quad r_{\rm comp}\leqslant\bar r_{\rm comp},$$

then, by a-priori assumption (APA3) (a) and Lemma 8.12, the component Σ has to be a central 2-sphere of a δ_0-neck in $\mathcal{M}_t$ at scale $ar_{\rm comp}$ for some $a\in[1,2]$ and we must have

$$0.9r_{\rm comp}<\rho_1(y)<2.1r_{\rm comp}.$$

By bounded curvature at bounded distance, Lemma 8.6, along with a-priori assumption (APA2), applied at x, and assuming

$$\varepsilon_{\rm can}\leqslant\bar\varepsilon_{\rm can}(\lambda,A),$$

we find that $\rho_1(x)<C'\rho_1(y)<2.1C'r_{\mathrm{comp}}$ for some $C'=C'(A)<\infty$. So

$$d_t(x,y)<10A\rho_1(x)<21C'Ar_{\mathrm{comp}}.$$

Let $Y_\#<\infty$ be a constant whose value we will fix at the end of the proof. By a-priori assumption (APA3) (c), we can pick a $\Lambda r_{\mathrm{comp}}$-thick point $z\in\mathcal{N}_{t_{J+1}-}$ in the same component of $\mathcal{N}_{t_{J+1}-}$ as $\Sigma(t_{J+1})$. By a-priori assumption (APA3) (a) and bounded curvature at bounded distance, Lemma 8.6, applied at all points on $\partial\mathcal{N}_{t_{J+1}-}$, and assuming

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \Lambda\geqslant\underline{\Lambda}(Y_\#)\quad\text{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(Y_\#),$$

we obtain that $d_{t_J}(z(t),\partial\mathcal{N}_{t_J+})>Y_\#r_{\mathrm{comp}}$.

By a-priori assumption (APA2) and a distance distortion estimate, it follows that then

$$d_t(z(t),\partial\mathcal{N}_t^{J+1})>e^{-C''\lambda^{-2}}Y_\#r_{\mathrm{comp}}$$

for some universal constant $C''<\infty$. We can then apply Lemma 8.11, assuming that

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad Y_\#\geqslant\underline{Y}_\#(\lambda,A)\quad\text{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(A),$$

to show that $\rho_1(x)\geqslant\frac{1}{10}ar_{\mathrm{comp}}\geqslant\frac{1}{10}r_{\mathrm{comp}}$. So, (12.52) holds. □

Next, we establish a-priori assumption (APA7).

LEMMA 12.17. (Verification of (APA7)) *If*

$$E\geqslant\underline{E},\quad F\geqslant\underline{F}(E),\quad H\geqslant\underline{H}(E),\quad \eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}}(E),$$
$$\nu\leqslant\bar{\nu}(E),\quad \delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda},\quad A\geqslant\underline{A}(E,W),\quad \Lambda\geqslant\underline{\Lambda},$$
$$\delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(E,\lambda,D_{\mathrm{cut}},A,\Lambda),\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(E,\lambda,D_{\mathrm{cut}},W,A,\Lambda),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

then a-priori assumption (APA7) *holds. In other words, for all*

$$x\in\mathcal{N}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}$$

for which $P(x,A\rho_1(x))\cap\mathcal{D}=\varnothing$ *for all* $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$, *we have*

$$Q(x)\leqslant\overline{Q}.\tag{12.53}$$

The strategy of the proof is the following: Near the neck-like boundary of $\mathcal{N}$, the bound (12.53) is a direct consequence of assertion (c) of Lemma 12.14. So, it remains to consider points that are far away from this neck-like boundary. If a relaxed bound of the

form $Q \leqslant 10\overline{Q}$ holds on a parabolic neighborhood of size comparable to $L(E)$ around such a point, either via a-priori assumption (APA7) or (12.40), then we can use the semi-local maximum principle, Proposition 9.1, and a-priori assumption (APA12) to improve this bound by a factor of $\frac{1}{10}$. On the other hand, points for which such a relaxed bound is absent in such a parabolic neighborhood must be close enough to a cut, and thus even farther from the neck-like boundary. In this case, we can guarantee a bound of the form $Q \leqslant 10W\overline{Q}$ by either (APA8) or (12.41) on an even larger parabolic neighborhood, of size comparable to A. The bound (12.53) then follows from the interior decay estimate, Proposition 9.2 and a-priori assumption (APA12), for large enough A.

Proof. Let $x \in \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \text{and assume that } P(x, A\rho_1(x)) \cap \mathcal{D} = \varnothing$ for all $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$. Set $t := \mathfrak{t}(x)$. Our goal will be to show that $Q(x) \leqslant \overline{Q}$. By a-priori assumption (APA7) from our induction hypothesis, we only need to consider the case $t > t_J$ and $x \in \mathcal{N}^{J+1}_{(t_J,t^{**}]}$.

Let $L = L(E) < \infty$ be the constant from Proposition 9.1. By Lemma 8.7 and a-priori assumption (APA2), and assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L(E), \lambda, A),$$

we can find a constant $A' = A'(L(E), A) < \infty$ with $A' \geqslant \max\{A, L\}$ such that $P(x, A'\rho_1(x))$ is unscathed and

$$P(y, 10A\rho_1(y)) \subset P(x, A'\rho_1(x)) \quad \text{for all } y \in P(x, L\rho_1(x)). \tag{12.54}$$

We now distinguish two cases.

Case 1. $B(x, L\rho_1(x)) \not\subset \mathcal{N}_t^{J+1}$.

The goal in this case will be to apply assertion (c) of Lemma 12.14. To do this, we first need to bound $\rho_1(x)$ from above. For this purpose, choose $z \in \partial\mathcal{N}_t^{J+1}$ such that $d_t(x, z) < L\rho_1(x)$. By a-priori assumption (APA3) (a) and Lemma 8.12, and assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}} \quad \text{and} \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

we know that z is a center of a sufficiently precise neck $U \subset \mathcal{M}_t$ at scale ar_{comp} for some $a \in [1, 2]$ such that $\rho_1(z) < 2.1r_{\mathrm{comp}}$. By bounded curvature at bounded distance, Lemma 8.6, and assuming

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L(E)),$$

we therefore obtain $\rho_1(x) < Cr_{\mathrm{comp}}$ for some $C = C(L(E)) < \infty$. Thus,

$$d_t(x, \partial\mathcal{N}_t^{J+1}) \leqslant d_t(x, z) < CLr_{\mathrm{comp}}.$$

We can now apply assertion (c) of Lemma 12.14, assuming

$$F > C(L(E))\cdot L(E),$$

and obtain that $Q(x)\leqslant\overline{Q}$.

Case 2. $B(x, L\rho_1(x))\subset\mathcal{N}_t^{J+1}$.
We distinguish two subcases.

Case 2a. $P(x, A'\rho_1(x))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$.

Recall that $P(x, L\rho_1(x))\subset P(x, A'\rho_1(x))$. So, by Lemma 8.18, we have

$$P(x, L\rho_1(x)) \subset \mathcal{N}_{[0,t^{**}]}\setminus \bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J} \mathcal{D}.$$

Using the assumption of Case 2a, (12.54), a-priori assumption (APA7) from the induction hypothesis and (12.40), we obtain $Q\leqslant 10\overline{Q}$ on $P(x, L\rho_1(x))$. By Lemma 12.16, we have $|h|\leqslant\eta_{\mathrm{lin}}$ on $P(x, L\rho_1(x))$. If $P(x, L\rho_1(x))$ intersects the initial time-slice $\mathcal{M}_0$, then a-priori assumption (APA12) also implies that $Q\leqslant\nu\overline{Q}$ on $P(x, L\rho_1(x))\cap\mathcal{M}_0$. So, by Proposition 9.1, a-priori assumption (APA2), and assuming that

$$E > 2, \quad H \geqslant \underline{H}(E), \quad \eta_{\mathrm{lin}} \leqslant \bar{\eta}_{\mathrm{lin}}(E), \quad \nu \leqslant \bar{\nu}(E) \quad \text{and} \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(E, \lambda),$$

we obtain the improved estimate $Q(x)\leqslant\overline{Q}$.

Case 2b. $P(x, A'\rho_1(x))\cap\mathcal{D}\neq\varnothing$ for some $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$.

Applying Lemma 8.19 with $A_0=A'$ and a-priori assumption (APA11), and assuming

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A'(E, A), \Lambda),$$
$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A'(E, A), \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

we find that $B(x, A\rho_1(x))\subset\mathcal{N}$. Recall moreover that by assumption of the lemma we have $P(x, A\rho_1(x))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$. Therefore, again by Lemma 8.18, we obtain that

$$P(x, A\rho_1(x)) \subset \mathcal{N}_{[0,t^{**}]}\setminus \bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J} \mathcal{D}.$$

By a-priori assumption (APA8) from the induction hypothesis and (12.41), we have $Q\leqslant 10W\overline{Q}$ on $P(x, A\rho_1(x))$. We will now apply Proposition 9.2 to $P(x, A\rho_1(x))$ in order to improve this estimate at x. To do this, observe that, by Lemma 12.16 we have $|h|\leqslant\eta_{\mathrm{lin}}$ on $P(x, A\rho_1(x))$ and if $P(x, A\rho_1(x))$ intersects the initial time-slice $\mathcal{M}_0$, then a-priori

assumption (APA12) implies that $Q \leqslant \nu \overline{Q}$ on $P(x, A\rho_1(x)) \cap \mathcal{M}_0$. We can therefore apply Proposition 9.2 to $P(x, A\rho_1(x))$, along with a-priori assumption (APA2), assuming that

$$E > 2, \quad H \geqslant \underline{H}(E), \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}(E), \quad \nu \leqslant \bar{\nu}(E),$$
$$A \geqslant \underline{A}(E, W), \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(E, \lambda, W),$$

and conclude that $Q(x) \leqslant \overline{Q}$. This finishes the proof. □

Next, we verify a-priori assumption (APA8).

LEMMA 12.18. (Verification of (APA8)) *If*

$$E \geqslant \underline{E}, \quad H \geqslant \underline{H}(E), \quad \eta_{\text{lin}} \leqslant \bar{\eta}_{\text{lin}}(E), \quad \nu \leqslant \bar{\nu}(E),$$
$$\delta_{\text{n}} \leqslant \bar{\delta}_{\text{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad W \geqslant \underline{W}(E, \lambda, D_{\text{cut}}) \quad A \geqslant \underline{A}(E), \quad \Lambda \geqslant \underline{\Lambda},$$
$$\delta_{\text{b}} \leqslant \bar{\delta}_{\text{b}}(\lambda, D_{\text{cut}}, A, \Lambda), \quad \varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}}(E, \lambda, D_{\text{cut}}, A, \Lambda), \quad r_{\text{comp}} \leqslant \bar{r}_{\text{comp}},$$

then a-priori assumption (APA8) *holds. In other words, we have*

$$Q \leqslant W\overline{Q} \quad \text{on } \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \text{Cut} \cup \text{Cut}^J} \mathcal{D}. \tag{12.55}$$

Note that a main aspect of this lemma is that W does not depend on A. Otherwise the inequality (12.55) would follow easily from (12.40) and (12.42). More specifically, at points whose distance to an extension cap is bounded in terms of A, we can only use (12.42) to obtain a bound on Q. However, the "conversion" factor between Q^* and Q at such a point depends on A. So, the bound (12.42) cannot be used directly to verify (12.55).

The idea of the proof is the following. We may focus on the time-slab $\mathcal{N}^{J+1}_{[t_J, t^{**}]}$, since the bound (12.55) follows from a-priori assumption (APA8) of the induction hypothesis. The bound (12.55) follows from (12.40) (the relaxed version of (APA7)) at points that are far away from the cuts, i.e. at distance comparable to A. For points that are close to the cuts, we distinguish two cases. The strategy in the first case is to deduce (12.55) from a-priori assumption (APA8) and its relaxed version (12.41) via the semi-local maximum principle (Proposition 9.1). This argument only works at points that are still sufficiently far away from the cuts, this time with separation comparable to $L(E) \ll A$. In the second case, we consider points that are close to cuts, comparable to $L(E)$. At these points (12.42) (the relaxed version of (APA9)) guarantees a bound of the form $Q^* \leqslant 10\overline{Q}^*$. This bound translates into a bound on Q and the conversion factor can be controlled in terms of $L(E)$, E, λ and D_{cut}. So, (12.55) follows as long we choose W larger than this conversion factor.

Proof. Consider a point $x\in\mathcal{N}^{J+1}_{[t_J,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}^J}\mathcal{D}$. Note that the case when

$$t:=\mathfrak{t}(x)=t_J$$

follows from the induction hypothesis, so we assume in the following that $t>t_J$.

We distinguish the following cases.

Case 1. $B(x,10A\rho_1(x))\not\subset\mathcal{N}^{J+1}_t$.

Then, we can apply Lemma 8.19 along with a-priori assumption (APA11), assuming that

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad \lambda\leqslant\bar{\lambda},\quad \Lambda\geqslant\underline{\Lambda},\quad \delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda,D_{\mathrm{cut}},A,\Lambda),$$
$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,D_{\mathrm{cut}},A,\Lambda),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}},$$

and obtain that $P(x,10A\rho_1(x))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$. So, by (12.40), we have $Q(x)\leqslant 10\overline{Q}$. Therefore, $Q(x)\leqslant W\overline{Q}$, as long as

$$W\geqslant 10.$$

Case 2. $B(x,10A\rho_1(x))\subset\mathcal{N}^{J+1}_t$.

Choose $L=L(E)$ from Proposition 9.1. We distinguish two subcases.

Case 2*a*. $P(x,L\rho_1(x))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$.

Assume that

$$10A\geqslant L(E).$$

So, $B(x,L\rho_1(x))\subset\mathcal{N}$ and thus, by Lemma 8.18,

$$P(x,L\rho_1(x))\subset\mathcal{N}^{J+1}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}.$$

Let us now apply Proposition 9.1 to $P(x,L\rho_1(x))$. To do this, note that by Lemma 12.16, a-priori assumption (APA8) from the induction hypothesis and (12.41), we know that $|h|\leqslant\eta_{\mathrm{lin}}$ and $Q\leqslant 10W\overline{Q}$ on $P(x,L\rho_1(x))$. If $P(x,L\rho_1(x))$ intersects $\mathcal{M}_0$, then, by a-priori assumption (APA12), we also have $Q\leqslant\nu\overline{Q}$ on the intersection. Lastly, by a-priori assumption (APA2), we have $\rho_1(x)>\lambda r_{\mathrm{comp}}$. So, assuming

$$E>2,\quad H\geqslant\underline{H}(E),\quad \nu\leqslant\bar{\nu}(E),\quad \eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}}(E)\quad\text{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(E,\lambda),$$

we obtain from Proposition 9.1 that $Q(x)\leqslant W\overline{Q}$, as desired.

Case 2b. $P(x, L\rho_1(x)) \cap \mathcal{D} \neq \varnothing$ for some $\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J$.

By a-priori assumptions (APA5) and (APA11), and assuming

$$\delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}),$$

we can conclude that there is a constant $C' = C'(\lambda, D_{\mathrm{cut}}) < \infty$ such that $\rho_1 \leqslant C' r_{\mathrm{comp}}$ on $\mathcal{D}$. Next, by bounded curvature at bounded distance, Lemma 8.6, applied at x, a-priori assumption (APA2), and assuming that

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L(E), \lambda),$$

we obtain a constant $C'' = C''(L(E)) < \infty$ such that

$$\rho_1(x) \leqslant C'' C' r_{\mathrm{comp}}.$$

Since $B(x, 10A\rho_1(x)) \subset \mathcal{N}_t^{J+1}$, we obtain from (12.42) that $Q^*(x) \leqslant 10\overline{Q}^*$. Assuming

$$E \geqslant 3 \quad \text{and} \quad \lambda \leqslant 1,$$

we can now convert this bound to a bound on $Q(x)$ as follows:

$$\begin{aligned} Q(x) = \rho_1^{E-3}(x) Q^*(x) &\leqslant \rho_1^{E-3}(x) 10 \overline{Q}^* \\ &\leqslant (C''(L(E)) C'(\lambda, D_{\mathrm{cut}}) r_{\mathrm{comp}})^{E-3} \eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3 \\ &\leqslant 10^{E+1} (C''(L(E)) C'(\lambda, D_{\mathrm{cut}}))^{E-3} 10^{-E-1} \eta_{\mathrm{lin}} r_{\mathrm{comp}}^E \\ &\leqslant 10^{E+1} (C''(L(E)) C'(\lambda, D_{\mathrm{cut}}))^{E-3} \overline{Q}. \end{aligned}$$

So, $Q(x) \leqslant W \overline{Q}$, as long as

$$W \geqslant \underline{W}(E, \lambda, D_{\mathrm{cut}}).$$

This finishes the proof. □

Lastly, we establish a-priori assumption (APA9).

LEMMA 12.19. (Verification of (APA9)) *If*

$$\begin{gathered} E \geqslant \underline{E}, \quad H \geqslant \underline{H}, \quad \eta_{\mathrm{lin}} \leqslant \bar{\eta}_{\mathrm{lin}}, \quad \nu \leqslant \bar{\nu}, \quad \delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \\ \eta_{\mathrm{cut}} \leqslant \bar{\eta}_{\mathrm{cut}}, \quad D_{\mathrm{cut}} \geqslant \underline{D}_{\mathrm{cut}}(\lambda), \quad A \geqslant \underline{A}(E, \lambda), \quad \Lambda \geqslant \underline{\Lambda}, \\ \delta_{\mathrm{b}} \leqslant \bar{\delta}_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A, \Lambda), \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(E, \lambda, D_{\mathrm{cut}}, A, \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}}(\lambda), \end{gathered}$$

then a-priori assumption (APA9) *holds. In other words, we have*

$$Q^*(x) \leqslant \overline{Q}^* = 10^{-1} \eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3 \tag{12.56}$$

for all $x \in \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}$ *for which* $B(x, A\rho_1(x)) \subset \mathcal{N}_{\mathfrak{t}(x)-}$.

Let us first summarize the strategy of the proof. As in the previous proofs, the semi-local maximum principle, Proposition 9.1, can be used to deduce (12.56) from a-priori assumption (APA9) or its relaxed version (12.42) at points that are sufficiently far away from the cuts and the neck-like boundary of $\mathcal{N}$. Now, consider points that are close to the neck-like boundary, but far enough (comparably to A) from this boundary such that the assertion does not become vacuous. At such points, we use the bound $Q \leqslant 10\overline{Q}$ from a-priori assumption (APA7) and its relaxed version (12.40) and the interior decay estimate, Proposition 9.2, to overcome the conversion factor between Q and Q^* for sufficiently large A. Lastly, consider points that are close to a cut. At such points, we invoke the semi-local maximum principle, Proposition 9.1 with initial condition, on a truncated parabolic neighborhood whose initial time-slice intersects the cut. We then use a-priori assumption (APA10) or (APA7) to deduce a very good bound for Q^* on this initial time-slice. Proposition 9.1 then implies (12.56).

Proof. Consider a point

$$x \in \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}$$

such that $B(x, A\rho_1(x)) \subset \mathcal{N}_{\mathfrak{t}(x)-}$ and set $t := \mathfrak{t}(x)$. The case $t \leqslant t_J$ follows from a-priori assumption (APA9) of the induction hypothesis. So, in the following we assume that $t > t_J$ and therefore that $B(x, A\rho_1(x)) \subset \mathcal{N}_t^{J+1}$.

Let $L = L(3)$ be the constant from Proposition 9.1 (for $E = 3$). Using Lemma 8.7, a-priori assumption (APA2), and assuming that

$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(L, \lambda, A),$$

we can find a constant $A' = A'(A) < \infty$ with $A' \geqslant A$ such that the parabolic neighborhood $P(x, A'\rho_1(x))$ is unscathed and such that

$$P(y, 10A\rho_1(y)) \subset P(x, A'\rho_1(x)) \quad \text{for all } y \in P(x, A\rho_1(x)). \tag{12.57}$$

Let us now distinguish three cases.

Case 1. We have

$$P(x, A'\rho_1(x), -(L\rho_1(x))^2) \subset \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}$$

and $P(x, A'\rho_1(x), -(L\rho_1(x))^2)$ does not intersect the initial time-slice $\mathcal{M}_0$.

So, assuming $A \geqslant L$, we have

$$P(x, L\rho_1(x)) \subset \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}.$$

Using (12.57), a-priori assumption (APA9), (12.42) and Lemma 12.16, we find that

$$Q^* \leqslant 10\overline{Q}^* \quad \text{and} \quad |h| \leqslant \bar{\eta}_{\mathrm{lin}}$$

on $P(x, L\rho_1(x))$. Since the exponent in the definition of Q^* is $3>2$, if

$$H \geqslant \underline{H}, \quad \eta_{\mathrm{lin}} \leqslant \bar{\eta}_{\mathrm{lin}} \quad \text{and} \quad \varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda),$$

we may apply the semi-local maximum principle, Proposition 9.1, to deduce that

$$Q^*(x) \leqslant \overline{Q}^*,$$

which finishes the proof in this case. Note that here we have used a-priori assumption (APA2).

Case 2. We have

$$B(x, A'\rho_1(x)) \not\subset \mathcal{N}.$$

By Lemma 8.19 and a-priori assumption (APA11), and assuming that

$$\delta_{\mathrm{n}} \leqslant \bar{\delta}_{\mathrm{n}}, \quad \lambda \leqslant \bar{\lambda}, \quad \Lambda \geqslant \underline{\Lambda}, \quad \delta_{\mathrm{b}} \leqslant \delta_{\mathrm{b}}(\lambda, D_{\mathrm{cut}}, A'(A), \Lambda),$$
$$\varepsilon_{\mathrm{can}} \leqslant \bar{\varepsilon}_{\mathrm{can}}(\lambda, D_{\mathrm{cut}}, A'(A), \Lambda), \quad r_{\mathrm{comp}} \leqslant \bar{r}_{\mathrm{comp}},$$

we find that

$$P(x, A'\rho_1(x)) \cap \mathcal{D} = \varnothing \quad \text{for all } \mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J. \tag{12.58}$$

So, by Lemma 8.18,

$$P(x, A\rho_1(x)) \subset \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J} \mathcal{D}. \tag{12.59}$$

Combining (12.58) with (12.57), we obtain that, for all $y \in P(x, A\rho_1(x))$,

$$P(y, 10A\rho_1(y)) \cap \mathcal{D} = \varnothing \quad \text{for all} \quad \mathcal{D} \in \mathrm{Cut} \cup \mathrm{Cut}^J.$$

Therefore, by a-priori assumption (APA7) and (12.40), we obtain that

$$Q \leqslant 10\overline{Q} \quad \text{on } P(x, A\rho_1(x)).$$

Let us now convert this bound into a bound on Q^*. There are two ways of doing this. One way would be to use Lemma 8.11, as in the proof of Lemma 12.16 leading up to (12.52) to show that $\rho_1(x)\geqslant\frac{1}{10}r_{\mathrm{comp}}$. In the following, however, we will use a different strategy, as it is technically easier.

Assuming

$$E\geqslant 3$$

and using a-priori assumption (APA2) and (12.59), we have, on $P(x,A\rho_1(x))$,

$$\begin{aligned}Q^*=\rho_1^{3-E}Q&\leqslant(\lambda r_{\mathrm{comp}})^{3-E}\cdot 10\overline{Q}=(\lambda r_{\mathrm{comp}})^{3-E}\cdot 10^{-E}\eta_{\mathrm{lin}}r_{\mathrm{comp}}^E\\&\leqslant\lambda^{-E}\eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3=10\lambda^{-E}\overline{Q}^*.\end{aligned}$$

We will now apply Proposition 9.2 to Q^* on $P(x,A\rho_1(x))$. To do this, observe that, by Lemma 12.16, we have $|h|\leqslant\eta_{\mathrm{lin}}$ on $P(x,A\rho_1(x))$. In addition, if $P(x,A\rho_1(x))$ intersects the initial time-slice $\mathcal{M}_0$, then by a-priori assumption (APA12) we have $Q^*\leqslant\nu\overline{Q}^*$ on the intersection. We also have $\rho_1(x)>\lambda r_{\mathrm{comp}}$ by a-priori assumption (APA2). So, if

$$H\geqslant\underline{H},\quad \eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}},\quad \nu\leqslant\bar{\nu},\quad A\geqslant\underline{A}(E,\lambda)\quad\text{and}\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(E,\lambda),$$

then we obtain that $Q^*(x)\leqslant\overline{Q}^*$, as desired.

Case 3. We have

$$B(x,A'\rho_1(x))\subset\mathcal{N},$$

and either

$$P(x,A'\rho_1(x),-(L\rho_1(x))^2)\not\subset\mathcal{N}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D},$$

or $P(x,A'\rho_1(x),-(L\rho_1(x))^2)$ intersects the initial time-slice $\mathcal{M}_0$.

In the following we will use the notation

$$\mathrm{Cut}^j=\{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J:\mathcal{D}\subset\mathcal{M}_{t_j}\}.$$

Choose $j_0\in\{1,...,J\}$ maximal with the property that

$$P(x,A'\rho_1(x),-(t-t_{j_0}))\not\subset\mathcal{N}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}.$$

If no such j_0 exists, then set $j_0:=0$. By Lemma 8.18, we have

$$P(x,A'\rho_1(x),-(t-t'))\subset\mathcal{N}_{[0,t^{**}]}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J}\mathcal{D}\quad\text{for all } t'\in(t_{j_0},t].$$

Letting $t' \to t_{j_0}$ and using the fact that $\mathcal{N}_{[0,t^{**}]}$ is a closed subset of $\mathcal{M}$, we obtain

$$P(x, A'\rho_1(x), -(t-t_{j_0})) \subset \mathcal{N}_{[0,t^{**}]} \setminus \bigcup_{\mathcal{D} \in \mathrm{Cut}^{j_0+1} \cup \ldots \cup \mathrm{Cut}^{J}} \mathcal{D}, \tag{12.60}$$

and either $j_0 = 0$ or there is a cut $\mathcal{D}_0 \in \mathrm{Cut}^{j_0}$ such that

$$P(x, A'\rho_1(x), -(t-t_{j_0})) \cap \mathcal{D}_0 \neq \varnothing. \tag{12.61}$$

Let

$$B_{t_{j_0}} := (B(x, L\rho_1(x)))(t_{j_0}).$$

In other words, $B_{t_{j_0}}$ is the initial time-slice of the parabolic neighborhood

$$P(x, L\rho_1(x), -(t-t_{j_0})).$$

Note that, by (12.61), the perturbation h is defined everywhere on

$$P(x, L\rho_1(x), -(t-t_{j_0})) \setminus B_{t_{j_0}},$$

and it can be smoothly extended to the entire parabolic neighborhood by setting $h = h_{t_{j_0}+}$ on $B_{t_{j_0}}$. Similarly, we can extend Q^* to the entire parabolic neighborhood

$$P(x, L\rho_1(x), -(t-t_{j_0}))$$

by setting $Q^* = Q^*_+$ on $B_{t_{j_0}}$.

We will now bound $Q^* = Q^*_+$ on $B_{t_{j_0}}$. Let $y \in B_{t_{j_0}}$. Then, the two cases indicated above lead to the following three subcases.

Case 3*a*. We have $j_0 = 0$ and therefore $y \in \mathcal{M}_0$.

In this case, by a-priori assumption (APA12), we have

$$Q^*(y) \leqslant \nu \overline{Q}^*.$$

Case 3*b*. We have $j_0 \geqslant 1$ and $y \in \mathcal{D}_0$.

In this case, a-priori assumption (APA10) yields

$$Q^*(y) \leqslant \eta_{\mathrm{cut}} \overline{Q}^*.$$

Case 3*c*. We have $j_0 \geqslant 1$ and $y \notin \mathcal{D}_0$.

Our strategy in this case is to use the bound on $Q(y)$ from a-priori assumption (APA7) and translate it into a bound on $Q^*(y)$. In order to do this, we need to ensure that a-priori assumption (APA7) apply at y (or slightly earlier) and that $\rho_1(y)$ is sufficiently large so that the Q-bound implies a good bound on Q^*.

Let $C_\#<\infty$ be a constant whose value we will determined at the end of the proof. We can now apply Lemma 8.20 and a-priori assumption (APA11), assuming

$$\delta_{\mathrm{n}}\leqslant\bar{\delta}_{\mathrm{n}},\quad D_{\mathrm{cut}}\geqslant\underline{D}_{\mathrm{cut}}(\lambda,C_\#),\quad \Lambda\geqslant\underline{\Lambda},\quad \delta_{\mathrm{b}}\leqslant\bar{\delta}_{\mathrm{b}}(\lambda,C_\#,D_{\mathrm{cut}},A'(A),\Lambda),$$
$$\varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda,D_{\mathrm{cut}},A'(A),\Lambda),\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}(C_\#),$$

to find that

$$\rho_1(y)\geqslant C_\# r_{\mathrm{comp}}$$

and $P(y,2A\rho_1(y))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}^1\cup\ldots\cup\mathrm{Cut}^{j_0-1}$. So, for any $t'\in[t_{j_0-1},t_{j_0})$, sufficiently close to t_{j_0}, we have $P(y(t'),A\rho_1(y(t')))\cap\mathcal{D}=\varnothing$ for all $\mathcal{D}\in\mathrm{Cut}\cup\mathrm{Cut}^J$. So, by a-priori assumption (APA7) we have $Q(y(t'))\leqslant\overline{Q}$. Letting $t'\to t_{j_0}$ yields $Q(y)\leqslant\overline{Q}$.

Assuming

$$E\geqslant 4,$$

we obtain

$$\begin{aligned}Q^*(y)&=\rho_1^{3-E}(y)Q(y)\leqslant(C_\# r_{\mathrm{comp}})^{3-E}\cdot\overline{Q}\\&=(C_\# r_{\mathrm{comp}})^{3-E}10^{-E-1}\eta_{\mathrm{lin}}r_{\mathrm{comp}}^E\leqslant C_\#^{-1}\lambda^{-3}10^{-E-1}\eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3\leqslant C_\#^{-1}\lambda^{-3}\overline{Q}^*.\end{aligned}$$

Summarizing the results of Cases (3a)–(3c), we obtain that

$$Q^*\leqslant(\nu+\eta_{\mathrm{cut}}+C_\#^{-1}\lambda^{-3})\overline{Q}^*\quad\text{on } B_{t_{j_0}}.$$

Similarly, as in Case (1), we can use (12.57) and (12.60) together with a-priori assumption (APA9) and (12.42) to show that $Q^*\leqslant 10\overline{Q}^*$ on $P(x,L\rho_1(x),-(t-t_{j_0}))$. By Lemma 12.16, we have $|h|\leqslant\eta_{\mathrm{lin}}$ on $P(x,L\rho_1(x),-(t-t_{j_0}))$. We can now apply Proposition 9.1 along with a-priori assumption (APA2), assuming

$$H\geqslant\underline{H},\quad \eta_{\mathrm{lin}}\leqslant\bar{\eta}_{\mathrm{lin}},\quad \nu\leqslant\bar{\nu},\quad C_\#\geqslant\underline{C}_\#(\lambda),$$
$$\eta_{\mathrm{cut}}\leqslant\bar{\eta}_{\mathrm{cut}},\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\lambda),$$

to show that $Q^*(x)\leqslant\overline{Q}^*$, as desired. □

We can finally finish the proof of Proposition 12.1. Lemmas 12.16–12.19 imply that

$$(\mathcal{N}_{[0,t^{**}]},\{\mathcal{N}^j,\mathcal{N}^{J+1}_{[t_J,t^{**}]}\}_{j=1}^J,\{t_j,t^{**}\}_{j=1}^J)$$

and

$$(\mathrm{Cut}\cup\mathrm{Cut}^J,\bar\phi|_{\mathcal{N}_{[0,t^{**}]}},\{\phi^j,\phi^{J+1}|_{\mathcal{N}^{J+1}_{[t_J,t^{**}]}}\}_{j=1}^J)$$

satisfy (APA6)–(APA9) whenever (12.40)–(12.42) hold, up to time t^{**}. So, by iterating Lemma 12.15, we may choose $t^{**}=t^*$. Since a-priori assumption (APA6) holds for the aforementioned comparison domain and comparison, we have $|h|\leqslant\eta_{\mathrm{lin}}$ on $\mathcal{N}^{J+1}_{t^*}$. So, by assertion (d) of Lemma 12.14, we obtain that $t^{**}=t^*=t_{J+1}$. So,

$$(\mathcal{N},\{\mathcal{N}^j\}_{j=1}^{J+1},\{t_j\}_{j=1}^{J+1})\quad\text{and}\quad(\mathrm{Cut}\cup\mathrm{Cut}^J,\bar\phi,\{\phi^j\}_{j=1}^{J+1})$$

satisfy (APA1) and (APA6)–(APA9). A-priori assumptions (APA10)–(APA13) follow from assertion (a) of Lemma 12.14. Recall that (APA2)–(APA5) hold by the assumptions of Proposition 12.1.

Lastly note that the auxiliary parameter F was assumed to be large depending only on E. So, it is straight forward to check that the assumptions of the parameters imposed in the course of this proof all follow from (12.1). This finishes the proof of Proposition 12.1.

13. Proofs of the main results

In this section we will combine the main results of §11 and §12 to prove the main result of the paper, Theorem 13.1. We then prove some corollaries, including several stability results and a uniqueness theorem, as presented in §1.3.

THEOREM 13.1. (Existence of comparison domain and comparison) *If*

$$\begin{gathered}T>0,\quad E\geqslant\underline{E},\quad H\geqslant\underline{H}(E),\quad \eta_{\mathrm{lin}}\leqslant\bar\eta_{\mathrm{lin}}(E),\\ \nu\leqslant\bar\nu(T,H,\eta_{\mathrm{lin}}),\quad \delta_{\mathrm{n}}\leqslant\bar\delta_{\mathrm{n}}(T,H,\eta_{\mathrm{lin}}),\quad \lambda\leqslant\bar\lambda(\delta_{\mathrm{n}}),\\ D_{\mathrm{cap}}\geqslant\underline{D}_{\mathrm{cap}}(\lambda),\quad \eta_{\mathrm{cut}}\leqslant\bar\eta_{\mathrm{cut}},\quad D_{\mathrm{cut}}\geqslant\underline{D}_{\mathrm{cut}}(D_{\mathrm{cap}},\eta_{\mathrm{cut}}),\\ W\geqslant\underline{W}(D_{\mathrm{cut}}),\quad A\geqslant\underline{A}(W),\quad \Lambda\geqslant\underline{\Lambda}(A),\quad \delta_{\mathrm{b}}\leqslant\bar\delta_{\mathrm{b}}(\Lambda),\\ \varepsilon_{\mathrm{can}}\leqslant\bar\varepsilon_{\mathrm{can}}(\delta_{\mathrm{b}}),\quad r_{\mathrm{comp}}\leqslant\bar r_{\mathrm{comp}}(\Lambda),\end{gathered}$$

then the following holds.

Consider two $(\varepsilon_{\mathrm{can}}r_{\mathrm{comp}},T)$-complete Ricci flow spacetimes $\mathcal{M}$ and $\mathcal{M}'$ that each satisfy the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}}r_{\mathrm{comp}},1)$.

Let $\zeta\colon\{x\in\mathcal{M}_0:\rho(x)>\lambda r_{\mathrm{comp}}\}\to\mathcal{M}'_0$ be a diffeomorphism onto its image that satisfies the following bounds:

$$\begin{aligned}|\zeta^*g'_0-g_0|&\leqslant\eta_{\mathrm{lin}},\\ e^{HT}\rho_1^E|\zeta^*g'_0-g_0|&\leqslant\nu\overline{Q}=\nu\cdot10^{-E-1}\eta_{\mathrm{lin}}r_{\mathrm{comp}}^E,\\ e^{HT}\rho_1^3|\zeta^*g'_0-g_0|&\leqslant\nu\overline{Q}^*=\nu\cdot10^{-1}\eta_{\mathrm{lin}}(\lambda r_{\mathrm{comp}})^3.\end{aligned}$$

Assume moreover that the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds at scales $(0,1)$ on the image of ζ.

Then, for any $J\geqslant 1$ with $Jr_{\mathrm{comp}}^2\leqslant T$, there is a comparison domain

$$(\mathcal{N},\{\mathcal{N}^j\}_{j=1}^J,\{t_j\}_{j=0}^J)$$

and a comparison $(\mathrm{Cut},\phi,\{\phi^j\}_{j=1}^J)$ *from $\mathcal{M}$ to $\mathcal{M}'$ defined on this domain such that a-priori assumptions* (APA1)–(APA6) *hold for the tuple of parameters*

$$(\eta_{\mathrm{lin}},\delta_{\mathrm{n}},\lambda,D_{\mathrm{cap}},\Lambda,\delta_{\mathrm{b}},\varepsilon_{\mathrm{can}},r_{\mathrm{comp}})$$

and a-priori assumptions (APA7)–(APA13) *hold for the tuple of parameters*

$$(T,E,H,\eta_{\mathrm{lin}},\nu,\lambda,\eta_{\mathrm{cut}},D_{\mathrm{cut}},W,A,r_{\mathrm{comp}}).$$

Moreover, $\phi_{0+}=\phi_0^1=\zeta|_{\mathcal{N}_0}$.

Proof of Theorem 13.1. The theorem follows from Propositions 11.1 and 12.1 by induction on J. Both propositions can be applied under restrictions on the parameters that follow from the restrictions stated in the beginning of this theorem. Note that in the first step of the induction one applies Proposition 11.1 to produce the first time slab $\mathcal{N}^1$ of the comparison domain. By a-priori assumption (APA2) we have $\mathcal{N}_0^1\subset X:=\{x\in\mathcal{M}_0:\rho(x)>\lambda r_{\mathrm{comp}}\}$. Assuming

$$\lambda\leqslant\bar{\lambda}(\delta_{\mathrm{n}}),\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}(\delta_{\mathrm{n}}),$$

by (APA3) and Lemma 8.6 it follows that the $\delta_{\mathrm{n}}^{-1}r_{\mathrm{comp}}$-tubular neighborhood around $\mathcal{N}_0$ is contained in X. Hence, the map ζ from the assumptions of Theorem 13.1 satisfies assumption (vi) of Proposition 12.1.

Note that we have simplified the restrictions on the parameters in the first part of this theorem by omitting arguments in parameter restrictions if they have already appeared in earlier restrictions. This simplification does not change the nature of these restrictions. For example, since we have imposed the restriction $H\geqslant\underline{H}(E)$, we can assume without loss of generality that $H\geqslant E$. Therefore, it is not necessary to list E in the restriction for $\nu\leqslant\bar{\nu}(H,\eta_{\mathrm{lin}},T)$, as ν already depends on H. □

Next, we prove Theorem 1.5. This theorem is similar to Theorem 13.1; however the parameters associated with the a-priori assumptions have been suppressed. The proof of Theorem 1.5 requires the following result.

LEMMA 13.2. *If*

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}} \quad \text{and} \quad r \leqslant \bar{r},$$

then the following holds.

Let $\mathcal{M}$ be an $(\varepsilon_{\text{can}} r, T)$-complete Ricci flow spacetime that satisfies the ε_{can}-canonical neighborhood assumption at scales $(\varepsilon_{\text{can}} r, 1)$. Let $x \in \mathcal{M}_{[0,T]}$ be a point with $\rho(x) > r$. Then, there is a continuous path $\gamma\colon [0,1] \to \mathcal{M}_{[0,T]}$ between x and a point in $\mathcal{M}_0$ such that $\mathfrak{t} \circ \gamma$ is non-increasing and such that $\rho(\gamma(s)) > .9r$ for all $s \in [0,1]$.

Proof. A slightly different version of this statement, which would also be adequate for our needs here, was proven in [KL2, Proposition 3.5]. For completeness, we provide an alternative argument.

Set $t_0 := \mathfrak{t}(x)$ and $r_0 := \rho_1(x) > r$. By Lemma 8.5, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}} \quad \text{and} \quad r \leqslant \bar{r},$$

we know that x survives until time $\max\{t_0 - r_0^2, 0\}$ and $\rho_1(x(t)) \geqslant .95 r_0 > .95 r$ for all $t \in [\max\{t_0 - r_0^2, 0\}, t_0]$. So, if $t_0 \leqslant r_0^2$, then we are done. Consider now the case $t_0 > r_0^2$. If $r_0 \leqslant \frac{1}{2}$ and $\rho(x(t_0 - r_0^2)) \leqslant \rho(x)$, then we can use Lemma 8.15, assuming

$$\varepsilon_{\text{can}} \leqslant \bar{\varepsilon}_{\text{can}},$$

to show that $(\mathcal{M}_{t_0}, x(t_0 - r_0^2))$ is close enough to $(M_{\text{Bry}}, g_{\text{Bry}}, x_{\text{Bry}})$ such that there is a point $y \in \mathcal{M}_{t_0 - r_0^2}$ with $\rho(y) > \rho(x)$ and such that $x(t_0 - r_0^2)$ can be connected with y by a continuous path inside $\mathcal{M}_{t_0 - r_0^2}$ whose image only consists of $.9r_0$-thick points.

So, summarizing our conclusions, each $x \in \mathcal{M}_{[0,T]}$ can be connected with a point $y \in \mathcal{M}_{[0,T]}$ by a path $\gamma\colon [0,1] \to \mathcal{M}_{[0,T]}$ such that $\mathfrak{t} \circ \gamma$ is non-increasing and $\rho(\gamma(s)) > .9 r_0$ for all $s \in [0,1]$, and one of the following holds:

(1) $y \in \mathcal{M}_0$,
(2) $\rho(y) \geqslant .95 \rho(x) > .95 \cdot \frac{1}{2}$ and $\mathfrak{t}(y) = \mathfrak{t}(x) - \rho_1^2(x)$,
(3) $\rho(y) > \rho(x)$ and $\mathfrak{t}(y) = \mathfrak{t}(x) - \rho_1^2(x)$.

Iterating this process yields a sequence of points $x_0 = x, x_1, x_2, \ldots \in \mathcal{M}_{[0,T]}$ such that x_i and x_{i+1} can be connected by a path with the desired properties. It now remains to show that this sequence terminates at some index i and that $x_i \in \mathcal{M}_0$. To see this, note that by (1)–(3) the sequence of times $\mathfrak{t}(x_i)$ is non-increasing and $\rho(x_i) > r$, assuming

$$r \leqslant .95 \cdot \tfrac{1}{4}.$$

Since

$$\mathfrak{t}(x_{i+1}) = \mathfrak{t}(x_i) - \rho_1^2(x_i) \leqslant \mathfrak{t}(x_i) - r^2$$

in cases (2) and (3), the sequence must terminate after a finite number of steps. □

Proof of Theorem 1.5. Since we will invoke Theorem 13.1 below, in order to make the estimates in Theorem 1.5 conform more closely with those in Theorem 13.1, it will be convenient to prove the theorem for E replaced by $\frac{1}{2}E$, ϕ replaced by ζ and $\hat{\phi}$ replaced by ϕ. So, we assume that $\zeta: U \to U'$ satisfies

$$|\zeta^* g'_0 - g_0| \leqslant \varepsilon \cdot r^E (|\mathrm{Rm}|+1)^{E/2},$$

and our goal is to construct $\phi: \widehat{U} \to \widehat{U}'$ such that

$$|\phi^* g'_0 - g_0| \leqslant \delta \cdot r^E (|\mathrm{Rm}|+1)^{E/2}.$$

We will first prove a slightly weaker version of the theorem in which we allow $\varepsilon_{\mathrm{can}}$ to also depend on T. We will mention how we can remove this dependence at the end of this proof.

Fix T and $E \geqslant \underline{E}$, where $\underline{E}$ is the constant from Theorem 13.1 and assume that $\underline{E} \geqslant 3$. Based on these choices, fix constants H, η_{lin}, ν, δ_{n}, λ, D_{cap}, η_{cut}, D_{cut}, W, A, Λ, δ_{b}, $\varepsilon_{\mathrm{can}}$ and $\bar{r}_{\mathrm{comp}}$ that satisfy the restrictions stated in Theorem 13.1. Without loss of generality, we may assume that $\bar{r}_{\mathrm{comp}} \leqslant 1$. Choose $r_{\mathrm{comp}} := \alpha r \cdot \bar{r}_{\mathrm{comp}}$, where $0 < \alpha = \alpha(\delta, T, E) \leqslant 1$ is a constant whose value will be determined in the course of this proof.

We now verify the assumptions of Theorem 13.1. In what follows, we will be imposing several upper bounds on the parameters α and ε. The upper bounds on α will only depend on δ, T and E, and the upper bounds on ε will only depend on δ, T, E and α. As α will not be chosen depending on ε, there will be no circular dependence. At a number of steps in the following proof, we will also assume that the constants $\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}$ and $\varepsilon_{\mathrm{can}}$ have been chosen smaller than some universal constant.

Assuming

$$\varepsilon \leqslant \varepsilon_{\mathrm{can}} \cdot \alpha \bar{r}_{\mathrm{comp}},$$

we get that $\mathcal{M}, \mathcal{M}'$ are $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, T)$-complete and satisfy the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at scales $(\varepsilon_{\mathrm{can}} r_{\mathrm{comp}}, 1)$. If

$$\varepsilon \leqslant c_1 \lambda \cdot \alpha \bar{r}_{\mathrm{comp}},$$

for some universal constant $c_1 > 0$, then

$$U^* := \{x \in \mathcal{M}_0 : \rho(x) > \lambda r_{\mathrm{comp}}\} \subset U.$$

So, without loss of generality, we can replace ζ by $\zeta|_{U^*}$.

Let us now verify the bounds on $h_0 := \zeta^* g'_0 - g_0$ in the assumptions of Theorem 13.1. For this purpose note that there is a universal constant $C_1 < \infty$ such that

$$C_1^{-1} \rho_1^{-2} \leqslant |\mathrm{Rm}| + 1 \leqslant C_1 \rho_1^{-2}.$$

Now by assumption of this theorem and the fact that

$$|\mathrm{Rm}|+1 \leqslant C_1\rho_1^{-2} < C_1\lambda^{-2}r_{\mathrm{comp}}^{-2} \leqslant C_1(\lambda\alpha\bar{r}_{\mathrm{comp}})^{-2}r^{-2},$$

on U^*, we have

$$|h_0| \leqslant \varepsilon\cdot r^E(|\mathrm{Rm}|+1)^{E/2} \leqslant \varepsilon\cdot C_1^{E/2}(\lambda\alpha\bar{r}_{\mathrm{comp}})^{-E} \leqslant \eta_{\mathrm{lin}},$$

as long as

$$\varepsilon \leqslant C_1^{-E/2}(\lambda\alpha\bar{r}_{\mathrm{comp}})^E\eta_{\mathrm{lin}}.$$

Similarly, we obtain that

$$\begin{aligned} e^{HT}\rho_1^E|h_0| &\leqslant C_1^{E/2}e^{HT}(|\mathrm{Rm}|+1)^{-E/2}|h_0| \leqslant C_1^{E/2}e^{HT}\cdot\varepsilon r^E \\ &\leqslant \nu\cdot 10^{-E-1}\eta_{\mathrm{lin}}(\alpha r\cdot\bar{r}_{\mathrm{comp}})^E, \end{aligned}$$

as long as

$$\varepsilon \leqslant C_1^{-E/2}e^{-HT}\cdot\nu\cdot 10^{-E-1}\eta_{\mathrm{lin}}(\alpha\bar{r}_{\mathrm{comp}})^E$$

and

$$\begin{aligned} e^{HT}\rho_1^3|h_0| &\leqslant C_1^{3/2}e^{HT}(|\mathrm{Rm}|+1)^{-3/2}|h_0| \\ &\leqslant C_1^{E/2}(\lambda\alpha\bar{r}_{\mathrm{comp}})^{3-E}r^{3-E}\cdot e^{HT}(|\mathrm{Rm}|+1)^{-E/2}|h_0| \\ &\leqslant C_1^{E/2}(\lambda\alpha\bar{r}_{\mathrm{comp}})^{3-E}e^{HT}\cdot\varepsilon r^3 \leqslant \nu\cdot 10^{-1}(\lambda\alpha r\cdot\bar{r}_{\mathrm{comp}})^3, \end{aligned}$$

as long as

$$\varepsilon \leqslant C_1^{-E/2}e^{-HT}\cdot\nu\cdot 10^{-1}(\lambda\alpha\bar{r}_{\mathrm{comp}})^E.$$

Note that the three bounds that we have imposed on ε in this paragraph depend only on δ, T and E, assuming that α can be chosen depending on these three constants.

Lastly, note by assumption of this theorem the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption holds on the image of ζ.

We can therefore apply Theorem 13.1 and obtain a comparison domain

$$(\mathcal{N}, \{\mathcal{N}^j\}_{j=1}^J, \{t_j\}_{j=0}^J)$$

and a comparison $(\mathrm{Cut}, \phi, \{\phi^j\}_{j=1}^J)$ that satisfy the a-priori assumptions (APA1)–(APA6) for the parameters

$$(\eta_{\mathrm{lin}}, \delta_{\mathrm{n}}, \lambda, D_{\mathrm{cap}}, \Lambda, \delta_{\mathrm{b}}, \varepsilon_{\mathrm{can}}, r_{\mathrm{comp}})$$

and a-priori assumptions (APA7)–(APA13) for the parameters

$$(T, E, H, \eta_{\mathrm{lin}}, \nu, \lambda, \eta_{\mathrm{cut}}, D_{\mathrm{cut}}, W, A, r_{\mathrm{comp}}).$$

Without loss of generality, we may assume that T is an integral multiple of r_{comp}^2, i.e. $t_J=T$; otherwise we may decrease r_{comp} or increase T slightly.

Let now $\widehat{U}\subset\mathcal{M}_{[0,T]}$ be the set of $C_2 r_{\mathrm{comp}}$-thick points, where $C_2=C_2(\Lambda)<\infty$ is a constant whose value will be determined at the end of this paragraph. We claim that $\widehat{U}\subset\mathcal{N}\setminus\bigcup_{\mathcal{D}\in\mathrm{Cut}}\mathcal{D}$. To see this, consider $x\in\widehat{U}$ and choose $j\in\{1,...,J\}$ such that $x\in\mathcal{M}_{[t_{j-1},t_j]}$. Then, by Lemma 8.4, if

$$C_3\geqslant\underline{C}_3,\quad C_2\geqslant\Lambda C_3,\quad \varepsilon_{\mathrm{can}}\leqslant\bar{\varepsilon}_{\mathrm{can}}^{\text{Lemma 8.4}}\quad\text{and}\quad r_{\mathrm{comp}}\leqslant\bar{r}_{\mathrm{comp}}^{\text{Lemma 8.4}}$$

then x survives until time t_j and $x(t_j)$ is $\Lambda r_{\mathrm{comp}}$-thick. (Here, we have used the notation $\bar{\varepsilon}_{\mathrm{can}}^{\text{Lemma 8.4}}$ and $\bar{r}_{\mathrm{comp}}^{\text{Lemma 8.4}}$ to avoid confusion with the upper bounds $\bar{\varepsilon}_{\mathrm{can}}$ and $\bar{r}_{\mathrm{comp}}$ from Theorem 13.1.) So, by a-priori assumption (APA3) (b), we have $x(t_j)\in\mathcal{N}_{t_j-}$, and thus $x\in\mathcal{N}$. Lastly, by a-priori assumption (APA3) (e), we have $x\notin\mathcal{D}$ for all $\mathcal{D}\in\mathrm{Cut}$.

By the choice of $\widehat{U}$, we have

$$|\mathrm{Rm}|+1\geqslant C_1^{-1}\rho_1^{-2}\geqslant C_1^{-1}C_2^{-2}r_{\mathrm{comp}}^{-2}=C_1^{-1}C_2^{-1}\bar{r}_{\mathrm{comp}}^{-2}\alpha^{-2}r^{-2}$$

on $\mathcal{M}_{[0,T]}\setminus\widehat{U}$. So, if

$$\alpha\leqslant\tfrac{1}{2}C_1^{-1/2}C_2^{-1/2}\bar{r}_{\mathrm{comp}}\,,$$

then $|\mathrm{Rm}|+1\geqslant 4r^{-2}\geqslant r^{-2}+1$ and therefore $|\mathrm{Rm}|\geqslant r^{-2}$ on $\mathcal{M}_{[0,T]}\setminus\widehat{U}$. In the notation of Theorem 1.5, we can now set $\hat{\phi}:=\phi|_{\widehat{U}}$ and $\widehat{U}':=\phi(\widehat{U})$.

We now need to verify the upper bound on $|\tilde{\phi}^*g'-g|$ on $\widehat{U}$, in the notation of Theorem 1.5. To this end, note that by a-priori assumption (APA8) we have

$$\begin{aligned}|h|&\leqslant e^{-H(T-\mathfrak{t})}\rho_1^{-E}\cdot W\cdot 10^{-E-1}\eta_{\mathrm{lin}}(\alpha r\bar{r}_{\mathrm{comp}})^E\\&\leqslant C_1^{E/2}(|\mathrm{Rm}|+1)^{E/2}W\eta_{\mathrm{lin}}\alpha^E r^E\bar{r}_{\mathrm{comp}}^E\leqslant\delta\cdot(|\mathrm{Rm}|+1)^{E/2}r^E,\end{aligned}$$

as long as

$$\alpha\leqslant\delta^{1/E}C_1^{-1/2}W^{-1/E}\eta_{\mathrm{lin}}^{-1/E}\bar{r}_{\mathrm{comp}}.$$

We now show that $|\mathrm{Rm}|\geqslant r^{-2}$ on $\mathcal{M}'_{[0,T]}\setminus\widehat{U}'$ for sufficiently small α and ε if we additionally assume that $|\mathrm{Rm}|\geqslant(\varepsilon r)^{-2}$ on $\mathcal{M}'_0\setminus U'$. To see this, assume $|\mathrm{Rm}|(x')<r^{-2}$ for some $x'\in\mathcal{M}'_{[0,T]}\setminus\widehat{U}'$. So, $\rho(x')\geqslant\frac{1}{2}C_1^{-1/2}r$. We can now apply Lemma 13.2, assuming that $\varepsilon_{\mathrm{can}}$ is smaller than some universal constant, to find a continuous path $\gamma\colon[0,1]\to\mathcal{M}'_{[0,T]}$ between x' and a point $y'\in\mathcal{M}'_0$ such that $\rho(\gamma(s))>C_3^{-1}r$ for all $s\in[0,1]$, where C_3 is some universal constant, and $\mathfrak{t}'\circ\gamma$ is non-increasing. On the other hand, we have $\rho_1<C_1^{1/2}|\mathrm{Rm}|^{-1/2}\leqslant C_1^{1/2}\varepsilon r$ on $\mathcal{M}'_0\setminus U'$. So, if

$$\varepsilon\leqslant C_1^{-1/2}C_3^{-1},$$

then $y' \in U'$. Set $y := \zeta^{-1}(y')$.

Our next goal is to show that $y' \in \widehat{U}'_0$. To see this, we will argue that $\rho(y) > C_2 r_{\mathrm{comp}}$. By Lemma 8.4, and assuming that $\varepsilon_{\mathrm{can}}$ is smaller than some universal constant, we can find a universal constant $c_2 > 0$ such that $\rho > \frac{1}{2} C_3^{-1} r$ on $B(y', c_2 r)$. So, as in the last paragraph, we obtain that $B(y', c_2 r) \subset U'$, assuming

$$\varepsilon \leqslant \tfrac{1}{2} C_1^{-1/2} C_3^{-1}.$$

If $\rho(y') < 1$, then we can use the $\varepsilon_{\mathrm{can}}$-canonical neighborhood assumption at y' to deduce bounds on higher curvature derivatives on $B(y', c_2 r)$ (as in Lemma 8.1), assuming that $\varepsilon_{\mathrm{can}}$ is smaller than some universal constant. On the other hand, if $\rho(y') \geqslant 1$, then we obtain an improved bound of the form $\rho > c_3$ on $B(y', c_2 r)$ for some universal constant $c_3 > 0$ (via Lemma 8.6). So, using Lemma 8.8, applied similarly as in the proof of Lemma 8.9, we obtain a universal constant $c_4 > 0$ such that $\rho(y) > c_4 r$, assuming that η_{lin}, $\varepsilon_{\mathrm{can}}$ and α are smaller than some universal constant. So, if

$$\alpha \leqslant c_4 C_2^{-1},$$

then $\rho(y) > C_2 r_{\mathrm{comp}}$ and therefore, by construction of $\widehat{U}$, we have $y \in \widehat{U}_0$. It follows that $y' = \phi(y) \in \widehat{U}'_0$.

Choose $s_0 \in [0, 1]$ minimal with the property that $\gamma((s_0, 1]) \subset \widehat{U}'$. As $y \in \widehat{U}'$, we know that $s_0 < 1$, and since $\widehat{U}'$ is open and $x' \notin \widehat{U}'$, we obtain $\gamma(s_0) \notin \widehat{U}'$. For any $s \in (s_0, 1]$, we have $\phi^{-1}(\gamma(s)) \in \widehat{U} \subset \mathcal{N}$. So, by Lemma 8.9, and assuming that η_{lin}, δ_{n}, $\varepsilon_{\mathrm{can}}$ and α were chosen smaller than some universal constant, we obtain

$$\rho_1(\phi^{-1}(\gamma(s))) \geqslant C_{\mathrm{SD}}^{-1} \rho_1(\gamma(s)) > C_{\mathrm{SD}}^{-1} \cdot C_3^{-1} r.$$

Therefore, if

$$\alpha \leqslant \tfrac{1}{2} C_{\mathrm{SD}}^{-1} C_2^{-1} C_3^{-1},$$

then

$$\rho(\phi^{-1}(\gamma(s))) > 2 C_2 r_{\mathrm{comp}}.$$

Using Lemma 8.4, Proposition A.10 and the uniform lower bounds on the scales of $\gamma(s)$ and $\phi^{-1}(\gamma(s))$, we obtain that $z := \lim_{s \nearrow s_0} \phi^{-1}(\gamma(s))$ exists. It follows that

$$\rho(z) \geqslant 2 C_2 r_{\mathrm{comp}}.$$

So, $z \in \widehat{U}$ and thus $\gamma(s_0) = \phi(z) \in \widehat{U}'$, contradicting the choice of s_0.

This finishes the proof of the theorem if we allow $\varepsilon_{\mathrm{can}}$ to depend on T. To see that $\varepsilon_{\mathrm{can}}$ can even be chosen independently of T, we revert back to the notation used in the

theorem, and we construct $\hat{\phi}$ successively on time-intervals of the form $[0,1]$, $[1,2]$, etc. More specifically, given $E\geqslant\underline{E}$ set $\varepsilon_{\mathrm{can}}:=\varepsilon_{\mathrm{can}}(1,E)$ (i.e. the value in the weaker version of the theorem for $T=1$). Now assume that $\delta>0$ and $T<\infty$ are given and assume without loss of generality that T is an integer. Set inductively $\varepsilon_0:=\delta$ and

$$\varepsilon_i:=\min\{\varepsilon(\varepsilon_{i-1},1,E),\varepsilon_{i-1},1\},$$

where $\varepsilon(\cdot,1,E)$ is as in the statement of the weaker version of the theorem in the $T=1$ case, as well as $r_0:=r$ and $r_i:=\varepsilon_i r_{i-1}$, for $i=1,...,T$. Assume now that $|\mathrm{Rm}|\geqslant(\varepsilon_1\dots\varepsilon_T r)^{-2}=(\varepsilon_T r_{T-1})^{-2}$ on $\mathcal{M}_0\backslash U$ (and possibly also on $\mathcal{M}_0'\backslash U'$) and

$$|\phi^*g_0'-g_0|\leqslant\varepsilon_T\cdot r_T^{2E}(|\mathrm{Rm}|+1)^E.$$

We can then apply the weaker version of the theorem for $r=r_{T-i}$ and $\delta=\varepsilon_{T-i}$ to find a sequence of subsets $\widehat{U}_i\subset\mathcal{M}_{[i-1,i]}$ and $\widehat{U}_i'\subset\mathcal{M}'_{[i-1,i]}$, and diffeomorphisms $\hat{\phi}_i\colon\widehat{U}_i\to\widehat{U}_i'$ such that

$$|\hat{\phi}_i^*g'-g|\leqslant\varepsilon_{T-i}\cdot r_{T-i}^{2E}(|\mathrm{Rm}|+1)^E$$

and $|\mathrm{Rm}|\geqslant r_{T-i}^{-2}$ on $\mathcal{M}_{[i-1,i]}$ (and possibly also on $\mathcal{M}_{[i-1,i]}\backslash\widehat{U}_i'$) for $i=1,...,T$. Moreover, $\hat{\phi}_{i-1}=\hat{\phi}_i$ on $\widehat{U}_{i-1}\cap\widehat{U}_i\subset\mathcal{M}_i$. Then, $\hat{\phi}$ can be constructed by combining the diffeomorphisms $\hat{\phi}_1,...,\hat{\phi}_T$ on the open subset

$$(\widehat{U}_1\cap\mathcal{M}_{[0,1)})\cup(\widehat{U}_1\cap\widehat{U}_2)\cup(\widehat{U}_2\cap\mathcal{M}_{(1,2)})\cup(\widehat{U}_2\cap\widehat{U}_3)\cup\dots\cup(\widehat{U}_T\cap\mathcal{M}_{(T-1,T]}).$$

This finishes the proof. □

Proof of Theorem 1.3. The theorem is a consequence of Theorem 1.5. To see this, assume $\delta\leqslant1$, choose $E:=\underline{E}$ and consider the constants

$$\varepsilon_{\mathrm{can}}=\varepsilon_{\mathrm{can}}(E)\quad\text{and}\quad\varepsilon=\varepsilon\left(\frac{1}{3^E}\delta,T,E\right)$$

from Theorem 1.5. Set $\varepsilon':=\min\{\delta^{2E}\varepsilon,\delta\varepsilon,\varepsilon_{\mathrm{can}}\}$ and $r:=\delta$.

We claim that Theorem 1.3 holds for $\varepsilon=\varepsilon'$. By the assumption of this theorem, we have

$$|\phi^*g_0'-g_0|\leqslant\varepsilon'\leqslant\varepsilon\cdot\delta^{2E}\leqslant\varepsilon\cdot r^{2E}(|\mathrm{Rm}|+1)^E.$$

We also have $|\mathrm{Rm}|\geqslant\varepsilon'^{-2}\geqslant(\varepsilon r)^{-2}$ on $\mathcal{M}_0\backslash U$. So, Theorem 1.5 can be applied and yields the existence of a time-preserving diffeomorphism $\hat{\phi}\colon\widehat{U}\to\widehat{U}'$ such that $\hat{\phi}=\phi$ on $U\cap\widehat{U}$ and

$$|\hat{\phi}^*g'-g|\leqslant\frac{1}{3^E}\delta\cdot\delta^{2E}(|\mathrm{Rm}|+1)^E<\delta,\tag{13.1}$$

on $\widehat{U}\cap\{|\mathrm{Rm}|<2\delta^{-2}\}$ and $|\mathrm{Rm}|\geqslant\delta^{-2}$ on $\mathcal{M}_{[0,T]}\backslash\widehat{U}$. We can now replace $\widehat{U}$ by

$$\widehat{U}\cap\{|\mathrm{Rm}|<2\delta^{-2}\}$$

and then replace $\widehat{U}'$ by $\hat{\phi}(\widehat{U})$. Then, (13.1) holds on all of $\widehat{U}$ and we still have $|\mathrm{Rm}|\geqslant\delta^{-2}$ on $\mathcal{M}_{[0,T]}\backslash\widehat{U}$. □

Proof of Addendum to Theorem 1.3. The bounds on the higher derivatives follow from Lemma A.2, combined with Lemma 8.4 and Shi's estimates, since $\hat{\phi}^* g' - g$ satisfies that Ricci–DeTurck equation. □

We now apply the stability theorem, Theorem 1.5, to prove Theorem 1.1, which asserts the uniqueness of the Ricci flow spacetimes with a given initial condition, under completeness and canonical neighborhood assumptions.

The idea of the proof is as follows. We first apply Theorem 1.5 to produce a sequence of maps $\hat{\phi}_i : U_i \to \mathcal{M}'_{[0,T]}$ such that $\bigcup_i U_i = \mathcal{M}_{[0,T]}$ and $|\hat{\phi}_i^* g' - g| \leqslant \delta_i \to 0$. We then show that the $\hat{\phi}_i$s converge locally smoothly to the desired diffeomorphism $\hat{\phi}$. To do this, we appeal to the drift bound in Proposition A.10 to propagate the region of convergence over time, and we use uniqueness of isometries of Riemannian manifolds to propagate the convergence within time-slices.

Proof of Theorem 1.1. We will prove the theorem in the case $T < \infty$. The case $T = \infty$ follows by letting $T \to \infty$. Choose $\underline{E}$ and $\varepsilon_{\text{can}} := \varepsilon_{\text{can}}(\underline{E})$ according to Theorem 1.5. Also, by parabolic rescaling, we may assume without loss of generality that $r = 1$.

By Theorem 1.5, we can find a sequence of open subsets $U_1 \subset U_2 \subset \ldots \subset \mathcal{M}_{[0,T]}$ such that

$$\bigcup_{i=1}^{\infty} U_i = \mathcal{M}_{[0,T]}$$

and a sequence of time-preserving diffeomorphisms onto their images $\hat{\phi}_i : U_i \to \mathcal{M}'_{[0,T]}$ that satisfy the harmonic map heat flow equation, such that

$$\hat{\phi}_i|_{U_i \cap \mathcal{M}_0} = \phi|_{U_i \cap \mathcal{M}_0}, \quad \mathfrak{t}' \circ \hat{\phi}_i = \mathfrak{t}$$

and

$$|\hat{\phi}_i^* g' - g| \leqslant \delta_i \to 0. \tag{13.2}$$

Let Y be the set of points $x \in \mathcal{M}_{[0,T]}$ such that the pointwise limit

$$\phi_\infty(x) := \lim_{i \to \infty} \hat{\phi}_i(x)$$

exists. Let

$$X = \{x \in \mathcal{M}_{[0,T]} : B(x, r) \subset Y \text{ for some } r > 0\},$$

so X is the set of points $x \in \mathcal{M}_{[0,T]}$ that belong to relative interior of $Y \cap \mathcal{M}_{\mathfrak{t}(x)}$ in $\mathcal{M}_{\mathfrak{t}(x)}$. Recall that $X_t = X \cap \mathcal{M}_t$ for $t \geqslant 0$. Our main goal is to show that $X = \mathcal{M}_{[0,T]}$ and that the pointwise limit $\hat{\phi}_\infty$ is smooth, preserves the metric, and time vector field. Obviously, $X_0 = Y_0 = \mathcal{M}_0$, since $\hat{\phi}_i = \phi$ on $U_i \cap \mathcal{M}_0$.

CLAIM 13.3. *For every $t\in[0,T]$, the following statements hold*:

(a) $\hat{\phi}_\infty|_{X_t}\to\mathcal{M}_t$ *is a smooth isometric immersion.*

(b) X_t *is a union of connected components of* $\mathcal{M}_t$.

(c) *For all* $x\in X_t$, $\rho(\hat{\phi}_\infty(x))=\rho(x)$.

Proof. Suppose that $t\in[0,T]$ and z is in the closure of X_t. Choose $r_0>0$ such that $\overline{B(z,6r_0)}$ is compact, and pick $x\in B(z,r_0)\cap X_t$. Hence, we have that $\overline{B(x,5r_0)}$ is compact and $z\in B(x,r_0)$. There is a sequence $L_i\to 1$ such that for large i we have $\overline{B(x,5r_0)}\subset U_i$, and $L_i^{-1}g\leqslant\hat{\phi}_i^*g'\leqslant L_ig$ on $\overline{B(x,5r_0)}$. An elementary Riemannian geometry argument gives, for large i, that $\hat{\phi}_i(\overline{B(x,5r_0)})\supset B(\hat{\phi}_i(x),4r_0)$ and the restriction of $\hat{\phi}_i$ to $B(x,r_0)$ is L_i-bilipschitz with respect to the Riemannian distance functions on $\mathcal{M}_t$ and $\mathcal{M}'_t$. Since $\hat{\phi}_i(x)\to\hat{\phi}_\infty(x)$, for large i we have

$$\hat{\phi}_i(B(x,r_0))\subset B(\hat{\phi}_i(x),2r_0)\subset B(\hat{\phi}_\infty(x),3r_0)\subset B(\hat{\phi}_i(x),4r_0)\subset\hat{\phi}_i(\overline{B(x,5r_0)}),$$

and therefore $\overline{B(\hat{\phi}_\infty(x),3r_0)}$ is compact.

Put $B:=B(x,r_0)$. Suppose that $\{\hat{\phi}_i|_B\}$ does not converge pointwise. Then, by the Arzela-Ascoli theorem, the sequence $\{\hat{\phi}_i|_B\}$ has two distinct subsequential limits

$$\psi,\psi'\colon B\longrightarrow\mathcal{M}'_t,$$

and since $L_i\to 1$, both maps preserve the distance functions on $\mathcal{M}$ and $\mathcal{M}'$. Hence, ψ and ψ' are smooth Riemannian isometries. They agree on a neighborhood of x in X_t, because $x\in X$, and since B is connected, they must coincide, contradicting $\psi\neq\psi'$. Thus $B\subset X_t$, and the pointwise limit $\hat{\phi}_\infty$ is a smooth Riemannian isometry on B. This shows that the closure of X_t is open, which implies assertion (b). Our proof also implies assertion (a), which implies assertion (c). □

CLAIM 13.4. *There is a universal constant $c>0$ such that, for every $x\in Y_t=Y\cap\mathcal{M}_t$ and $\tau_x:=c\rho_1^2(x)$, the following holds for all $t'\in[t-\tau_x,t+\tau_x]\cap[0,T]$*:

(a) x *survives until* t', *and* $x(t')\in Y$.

(b) $\hat{\phi}_\infty(x)$ *survives until* t', *and* $(\hat{\phi}_\infty(x))(t')=\hat{\phi}_\infty(x(t'))$.

Proof. The claim follows from Proposition A.10 via a continuity argument. Let $x\in X_t$, and set $x':=\hat{\phi}_\infty(x)$.

Using Claim 13.3 (c) and Lemma 8.4, assuming that ε_{can} is smaller than some universal constant, we can find a universal constant $c>0$ such that, for $r_0:=c^{1/2}\rho_1(x)$, the following holds: For all $t_0\geqslant 0$ with $|t-t_0|\leqslant r_0^2$ the parabolic neighborhoods $P(x(t_0),r_0,2r_0^2)$ and $P(x'(t_0),100r_0,2r_0^2)$ are unscathed and $|\mathrm{Rm}|\leqslant r_0^{-2}$ on both. By compactness, we moreover find a constant $1\leqslant A_x<\infty$, which may depend on x, such that

$$|\nabla^m\mathrm{Rm}|\leqslant A_xr_0^{-2-m}\quad\text{for } m=0,1,...,3$$

on both parabolic neighborhoods. Moreover, lengths of curves inside these parabolic neighborhoods are distorted by at most a factor of 2 under the Ricci flow.

For each i choose $t^*_{-,i}$ minimal and $t^*_{+,i}$ maximal, with $0 \leqslant t^*_{-,i} \leqslant t \leqslant t^*_{+,i} \leqslant T$ and

$$|t^*_{\pm,i} - t| \leqslant r_0^2,$$

such that $d_{t_0}(\hat{\phi}_i(x(t_0)), (\hat{\phi}_i(x))(t_0)) < r_0$ for all $t_0 \in (t^*_{-,i}, t^*_{+,i})$. Since $x \in Y_t$, we have

$$d_t(\hat{\phi}_i(x), x') < r_0$$

for large i. So, by the length distortion bound on $P(x'(t_0), 100r_0, 2r_0^2)$, we have

$$d_{t_0}((\hat{\phi}_i(x(t')))(t_0), x'(t_0)) < 4r_0$$

for all $t_0 \in (t^*_{-,i}, t]$ and $t' \in [t_0, t^*_{+,i})$ if i is large (we use the convention $[t,t) = (t,t] = \{t\}$ here). By (13.2) and the distance distortion bounds on

$$P(x(t_0), r_0, 2r_0^2) \quad \text{and} \quad P(x'(t_0), 100r_0, 2r_0^2),$$

we therefore obtain that, for large i and $t_0 \in (t^*_{-,i}, t]$,

$$\hat{\phi}_i(P(x(t_0), r_0, t^*_{+,i} - t_0)) \subset P(x'(t_0), 100r_0, 2r_0^2).$$

We can therefore apply Proposition A.10 for

$$M = B(x(t_0), r_0), \quad M' = B(x'(t_0), 100r_0), \quad r = r_0 \quad \text{and} \quad A = A_x,$$

along with (13.2) to find that there is a sequence $\varepsilon_i \to 0$ such that, for large i, we have

$$d_{t_0}((\hat{\phi}_i(x(t')))(t_0), (\hat{\phi}_i(x))(t_0)) \leqslant \varepsilon_i r_0$$

for all $t_0 \in (t^*_{-,i}, t]$ and $t' \in [t, t^*_{+,i})$. By the distance distortion bound on

$$P(x'(t_0), 100r_0, 2r_0^2)$$

this implies that

$$d_{t'}(\hat{\phi}_i(x(t')), (\hat{\phi}_i(x))(t')) \leqslant 2\varepsilon_i r_0 \tag{13.3}$$

for all such t_0 and t'. The bound (13.3) implies that for large i we have $t^*_{-,i} = t - r_0^2$ or $t^*_{-,i} = 0$ and $t^*_{+,i} = t + r_0^2$ or $t^*_{+,i} = T$, due to their minimal and maximal choice. So, (13.3) implies assertions (a) and (b). □

CLAIM 13.5. (a) *X is (relatively) open and closed in $\mathcal{M}_{[0,T]}$.*
(b) *$\hat{\phi}_\infty$ is smooth, and $(\hat{\phi}_\infty)_*(\partial_{\mathfrak{t}})=\partial_{\mathfrak{t}'}$.*
(c) *$X=\mathcal{M}_{[0,T]}$.*

Proof. Suppose $z\in\mathcal{M}_t$ and z belongs to the closure of X. For $r>0$ sufficiently small, by Claim 13.3 (b), there exists $t'\in[t-r^2,t+r^2]\cap[0,T]$ such that $(B(z,r))(t')$ is contained in X. Shrinking r if necessary, we may assume that, for all $x\in(B(z,r))(t')$, we have $\tau_x\geqslant 2r^2$, where τ_x is as in Claim 13.4. Thus, by Claim 13.4 (a), we conclude that $(B(x,r))(\bar{t})\subset X$ for all $\bar{t}\in[t-r^2,t+r^2]\cap[0,T]$. This implies that the closure of X in $\mathcal{M}_{[0,T]}$ is open, which implies assertion (a).

By Claim 13.4 (b), it follows that $\hat{\phi}_\infty$ locally commutes with the flows of the time vector fields $\partial_{\mathfrak{t}}$ and $\partial_{\mathfrak{t}'}$ on $\mathcal{M}$ and $\mathcal{M}'$, respectively. Combining this with assertion (a) of Claim 13.3, we obtain assertion (b). By assertion (a), it follows that X is a union of connected components of $\mathcal{M}_{[0,T]}$. Assertion (c) now follows from Lemma 13.2, assuming that $\varepsilon_{\mathrm{can}}$ is smaller than some universal constant, and the fact that $\mathcal{M}_0\subset X$. □

By Claim 13.5, we have constructed a smooth map $\hat{\phi}_\infty\colon\mathcal{M}_{[0,T]}\to\mathcal{M}'_{[0,T]}$ such that

$$\hat{\phi}_\infty^* g'=g,\quad \hat{\phi}_\infty|_{\mathcal{M}_0}=\phi,\quad (\hat{\phi}_\infty)_*\partial_{\mathfrak{t}}=\partial_{\mathfrak{t}'}\quad\text{and}\quad \mathfrak{t}'\circ\hat{\phi}_\infty=\mathfrak{t}. \tag{13.4}$$

We now claim that the map $\hat{\phi}_\infty$ is uniquely characterized by (13.4). To see this, consider two such maps $\hat{\phi}_\infty$ and $\hat{\phi}'_\infty$. As both maps satisfy the harmonic map heat flow equation, we can apply the conclusions of our proof to up to this point to the sequences $U_i=\mathcal{M}_{[0,T]}$, $\hat{\phi}_{2i-1}=\hat{\phi}_\infty$ and $\hat{\phi}_{2i}=\hat{\phi}'_\infty$. It follows that $\hat{\phi}_i$ converges pointwise, and therefore we must have $\hat{\phi}_\infty=\hat{\phi}'_\infty$ as asserted.

It remains to show that $\hat{\phi}_\infty$ is bijective. To see this, we can interchange the roles of $\mathcal{M}$ and $\mathcal{M}'$, and apply our discussion to obtain a map $\hat{\psi}_\infty\colon\mathcal{M}'_{[0,T]}\to\mathcal{M}_{[0,T]}$ such that

$$\hat{\psi}_\infty^* g=g',\quad \hat{\psi}_\infty|_{\mathcal{M}'_0}=\psi:=\phi^{-1},\quad (\hat{\psi}_\infty)_*\partial_{\mathfrak{t}'}=\partial_{\mathfrak{t}}\quad\text{and}\quad \mathfrak{t}\circ\hat{\psi}_\infty=\mathfrak{t}'.$$

Now consider the composition $\alpha:=\hat{\psi}_\infty\circ\hat{\phi}_\infty$ such that

$$\alpha^* g=g,\quad \alpha|_{\mathcal{M}_0}=\mathrm{id}_{\mathcal{M}_0},\quad \alpha_*\partial_{\mathfrak{t}}=\partial_{\mathfrak{t}}\quad\text{and}\quad \mathfrak{t}\circ\alpha=\mathfrak{t},$$

By the uniqueness property, as discussed in the previous paragraph (for $\mathcal{M}=\mathcal{M}'$), we obtain that $\hat{\psi}_\infty\circ\hat{\phi}_\infty=\alpha=\mathrm{id}_{\mathcal{M}_{[0,T]}}$. Similarly, we obtain that $\hat{\phi}_\infty\circ\hat{\psi}_\infty=\mathrm{id}_{\mathcal{M}'_{[0,T]}}$. This shows that $\hat{\phi}_\infty$ is bijective, finishing the proof. □

Appendix A. Ricci–DeTurck flow and harmonic map heat flow

In this section we discuss the main estimates for harmonic map heat flow and Ricci–DeTurck flow that will be needed in the paper. While the general methodology is fairly standard, we were unable to find suitable references in the general PDE literature for these results.

A.1. The main equations

In this subsection we derive the general equations for harmonic map heat flow with time dependent metrics on the source and target, and the associated Ricci–DeTurck flow. Most of the ideas presented in this subsection go back to DeTurck [DT] and Hamilton [Ha4].

Consider two n-dimensional manifolds M and M', each equipped with a smooth family of Riemannian metrics $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$. Let moreover $(\chi_t)_{t\in[0,T]}$, $\chi_t\colon M'\to M$ be a smooth family of maps.

Definition A.1. We say that the family $(\chi_t)_{t\in[0,T]}$ moves by *harmonic map heat flow between* (M',g'_t) *and* (M,g_t) if the family satisfies the following evolution equation:

$$\partial_t\chi_t=\Delta_{g'_t,g_t}\chi_t=\sum_{i=1}^n(\nabla^{g_t}_{d\chi_t(e_i)}d\chi_t(e_i)-d\chi_t(\nabla^{g'_t}_{e_i}e_i)),\tag{A.1}$$

where $\{e_i\}_{i=1}^n$ is a local frame field on M' that is orthonormal with respect to g'_t.

Assume now for the remainder of this subsection that all the maps χ_t are diffeomorphisms and consider their inverses χ_t^{-1}. Let

$$h_t:=(\chi_t^{-1})^*g'_t-g_t\tag{A.2}$$

be the *associated perturbation*. The pullback $(\chi_t^{-1})^*g'_t=g_t+h_t$ evolves by the following equation

$$\begin{aligned}\partial_t((\chi_t^{-1})^*g'_t)&=(\chi_t^{-1})^*\partial_tg'_t-\mathcal{L}_{\partial_t\chi_t\circ\chi_t^{-1}}((\chi_t^{-1})^*g'_t)\\&=(\chi_t^{-1})^*(\partial_tg'_t+2\operatorname{Ric}(g'_t))-2(\chi_t^{-1})^*\operatorname{Ric}(g'_t)-\mathcal{L}_{\partial_t\chi_t\circ\chi_t^{-1}}((\chi_t^{-1})^*g'_t)\\&=(\chi_t^{-1})^*(\partial_tg'_t+2\operatorname{Ric}(g'_t))-2\operatorname{Ric}(g_t+h_t)-\mathcal{L}_{\partial_t\chi_t\circ\chi_t^{-1}}(g_t+h_t)\\&=(\chi_t^{-1})^*(\partial_tg'_t+2\operatorname{Ric}(g'_t))-2\operatorname{Ric}(g_t)+\mathcal{X}_t,\end{aligned}\tag{A.3}$$

where $\mathcal{X}_t$ can be expressed as follows (in the following identity, covariant derivatives and curvature quantities are taken with respect to g_t and the time-index t is suppressed)

$$\begin{aligned}\mathcal{X}_{ij}=(g+h)^{pq}(\nabla^2_{pq}h_{ij}+R_{pij}{}^uh_{uq}+R_{pji}{}^uh_{uq}-R_{ipq}{}^uh_{uj}-R_{jpq}{}^uh_{iu})\\-\tfrac12(g+h)^{pq}(g+h)^{uv}(-\nabla_ih_{pu}\nabla_jh_{qv}-2\nabla_uh_{ip}\nabla_qh_{jv}\\+2\nabla_uh_{ip}\nabla_vh_{jq}+2\nabla_ph_{iv}\nabla_jh_{qu}+2\nabla_ih_{pu}\nabla_qh_{jv}).\end{aligned}$$

We will now use (A.3) to derive an evolution equation for h_t. First observe that

$$\partial_t h_t = \partial_t((\chi_t^{-1})^* g'_t) - \partial_t g_t.$$

Similarly as in Uhlenbeck's trick, we define (we will suppress the time-index again wherever it interferes with the index notation)

$$(\nabla_{\partial_t} h_t)_{ij} = (\partial_t h_t)_{ij} - \tfrac{1}{2} g^{pq}(h_{pj}\partial_t g_{qi} + h_{ip}\partial_t g_{qj}).$$

Then,

$$\begin{aligned}(\nabla_{\partial_t} h_t)_{ij} &= ((\chi_t^{-1})^*(\partial_t g'_t + 2\operatorname{Ric}(g'_t)))_{ij} - (\partial_t g_t + 2\operatorname{Ric}(g_t))_{ij} \\ &\quad - \tfrac{1}{2} g^{pq}(h_{pj}\partial_t g_{qi} + h_{ip}\partial_t g_{qj}) + \mathcal{X}_{ij} \\ &= ((\chi_t^{-1})^*(\partial_t g'_t + 2\operatorname{Ric}(g'_t)))_{ij} - (\partial_t g_t + 2\operatorname{Ric}(g_t))_{ij} \\ &\quad - \tfrac{1}{2} g^{pq}(\partial_t g_t + 2\operatorname{Ric}(g_t))_{pi} h_{qj} - \tfrac{1}{2} g^{pq}(\partial_t g_t + 2\operatorname{Ric}(g_t))_{pj} h_{iq} + \mathcal{Y}_{ij},\end{aligned} \tag{A.4}$$

where

$$\begin{aligned}\mathcal{Y}_{ij} &= \mathcal{X}_{ij} + g^{pq}\operatorname{Ric}_{ip} h_{qj} + g^{pq}\operatorname{Ric}_{jp} h_{qi} \\ &= (g+h)^{pq}(\nabla^2_{pq} h_{ij} + R_{pij}{}^u h_{uq} + R_{pji}{}^u h_{uq}) \\ &\quad + (g^{pq} - (g+h)^{pq})(R_{ipq}{}^u h_{uj} + R_{jpq}{}^u h_{iu}) \\ &\quad - \tfrac{1}{2}(g+h)^{pq}(g+h)^{uv}(-\nabla_i h_{pu}\nabla_j h_{qv} - 2\nabla_u h_{ip}\nabla_q h_{jv} \\ &\qquad + 2\nabla_u h_{ip}\nabla_v h_{jq} + 2\nabla_p h_{iv}\nabla_j h_{qu} + 2\nabla_i h_{pu}\nabla_q h_{jv}).\end{aligned}$$

In the following, we will focus on the case in which h_t is small and in which the families of metrics $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$ almost satisfy the Ricci flow equation in the following sense. For parameters $0<\eta<0.1$ and $\delta>0$ we assume that, for all $t\in[0,T]$,

$$-\eta g_t \leqslant h_t \leqslant \eta g_t$$

and

$$-\delta g'_t \leqslant \partial_t g'_t + 2\operatorname{Ric}(g'_t) \leqslant \delta g'_t \quad \text{and} \quad -\delta g_t \leqslant \partial_t g_t + 2\operatorname{Ric}(g_t) \leqslant \delta g_t. \tag{A.5}$$

If we now multiply (A.4) by $2g^{iu}g^{jv}h_{uv}$, then we obtain that, for some dimensional constant $C_0<\infty$,

$$\begin{aligned}\partial_t|h|^2 \leqslant (g+h)^{ij}\nabla^2_{ij}|h|^2 - 2(g+h)^{ij}g^{pq}g^{uv}\nabla_i h_{pu}\nabla_j h_{qv} \\ + C_0\delta\cdot|h| + C_0|\operatorname{Rm}_g|\cdot|h|^2 + C_0|h|\cdot|\nabla h|^2.\end{aligned}$$

We will later consider the case $\eta<\min\{0.1, C_0^{-1}\}$. Note that then

$$\partial_t|h|^2 \leqslant (g+h)^{ij}\nabla^2_{ij}|h|^2+C_0\delta\cdot|h|+C_0|\mathrm{Rm}_g|\cdot|h|^2. \tag{A.6}$$

Next, let us consider the case in which $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$ both satisfy the (exact) Ricci flow equation. Then, (A.3) implies the *Ricci–DeTurck equation* for the pullback metric $g_t+h_t=(\chi_t^{-1})^*g'_t$:

$$\partial_t(g_t+h_t) = -2\,\mathrm{Ric}(g_t+h_t)-\mathcal{L}_{X_{g_t}(g_t+h_t)}(g_t+h_t),$$

where the vector field $X_{g_t}(g_t+h_t)$ is defined by

$$X_g(g^*) := \Delta_{g^*,g}\,\mathrm{id}_M = \sum_{i=1}^n(\nabla^g_{e_i}e_i-\nabla^{g^*}_{e_i}e_i), \tag{A.7}$$

for a local frame $\{e_i\}_{i=1}^n$ that is orthonormal with respect to g^*. Note that

$$X_{g_t}(g_t+h_t) = \partial_t\chi_t\circ\chi_t^{-1}. \tag{A.8}$$

From an analytical point of view, (A.4) implies that the Ricci–DeTurck equation can be expressed as follows in terms of the perturbations h_t (also referred to as the *Ricci–DeTurck perturbation equation* here):

$$\nabla_{\partial_t}h_t = \Delta_{g_t}h_t+2\,\mathrm{Rm}_{g_t}(h_t)+\mathcal{Q}_{g_t}[h_t]. \tag{A.9}$$

Here the expression on the left-hand side now denotes the conventional Uhlenbeck trick:

$$(\nabla_{\partial_t}h_t)_{ij} = (\partial_t h_t)_{ij}+g_t^{pq}((h_t)_{pj}\,\mathrm{Ric}_{qi}+(h_t)_{ip}\,\mathrm{Ric}_{qj}).$$

Moreover,

$$(\mathrm{Rm}_{g_t}(h_t))_{ij} = g^{pq}R_{pij}{}^uh_{qu}$$

and $\mathcal{Q}_{g_t}[h_t]=\mathcal{Q}^{(1)}_{g_t}[h_t]$, where

$$\begin{aligned}(\mathcal{Q}^{(\alpha)}_{g_t}[h_t])_{ij} = {}&((g+\alpha h)^{pq}-g^{pq})(\nabla^2_{pq}h_{ij}+R_{pij}{}^uh_{uq}+R_{pji}{}^uh_{uq})\\ &+(g^{pq}-(g+\alpha h)^{pq})(R_{ipq}{}^uh_{uj}+R_{jpq}{}^uh_{iu})\\ &-\tfrac{1}{2}\alpha(g+\alpha h)^{pq}(g+\alpha h)^{uv}(-\nabla_ih_{pu}\nabla_jh_{qv}-2\nabla_uh_{ip}\nabla_qh_{jv}\\ &\quad+2\nabla_uh_{ip}\nabla_vh_{jq}+2\nabla_ph_{iv}\nabla_jh_{qu}+2\nabla_ih_{pu}\nabla_qh_{jv}).\end{aligned} \tag{A.10}$$

In this paper, we also consider the *rescaled Ricci–DeTurck equation* for perturbations of the form $\tilde{h}_t:=\alpha^{-1}h_t$ (we will be interested in the case $\alpha\leqslant 1$ mostly):

$$\nabla_{\partial_t}\tilde{h}_t = \Delta_{g_t}\tilde{h}_t+2\,\mathrm{Rm}_{g_t}(\tilde{h}_t)+\mathcal{Q}^{(\alpha)}_{g_t}[\tilde{h}_t]. \tag{A.11}$$

Note that $Q^{(0)}[h_t]=0$. So, for $\alpha\to 0$, the equation (A.11) converges to the *linearized Ricci–DeTurck equation*

$$\nabla_{\partial_t}\tilde{h}_t = \Delta_{g_t}\tilde{h}_t+2\,\mathrm{Rm}_{g_t}(\tilde{h}_t). \tag{A.12}$$

A.2. Local derivative estimates

In the following, we will derive local bounds on derivatives of the Ricci–DeTurck equation and the harmonic map heat flow equation. Let us first consider the Ricci–DeTurck perturbation equation. We obtain the following local derivative bounds.

LEMMA A.2. (Local derivative estimates for Ricci–DeTurck flow) *For any $m, n \geqslant 1$ there are constants $\eta_m = \eta_m(n) > 0$ and $C_m = C_m(n) < \infty$ such that the following holds.*

Consider a Ricci flow $(g_t)_{t\in[0,r^2]}$ on an n-dimensional manifold M. Let $p \in M$ be a point, $r > 0$ and assume that the time-zero ball $B(p,0,r) \subset M$ is relatively compact and that $|\nabla^{m'} \mathrm{Rm}| \leqslant r^{-2-m'}$ on $B(p,0,r)\times[0,r^2]$ for all $m' = 0, ..., m+2$.

Consider a solution $(h_t)_{t\in[0,r^2]}$ on $(M,(g_t)_{t\in[0,r^2]})$ to the Ricci-DeTurck perturbation equation (A.9). Then, the following holds:

(a) *If*

$$H := \sup_{B(x,0,r)\times[0,r^2]} |h_t|_{g_t} \leqslant \eta_m,$$

then

$$|\nabla^{m'} h_t|_{g_t} \leqslant C_m H t^{-m'/2}$$

on $B\big(p,0,\frac{1}{2}r\big)\times(0,r^2]$ for all $m' = 1, ..., m$.

(b) *If*

$$H_0 := \sup_{B(x,0,r)\times[0,r^2]} |h_t|_{g_t} + \max_{0\leqslant m'\leqslant m+1} \sup_{B(x,0,r)} r^{m'} |\nabla^{m'} h_t|_{g_t} \leqslant \eta_m,$$

then

$$|\nabla^{m'} h_t|_{g_t} \leqslant C_m H_0 r^{-m'}$$

on $B\big(p,0,\frac{1}{2}r\big)\times[0,r^2]$ for all $m' = 1, ..., m$.

Proof. This follows directly from [Ba1, Proposition 2.5], [Ap1, Lemma 4.4] and (A.9). □

Next, we discuss similar local derivative bounds for the harmonic map heat flow. To this end, consider families of metrics $(g_t)_{[0,T]}$ on M and $(g'_t)_{[0,T]}$ on M' and a solution $(\chi_t : M' \to M)_{t\in[0,T]}$ to the harmonic map heat flow equation (A.1) between M' and M. Choose local coordinates $(x^1, ..., x^n)$ on $U \subset M$ and $(y^1, ..., y^n)$ on $V \subset M'$ such that $\chi_t(V) \subset U$ for all $t \in [0,T]$. Express the families of metric $(g_t)_{t\in[0,T]}$ on U and $(g'_t)_{[0,T]}$ on V as

$$g_t = g_{t,ij}\, dx^i dx^j \quad \text{and} \quad g'_t = g'_{t,ij}\, dy^i dy^j.$$

The maps χ_t can be expressed on V as an n-tuple of functions

$$(\chi_t^1(y^1, ..., y^n), ..., \chi_t^n(y^1, ..., y^n))$$

and the harmonic map heat flow equation takes the form (we again suppress the t-index)

$$\partial_t \chi^k = g'^{ij} \frac{\partial^2 \chi^k}{\partial y^i \partial y^j} - g'^{ij} g'^{uv} \left(2 \frac{\partial g'_{iu}}{\partial y^j} - \frac{\partial g'_{ij}}{\partial y^u} \right) \frac{\partial \chi^k}{\partial y^v} + g'^{ij} g^{kl}(\chi^1, ..., \chi^n) \frac{\partial \chi^u}{\partial y^i} \frac{\partial \chi^v}{\partial y^j} \left(2 \frac{\partial g_{ul}}{\partial x^v} - \frac{\partial g_{uv}}{\partial x^l} \right)$$

Using this notation, we can now state the following local regularity result.

LEMMA A.3. (Local gradient bounds for harmonic map heat flow) *For any $m, n \geqslant 1$ and $A < \infty$ there are constants $\alpha_m = \alpha_m(A,n)$ and $C_m = C_m(A,n) < \infty$ such that the following holds.*

Choose $r > 0$ such that $r^2 \leqslant T$ and $p \in V$. Assume that the Euclidean ball $B(q,r) \subset V$ is relatively compact and that, on $U \times [0, r^2]$ and $V \times [0, r^2]$, we have

$$|\partial^{m_1} \partial_t^{m_2} (g_{ij} - \delta_{ij})| \leqslant \alpha_m r^{-m_1 - 2m_2} \quad \text{and} \quad |\partial^{m_1} \partial_t^{m_2} (g'_{ij} - \delta_{ij})| \leqslant \alpha_m r^{-m_1 - 2m_2}$$

for all $0 \leqslant m_1 + 2m_2 \leqslant m+1$ (here ∂^{m_1} denotes spatial derivatives). Assume moreover that there is a $p \in U$ such that $\chi_t(B(q,r)) \subset B(p, Ar)$ for all $t \in [0, r^2]$.

Then, the following holds:

(a) *We have*

$$|\partial^{m_1} \partial_t^{m_2} \chi^k| \leqslant C_m t^{-(m_1 + 2m_2 - 1)/2}. \tag{A.13}$$

on $B\big(q, \frac{1}{2} r\big) \times (0, r^2]$ for all $0 < m_1 + 2m_2 \leqslant m$.

(b) *If moreover for all $0 < m_1 \leqslant m+1$ we have*

$$|\partial^{m_1} \chi^k| \leqslant A r^{1 - m_1} \quad \text{on } B(p,r) \times \{0\}$$

(for $t = 0$), then we even have

$$|\partial^{m_1} \partial_t^{m_2} \chi^k| \leqslant C_m r^{-(m_1 + 2m_2 - 1)}$$

on $B\big(q, \frac{1}{2} r\big) \times [0, r^2]$ for all $0 < m_1 + 2m_2 \leqslant m$.

Proof. Without loss of generality, we may assume via parabolic rescaling and translating that $r = 1$ and $p = q = 0$. The constant α_m will be chosen in the course of this proof. We will always assume that $\alpha_m < 0.1$.

Let $\beta > 0$ be a constant whose value we determine in the course of this proof. Set $\widetilde{\chi}^k_t := \beta \cdot \chi^k_t$. Then, $\widetilde{\chi}$ satisfies an equation of the form

$$\partial_t \widetilde{\chi} = g'^{ij} \partial^2_{ij} \widetilde{\chi} + f_1(x^1, ..., x^n, t) * \partial \widetilde{\chi} + \alpha_m \beta^{-1} \cdot f_2(x^1, ..., x^n, \beta^{-1} \widetilde{\chi}^1, ..., \beta^{-1} \widetilde{\chi}^n, t) * \partial \widetilde{\chi} * \partial \widetilde{\chi},$$

where f_1 and f_2 are functions with

$$|\partial^{m_1}\partial_t^{m_2} f_i| \leqslant C(m,n) \tag{A.14}$$

on $B(p,0,1)\times[0,1]$ for all $0\leqslant m_1+2m_2\leqslant m$ (note that f_2 has $2n$ spatial components). Assume for the remainder of the proof that $\alpha_m\leqslant\beta^{m+1}$. Then,

$$f_3(x^1,...,x^n,z^1,...,z^n,t) := \alpha_m\beta^{-1} f_2(x^1,...,x^n,\beta^{-1}z^1,...,\beta^{-1}z^n,t)$$

also satisfies a bound of the form (A.14).

Next, note that

$$|\widetilde{\chi}| \leqslant \beta A \quad \text{on } B(p,0,1)\times[0,t_0].$$

So, if β is chosen small, depending on A, m and n, then we can again use [Ba1, Proposition 2.5] in assertion (a) to derive bounds for $|\partial^{m_1}\partial_t^{m_2}\widetilde{\chi}^k_{ij}|$ on $B\big(p,0,\frac{1}{2}\big)\times(0,1]$ that depend only on A, m and n. These bounds imply (A.13). For assertion (b) we can use [Ap1, Lemma 4.4]. □

Using this local gradient estimate, we can now prove the following drift bound.

LEMMA A.4. (Drift bound in local coordinates) *For every $n\geqslant 1$ and $A<\infty$ there are constants $\tau=\tau(A,n)$, $\alpha=\alpha(A,n)>0$ and $C=C(A,n)<\infty$ such that the following holds.*

Let $r>0$ and let $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,r^2]}$ be smooth families of Riemannian metrics on n-dimensional manifolds M and M'. Assume that $(\chi_t)_{t\in[0,r^2]}$ is a solution to the harmonic map heat flow equation (A.1) with the property that χ_t is A-Lipschitz for all $t\in[0,r^2]$.

Let $q\in M'$ and $p:=\chi_0(q)\in M$. Assume that we have the bounds

$$|\nabla^m \operatorname{Rm}(g_t)|, |\nabla^m \operatorname{Rm}(g'_t)| \leqslant \alpha r^{-2-m} \quad \text{for } m=0,...,3$$

and

$$|\nabla^m \partial_t g_t|, |\nabla^m \partial_t g'_t| \leqslant \alpha r^{-2-m} \quad \text{for } m=0,1$$

on $B(p,0,r)\times[0,r^2]$ and $B(q,0,r)\times[0,r^2]$. Assume also that $B(p,0,r)$ and $B(q,0,r)$ are relatively compact in M and M', respectively.

Then, $d_0(\chi_t(q),p)\leqslant Ct^{1/2}$ for all $t\in[0,\tau r^2]$.

Proof. Without loss of generality, we may assume that $r=1$. Choose $t^*\in[0,1]$ maximal with the property that $d_0(\chi_t(q),p)\leqslant\frac{1}{10}$ for all $t\in[0,t^*]$. Obviously, $t^*>0$. In the following we will find a lower bound on t^* in terms of A and n.

Assuming α to be sufficiently small, we can use the A-Lipschitz bound on χ_t to conclude that we have, for all $t\in[0,t^*]$,

$$\chi_t(B(q,0,(2A)^{-1}))\subset B(p,0,1)\subset U.$$

We will now apply Lemma A.3 for $r=r^*:=\frac{1}{2}(2A)^{-1}$. To do this, consider the exponential map

$$\exp_{q,g_0'}: T_qM' \supset B(0,r^*) \longrightarrow B(q,0,r^*) \subset M'$$

based at q with respect to the metric g_0'. Then, the family of pullback metrics $(\exp_{q,g_0'})^* g_t'$ on $B(0,r^*)$ satisfies a bound of the form (A.13), for $r=r^*$ and α replaced by $C(A,n)\alpha$. A similar bound holds for the family of pullback $(\exp_{p,g_0})^* g_t$ on $B(0,1)$. The family of maps

$$\chi_t \circ \exp_{q,g_0'}: B(0,r^*) \longrightarrow B(p,0,1)$$

can be lifted to a family of maps

$$\widetilde{\chi}_t: B(0,r^*) \longrightarrow B(0,1)$$

with $\chi_t \circ \exp_{q,g_0'} = \exp_{p,g_0} \widetilde{\chi}_t$ and $\widetilde{\chi}_0(0)=0$.

We can now apply Lemma A.3 for $\widetilde{\chi}_t$ and assuming that α is sufficiently small, and obtain that

$$|\partial_t \widetilde{\chi}_t| \leqslant C' t^{-1/2}$$

for some $C'=C'(A,n)<\infty$. Integrating this bound yields

$$d_0(\chi_t(q),p)) \leqslant d_0(\widetilde{\chi}_t(0),0) \leqslant C t^{1/2}$$

for all $t\in[0,t^*]$, where $C=C(A,n)<\infty$.

Set $\tau:=\min\{(100C)^{-2},1\}$. If $t^*<\tau$, then $d_0(\chi_t(q),p))<\frac{1}{10}$ for all $t\in[0,t^*]$, in contradiction to the maximal choice of t^*. So, $t^*\geqslant\tau$, which finishes the proof. □

A.3. Short-time existence

In this subsection, we prove our main short-time existence result, Proposition A.9, for the harmonic map heat flow. The main technical challenges come from the fact that we will work in the non-compact setting and that the background metrics on domain and target are time-dependent and may not strictly satisfy the Ricci flow equation.

We first derive the following bound for solutions of the harmonic map heat flow, which is a consequence of (A.6).

LEMMA A.5. *For every $n\geqslant 1$ there exists a constant $\bar{\eta}_n>0$ such that, for any $0<\eta_0<\eta_1<\bar{\eta}_n$ and $0<\delta,C<\infty$, there is a constant $\tau=\tau(\eta_0,\eta_1,\delta,C,n)>0$ such that the following holds.*

Consider smooth families of metrics $(g_t)_{t\in[0,T]}$ and $(g_t')_{t\in[0,T]}$ on n-dimensional manifolds M and M' such that (A.5) holds. Assume moreover that (M,g_t) and (M',g_t')

are complete and $|\mathrm{Rm}(g_t)|, |\mathrm{Rm}(g'_t)| \leqslant C$ *for all* $t \in [0,T]$ *and that* $|\nabla^{g_t} \partial_t g_t|$ *is uniformly bounded on* $M \times [0,T]$ *(by some constant that may be independent of* C*).*

Let $(\chi_t)_{t\in[0,T]}$ *be a smooth family of diffeomorphisms between* M' *and* M *moving by harmonic map heat flow (A.1) and set* $h_t := (\chi_t^{-1})^* g'_t - g_t$*. Assume that* $|h_0| \leqslant \eta_0$ *and that* $|\partial_t h_t| < C' t^{-1/2}$ *on* $M \times (0,T]$ *for some finite constant* C'*.*

Then, for all $t \in [0, \min\{\tau, T\}]$*, we have* $|h_t| \leqslant \eta_1$*.*

Note that in this lemma the constants η_0, η_1 and δ can be chosen independently of C.

Proof. By (A.6) we have

$$\partial_t |h|^2 \leqslant (g+h)^{ij} \nabla^2_{ij} |h|^2 + C_0^2 \delta^2 + |h|^2 + C_0 C \cdot |h|^2,$$

as long as $|h| \leqslant \bar{\eta}_n$ for some universal $\bar{\eta}_n > 0$. So, by the weak maximum principle applied to (A.6), we obtain

$$|h_t|^2 \leqslant \eta_0^2 e^{(C_0 C+1)t} + \frac{C_0^2 \delta^2}{C_0 C + 1} (e^{(C_0 C+1)t} - 1).$$

Note that for the application of the weak maximum principle we need to use the fact that $\partial_t g_t$ and $|\nabla^{g_t} \partial_t g_t|$ are uniformly bounded on $M \times [0,T]$. The bound on the first quantity follows from (A.5) and the curvature bound and the second quantity is bounded by assumption.

The lemma now follows immediately by a continuity argument. Observe here that the condition $|h_t| \leqslant \bar{\eta}_n$ always holds on a slightly larger time-interval than the condition $|h_t| \leqslant \eta_1$, due to the bound on $|\partial_t h_t|$ and the fact that $C' t^{-1/2}$ is integrable. □

We first discuss the existence theory of the harmonic map heat flow in the case in which the domain M' is compact and we will derive a lower bound on the time of existence.

LEMMA A.6. (Short-time existence of harmonic map heat flow, compact case) *For every* $n \geqslant 1$ *and* $C < \infty$ *there are constants* $\tau = \tau(C,n) > 0$ *and* $C^* = C^*(C,n) < \infty$ *such that the following holds.*

Let $(g_t)_{t\in[0,T]}$ *and* $(g'_t)_{t\in[0,T]}$ *be two smooth families of Riemannian metrics on* n*-dimensional manifolds* M *and* M' *and* $\bar{\chi}: M' \to M$ *a smooth map such that*

(i) (M, g_0) *is complete and* M' *is compact.*

(ii) *We have* $|\nabla^m_{g_t} \mathrm{Rm}(g_t)|, |\nabla^m_{g'_t} \mathrm{Rm}(g'_t)|, |\nabla^m_{g_t} \partial_t g_t|, |\nabla^m_{g'_t} \partial_t g'_t| \leqslant C$ *on* M *and* M' *for all* $t \in [0,T]$ *and* $m = 0, \ldots, 3$*.*

(iii) $\bar{\chi}$ *is* C*-Lipschitz.*

Then, the harmonic map heat flow equation

$$\partial_t \chi_t = \Delta_{g'_t, g_t} \chi_t, \quad \chi_0 = \bar{\chi}, \tag{A.15}$$

has a smooth solution on the time-interval $[0, \min\{\tau, T\})$ *and* χ_t *is* C^**-Lipschitz for all* $t \in [0, \min\{\tau, T\}]$.

Proof. By standard parabolic theory, we find that (A.15) has a solution $(\chi_t)_{t\in[0,T^*)}$ for some maximal $0<T^*\leqslant T$. If $T^*<T$, then this solution does not extend smoothly until time T. It remains to deduce a lower bound on T^* and a Lipschitz bound that only depend on C, n.

As explained in [BB], the norm of the differential

$$d\chi_t \in C^\infty(M'; T^*M' \otimes \chi_t^* TM)$$

satisfies an evolution inequality of the form

$$\begin{aligned}\partial_t |d\chi_t|^2 \leqslant \Delta_{g'_t} |d\chi_t|^2 + \partial_t g'_t * d\chi_t * d\chi_t + \mathrm{Ric}(g'_t) * d\chi_t * d\chi_t \\ + \partial_t g_t * d\chi_t * d\chi_t + \mathrm{Rm}(g_t) * d\chi_t * d\chi_t * d\chi_t * d\chi_t.\end{aligned}$$

So, for some $C'=C'(C,n)<\infty$ we have

$$\partial_t |d\chi_t|^2 \leqslant \Delta_{g'_t} |d\chi_t|^2 + C' |d\chi_t|^2 + C' |d\chi_t|^4.$$

So, using assumption (iii) and the weak maximum principle, we can find constants $\tau=\tau(C,n)>0$ and $C''=C''(C,n)<\infty$ such that $|d\chi_t|^2 \leqslant C''$ for all $t \in [0, \min\{\tau, T^*\})$. So, χ_t remains C^*-Lipschitz for all $t \in [0, \min\{\tau, T^*\})$ for some $C^*=C^*(C,n)<\infty$.

We can now use Lemma A.4, followed by Lemma A.3, to derive bounds on higher derivative of χ_t that are independent of t. This shows that χ_t extends smoothly to time $\min\{\tau, T^*\}$. We therefore obtain a contradiction to the maximality of T^* in the case in which $T^*<\min\{\tau, T\}$. □

Next, we remove the compactness assumption on M', but assume that the injectivity radius of M' is positive.

LEMMA A.7. (Short-time existence of harmonic map heat flow, non-compact case, positive injectivity radius) *Lemma A.6 continues to hold if we modify the assumptions by replacing* (i) *and* (ii) *by the following conditions*:

(i′) (M, g_0) *and* (M', g'_0) *are complete*,

(ii′) $|\nabla^m_{g_t} \mathrm{Rm}(g_t)|, |\nabla^m_{g'_t} \mathrm{Rm}(g'_t)|, |\nabla^m_{g_t} \partial_t g_t|, |\nabla^m_{g'_t} \partial_t g'_t| \leqslant C$ *on* M *and* M' *for all* $t \in [0, T]$ *and* $m=0, ..., 7$,

and if we assume in addition that

(iv′) *The injectivity radius of* (M, g_0) *and* (M', g_0') *is uniformly bounded from below by a positive constant.*

Proof. We will reduce the non-compact case to the compact case via a standard doubling construction. By [CG] we can find a sequence $N^{(1)} \Subset N^{(2)} \Subset ... \Subset M'$ of domains with smooth boundary such that $\bigcup_{i=1}^{\infty} \operatorname{Int} N^{(i)} = M'$ and such that the second fundamental form of $\partial N^{(i)}$ is bounded by some constant $C' = C'(C, n) < \infty$. Let $M'^{(i)}$ be the manifold that arises by identifying two copies of $N^{(i)}$ along their boundary and define $\bar{\chi}^{(i)} : M'^{(i)} \to M$ to be equal to $\bar{\chi}|_{N^{(i)}}$ on each copy of $N^{(i)}$. By a smoothing construction, and using assumption (iv′), we can find families of metrics $(g_t'^{(i)})_{t \in [0,T]}$ on $M'^{(i)}$ that agree with $g^{(i)}$ away from a 1-tubular neighborhood of $\partial N^{(i)}$ and such that the bounds in assumption (ii′) continue to hold for all $m = 0, ..., 3$. Moreover, by modifying $\bar{\chi}^{(i)}$ in a 1-tubular neighborhood of $\partial N^{(i)}$, we can construct maps $\bar{\chi}'^{(i)} : M'^{(i)} \to M$ that are C''-Lipschitz, for some $C'' < \infty$ that is independent of i. Note that C'' may, however, depend on the injectivity radius bound in assumption (iv′).

Using Lemma A.6, we can evolve $\bar{\chi}'^{(i)}$ by the harmonic map heat flow to some family $(\chi_t'^{(i)})_{t \in [0,T^*]}$ for some $T^* > 0$ that is independent of i, but may depend on the injectivity radius bound in assumption (iv′). Moreover, the maps $\chi_t'^{(i)}$ are C'^*-Lipschitz for some uniform $C'^* < \infty$.

Using Lemmas A.4 and A.3, we obtain uniform local derivative bounds on the families $(\chi_t'^{(i)})_{t \in [0,T^*]}$. So, after passing to a subsequence, these families converge to a solution $\chi_t : M' \to M$ of the harmonic map heat flow on the time-interval $[0, T^*]$.

By the same maximum principle argument as used in the proof of Lemma A.6, we obtain a Lipschitz bound on χ_t of the form $C^*(C, n)$ that holds up to time $\min\{\tau, T\}$, where $\tau = \tau(C, n) > 0$ is a constant that does not depend on the injectivity radius bound in (iv′). Assume now that $T^* < T$ is chosen maximal with the property that the harmonic map heat flow exists on $[0, T^*)$. If $T^* < \min\{\tau, T\}$, then we can argue as in the proof of Lemma A.6 that the flow extends smoothly to time T^* and then restart the flow at time T^*. This would contradict the maximal choice of T^*. Therefore, $T^* \geqslant \min\{\tau, T\}$. □

Using a similar construction, we can remove the assumption on the positivity of the injectivity radius.

LEMMA A.8. *Lemma A.7 continues to hold if we remove assumption* (iv′) *and replace assumption* (ii′) *by*

(ii″) $|\nabla_{g_t}^m \operatorname{Rm}(g_t)|, |\nabla_{g_t'}^m \operatorname{Rm}(g_t')|, |\nabla_{g_t}^m \partial_t g_t|, |\nabla_{g_t'}^m \partial_t g_t'| \leqslant C$ *on* M *and* M' *for all* t *in the interval* $[0, T]$ *and all* $m = 0, ..., 10$.

Proof. The solution (χ_t) arises again via a limit argument, by Lemma A.7. For

this purpose we represent (M, g_0) and (M', g_0') as a limit of Riemannian manifolds with positive injectivity radius. The method used here can also be found in [CZ].

Choose $p' \in M'$ and $p := \bar{\chi}(p') \in M$, and denote by $r := d_0(p, \cdot)$ and $r' := d_0(p', \cdot)$ the distance functions to the respective basepoints. Due to assumptions (i′) and (ii′), we have $\mathrm{InjRad} > c e^{-C_1' r}$ on M and M' for some $c = c(C) > 0$ and $C_1' = C_1'(C) < \infty$.

Let $i \geqslant 1$. By mollification of the functions

$$\zeta_i^* := \max\{0, r - i - 1\} \quad \text{and} \quad \zeta_i'^* := \max\{0, Cr' - i - 1\}$$

(for example by application of the heat flow for some uniform time and composition with a cutoff function), we obtain approximations $\zeta_i \in C^\infty(M)$ and $\zeta_i' \in C^\infty(M')$ such that

(i) $\zeta_i \equiv 0$ on $B(p, 0, i)$ and $\zeta_i' \equiv 0$ on $B(p', 0, C^{-1} i)$.

(ii) $\zeta_i > r - i$ and $\zeta_i' > Cr' - i$.

(iii) $\zeta_i \circ \bar{\chi} < \zeta_i' + 10$.

(iv) $|\nabla^m \zeta_i|, |\nabla^m \zeta_i'| < C'$ for all $m = 1, ..., 9$, for some $C_2' = C_2'(C) < \infty$.

Set $g_t^{(i)} := \exp(2C_1' \zeta_i) g_t$ and $g_t'^{(i)} := \exp(2C_1' \zeta_i') g_t'$. By property (4), assumption (ii′) of Lemma A.7 holds for g_t and g_t' replaced by $g_t^{(i)}$ and $g_t'^{(i)}$, and C replaced by some constant $C_3' = C_3'(C) < \infty$. Moreover, the injectivity radius on $(M, g_0^{(i)})$ and $(M', g_0'^{(i)})$ is uniformly bounded from below, by a constant that may depend on i. By property (3), the map $\bar{\chi}$ is moreover C_4'-Lipschitz for some $C_4'(C) < \infty$.

We can now use Lemma A.7 to solve the harmonic map heat flow starting from $\bar{\chi}$ with the background metrics $(g_t'^{(i)})_{t \in [0,T]}$ and $(g_t^{(i)})_{t \in [0,T]}$, on a time-interval of the form $[0, \min\{\tau, T\}]$ for some $\tau = \tau(C, n) > 0$. Similarly as in the proof of Lemma A.7, these solutions then subsequentially converge to the desired solution of the harmonic map heat flow with background metrics $(g_t')_{t \in [0,T]}$ and $(g_t)_{t \in [0,T]}$. □

Using Lemmas A.5 and A.8, we can finally prove the main short-time existence result that is used in §12.

PROPOSITION A.9. (Short-time existence of harmonic map heat flow, general form) *For every $n \geqslant 1$ there exists a constant $\bar{\eta}_n > 0$ such that, for any $0 < \eta_0 < \eta_1 < \bar{\eta}_n$ and any $0 < \delta, C < \infty$, there is a constant $\tau = \tau(\eta_0, \eta_1, \delta, C, n) > 0$ such that the following holds.*

Let $(g_t)_{t \in [0,T]}$ and $(g_t')_{t \in [0,T]}$ be smooth families of Riemannian metrics on n-dimensional manifolds M and M', and consider a smooth map $\bar{\chi} : M' \to M$ such that the following holds for some $C' < \infty$.

(i) *(M, g_0) and (M, g_0') are complete.*

(ii) *$|\mathrm{Rm}(g_t)|, |\mathrm{Rm}(g_t')| \leqslant C$ on M and M' for all $t \in [0, T]$.*

(iii) *$|\nabla_{g_t}^m \mathrm{Rm}(g_t)|, |\nabla_{g_t'}^m \mathrm{Rm}(g_t')|, |\nabla_{g_t} \partial_t g_t|, |\nabla_{g_t'}^m \partial_t g_t'| \leqslant C' t^{-m/2}$ on M and M' for all $t \in (0, T]$ and all $m = 0, ..., 10$.*

(iv) $-\delta g_t \leqslant \partial_t g_t + 2\,\mathrm{Ric}(g_t) \leqslant \delta g_t$ *for all* $t \in [0,T]$.
(v) $-\delta g'_t \leqslant \partial_t g'_t + 2\,\mathrm{Ric}(g'_t) \leqslant \delta g'_t$ *for all* $t \in [0,T]$.
(vi″) $\bar{\chi}$ *is a diffeomorphism and*

$$|(\bar{\chi}^{-1})^* g'_0 - g_0| \leqslant \eta_0.$$

Then, the harmonic map heat flow equation

$$\partial_t \chi_t = \Delta_{g'_t, g_t} \chi_t, \quad \chi_0 = \bar{\chi}$$

has a smooth solution on the time-interval $[0,T^*]$ *for some* $\min\{\tau,T\} \leqslant T^* \leqslant T$ *and for* $h_t := (\chi_t^{-1})^* g'_t - g_t$ *we have* $|h_t| \leqslant \eta_1$ *for all* $t \in [0,T^*]$. *For all* $t \in [0,T^*]$, *the map* χ_t *is a diffeomorphism. Moreover, if* $|h_{T^*}| \leqslant \eta' < \eta_1$, *then* $T^* = T$.

Proof. Fix some sequence $\theta_i \to 0$. By assumptions (iii) and (vi″), we can find a sequence $\eta_0^{(i)} \to \eta_0$ such that

$$|(\bar{\chi}^{-1})^* g'_{\theta_i} - g_{\theta_i}| \leqslant \eta_0^{(i)}.$$

Let $\tau = \tau\big(\frac{1}{2}(\eta_0+\eta_1), \eta_1, \delta, C, n\big)$ be the constant from Lemma A.5. Fix some large i and assume that $\eta_0^{(i)} \leqslant \frac{1}{2}(\eta_0+\eta_1)$. By Lemma A.8 and assumptions (i)–(iii) and (vi″), we can solve the harmonic map heat flow equation

$$\partial_t \chi_t^{(i)} = \Delta_{g'_t, g_t} \chi_t^{(i)}, \quad \chi_{\theta_i}^{(i)} = \bar{\chi}. \tag{A.16}$$

on a time-interval of the form $[\theta_i, T_i]$, where $\theta_i < T_i \leqslant T$. Assume $T_i \leqslant T$ is chosen maximally with the property that (A.16) has a solution on $[\theta_i, T_i)$ and that, for

$$h_t^{(i)} := (\chi_t^{-1})^* g'_t - g_t,$$

we have $|h_t| \leqslant \eta_1$ for all $t \in [\theta_i, T_i)$. By Lemmas A.5 and A.8 we find that

$$T_i \geqslant \min\{\theta_i + \tau, T\}.$$

Note that Lemma A.5 requires a bound on $|\partial_t h_t|$, which we can obtain from assumption (iii) and Lemmas A.4 and A.3. The same bound combined with assumption (vi″) also allows us to argue that $\chi_t^{(i)}$ is a diffeomorphism for all $t \in [\theta_i, T_i)$.

We now show that we have smooth convergence of the $(\chi_t^{(i)})$ to a harmonic map heat flow (χ_t) on $[0, \min\{\tau, T\})$, after passing to a subsequence. Consider some point $q \in M'$ and set $p := \bar{\chi}(q)$. By smoothness of (g_t) and (g'_t), we can find some constant $r_q > 0$ such that Lemmas A.4 and A.3 are applicable at scale r_q near q and p and at

time θ_i, for all i. So, after passing to a subsequence, χ_t^i converges locally in the C^{10}-sense in the neighborhood of every point $(q,0)\in M'\times\{0\}$. Moreover, by assumption (iii), for any $t_0>0$, Lemma A.4 is applicable at any point of $M\times[t_0,T_i]$ and $M'\times[t_0,T_i]$ at some uniform scale $r_0=r_0(t_0,C',n)>0$, which is independent of i. Iterating this fact, and using Lemma A.3, yields local bounds on $\chi_t^{(i)}$ for any $t\in[0,\tau)$. So, after passing to a subsequence, the families of maps $(\chi_t^{(i)})$ indeed converge to a solution of the harmonic map heat flow (χ_t) on $[0,\min\{\tau,T\})$, in the C^{10}-sense. Repeated application of Lemma A.3 yields higher-derivative bounds and implies that the convergence is smooth, after passing to another subsequence. The bounds in part (b) of Lemma A.3 imply that $\chi_0=\bar{\chi}$. The fact that χ_t is a diffeomorphism for all $t\in[0,\min\{\tau,T\})$ follows from the fact that it is a limit of diffeomorphisms $\chi_t^{(i)}$ that are uniformly bilipschitz.

It remains to show the last assertion. So, assume by contradiction that $|h_{T^*}|\leqslant\eta'<\eta_1$, but $T^*<T$. Then, by Lemma A.8 we can extend the flow past time T^*, and by Lemma A.5 with $\eta_0=\eta'$ we have $|h_t|\leqslant\eta_1$ for t close to T^*, contradicting our choice of T^*. □

A.4. Further results

In the following, we will prove several analytical results that are needed in §12. The results will mostly build on the computations of §A.1.

The following proposition provides a bound on the drift of a solution to the harmonic map heat flow, whenever the associated perturbation h is small.

PROPOSITION A.10. (Drift control) *For any $n\geqslant 1$, $\delta>0$ and $A<\infty$ there is a constant $\eta=\eta(\delta,A,n)>0$ such that the following holds.*

Let $r>0$ and $T\leqslant Ar^2$ and consider Ricci flows $(g_t)_{t\in[0,T]}$ and $(g'_t)_{t\in[0,T]}$ on n-dimensional manifolds M and M'. Let $(\phi_t)_{t\in[0,T]}$, $\phi_t\colon M\to M'$, be a smooth family of diffeomorphisms onto their images whose inverses $\phi_t^{-1}\colon\phi_t(M)\to M'$ satisfy the harmonic map heat flow equation

$$\partial_t\phi_t^{-1}=\Delta_{g'_t,g_t}\phi_t^{-1}.$$

Let $x\in M$ and assume that for $x':=\phi_0(x)$, the following statements hold:

(i) *$B(x,0,r)$ is relatively compact in M.*

(ii) *$|\nabla^m\mathrm{Rm}|\leqslant Ar^{-2-m}$ on $B(x,0,r)\times[0,T]$ for all $m=0,1,2,3$.*

(iii) *$|\mathrm{Rm}|\leqslant Ar^{-2}$ on $B(x',0,r)\times[0,T]$.*

(iv) *$-\eta g'_t\leqslant h_t=\phi_t^*g_t-g'_t\leqslant\eta g'_t$ for all $t\in[0,T]$.*

Then, for all $t\in[0,T]$,

$$d_0(\phi_t(x),\phi_0(x))<\delta r.$$

We note that this proposition is similar to Lemma A.4. In fact, this lemma could be used in lieu of Proposition A.10. Nevertheless, we have included this proposition since its proof is somewhat shorter and does not use local coordinates.

Proof. By parabolic rescaling, we may assume without loss of generality that $r=1$.

Using Lemma A.2, we obtain that if η is sufficiently small depending on A and n, then for all $t\in[0,T]$ we have

$$|\nabla h_t|(x)\leqslant C_1\eta t^{-1/2},$$

where $C_1=C_1(A,n)<\infty$. So, by (A.8) and (A.7), we obtain

$$|\partial_t\phi_t^{-1}(\phi_t(x))|=|X_{g_t}(g_t+h_t)|\leqslant C_2\eta t^{-1/2},$$

where $C_2=C_2(A,n)<\infty$. As $\partial_t\phi_t=-d\phi_t(\partial_t\phi_t^{-1}\circ\phi_t)$, we obtain that

$$|\partial_t\phi_t(x)|\leqslant C_3\eta t^{-1/2},$$

where $C_3=C_3(A,n)<\infty$. Integrating this bound, and taking into account the distance distortion in M' via assumption (iii), we obtain that

$$d_0(\phi_t(x),\phi_0(x))<C_4\eta t^{1/2}$$

for some $C_4=C_4(A,n)<\infty$, as long as $C_4\eta t^{1/2}<1$. So, the proposition follows if $\eta\leqslant C_4^{-1}\min\{\delta,(2A)^{-1/2}\}$. □

Next, we derive short-time bounds for solutions to the Ricci–DeTurck equation. To do this, we first establish the following barrier-type estimate.

LEMMA A.11. *For any $n\geqslant 1$ and $A<\infty$ there is a constant $C=C(A,n)<\infty$ such that the following holds.*

Let $r>0$. Consider smooth families of metrics $(g_t)_{t\in[0,r^2]}$ and $(g_t+h_t)_{t\in[0,r^2]}$ on an n-dimensional manifold M, a point $x\in M$ and a smooth function $u\in C^\infty(B(x,r)\times[0,r^2])$ such that

(i) *$B(x,0,r)$ is relatively compact in M.*

(ii) *$\frac{1}{2}g_t\leqslant h_t\leqslant 2g_t$ on $B(x,0,r)$ for all $t\in[0,r^2]$.*

(iii) *u satisfies the inequality*

$$\partial_t u\leqslant(g_t+h_t)^{ij}\nabla^{2,g_t}_{ij}u. \tag{A.17}$$

(iv) *$|\mathrm{Rm}(g_t)|_{g_t},|\partial_t g_t|_{g_t}\leqslant Ar^{-2}$ on $B(x,0,r)$ for all $t\in[0,r^2]$.*

(v) *$|\nabla^{g_t}\partial_t g_t|_{g_t}\leqslant Ar^{-2}t^{-1/2}$ on $B(x,0,r)$ for all $t\in[0,r^2]$.*

(vi) *$|u|\leqslant 1$ on $B(x,0,r)\times[0,r^2]$.*

Then, for all $t\in[0,r^2]$ *we have*

$$u(x,t)\leqslant Ctr^{-2}+\sup_{B(x,0,r)}u_0. \tag{A.18}$$

Proof. Without loss of generality we may assume that $r=1$. Fix a function

$$f\colon[0,1]\longrightarrow[0,\infty)$$

such that $f\equiv 0$ on $[0,\frac{1}{4}]$ and $f\big(\frac{1}{2}\big)>1$ and $f'\geqslant 0$ everywhere. Set $w(y):=f(d_0(x,y))$ for all $y\in B(x,0,1)$. By Hessian comparison and assumptions (i) and (ii), there is a constant $C_1'=C_1'(n,A)<\infty$ such that $\nabla^{2,g_0}w\leqslant C_1'g_0$ on $B(x,0,1)$ in the barrier sense. By a local smoothing procedure (see for example [GW]) we can construct a smooth, non-negative $w'\in C^\infty(B(x,0,1))$ such that, for some $C_2'=C_2'(n,A)<\infty$, we have

$$|\nabla^{g_0}w'|\leqslant C_2',\quad \nabla^{2,g_0}w'\leqslant C_2'g_0\quad\text{and}\quad w'\equiv 0$$

on $B\big(x,0,\frac{1}{2}\big)$ and $w'>1$ on $B(x,0,1)\setminus B\big(x,0,\frac{1}{2}\big)$.

For any vector $v\in T_yM$, $y\in B(x,0,1)$, we have, by assumptions (ii) and (iii),

$$\begin{aligned}\left|\frac{d}{dt}\nabla^{2,g_t}_{v,v}w'\right|&=\left|dw'\left(\left(\frac{d}{dt}\nabla^{g_t}\right)(v,v)\right)\right|\\&=|dw'|_{g_t}\left|g_t^{-1}((\nabla^{g_t}_v\partial_tg_t)(v,\cdot)-\frac{1}{2}(\nabla^{g_t}_\cdot\partial_tg_t)(v,v))\right|_{g_t}\leqslant C_3't^{-1/2}\end{aligned}$$

for some $C_3'=C_3'(n,A)<\infty$. Integrating this bound over t and tracing in v, we conclude, using assumption (ii), that there is a constant $C_4'=C_4'(n,A)<\infty$ such that

$$(g_t+h_t)^{ij}\nabla^{2,g_t}_{ij}w'<C_4'\quad\text{on } B(x,0,1)\quad\text{for all } t\in[0,1]. \tag{A.19}$$

We now show that for any $\varepsilon>0$ we have

$$u_t<w'+C_4't+\sup_{B(x,0,1)}u_0+\varepsilon \tag{A.20}$$

on $B(x,0,1)$ for all $t\in[0,1]$. Evaluating this bound at x and letting $\varepsilon\to 0$ will then imply (A.18).

Note that (A.20) trivially holds for $t=0$ and for $t>0$ it can only fail on $B(x,0,\frac{1}{2})$ due to assumption (vi), since $w'>1$ on $B(x,0,1)\setminus B\big(x,0,\frac{1}{2}\big)$. Assume by contradiction that (A.20) fails for some $t\in[0,1]$. As $B\big(x,0,\frac{1}{2}\big)$ is relatively compact in M, we may assume that t is chosen minimal with this property. Then, $t>0$ and there is a point $y\in B(x,0,\frac{1}{2})$ at which equality holds in (A.20). It follows that at y we have, using (A.19),

$$\partial_tu_t-(g_t+h_t)^{ij}\nabla^{2,g_t}_{ij}u_t\geqslant C_4'-(g_t+h_t)^{ij}\nabla^{2,g_t}_{ij}w'>0.$$

This, however, contradicts (A.17).

Therefore, (A.20) holds on $B(x,0,1)$ for all $t\in[0,1]$ and all $\varepsilon>0$, which finishes the proof. □

Using Lemma A.11, we can prove the following short-time bounds for the Ricci–DeTurck flow.

PROPOSITION A.12. (Short-time bounds for Ricci–DeTurck flow) *For any $n\geqslant 1$ there is a constant $\eta_0=\eta_0(n)>0$ and for any $A<\infty$ there is a constant $C=C(A,n)<\infty$ such that the following holds.*

Let $(g_t)_{t\in[0,r^2]}$, $r>0$, be a Ricci flow on an n-dimensional manifold M and let $(h_t)_{t\in[0,r^2]}$ be a Ricci–DeTurck flow with background metric $(g_t)_{t\in[0,r^2]}$. Let $x\in M$ be a point and assume that

(i) *$B(x,0,r)$ is relatively compact in M.*

(ii) *$|\mathrm{Rm}(g_t)|\leqslant Ar^{-2}$ on $B(x,0,r)$ for all $t\in[0,r^2]$.*

(iii) *$|h_t|\leqslant\eta\leqslant\eta_0$ on $B(x,0,r)$ for all $t\in[0,r^2]$.*

Then, for all $t\in[0,r^2]$ we have

$$|h(x,t)|^2\leqslant C\eta^2tr^{-2}+\sup_{B(x,0,r)}|h_0|^2.$$

Proof. By (A.6), if η is smaller than some dimensional constant, then

$$\partial_t|h_t|^2\leqslant(g_t+h_t)^{ij}\nabla^2_{ij}|h_t|^2+C_1'|h_t|^2$$

for some constant $C_1'=C_1'(A,n)<\infty$. So, the proposition follows from Lemma A.11 by setting $u_t:=\eta^{-2}e^{-C_1't}|h_t|^2$. Note that assumption (ii) in Lemma A.11 is guaranteed if we choose η_0 sufficiently small and assumption (v) follows using Shi's estimates. □

Lastly, we prove that solutions to the Ricci–DeTurck perturbation equation remain small in a parabolic neighborhood if they are small on a larger ball at time zero. The following proposition also holds for perturbations that arise from almost Ricci flows, as discussed in §A.1.

PROPOSITION A.13. (Smallness of h at time zero implies smallness of h at later times) *For any $n\geqslant 1$ there is a constant $\eta=\eta(n)>0$ such that, for any $\varepsilon>0$ and $A<\infty$, there is a constant $0<\delta=\delta(\varepsilon,A,n)<1$ such that the following holds.*

Let $r>0$. Consider smooth families of metrics $(g_t)_{t\in[0,r^2]}$ and $(g'_t)_{t\in[0,r^2]}$ on n-dimensional manifolds M and M', respectively, as well as a smooth family of diffeomorphisms $(\chi_t)_{t\in[0,r^2]}$ between M' and M that satisfies the harmonic map heat flow equation

$$\partial_t\chi_t=\Delta_{g'_t,g_t}\chi_t.$$

Set $h_t:=(\chi_t^{-1})^*g'_t-g_t$ as in (A.2), and assume that, for some $x\in M$, the following bounds hold on $B(x,0,\delta^{-1}r)$ for all $t\in[0,r^2]$:

(i) *$B(x,0,\delta^{-1}r)$ is relatively compact in M.*

(ii) $|\mathrm{Rm}(g_t)|\leqslant Ar^{-2}$.
(iii) $|\nabla^{g_t}\partial_t g_t|\leqslant Ar^{-1}$.
(iv) $-\delta g_t\leqslant\partial_t g_t+2\,\mathrm{Ric}(g_t)\leqslant\delta g_t$.
(v) $-\delta g'_t\leqslant\partial_t g'_t+2\,\mathrm{Ric}(g'_t)\leqslant\delta g'_t$.
(vi) $|h_t|\leqslant\eta$.
(vii) $|h_0|\leqslant\delta$.
Then, $|h(x,t)|<\varepsilon$ *for all* $t\in[0,r^2]$.

Proof. By parabolic rescaling, we may assume $r=1$. The constant δ will be chosen in the course of the proof. In the following, we will always assume that $\delta\leqslant 1$.

By assumptions (ii), (iv)–(vi) and (A.6), and assuming that η is smaller than some dimensional constant, we can find a constant $C_1=C_1(A,n)<\infty$ such that h_t satisfies the evolution inequality

$$\partial_t|h_t|^2\leqslant(g_t+h_t)^{ij}\nabla^2_{ij}|h_t|^2+C_1\delta+C_1|h_t|^2$$

on $B(x,0,\delta^{-1})\times[0,1]$. So,

$$u_t:=e^{-C_1t}(|h_t|^2+\delta) \tag{A.21}$$

satisfies the evolution inequality

$$\partial_t u_t\leqslant(g_t+h_t)^{ij}\nabla^2_{ij}u.$$

We will now derive a bound on u by an argument that is analogous to the proof of Lemma A.11. Fix a non-decreasing function $f\colon[0,1]\to[0,\infty)$ such that $f\equiv 0$ on $\left[0,\frac{1}{4}\right]$ and $f\left(\frac{1}{2}\right)>1$. Set $w(y):=f(\delta\cdot d(x,y))$. By Hessian comparison and assumption (ii), we obtain that at any $y\in B(x,0,\delta^{-1})$ in the barrier sense, for $d:=d_0(x,y)$,

$$\nabla^{2,g_0}w\leqslant C_3\left(\delta^2f''(\delta\cdot d)+C_2\frac{\cosh(C_2d)}{\sinh(C_2d)}\cdot\delta f'(\delta\cdot d)\right)g_0\leqslant C_4\cdot\delta,$$

where $C_i=C_i(A,n)<\infty$ for $i=2,3,4$.

As $|\nabla^{g_0}w|\leqslant C_5\delta$ for some $C_5=C_5(A,n)<\infty$, we can argue similarly as in the proof of Lemma A.11, using assumption (iii), that there is a smoothing $w'\in C^\infty(B(x,0,\delta^{-1}))$ of w such that $w'>1$ on $B(x,0,\delta^{-1})\setminus B\left(x,0,\frac{1}{2}\delta^{-1}\right)$ and

$$(g_t+h_t)^{ij}\nabla^{2,g_t}_{ij}w'<C_6\cdot\delta\quad\text{on } B(x,0,\delta^{-1})\quad\text{for all } t\in[0,1],$$

where $C_6=C_6(A,n)<\infty$. Compare this inequality with (A.19). We can now argue similarly as for (A.20) to show that, for all $\varepsilon'>0$ and $t\in[0,1]$,

$$u_t<w'+C_6\delta\cdot t+\sup_{B(x,0,\delta^{-1})}u_0+\varepsilon'.$$

So, by assumption (vii) and letting $\varepsilon'\to 0$, we obtain that, for all $t\in[0,1]$

$$u(x,t)\leqslant C_5\delta+\delta+C_1\delta\leqslant C_6\delta.$$

The proposition now follows using (A.21) if δ is chosen small depending on A and n. □

Appendix B. Properties of Bryant solitons

In this appendix we discuss properties of the (normalized) Bryant soliton $(M_{\mathrm{Bry}}, g_{\mathrm{Bry}})$ that are needed in this paper. In the following, we denote by $x_{\mathrm{Bry}} \in M_{\mathrm{Bry}}$ the tip of the Bryant soliton, i.e. the center of rotational symmetry, and denote by $\sigma := d_{g_{\mathrm{Bry}}}(\cdot, x_{\mathrm{Bry}})$ the distance function from the tip.

LEMMA B.1. (Properties of the Bryant soliton) *There is a rotationally symmetric potential function $f \in C^\infty(M_{\mathrm{Bry}})$ such that*

$$\mathrm{Ric} + \nabla^2 f \equiv 0, \tag{B.1}$$

$$R + |\nabla f|^2 \equiv R(x_{\mathrm{Bry}}), \tag{B.2}$$

$$dR = 2\,\mathrm{Ric}(\nabla f, \cdot). \tag{B.3}$$

Moreover, there is a constant $C_B < \infty$ such that the following holds: If $\sigma > C_B$, then

$$C_B^{-1}\sigma^{-1} < R < C_B\sigma^{-1}, \tag{B.4}$$

$$\mathrm{Ric} > C_B^{-1}\sigma^{-2} g_{\mathrm{Bry}}, \quad \textit{as quadratic forms}, \tag{B.5}$$

$$-\partial_\sigma R > C_B^{-1}\sigma^{-2}, \tag{B.6}$$

$$|\nabla f| < C_B, \tag{B.7}$$

$$|\nabla R|, |\nabla^2 \mathrm{Rm}\,|, |\nabla^3 \mathrm{Rm}\,| < C_B. \tag{B.8}$$

The metric g_{Bry} is a warped product of the form $g_{\mathrm{Bry}} = d\sigma^2 + w^2(\sigma) g_{S^2}$ such that, for $\sigma > C_B$,

$$C_B^{-1}\sqrt{\sigma} < w(\sigma) < C_B\sqrt{\sigma}. \tag{B.9}$$

Moreover, for any $\sigma_0 > C_B$, if we consider the normalized function and parameter

$$\bar{w} := \frac{w}{w(\sigma_0)} \quad \textit{and} \quad \bar{\sigma} := \frac{\sigma - \sigma_0}{w(\sigma_0)},$$

and if we express $\bar{w}$ in terms of $\bar{\sigma}$, then, for all $\bar{\sigma} \in [-1, 1]$,

$$|\bar{w}(\bar{\sigma}) - 1|, \left|\frac{d\bar{w}}{d\bar{\sigma}}\right|, \left|\frac{d^2\bar{w}}{d\bar{\sigma}^2}\right| < C_B\sigma_0^{-1/2}. \tag{B.10}$$

In other words, $w^{-2}(\sigma_0) g_{\mathrm{Bry}}$ is geometrically C^2-close to a piece of a round cylinder on

$$M_{\mathrm{Bry}}(\sigma_0 - w(\sigma_0), \sigma_0 + w(\sigma_0)).$$

Proof. Identities (B.2) and (B.3) are standard bounds for a complete gradient steady soliton with bounded curvature, where f satisfies the steady gradient soliton potential equation (B.1). For (B.2) observe that the left-hand side is constant and $|\nabla f|=0$ at x_{Bry}. Identity (B.7) is a direct consequence of (B.2).

By [Bry, Theorem 1] we know that $w\sim c_1\sqrt{\sigma}$ for large σ and that the radial and orbital sectional curvatures, K_R and K_O, behave like

$$K_R\sim c_2\sigma^{-2} \quad \text{and} \quad K_O\sim c_3\sigma^{-1}.$$

Here c_1, c_2 and c_3 are positive constants that depend on the normalization of g_{Bry}. The bounds (B.4), (B.9) and (B.5) follow immediately. It also follows that R decays to zero at infinity and, therefore, by (B.2), we have $|\nabla f|^2\to R(x_{\mathrm{Bry}})$ at infinity. Combining this with (B.3) yields (B.6). The bound (B.8) follows by Shi's estimates.

Lastly, by the Jacobi equation, we obtain

$$w''=-K_Rw\sim -c_1c_3\sigma^{-3/2}.$$

Integrating this bound and using (B.9), we obtain that $w'\sim 2c_1c_3\sigma^{-1/2}$. Rescaling both bounds by $w(\sigma_0)$ implies the bounds on the second and third term in (B.10). The first bound follows by integration over $[-1,1]$ and observing that $\bar{w}(0)=1$. □

Appendix C. Properties of κ-solutions

In this section we discuss properties of κ-solutions (Definition 5.4) that are needed in the paper. We remind the reader that we are using the curvature scale function ρ from Definition 6.1.

Lemma C.1. *3-dimensional κ-solutions have the following properties*:

(a) *There is a $\kappa_0>0$ such that every 3-dimensional κ-solution $(M^3,(g_t)_{t\in(-\infty,0]})$ is either a κ_0-solution or a shrinking round spherical space form.*

(b) (Compactness) *For any $\underline{\kappa}>0$, the collection of pointed κ-solutions*

$$(M^3,(g_t)_{t\in(-\infty,0]},x),$$

with $\kappa\geqslant\underline{\kappa}$ and $R(x,0)=1$, is compact in the pointed smooth topology.

(c) *For every $A<\infty$ there is a constant $C=C(A)<\infty$ such that, for any point (x,t) in a κ-solution $(M^3,(g_t)_{t\in(-\infty,0]})$, we have*

$$C^{-1}\rho(x,t)\leqslant\rho\leqslant C\rho(x,t)$$

on the parabolic neighborhood $P(x,t,A\rho(x,t))$.

(d) *For every* $k,l,A<\infty$ *there is a constant* $C=C(A,k,l)<\infty$ *such that for any point* (x,t) *in a* κ*-solution* $(M^3,(g_t)_{t\in(-\infty,0]})$ *we have*

$$|\partial_t^k\nabla^l \operatorname{Rm}| \leqslant C_{k,l}\rho^{-2-2k-l}(x,t)$$

on the parabolic neighborhood $P(x,t,A\rho(x,t))$.

(e) *We have* $\partial_t R\geqslant 0$ *and* $-C\leqslant\partial_t R^{-1}\leqslant 0$ *on every* κ*-solution for some universal constant* $C<\infty$.

(f) *The shrinking round cylinder is the only* 3*-dimensional* κ*-solution with more than one end.*

Proof. For assertions (a), (b), (f), and the first part of (e), see [P1, §11] or [KL1, §§38–51]. Assertions (c) and (d), and the second part of (e) follow immediately from the compactness assertion (b). □

The following lemma is a variation on the geometric definition of canonical neighborhoods used by Perelman in [P2, §1.5, §4.1].

LEMMA C.2. *For every* $\delta>0$ *there is a constant* $C_0=C_0(\delta)<\infty$ *such that the following holds.*

If $(M,(g_t)_{t\in(-\infty,0]})$ *is a* 3*-dimensional* κ*-solution and* $(x,t)\in M\times(-\infty,0]$, *then one of the following holds*:

(a) *The point* (x,t) *is the center of a* δ*-neck at scale* $\rho(x,t)$.

(b) *There is a compact, connected domain* $V\subset M$ *with connected (possibly empty) boundary such that the following holds*:

(1) $B(x,t,\delta^{-1}\rho(x))\subset V$.

(2) $\rho(y_1,t)<C_0\rho(y_2,t)$ *for all* $y_1,y_2\in V$.

(3) $\operatorname{diam}_t V<C_0\rho(x,t)$.

(4) *If* $\partial V\neq\varnothing$, *then*

(i) ∂V *is a central* 2*-sphere of a* δ*-neck at time* t.

(ii) *Either* V *is a* 3*-disk and* (M,g_t) *has strictly positive sectional curvature, or* V *is diffeomorphic to a twisted interval bundle over* $\mathbb{R}P^2$ *and* (M,g_t) *is a* $\mathbb{Z}_2$*-quotient of the shrinking round cylinder.*

(iii) *Any two points* $z_1,z_2\in\partial V$ *can be connected by a continuous path inside* ∂V *whose length is less than*

$$\min\{d_t(z_1,x),d_t(x,z_2)\}-110\rho(x,t).$$

Proof. Suppose the lemma were false for some $\delta>0$. Then, there exists a sequence $(M_k,(g_{k,t})_{t\in(-\infty,0]},x_k)$ of pointed κ_k-solutions and a sequence $C_k\to\infty$ such the conclusion of the lemma fails at time zero for all k, where C_0 is replaced by C_k. Without loss of generality, we may assume that $\rho(x_k,0)=1$ for all k. Since the conclusion holds for shrinking round spherical space forms when C_k is sufficiently large, we may assume by assertion (a) of Lemma C.1 that $\kappa_k\geqslant\kappa_0>0$ for large k. So, by assertion (b) of Lemma C.1, we may assume that, after passing to a subsequence, the sequence $\{(M_k,(g_{k,t})_{t\in(-\infty,0]},x_k)\}$ converges to a pointed κ-solution $(M_\infty,(g_{\infty,t})_{t\in(-\infty,0]},x_\infty)$ in the pointed smooth topology. Note that M_∞ must non-compact, and $(M_\infty,(g_t)_{t\in(-\infty,0]})$ cannot be a shrinking round cylinder, since otherwise assertion (a) or (b) will hold for large k, contradicting our assumptions. Now, by [KL1, Lemma 59.1], its proof, and the discussing preceding the statement of that lemma, there is a compact manifold with boundary $V_\infty\subset M_\infty$ such that $B(x_\infty,0,2\delta^{-1}\rho(x_\infty,0))\subset V_\infty$, the boundary ∂V_∞ is the central 2-sphere of a $\frac{1}{2}\delta$-neck at time t, and either V_∞ is diffeomorphic to a 3-disk and $(M_\infty,(g_t)_{t\in(-\infty,0]})$ has strictly positive sectional curvature, or V_∞ is diffeomorphic to a twisted interval bundle over $\mathbb{R}P^2$ and $(M_\infty,(g_t)_{t\in(-\infty,0]})$ is isometric to a $\mathbb{Z}_2$-quotient of a shrinking round cylinder. Now, for large k, the domain V_∞ yields a compact domain with boundary satisfying assertions (i)-(iii). This is a contradiction. □

Proposition C.3. *Let $(M,(g_t)_{t\in(-\infty,0]})$ be a 3-dimensional κ-solution. If*

$$\partial_t R(p,0)=0 \quad \textit{for some } p\in M,$$

then, modulo parabolic rescaling, there is a pointed isometry of Ricci flows

$$(M,(g_t)_{t\in(-\infty,0]},p)\longrightarrow(M_{\mathrm{Bry}},(g_{\mathrm{Bry},t})_{t\in(-\infty,0]},x_{\mathrm{Bry}}).$$

Proof. The fact that $(M,(g_t)_{t\in(-\infty,0]})$ is a steady gradient soliton was shown by Hamilton in [Ha3], where he analyzed the equality case of his matrix Harnack inequality for the Ricci flow (see [Ha2]). Below we have included an alternate proof of this that incorporates several simplifications. Brendle showed that up to homothety the Bryant soliton is the only κ-solution that is a gradient steady soliton (see [Bre2]). Finally, for gradient steady solitons, we have

$$\partial_t R=dR(\nabla f)=2\,\mathrm{Ric}(\nabla f,\nabla f),$$

where f is the soliton potential. Since $\mathrm{Ric}>0$ on the Bryant soliton, we have $\nabla f(p)=0$. Because x_{Bry} is the unique critical point of the soliton potential of g_{Bry}, this forces the homothety $(M,g_0)\to(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$ to map p to x_{Bry}. □

In the remainder of this appendix, we will give a simplified proof of the first part of Proposition C.3, which was shown by Hamilton in [Ha3]. The proof is based on his matrix Harnack inequality (see [Ha2]) and Brendle's strong maximum principle in vector bundles (see [BS, §2]). The reader may also consult a more general treatment of Hamilton's Harnack inequality due to Brendle (see [Bre1]), as we will mainly rely on the terminology developed in this work. As a preparation, we briefly recall the main ideas of Hamilton's proof. The bound $\partial_t R \geqslant 0$ follows from the following theorem after passing to the limit $T \to \infty$.

THEOREM C.4. *Let $(M, (g_t)_{t\in(-T,0]})$ be a 3-dimensional Ricci flow with complete time-slices and bounded, non-negative sectional curvature. Then,*

$$\partial_t R \geqslant -\frac{R}{T+t}. \tag{C.1}$$

The proof of this bound follows from the following matrix Harnack estimate: Consider the bundle $E = TM \oplus \Lambda_2 TM$ over M. We introduce the following (time-dependent) generalized curvature quantity $S \in \mathrm{Sym}_2 E^*$:

$$S_t((x, u_1 \wedge u_2), (y, v_1 \wedge v_2)) = W(x,y) + P(u_1 \wedge u_2, y) + P(v_1 \wedge v_2, x) + R(u_1, u_2, v_2, v_1),$$

where

$$W(x,y) = (\Delta \operatorname{Ric})(x,y) - \frac{1}{2}\nabla^2_{x,y} R + 2\sum_{i,j=1}^{3} R(x, e_i, e_j, y) \operatorname{Ric}(e_i, e_j) - \sum_{i=1}^{3} \operatorname{Ric}(x, e_i) \operatorname{Ric}(e_i, y).$$

and

$$P(u_1, u_2, y) = (\nabla_{u_1} \operatorname{Ric})(u_2, y) - (\nabla_{u_2} \operatorname{Ric})(u_1, y).$$

Hamilton (see also [Bre1] for the terminology used here) observed that this generalized curvature quantity satisfies an evolution equation that is similar to the evolution equation for the curvature tensor:

$$\widetilde{D}_{\partial_t} S = \tilde{\Delta} S + Q(S).$$

Here $\tilde{\Delta} = \sum_{i=1}^{3} \widetilde{\nabla}_{e_i} \widetilde{\nabla}_{e_i}$ is the connection Laplacian on $\mathrm{Sym}_2 E^*$ with respect to the connection $\widetilde{\nabla}$ that is induced by the following connection on E:

$$\widetilde{\nabla}_z (X, \alpha) = (\nabla_z X, \nabla_z \alpha + \operatorname{Ric}(z) \wedge X)$$

and

$$\widetilde{D}_{\partial_t}(X, U_1 \wedge U_2) = \big(\partial_t X + \mathrm{Ric}(X) - \tfrac{1}{2}(U_1 \wedge U_2)(\nabla R, \cdot),\\ \partial_t(U_1 \wedge U_2) + \mathrm{Ric}(U_1) \wedge U_2 + U_1 \wedge \mathrm{Ric}(U_2)\big).$$

(X, U_1, U_2 and α denote time-dependent local sections of TM and $\Lambda_2 TM$, respectively.) The quadratic part $Q(S)$ is non-negative definite, whenever S is non-negative definite. By a more general approach, which takes into account the case in which S is indefinite, but bounded from below, Hamilton deduces the following lower bound for the quadratic form S: For all $(x, \alpha) \in E$ we have

$$S((x, \alpha), (x, \alpha)) \geqslant -\frac{1}{2(T-t)} \mathrm{Ric}(x, x). \tag{C.2}$$

This implies that

$$W(x, x) \geqslant -\frac{1}{2(T-t)} \mathrm{Ric}(x, x).$$

Tracing this equation in x yields

$$\frac{1}{2}\partial_t R = \Delta R - \frac{1}{2}\Delta R + |\mathrm{Ric}|^2 = \mathrm{tr}\, W \geqslant -\frac{1}{2(T-t)} R, \tag{C.3}$$

which implies (C.1).

We can now present the proof of Proposition C.3.

Proof of Proposition C.3. It remains to consider the case in which $\partial_t R(p, 0) = 0$ for some $p \in M$. We first argue that the all sectional curvatures on $M \times (-\infty, 0]$ are positive. If not, then by a standard strong maximum principle argument, this implies that the flow locally splits off an $\mathbb{R}$-factor. It follows that the universal covering flow is homothetic to the round shrinking cylinder, in contradiction to $\partial_t R(p, 0) = 0$.

Letting $T \to \infty$ in (C.2) we obtain that S, and hence W are non-negative definite everywhere on $M \times (-\infty, 0]$. As $\partial_t R(p, 0) = 0$, we obtain from (C.3) that $W(p, 0) = 0$. So, $S(p, 0)$ has nullity of at least 3. On the other hand, $S(p, 0)$ restricted to $0 \oplus \Lambda_2 TM$ is strictly positive definite, as the sectional curvatures at $(p, 0)$ are positive.

So, the nullity of $S(p, 0)$ is equal to 3. We can now apply the strong maximum principle due to Brendle (see [BS, §2]) and conclude that, for all $(q, t) \in M \times (-\infty, 0]$, the nullity of $S(q, t)$ is 3 and the nullspace $N_{q,t}$ of $S(q, t)$ forms a time-dependent sub-bundle in E that is invariant under parallel transport with respect to $\widetilde{\nabla}$ (in space) and $\widetilde{D}_{\partial_t}$ (in time).

Next, observe that since all sectional curvatures on $M \times (-\infty, 0]$ are positive, the sub-bundle $0 \oplus \Lambda_2 TM \subset M$ intersects N only in the origin over every $(q, t) \in M \times (-\infty, 0]$.

So, at every time $t\leqslant 0$, the vector bundle E is the direct sum of the subbundles $N_{\cdot,0}$ and $0\oplus\Lambda_2 TM$. It follows that there is a smooth, time-dependent, section $(F_t)_{t\in(-\infty,0]}$ of the endomorphism bundle $\mathrm{End}(TM,\Lambda_2 TM)$ such that, for all $t\leqslant 0$,

$$N_{\cdot,t}=\{(x,F_t(x)):x\in TM\}.$$

Let us now express the fact that N is parallel with respect to $\widetilde{\nabla}$ in terms of F, at some fixed time $t\leqslant 0$. To do this, let $q\in M$ and $w\in T_qM$, and consider a locally defined vector field X such that, at q,

$$0=\widetilde{\nabla}_z(X,F(X))=(\nabla_z X,(\nabla_z F)(X)+F(\nabla_z X)+\mathrm{Ric}(z)\wedge X).$$

It follows that $\nabla_z X=0$ and

$$(\nabla_z F)(X)+\mathrm{Ric}(z)\wedge X=0. \tag{C.4}$$

Let $A:=\mathrm{tr}_{12}F$ be the trace of the first two factors of F viewed as a section of

$$T^*M\otimes TM\otimes TM.$$

Tracing (C.4) yields

$$\nabla_z A-2\,\mathrm{Ric}(z)=\nabla_z A+\sum_{i=1}^{3}(\langle\mathrm{Ric}(z),e_i\rangle e_i-\langle e_i,e_i\rangle\,\mathrm{Ric}(z))=0.$$

So

$$(\mathcal{L}_A g)(x,y)=\langle\nabla_x A,y\rangle+\langle x,\nabla_y A\rangle=4\,\mathrm{Ric}(x,y),$$

which implies that (M,g_t) is a steady soliton. As

$$\langle\nabla_x A,y\rangle=2\,\mathrm{Ric}(x,y)$$

is symmetric in x and y, the vector field A is a gradient vector field if M is simply connected.

We can now apply Brendle's result (see [Bre2]) and conclude that the universal cover of (M,g_t) is homothetic to $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$ for all $t\leqslant 0$. Since all isometries of M_{Bry} leave x_{Bry} invariant, it follows that (M,g_t) is homothetic to $(M_{\mathrm{Bry}},g_{\mathrm{Bry}})$ for all $t\leqslant 0$. So, by uniqueness of Ricci flows with bounded curvature, the flow $(M,(g_t)_{t\in(-\infty,0]})$ has to be homothetic to the Bryant soliton. □

Appendix D. Smoothing maps

LEMMA D.1. (Smoothing bilipschitz maps between cylinders) *For every $\varepsilon>0$ there is a constant $\delta>0$ such that the following holds.*

Let

$$\phi\colon S^2\times(0,3)\longrightarrow S^2\times\mathbb{R}$$

be a $(1+\delta)$-bilipschitz map, where both cylinders are considered to be round and of the same scale. Then, there is a $(1+\varepsilon)$-bilipschitz map

$$\tilde{\phi}\colon S^2\times(0,3)\longrightarrow S^2\times\mathbb{R}$$

such that $\tilde{\phi}=\phi$ on $S^1\times(0,1)$ and such that $\tilde{\phi}|_{S^2\times(1,2)}$ is an isometry.

Proof. Let $\alpha>0$ be a small constant whose value we will determine in the course of the proof, depending on ε. A limit argument implies that, if

$$\delta\leqslant\bar{\delta}(\alpha), \tag{D.1}$$

then there is an isometric embedding $\psi\colon S^2\times(0.1,2.9)\to S^2\times\mathbb{R}$ such that

$$d(\phi(x),\chi(x))<\alpha\quad\text{for all } x\in S^2\times(0.1,2.9).$$

After replacing ϕ with $\bar{\chi}^{-1}\circ\phi$, where $\bar{\chi}\colon S^2\times\mathbb{R}\to S^2\times\mathbb{R}$ is the isometric extension of χ, we may assume without loss of generality that

$$d(\phi(x),x)<\alpha\quad\text{for all } x\in S^2\times(0.1,2.9). \tag{D.2}$$

Next, we will carry out a mollification procedure on $S^2\times(0,3)$ producing a family of bilipschitz maps $\phi'_\beta\colon S^2\times(0,2.9)\to S^2\times\mathbb{R}$ such that $\phi'_\beta=\phi$ on $S^2\times(0,1)$ and such that ϕ'_β has improved regularity on $S^2\times(1.5,2.9)$. This would be a completely standard mollification procedure, except for the fact that the scale of the mollification varies slowly. For this purpose, we fix a smooth cutoff function $\rho\colon S^2\times(0,3)\to[0,1]$, depending only on the $(0,3)$-factor, such that $\rho\equiv0$ on $S^2\times(0,1)$ and $\rho\equiv1$ on $S^2\times(1.5,3)$. Let moreover, $0<\beta<0.01$ be a constant whose value we will fix in the course of the proof. The function $\beta\rho$ will determine the scale at which we mollify ϕ.

Let

$$X:=S^2\times\mathbb{R}\subset\mathbb{R}^3\times\mathbb{R}=\mathbb{R}^4$$

be the standard embedding. Our mollification construction is similar to that in [K]. However, in our case we can simplify the construction by using the embedding $X\subset\mathbb{R}^4$ and the nearest point projection

$$\mathrm{proj}_X\colon\mathbb{R}^4\setminus(\{(0,0,0)\}\times\mathbb{R})\longrightarrow X.$$

Let $\psi\colon[0,\infty)\to[0,1]$ be a smooth cutoff function such that $\psi\equiv 1$ on $[0,\frac{1}{2}]$ and $\psi\equiv 0$ on $[1,\infty)$. Set

$$a := \int_{\mathbb{R}^3} \psi(|v|)\,dv.$$

Then, we can define $\phi'_\beta\colon S^2\times(0,2.8)\to S^2\times\mathbb{R}$ as follows:

$$\phi'_\beta(x) := \mathrm{proj}_X\left(\int_{T_xX} \phi(\exp_x(\beta\rho v))a^{-1}\psi(|v|)\,dv\right). \tag{D.3}$$

Claim. (a) *ϕ'_β is smooth.*

(b) *$\phi'_\beta\equiv\phi$ on $\{\rho=0\}$ along with all higher derivatives.*

(c) *$d(\phi'_\beta(x),x)<3(\alpha+2\beta)$, for all $x\in S^2\times(0.1,2.9)$, assuming $\alpha,\delta<0.01$.*

(d) *For any $\varepsilon'>0$, the following holds if $\beta\leqslant\bar\beta(\varepsilon')$, $\delta\leqslant\bar\delta(\varepsilon')$ and $\alpha\leqslant\bar\alpha(\varepsilon',\beta)$. The map ϕ'_β is $(1+\varepsilon')$-bilipschitz and, for all $x\in S^2\times(1.5,2.8)$, we have*

$$|(d\phi'_\beta)_x-(d\,\mathrm{id}_{S^2\times\mathbb{R}})_x|<\varepsilon'.$$

Here, we compare both differentials within the ambient space $\mathbb{R}^4$.

Proof. Assertion (a) follows from the definition of ϕ'_β and assertion (b) holds since all derivatives of ρ vanish on $\{\rho=0\}$.

For assertion (c), observe that

$$\int_{T_xX} \phi(\exp_x(\beta\rho v))a^{-1}\psi(|v|)\,dv$$

is the center of mass of a distribution that is supported on

$$\phi(B_X(x,\beta))\subset B_{\mathbb{R}^4}(\phi(x),2\beta)\subset B_{\mathbb{R}^4}(x,\alpha+2\beta).$$

Due to the convexity of the latter ball, the center of mass must be contained in the same ball, and hence the nearest point projection lies in $B_{\mathbb{R}^4}(x,2(\alpha+2\beta))$. Since $\beta<0.01$, we have

$$B_{\mathbb{R}^4}(x,2(\alpha+2\beta))\cap X\subset B_X(x,3(\alpha+2\beta)).$$

We now prove assertion (d) using a contradiction argument. Assume that assertion (d) was false for some fixed $\varepsilon'>0$. Choose a sequence $\delta_i,\alpha_i,\beta_i\to 0$ such that $\alpha_i/\beta_i\to 0$. Then, we can find a sequence of $(1+\delta_i)$-bilipschitz maps satisfying (D.2) for $\alpha=\alpha_i$ and points $x_i\in S^2\times(0.1,2.8)$ such that one of the following holds:

(A) We have

$$\big|\,|(d\phi'_{i,\beta_i})_{x_i}(v_i)|-1\big|\geqslant\varepsilon' \tag{D.4}$$

for some unit tangent vector v_i at x_i.

(B) We have $x_i\in S^2\times(1.5,2.9)$ and

$$|(d\phi'_{i,\beta_i})_{x_i}-(d\,\mathrm{id}_{S^2\times\mathbb{R}})_{x_i}|\geqslant\varepsilon'. \tag{D.5}$$

By assertion (b), we have $\rho(x_i)>0$ for large i in case (A) and $\rho(x_i)=1$ in the case (B), by the definition of ρ. Moreover, by passing to a subsequence, we may assume that one of the above cases holds for all i.

Consider the rescaled metric $g_i:=(\beta_i\rho(x_i))^{-2}g_{S^2\times\mathbb{R}}$. Then, the sequences of pointed manifolds

$$\{(S^2\times(0,3),g_i,x_i)\} \quad \text{and} \quad \{(S^2\times(0,3),g_i,\phi_i(x_i))\}$$

converge in the pointed smooth topology to pointed Euclidean space. Moreover, with respect to the corresponding rescaling, the maps ϕ_i converge in the pointed topology to maps $\phi_\infty\colon\mathbb{R}^3\to\mathbb{R}^3$, after passing to a subsequence. As the ϕ_i are $(1+\delta_i)$-bilipschitz, their limit ϕ_∞ must be a Euclidean isometry. Furthermore, note that in case (B) we have $d_{g_i}(\phi_i(y),y)\leqslant\alpha_i/\beta_i\to0$, so in this case we even have $\phi_\infty=\mathrm{id}_{\mathbb{R}^3}$.

On the other hand, due to the mollification procedure (D.3), the maps ϕ'_{i,β_i} converge smoothly to a map ϕ'_∞ with

$$\phi'_\infty(x)=\int_{\mathbb{R}^3}a^{-1}\psi(|v|)\phi_\infty(x+v)\,dv.$$

This implies that ϕ'_∞ is also a Euclidean isometry, and in case (B) we even have $\phi'_\infty=\mathrm{id}_{\mathbb{R}^3}$. This, however, contradicts the smooth convergence and (D.4) or (D.5). □

We now apply a standard gluing procedure on $S^2\times(1.5,2)$ to obtain a map

$$\tilde\phi\colon S^2\times(0,3)\longrightarrow S^2\times\mathbb{R}$$

that agrees with ϕ'_β on $S^2\times(0,1.5)$ and with $\mathrm{id}_{S^2\times\mathbb{R}}$ on $S^2\times(2,3)$. In order to ensure that this map is $(1+\varepsilon)$-Lipschitz, we use assertions (c) and (d) of the Claim and assume that

$$\alpha+2\beta\leqslant c(\varepsilon) \quad \text{and} \quad \varepsilon'\leqslant c(\varepsilon)$$

for some constant $c(\varepsilon)>0$. We now verify that we can choose α, β and δ such that these bounds, the conditions of assertions (c) and (d) of the claim and (D.2) hold. Choose $\varepsilon'\leqslant c(\varepsilon)$ and then $\beta\leqslant\min\{\bar\beta(\varepsilon'),\frac14 c\}$, where $\bar\beta(\varepsilon')$ denotes the upper bound from assertion (d) of the claim. Next, choose $\alpha\leqslant\min\{\bar\alpha(\varepsilon',\beta),\frac14 c\}$, where $\bar\alpha(\varepsilon',\beta)$ denotes the upper bound from assertion (d) of the claim. Lastly, we choose $\delta\leqslant\bar\delta(\varepsilon')$ according to assertion (d) of the claim, and $\delta\leqslant\bar\delta(\alpha)$ according to (D.1). □

References

[Al] Almgren, F. J., Jr., Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem. *Ann. of Math.*, 84 (1966), 277–292.

[AC] Anderson, G. & Chow, B., A pinching estimate for solutions of the linearized Ricci flow system on 3-manifolds. *Calc. Var. Partial Differential Equations*, 23 (2005), 1–12.

[ACK] Angenent, S. B., Caputo, M. C. & Knopf, D., Minimally invasive surgery for Ricci flow singularities. *J. Reine Angew. Math.*, 672 (2012), 39–87.

[AIK] Angenent, S. B., Isenberg, J. & Knopf, D., Degenerate neckpinches in Ricci flow. *J. Reine Angew. Math.*, 709 (2015), 81–117.

[AK] Angenent, S. B. & Knopf, D., Precise asymptotics of the Ricci flow neckpinch. *Comm. Anal. Geom.*, 15 (2007), 773–844.

[Ap1] Appleton, A., Scalar curvature rigidity and Ricci DeTurck flow on perturbations of Euclidean space. *Calc. Var. Partial Differential Equations*, 57 (2018), Paper No. 132, 23 pp.

[Ap2] — A family of non-collapsed steady Ricci solitons in even dimensions greater or equal to four. Preprint, 2017. arXiv:1708.00161[math.DG].

[Ba1] Bamler, R. H., Stability of hyperbolic manifolds with cusps under Ricci flow. *Adv. Math.*, 263 (2014), 412–467.

[Ba2] — Long-time behavior of 3-dimensional Ricci flow: introduction. *Geom. Topol.*, 22 (2018), 757–774.

[Ba3] — Long-time behavior of 3-dimensional Ricci flow: A: Generalizations of Perelman's long-time estimates. *Geom. Topol.*, 22 (2018), 775–844.

[Ba4] — Long-time behavior of 3-dimensional Ricci flow: B: Evolution of the minimal area of simplicial complexes under Ricci flow. *Geom. Topol.*, 22 (2018), 845–892.

[Ba5] — Long-time behavior of 3-dimensional Ricci flow: C: 3-manifold topology and combinatorics of simplicial complexes in 3-manifolds. *Geom. Topol.*, 22 (2018), 893–948.

[Ba6] — Long-time behavior of 3-dimensional Ricci flow: D: Proof of the main results. *Geom. Topol.*, 22 (2018), 949–1068.

[BB] Bamler, R. H. & Brendle, S., A comparison principle for solutions to the Ricci flow. *Math. Res. Lett.*, 22 (2015), 983–988.

[BK1] Bamler, R. H. & Kleiner, B., Ricci flow and diffeomorphism groups of 3-manifolds. To appear in *J. Amer. Math. Soc.*

[BK2] — Ricci flow and contractibility of spaces of metrics. Preprint, 2019. arXiv:1909.08710[math.DG].

[BK3] — Diffeomorphism groups of prime 3-manifolds. Preprint, 2021. arXiv:2108.03302[math.DG].

[BBM+] Bessières, L., Besson, G., Maillot, S., Boileau, M. & Porti, J., *Geometrisation of 3-Manifolds*. EMS Tracts in Mathematics, 13. Eur. Math. Soc. (EMS), Zürich, 2010.

[Bra] Brakke, K. A., *The Motion of a Surface by its Mean Curvature*. Math. Notes, 20. Princeton Univ. Press, Princeton, NJ, 1978.

[Bre1] Brendle, S., A generalization of Hamilton's differential Harnack inequality for the Ricci flow. *J. Differential Geom.*, 82 (2009), 207–227.

[Bre2] — Rotational symmetry of self-similar solutions to the Ricci flow. *Invent. Math.*, 194 (2013), 731–764.

[BH] Brendle, S. & Huisken, G., Mean curvature flow with surgery of mean convex surfaces in $\mathbb{R}^3$. *Invent. Math.*, 203 (2016), 615–654.

[BS] Brendle, S. & Schoen, R. M., Classification of manifolds with weakly 1/4-pinched curvatures. *Acta Math.*, 200 (2008), 1–13.

[Bry] Bryant, R. L., Ricci flow solitons in dimension three with SO(3)-symmetries. Unpublished paper, 2005. http://www.math.duke.edu/~bryant/3DRotSymRicciSolitons.pdf.

[CG] Cheeger, J. & Gromov, M., Chopping Riemannian manifolds, in *Differential Geometry*, Pitman Monogr. Surveys Pure Appl. Math., 52, pp. 85–94. Longman, Harlow, 1991.

[CZ] Chen, B.-L. & Zhu, X.-P., Uniqueness of the Ricci flow on complete noncompact manifolds. *J. Differential Geom.*, 74 (2006), 119–154.

[CGG] Chen, Y. G., Giga, Y. & Goto, S., Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33 (1991), 749–786.

[CS] Chen, Y. M. & Struwe, M., Existence and partial regularity results for the heat flow for harmonic maps. *Math. Z.*, 201 (1989), 83–103.

[CM] Colding, T. H. & Minicozzi, W. P., II, The singular set of mean curvature flow with generic singularities. *Invent. Math.*, 204 (2016), 443–471.

[DT] DeTurck, D. M., Deforming metrics in the direction of their Ricci tensors. *J. Differential Geom.*, 18 (1983), 157–162.

[ES] Evans, L. C. & Spruck, J., Motion of level sets by mean curvature. I. *J. Differential Geom.*, 33 (1991), 635–681.

[EGZ] Eyssidieux, P., Guedj, V. & Zeriahi, A., Weak solutions to degenerate complex Monge-Ampère flows II. *Adv. Math.*, 293 (2016), 37–80.

[FIK] Feldman, M., Ilmanen, T. & Knopf, D., Rotationally symmetric shrinking and expanding gradient Kähler–Ricci solitons. *J. Differential Geom.*, 65 (2003), 169–209.

[GGM] Germain, P., Ghoul, T.-E. & Miura, H., On uniqueness for the harmonic map heat flow in supercritical dimensions. *Comm. Pure Appl. Math.*, 70 (2017), 2247–2299.

[GR] Germain, P. & Rupflin, M., Selfsimilar expanders of the harmonic map flow. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 28 (2011), 743–773.

[GW] Greene, R. E. & Wu, H., On the subharmonicity and plurisubharmonicity of geodesically convex functions. *Indiana Univ. Math. J.*, 22 (1972/73), 641–653.

[GZ] Gu, H.-L. & Zhu, X.-P., The existence of type II singularities for the Ricci flow on S^{n+1}. *Comm. Anal. Geom.*, 16 (2008), 467–494.

[Ha1] Hamilton, R. S., Three-manifolds with positive Ricci curvature. *J. Differential Geometry*, 17 (1982), 255–306.

[Ha2] — The Harnack estimate for the Ricci flow. *J. Differential Geom.*, 37 (1993), 225–243.

[Ha3] — Eternal solutions to the Ricci flow. *J. Differential Geom.*, 38 (1993), 1–11.

[Ha4] — The formation of singularities in the Ricci flow, in *Surveys in Differential Geometry*, Vol. II (Cambridge, MA, 1993), pp. 7–136. Int. Press, Cambridge, MA, 1995.

[HK] Haslhofer, R. & Kleiner, B., Mean curvature flow with surgery. *Duke Math. J.*, 166 (2017), 1591–1626.

[He] Head, J., On the mean curvature evolution of two-convex hypersurfaces. *J. Differential Geom.*, 94 (2013), 241–266.

[HW] Hershkovits, O. & White, B., Nonfattening of mean curvature flow at singularities of mean convex type. *Comm. Pure Appl. Math.*, 73 (2020), 558–580.

[HS] Huisken, G. & Sinestrari, C., Mean curvature flow with surgeries of two-convex hypersurfaces. *Invent. Math.*, 175 (2009), 137–221.

[I1] Ilmanen, T., Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108 (1994), 90 pp.

[I2] — Lectures on mean curvature flow and related equations. Unpublished notes, 1995.

[K] Karcher, H., Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.*, 30 (1977), 509–541.

[KL1] Kleiner, B. & Lott, J., Notes on Perelman's papers. *Geom. Topol.*, 12 (2008), 2587–2855.

[KL2] — Singular Ricci flows I. *Acta Math.*, 219 (2017), 65–134.

[KL3] — Singular Ricci flows II, in *Geometric Analysis*, Progr. Math., 333, pp. 137–155. Birkhäuser/Springer, Cham, 2020.

[L] Lauer, J., Convergence of mean curvature flows with surgery. *Comm. Anal. Geom.*, 21 (2013), 355–363.

[MT] Morgan, J. & Tian, G., *Ricci Flow and the Poincaré Conjecture*. Clay Mathematics Monographs, 3. Amer. Math. Soc., Providence, RI; Clay Mathematics Institute, Cambridge, MA, 2007.

[P1] Perelman, G., The entropy formula for the Ricci flow and its geometric applications. Preprint, 2002. arXiv:math/0211159[math.DG].

[P2] — Ricci flow with surgery on three-manifolds. Preprint, 2003. arXiv:math/0303109[math.DG].

[SU] Schoen, R. & Uhlenbeck, K., A regularity theory for harmonic maps. *J. Differential Geometry*, 17 (1982), 307–335. Correction in *J. Differential Geometry* 18 (1983), 329.

[Si] Simons, J., Minimal varieties in riemannian manifolds. *Ann. of Math.*, 88 (1968), 62–105.

[ST] Song, J. & Tian, G., The Kähler–Ricci flow through singularities. *Invent. Math.*, 207 (2017), 519–595.

[St] Struwe, M., On the evolution of harmonic mappings of Riemannian surfaces. *Comment. Math. Helv.*, 60 (1985), 558–581.

[W1] White, B., The size of the singular set in mean curvature flow of mean-convex sets. *J. Amer. Math. Soc.*, 13 (2000), 665–695.

[W2] — Evolution of curves and surfaces by mean curvature, in *Proceedings of the International Congress of Mathematicians*, Vol. I (Beijing, 2002), pp. 525–538. Higher Ed. Press, Beijing, 2002.

[W3] — The nature of singularities in mean curvature flow of mean-convex sets. *J. Amer. Math. Soc.*, 16 (2003), 123–138.

[W4] — A local regularity theorem for mean curvature flow. *Ann. of Math.*, 161 (2005), 1487–1519.

Richard H. Bamler
Department of Mathematics
University of California, Berkeley
Berkeley, CA 94720
U.S.A.
rbamler@berkeley.edu

Bruce Kleiner
Courant Institute of Mathematical Sciences
New York University
251 Mercer St.
New York, NY 10012
U.S.A.
bkleiner@cims.nyu.edu

Received April 10, 2019