# ANCIENT LOW ENTROPY FLOWS, MEAN CONVEX NEIGHBORHOODS, AND UNIQUENESS

KYEONGSU CHOI, ROBERT HASLHOFER, OR HERSHKOVITS

ABSTRACT. In this article, we prove the mean convex neighborhood conjecture for the mean curvature flow of surfaces in $\mathbb{R}^3$. Namely, if the flow has a spherical or cylindrical singularity at a space-time point $X = (x, t)$, then there exists a positive $\varepsilon = \varepsilon(X) > 0$ such that the flow is mean convex in a space-time neighborhood of size $\varepsilon$ around $X$. The major difficulty is to promote the infinitesimal information about the singularity to a conclusion of macroscopic size. In fact, we prove a more general classification result for all ancient low entropy flows that arise as potential limit flows near $X$. Namely, we prove that any ancient, unit-regular, cyclic, integral Brakke flow in $\mathbb{R}^3$ with entropy at most $\sqrt{2\pi/e} + \delta$ is either a flat plane, a round shrinking sphere, a round shrinking cylinder, a translating bowl soliton, or an ancient oval. As an application, we prove the uniqueness conjecture for mean curvature flow through spherical or cylindrical singularities. In particular, assuming Ilmanen's multiplicity one conjecture, we conclude that for embedded two-spheres the mean curvature flow through singularities is well-posed.

## CONTENTS

## 1. Introduction

A family of surfaces $M_t \subset \mathbb{R}^3$ moves by mean curvature flow if the normal velocity at each point is given by the mean curvature vector,[1]

$$(1) \qquad (\partial_t x)^\perp = \mathbf{H}(x) \qquad (x \in M_t).$$

Given any smooth initial surface $M \subset \mathbb{R}^3$, say closed and embedded, by classical theory [Hui84, HP99] there exists a unique smooth solution $\mathcal{M} = \{M_t\}_{t \in [0,T)}$ with initial condition $M$ defined on a maximal time interval $[0,T)$. The first singular time $T < \infty$ is characterized by the fact that the curvature blows up, i.e.

$$(2) \qquad \lim_{t \nearrow T} \max_{x \in M_t} |A|(x,t) = \infty,$$

where $|A|$ denotes the norm of the second fundamental form. The main task in the study of mean curvature flow, both from the theoretical point of view and also in order to facilitate the most striking applications, is then to understand the structure of singularities, to find ways to continue the flow beyond the first singular time, and to analyze its properties.

In the mean convex case, i.e. when the mean curvature vector at every point on the surface points inwards, there is a highly developed theory. On the one hand, the flow can be continued smoothly as a surgical solution as constructed by Brendle-Huisken [BH16] and Haslhofer-Kleiner [HK17b]. This in turn facilitates topological and geometric applications, see e.g. [BHH16, HK18]. On the other hand, the flow can also be continued uniquely as a weak (generalized) solution. Weak solutions can be described either as level set solutions as in Evans-Spruck [ES91] and Chen-Giga-Goto [CGG91], or in the framework of geometric measure theory using Brakke

---

[1]The equation of course also makes sense in higher dimension and co-dimension and in other ambient manifolds. In this paper, however, we focus on evolving surfaces in $\mathbb{R}^3$.

solutions [Bra78].[2] By the deep structure theory of White [Whi00, Whi03] (see also [HK17a]) the space-time dimension of the singular set is at most one, and all blowup limits are smooth and convex. In fact, by a result of Colding-Minicozzi [CM16] the space-time singular set is contained in finitely many compact embedded Lipschitz curves together with a countable set of point singularities. Moreover, the recent work of Brendle-Choi [BC17] and Angenent-Daskalopoulos-Sesum [ADS18] provides a short list of all potential blowup limits (singularity models) in the flow of mean convex surfaces: the round shrinking sphere, the round shrinking cylinder, the translating bowl soliton [AW94], and the ancient ovals [Whi00, HH16].

In stark contrast to the above, when the initial surface is not mean convex, the theory is much more rudimentary. This is, to some extent, an unavoidable feature of the equation. In particular, as already pointed out in the pioneering work of Brakke [Bra78] and Evans-Spruck [ES91] there is the phenomenon of non-uniqueness or fattening. Angenent-Ilmanen-Chopp [AIC95] and Ilmanen-White [Whi02] gave examples of smooth embedded surfaces $M \subset \mathbb{R}^3$ whose level set flow $F_t(M)$ develops a non-empty interior at some positive time. In particular, $F_t(M)$ does not look at all like a two-dimensional evolving surface. These examples also illustrate, in a striking way, the non-uniqueness of (enhanced) Brakke flows.

In the present paper, we make some progress towards decreasing the gap between the theory in the mean convex case and the theory in the general case without curvature assumptions. Most importantly, in Theorem 1.7 (see also Theorem 1.6), we prove the mean convex neighborhood conjecture. As an application, we prove the nonfattening conjecture for mean curvature flow through cylindrical or spherical singularities (see Theorem 1.9). In particular, assuming Ilmanen's multiplicity one conjecture, we conclude that for embedded two-spheres the mean curvature flow through singularities is well-posed (see Theorem 1.12). The mean convex neighborhood conjecture is in turn a consequence of a general classification result (Theorem 1.2) for ancient low entropy flows that arise as potential limit flows near spherical or cylindrical singularities.

The rest of this introduction is organized as follows. In Section 1.1, we describe our general classification result for ancient low entropy flows. In Section 1.2, we state our results establishing the mean convex neighborhood conjecture and the uniqueness conjectures. In Section 1.3, we give an outline of the proofs.

---

[2]In the mean convex setting these notions of weak solutions are essentially equivalent [Whi00, MS08].

1.1. **Classification of ancient low entropy flows.** Let $M \subset \mathbb{R}^3$ be a surface. The entropy, introduced by Colding-Minicozzi [CM12], is defined as the supremum of the Gaussian area over all centers and all scales, namely

$$(3) \qquad \mathrm{Ent}[M] = \sup_{y \in \mathbb{R}^3, \lambda > 0} \int_M \frac{1}{4\pi\lambda} e^{-\frac{|x-y|^2}{4\lambda}} \, dA(x).$$

The entropy measures, in a certain sense, the complexity of the surface. For example, the values for a plane, sphere and cylinder are

$$(4) \quad \mathrm{Ent}[\mathbb{R}^2] = 1, \quad \mathrm{Ent}[\mathbb{S}^2] = \frac{4}{e} \sim 1.47, \quad \mathrm{Ent}[\mathbb{S}^1 \times \mathbb{R}] = \sqrt{\frac{2\pi}{e}} \sim 1.52.$$

If $\mathcal{M} = \{M_t\}_{t \in I}$ evolves by mean curvature flow, then $t \mapsto \mathrm{Ent}[M_t]$ is nonincreasing by Husiken's monotonicity formula [Hui90], hence

$$(5) \qquad \mathrm{Ent}[\mathcal{M}] := \sup_{t \in I} \mathrm{Ent}[M_t] = \lim_{t \to \inf(I)} \mathrm{Ent}[M_t].$$

For example, the entropy of a flat plane $\mathcal{P}$, a round shrinking sphere $\mathcal{S}$, a round shrinking cylinder $\mathcal{Z}$, a translating bowl soliton $\mathcal{B}$, and an ancient oval $\mathcal{O}$, are given by

$$\mathrm{Ent}[\mathcal{P}] = 1, \quad \mathrm{Ent}[\mathcal{S}] = \frac{4}{e} \sim 1.47$$

$$(6) \qquad \mathrm{Ent}[\mathcal{Z}] = \mathrm{Ent}[\mathcal{B}] = \mathrm{Ent}[\mathcal{O}] = \sqrt{\frac{2\pi}{e}} \sim 1.52.$$

By a beautiful classification result of Bernstein-Wang [BW17] any self-similarly shrinking mean curvature flow in $\mathbb{R}^3$ with entropy at most $\sqrt{2\pi/e}$ is either a flat plane, round shrinking sphere, or round shrinking cylinder. By a recent uniqueness result of Hershkovits [Her18] any self-similarly translating flow in $\mathbb{R}^3$ with entropy at most $\sqrt{2\pi/e}$ is a bowl soliton.

For the study of singularities, it is important that these concepts are also available in the non-smooth setting. Singular surfaces in Euclidean space are described most easily by two-rectifiable Radon measures, which generalize the area measure of smooth two-dimensional surfaces, see e.g. [Sim83, Ilm94]. Recall that a two-rectifiable Radon measure $\mu$ in $\mathbb{R}^3$ is a Radon measure that has a two-dimensional tangent plane of positive multiplicity at almost every point. The entropy of $\mu$ is defined as

$$(7) \qquad \mathrm{Ent}[\mu] = \sup_{y \in \mathbb{R}^3, \lambda > 0} \int \frac{1}{4\pi\lambda} e^{-\frac{|x-y|^2}{4\lambda}} \, d\mu(x).$$

As in [Bra78, Ilm94] we consider Brakke flows $\mathcal{M} = \{\mu_t\}_{t \in I}$ that are given by a family of Radon measures in $\mathbb{R}^3$ that is two-rectifiable for almost all times and satisfies

$$(8) \qquad \frac{d}{dt} \int \varphi \, d\mu_t \leq \int \left( -\varphi \mathbf{H}^2 + (\nabla\varphi)^\perp \cdot \mathbf{H} \right) d\mu_t$$

for all nonnegative test functions $\varphi$, see Section 2.1 for details. The class of all Brakke flows, as originally defined in [Bra78], is too large for most practical purposes. For example, by the very nature of the definition via the inequality (8), Brakke flows can suddenly vanish without any cause. Also, given that the initial surface is embedded, it is useful to keep track of its inside and outside. For these and other reasons the definition of Brakke flows has been refined over the years by Ilmanen [Ilm94], White [Whi05, Whi09], and their collaborators. For our purpose it is most appropriate to consider only Brakke flows that are integral (which prevents non-integer multiplicity), unit-regular (which to a certain extent prevents sudden vanishing), and cyclic (which partly keeps track of the inside and outside), see Section 2.1 for definitions and details. In particular, all Brakke flows starting at any closed embedded surface $M \subset \mathbb{R}^3$ that are constructed via Ilmanen's elliptic regularization [Ilm94] are integral, unit regular, and cyclic, and all these properties are preserved under passing to limits of sequences of Brakke flows. The entropy of a Brakke flow is defined by

$$\text{(9)} \qquad \qquad \text{Ent}[\mathcal{M}] = \sup_{t \in I} \text{Ent}[\mu_t].$$

Finally, let us recall that an *ancient* solution of a parabolic PDE, such as the mean curvature flow, is a solution that is defined for all $t \in (-\infty, T)$, where $T \leq \infty$. Ancient solutions are on the one hand some of the most interesting solutions by themselves, and on the other hand crucial for the analysis of singularities. In particular, every blowup limit is an ancient solution. A systematic study of ancient solutions of geometric flows has been pursued over the last decade by Daskalopoulos, Sesum, and their collaborators, see [Das14] for a nice overview of the main problems and key results.

We consider the following class of Brakke flows:

**Definition 1.1** (ancient low entropy flows)**.** The class of *ancient low entropy flows* consists of all ancient, unit-regular, cyclic, integral Brakke flows $\mathcal{M} = \{\mu_t\}_{t \in (-\infty, T_E(\mathcal{M})]}$ in $\mathbb{R}^3$ with

$$\text{(10)} \qquad \qquad \text{Ent}[\mathcal{M}] \leq \sqrt{\frac{2\pi}{e}},$$

where $T_E(\mathcal{M}) \leq \infty$ denotes the extinction time.

Our main theorem provides a complete classification:

**Theorem 1.2** (classification of ancient low entropy flows)**.** *Any ancient low entropy flow in* $\mathbb{R}^3$ *is either*

- *a flat plane, or*
- *a round shrinking sphere, or*
- *a round shrinking cylinder, or*
- *a translating bowl soliton, or*
- *an ancient oval.*

To explain the scope of our main classification result (Theorem 1.2) let us start by pointing out that the assumptions are essentially sharp.[3] If one removes the assumption of being unit-regular the flow can suddenly disappear, e.g. one could have a solution that looks like a translating bowl soliton until some time, but is not eternal. If one removes the assumption of being cyclic, then one can have a static or quasistatic configuration of three half-planes meeting at 120-degree angles, which have entropy 3/2. If one increases the entropy too much one can get other ancient solutions, e.g. the selfsimilarly shrinking Angenent torus [Ang92] has entropy bigger than $\sqrt{2\pi/e}$ but less than 2. Finally, removing the integer multipicity assumption would give rise to a zoo of solutions, in particular the infinite collection of shrinkers constructed by Kapouleas-Kleene-Moller and Nguyen [KKM18, Ngu14].

Let us discuss some of the most important prior classification results for solutions of the mean curvature flow of surfaces. Wang proved uniqueness of the bowl soliton among entire convex translating graphs [Wan11]. This was improved recently by Spruck-Xiao [SX17], who showed that every entire mean convex translating graph is in fact convex.[4] Without assuming that the translator is graphical, Martin, Savas-Halilaj and Smoczyk proved a uniqueness result for the bowl soliton under rather strong asymptotic assumptions [MSHS15]. This was improved by Haslhofer [Has15], who showed uniqueness of the bowl among translators that are convex and noncollapsed, and more recently by Hershkovits [Her18], who proved uniqueness among translators, not necessarily convex, that are asymptotic to a cylinder. Most closely related to the present article are the recent result by Brendle-Choi [BC17], which proves uniqueness of the bowl among noncompact ancient solutions that are noncollapsed and convex, and the recent result by Angenent-Daskalopoulos-Sesum [ADS18], which proves uniqueness of the ancient ovals among compact ancient solutions that are noncollapsed and convex. Other highly important results that play a direct role in the present paper are the classification of low entropy shrinkers by Bernstein-Wang [BW17], and the classification of genus zero shrinkers by Brendle [Bre16].

All prior classification results assume either convexity or self-similarity (or both), and the story is similar for other equations such as the Ricci flow, see e.g. [Bre13]. Our classification result seems to be the first one which assumes neither self-similarity nor convexity. To wit, convexity is not an assumption but a consequence:

**Corollary 1.3** (convexity)**.** *Any ancient low entropy flow in $\mathbb{R}^3$ is convex.*

More generally, a key feature of our classification theorem (Theorem 1.2) is that besides the entropy bound we assume almost nothing, essentially

---

[3]The entropy assumption can be relaxed, see Corollary 1.4 below.

[4]In the setting of noncollapsed solutions it was known before that every ancient noncollapsed mean convex solution of the mean curvature flow is in fact convex [HK17a].

only that the solution is ancient. In particular, a priori, an ancient low entropy flow could be quite singular and could have pathological behaviour caused by spatial infinity, e.g. there could be topological changes caused by "neckpinches at infinity", or there could be "contracting cusps" (see e.g. [Top11]), or the flow could "escape to spatial infinity" (see e.g. [IW16]).

The entropy assumption in our main classification result (Theorem 1.2) can be relaxed. To state the sharp result, let

$$(11) \quad \Upsilon := \inf\{\text{Ent}[S] \,|\, S \text{ is a smooth embedded shrinker in } \mathbb{R}^3,$$
$$\text{which is not a plane, sphere of cylinder}\}.$$

By Bernstein-Wang [BW17], the difference $\delta := \Upsilon - \sqrt{2\pi/e} > 0$ is strictly positive. Theorem 1.2 immediately implies the following corollary.

**Corollary 1.4** (sharp ancient low entropy classification). *Any ancient, unit-regular, cyclic, integral Brakke flows $\mathcal{M}$ in $\mathbb{R}^3$ with*

$$(12) \qquad\qquad\qquad Ent[\mathcal{M}] < \Upsilon$$

*is either a flat plane, a round shrinking sphere, a round shrinking cylinder, a translating bowl soliton, or an ancient oval.*

The proof of Theorem 1.2 involves, not surprisingly, many steps. For an overview of the main steps and ideas, please see Section 1.3.

1.2. **Mean convex neighborhoods and uniqueness.** The main conjecture towards reducing the gap between the theory in the mean convex case and the general case, is the following:

**Conjecture 1.5** (mean convex neighborhood conjecture[5]). *If the mean curvature flow of closed embedded surfaces has a spherical or cylindrical singularity at $(x, t)$, then there is a space-time neighborhood of $(x, t)$ in which the flow is mean convex.*

Let us explain some background. Let $\mathcal{M} = \{M_t\}$ be a Brakke flow starting at a closed embedded surface $M \subset \mathbb{R}^3$, say constructed via Ilmanen's elliptic regularization [Ilm94]. Given any space-time point $X = (x, t) \in \mathcal{M}$ and any factor $\lambda > 0$, we denote by $\mathcal{M}_{X,\lambda}$ the Brakke flow which is obtained from $\mathcal{M}$ by translating $X$ to the space-time origin and parabolically rescaling by $\lambda^{-1}$. By Huisken's monotonicity formula [Hui90] and Ilmanen's compactness theorem [Ilm94] for any sequence $\lambda_i \to 0$ one can find a subsequence $\lambda_{i'}$ such that $\mathcal{M}_{X,\lambda_{i'}}$ converges in the sense of Brakke flows to a limit $\hat{\mathcal{M}}_X$, called a tangent flow at $X$, and any tangent flow is backwardly selfsimilar. The only backwardly selfsimilar solutions of positive mean curvature are the round

---

[5]See e.g. Problem 4 on Ilmanen's problem list [Ilm03], Conjecture 10.2 in the survey of Colding-Minicozzi-Pedersen [CMP15], and the paragraph between Theorem 3.5 and Remark 3.6 in Hershkovits-White [HW18].

shrinking sphere and the round shrinking cylinder [Hui93, Whi03, CM12]. One precise formulation of the conjecture says that if some (and thus every [CIM15]) tangent flow $\hat{\mathcal{M}}_X$ is either a round shrinking sphere or a round shrinking cylinder of multiplicity one, then there exists some $\varepsilon = \varepsilon(X) > 0$ such that the flow $\mathcal{M}$ is, possibly after flipping the orientation, mean convex in the parabolic ball

$$(13) \qquad P(X, \varepsilon) = \{(x', t') \in \mathbb{R}^3 \times \mathbb{R} : |x' - x| < \varepsilon, \, t - \varepsilon^2 < t' \leq t\}.$$

The major challenge in proving the mean convex neighborhood conjecture is to upgrade the infinitesimal information about the mean convex tangent flow to a conclusion at some macroscopic scale $\varepsilon > 0$. Specifically, there are certain singularities that are not fully captured by just looking at the tangent flows. Indeed, it has been clear since the work of White [Whi00, Whi03], and reinforced in the work of Perelman [Per02, Per03], Bamler-Kleiner [BK17] and Hershkovits-White [HW18], that in order to fully capture all singularities and regions of high curvature one needs to understand all limit flows and not just all tangent flows. Given $X \in \mathcal{M}$ a limit flow is any subsequential limit of $\mathcal{M}_{X_i, \lambda_i}$ where $X_i \to X$ and $\lambda_i \to 0$. For example, in the case of the degenerate neckpinch [AV97] one needs to carefully choose a varying sequence of space-time points $X_i$ that "follow the tip" in order to see the translating bowl soliton in the limit, see [Ham95]. Potentially there could be very complex limit flows that are not self-similar, see e.g. [Ilm03, Problem 6] or [Whi03, Conjecture 3].

Some special cases of the mean convex neighborhood conjecture (Conjecture 1.5) have been established by Colding-Ilmanen-Minicozzi [CIM15] for selfsimilarly shrinking limit flows and by Hershkovits [Her18] for selfsimilarly translating limit flows. Without assuming selfsimilarity, in a very interesting recent preprint based on hard PDE techniques, Gang [Gan18] verified the mean convex neighborhood conjecture in certain regimes. However, these regimes exclude by assumption some of the main challenges of the conjecture, e.g. the potential scenario of a degenerate neckpinch with a non-convex cap.

In Theorem 1.7, we establish the mean convex neighborhood conjecture for the mean curvature flow of surfaces in $\mathbb{R}^3$ at all times.[6] Since it is somewhat easier to state, let us first give the result at the first singular time.

**Theorem 1.6** (mean convex neighborhoods at the first singular time)**.** *Let* $\mathcal{M} = \{M_t\}_{t \in [0,T)}$ *be a mean curvature flow of closed embedded surfaces in* $\mathbb{R}^3$, *where $T$ is the first singular time. If some tangent flow at $X = (x, T)$ is a sphere or cylinder with multilicity one, then there exists an $\varepsilon = \varepsilon(X) > 0$ such that, possibly after flipping the orientation, the flow $\mathcal{M}$ is mean convex*

---

[6]Of course our results carry over immediately to other ambient manifolds.

*in the parabolic ball $P(X, \varepsilon)$. Moreover, any limit flow at $X$ is either a round shrinking sphere, a round shrinking cylinder, or a translating bowl soliton.*[7]

Theorem 1.6 follows, via a short argument by contradiction, from our classification of ancient low entropy flows (Theorem 1.2).

Before stating the solution of the mean convex neighborhoods conjecture at subsequent singularities, let us explain better what it actually means to continue the mean curvature flow through singularities. We start by recalling some facts that are explained in more detail in [Ilm98, HW18]. For any closed set $K \subset \mathbb{R}^3$, its level set flow $F_t(K)$ is the maximal family of closed sets starting at $K$ that satisfies the avoidance principle. Now, given any closed embedded surface $M \subset \mathbb{R}^3$ there are at least three quite reasonable ways to evolve it through singularities, namely the *level set flow* $F_t(M)$, the *outer flow* $M_t$ and the *inner flow* $M_t'$. The latter two are defined as follows. Let $K$ be the compact domain bounded by $M$, and let $K' := \overline{\mathbb{R}^3 \setminus K}$. Denote the corresponding level set flows by

(14) $$K_t = F_t(K), \quad \text{and} \quad K_t' = F_t(K').$$

Let $\mathcal{K}$ and $\mathcal{K}'$ be their space-time tracks, namely

(15) $$\mathcal{K} = \{(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+ \,|\, x \in K_t\}$$
$$\mathcal{K}' = \{(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+ \,|\, x \in K_t'\}.$$

The outer flow and inner flow are then defined by

(16) $$M_t = \{x \in \mathbb{R}^3 \,|\, (x, t) \in \partial\mathcal{K}\}$$
$$M_t' = \{x \in \mathbb{R}^3 \,|\, (x, t) \in \partial\mathcal{K}'\}.$$

As long as the evolution is smooth, all three flows are of course the same.

Let $T$ be the first singular time, and for $x \in M_T$ and $\lambda > 0$, denote by $\mathcal{K}_{X,\lambda}$ and $\mathcal{K}_{X,\lambda}'$ the flows that are obained from $\mathcal{K}$ and $\mathcal{K}'$, respectively, by shifting $X = (x, T)$ to the origin, and parabolically rescaling by $\lambda^{-1}$. Observe that the assumption of Theorem 1.6 is equivalent to the assumption that for $\lambda \to 0$ either $\mathcal{K}_{X,\lambda}$ or $\mathcal{K}_{X,\lambda}'$ converges smoothly with multiplicity one to a round shrinking solid cylinder or round shrinking solid ball. Also observe that the conclusion of Theorem 1.6 is equivalent to the assertion that either

(17) $$K_{t_2} \cap B(x, \varepsilon) \subseteq K_{t_1} \setminus M_{t_1} \quad \text{or} \quad K_{t_2}' \cap B(x, \varepsilon) \subseteq K_{t_1}' \setminus M_{t_1}'$$

for all $T - \varepsilon^2 < t_1 < t_2 \leq T$. This reformulation has the advantage that it generalizes not just beyond the first singular time, but even beyond any potential fattening-time (see equation (20) below). We prove:

---

[7]It is easy to see that ancient ovals cannot arise as limit flows at the first singular time. It is unknown, whether or not ancient ovals can arise as limit flows at subsequent singularities. This is related to potential accumulations of neckpinches, see e.g. [CM17].

**Theorem 1.7** (mean convex neighborhoods at all times). *Assume $X = (x, t)$ is a space-time point such that $\mathcal{K}_{X,\lambda}$ converges for $\lambda \to 0$ smoothly with multiplicity one to a round shrinking solid cylinder or a round shrinking solid ball. Then there exists an $\varepsilon = \varepsilon(X) > 0$ such that*

$$(18) \qquad K_{t_2} \cap B(x, \varepsilon) \subseteq K_{t_1} \setminus M_{t_1}$$

*for all $T - \varepsilon^2 < t_1 < t_2 < T + \varepsilon^2$. Similarly, if $\mathcal{K}'_{X,\lambda}$ converges for $\lambda \to 0$ smoothly with multiplicity one to a round shrinking solid cylinder or a round shrinking solid ball, then for some $\varepsilon = \varepsilon(X) > 0$ we have*

$$(19) \qquad K'_{t_2} \cap B(x, \varepsilon) \subseteq K'_{t_1} \setminus M'_{t_1}$$

*for all $T - \varepsilon^2 < t_1 < t_2 < T + \varepsilon^2$. Furthermore, in both cases, any limit flow at $X$ is either a round shrinking sphere, a round shrinking cylinder, a translating bowl soliton, or an ancient oval.*

Note that Theorem 1.7 establishes the strongest version of the mean convex neighborhood conjecture, where the parabolic neighborhood is backward and forward in time. The proof of Theorem 1.7 is again based on our classification of ancient low entropy flows (Theorem 1.2). During the course of the proof, we also construct a unit-regular, cyclic, integral Brakke flow whose support is $M_t$ or $M'_t$, respectively, and which has density less than 2 in $P(X, \delta)$, for some $\delta > 0$. This highlights another advantage of our formulation of Theorem 1.7: its applicability even without assuming a priori that a Brakke flow starting from $M$ has multiplicity one almost everywhere.

Let us now discuss the applications concerning uniqueness of mean curvature flow through singularities. We recall from Ilmanen-White [Whi02] that there are examples of singularities with curvature of mixed signs, which cause non-uniqueness, i.e. fattening. The main conjecture regarding uniqueness at cylindrical or spherical singularities can be phrased as follows:

**Conjecture 1.8** (nonfattening conjecture[8]). *If all singularities of the level set flow are either cylindrical or spherical then the level set flow does not fatten.*

We recall that the *fattening time* is defined by

$$(20) \qquad T_{\mathrm{fat}} = \inf\{t > 0 \,|\, F_t(M) \text{ has non-empty interior}\},$$

and the *discrepancy time* is defined by

$$(21) \qquad T_{\mathrm{disc}} = \inf\{t > 0 \,|\, M_t, \, M'_t, \text{ and } F_t(M) \text{ are not all equal}\},$$

see [HW18]. It follows directly from the definitions that

$$(22) \qquad T_{\mathrm{fat}} \geq T_{\mathrm{disc}}.$$

---

[8]The conjecture has been indicated first in White's ICM lecture [Whi02], see also the discussion in the recent work of Hershkovits-White [HW18] (indeed, the nonfattening conjecture is a combination of [HW18, Conj. 1.2] and the "very interesting open problem" described in the paragraph between Theorem 3.5 and Remark 3.6 in [HW18]).

It is unknown whether or not $T_{\text{fat}} = T_{\text{disc}}$ holds, see [HW18, Ex. 2.5]. In any case, if one proves the (potentially) stronger result that there is no discrepancy, then one obtains nonfattening as a consequence. We prove:

**Theorem 1.9** (nonfattening). *Suppose that* $0 < T \leq T_{disc}$, *and suppose that all the backward singularities of the outer flow* $\{M_t\}$ *at time* $T$ *are cylindrical or spherical. Then* $T < T_{disc}$. *In particular, the level set flow does not fatten as long as all singularities are cylindrical or spherical.*

Theorem 1.9 follows immediately by combining the main theorem of Hershkovits-White [HW18], which establishes nondiscrepancy assuming the existence of mean convex neighborhoods around all singularities a priori, and Theorem 1.7, which proves the existence of such mean convex neighborhoods.

Our final application in this paper concerns the well-posedness problem for the mean curvature flow of embedded two-spheres through singularities. The main conjecture in this regard is the following.

**Conjecture 1.10** (evolution of embedded two-spheres[9]). *The mean curvature flow of embedded two-spheres through singularities is unique.*

One reason why Conjecture 1.10 is highly intriguing, is that the mean curvature flow of embedded two-spheres can be viewed as (somewhat more approachable) cousin of the Ricci flow of four-spheres, whose analysis seems to be out of reach with current technology.[10] A solution of Conjecture 1.10 would also facilitate several topological and geometric applications. E.g. there is a proposal of Yau for a flow proof of the Smale conjecture [Rub13] (see also the work of Bamler-Kleiner [BK17] and Buzano-Haslhofer-Hershkovits [BHH16]). Moreover, a well-posed flow of embedded-two spheres would also be useful for the Lusternik-Schnirelman problem of finding four embedded minimal two-spheres in three-spheres equipped with an arbitrary Riemannian metric (the first such minimal two-spheres has been found by Simon-Smith [Smi82] in 1983, and a second one has been found recently by Haslhofer-Ketover [HK18] using combined efforts of mean curvature flow and min-max theory).

Some very important recent progress towards Conjecture 1.10 has been made by Brendle [Bre16], who proved that the only nontrivial shrinkers of genus zero are the round shrinking sphere and the round shrinking cylinder.

---

[9]See e.g. White's ICM lecture [Whi02], and the introduction of Bamler-Kleiner [BK17].

[10]When comparing mean curvature flow and Ricci flow one has to multiply the dimension by two. For example, convergence to a round limit holds for curve-shortening flow and two-dimensional Ricci flow, respectively. Also, Huisken's monotonicity formula combined with Gauss-Bonnet gives local bounds for the $L^2$-norm of the second fundamental form, see [Ilm95], which is scale invariant in dimension two, while on the other hand Perelman's monotonicity formula combined with Gauss-Bonnet gives local bounds for the $L^2$-norm of the Riemann tensor, see [HM11], which is scale invariant in dimension four.

Combining Brendle's uniqueness result with our solution of the nonfattening conjecture (Theorem 1.9), we can prove Conjecture 1.10 modulo (a special case of) Ilmanen's multiplicity one conjecture:

**Conjecture 1.11** (multiplicity one conjecture for two-spheres[11]). *For the mean curvature flow starting at any closed embedded two-sphere $M$,[12] all tangent flows have multiplicity one.*

**Theorem 1.12** (evolution of embedded two-spheres). *Assuming the multiplicity one conjecture (Conjecture 1.11), mean curvature flow of embedded two-spheres through singularities is unique.*

Let us conclude this section by sketching a heuristic picture for the mean curvature flow of embedded surfaces of higher genus: By Ilmanen's strict genus reduction conjecture [Ilm03, Conjecture 13] one expects only a controlled number (bounded by the genus of the initial surface) of singularities modelled on asymptotically conical shrinkers of higher genus, see in particular the important work of Brendle [Bre16] and Wang [Wan16]. At these singularities the evolution can be non-unique, and one has to decide how to flow out of the conical singularities, either by hand or via stability or via a stochastic selection principle. Our solution of the nonfattening conjecture (Theorem 1.9) suggests that the evolution is determined completely by these finitely many choices.

1.3. **Outline of the proofs.** For the outline of the proof of the classification theorem for ancient low entropy flows (Theorem 1.2), let $\mathcal{M}$ be an ancient low entropy flow (see Definition 1.1) which is not a flat plane, round shrinking sphere, or round shrinking cylinder. The task is to prove that $\mathcal{M}$ is either a translating bowl soliton or an ancient oval.

In Section 3, we start by establishing several coarse properties:

First, we prove a partial regularity result (Theorem 3.2), which shows that $\mathcal{M}$ is smooth except for at most countably many spherical singularities. This is mostly based on ideas from Bernstein-Wang [BW17], and uses the assumptions that our Brakke flow is integral, unit-regular and cyclic.

Next, given any space-time point $X = (x, t) \in \mathcal{M}$ we consider the flow on backwards parabolic annuli around $X$ with radii $r_j = 2^j$. We prove a rough neck theorem (Theorem 3.5), which shows that there exists a scale $Z(X) = 2^{J(X)}$, the *cylindrical scale*, at which the flow starts to look $\varepsilon$-cylindrical, and a controlled number $N = N(\varepsilon) < \infty$, such that the flow looks $\varepsilon$-cylindrical

---

[11]Multiplicity one was called an hypothesis in Brakke's pioneering monograph [Bra78], and promoted to a conjecture in Ilmanen's paper [Ilm95], where he proved that for surfaces one can always find smooth tangent flows; see also the approach by Ecker [Eck07].

[12]To be concrete, we mean by that a unit-regular, cyclic, integral Brakke flow, whose support is the outer flow, and whose initial data is $\mathcal{H}^n \lfloor M$. A posteriori, this will turn out to imply that the (nonfattening) matching motion has only multiplicity one tangents.

at all scales $j \geq J(X) + N$. The proof uses the Bernstein-Wang classification of low entropy shrinkers [BW17], Huisken's monotonicity formula [Hui90], and ideas from quantitative differentiation, see e.g. [CHN13].

Third, we prove the vanishing asymptotic slope theorem (Theorem 3.9), which shows that at spatial infinity, the surfaces $M_t$ open up slower than any cone of positive angle. The proof is based on a blowdown argument, which shows that the cylindrical scale grows sublinearly. The result facilitates barrier arguments in later sections.

In Section 4, which is the longest section of the proof, we carry out a fine neck analysis. Given any point $X = (x_0, t_0) \in \mathcal{M}$, we consider the rescaled flow

$$(23) \qquad \bar{M}_\tau^X = e^{\frac{\tau}{2}} \left( M_{-e^{-\tau}} - x_0 \right),$$

where $\tau = -\log(t_0 - t)$. The surfaces $\bar{M}_\tau^X$ can be written as the graph of a function $u_X(\cdot, \tau)$ with small norm over a cylindrical domain of length $\rho(\tau)$, where $\rho(\tau) \to \infty$ as $\tau \to -\infty$. By Colding-Minicozzi [CM15], we can assume that the axis of each cylinder is in $z$-direction. The goal is to derive very sharp asymptotic estimates for the function $u_X(\cdot, \tau)$.

To get started, in Section 4.1 we set up the fine neck analysis similarly as in Angenent-Daskalopoulos-Sesum [ADS15] and Brendle-Choi [BC17]. The analysis is governed by the linear operator

$$(24) \qquad \mathcal{L} = \Delta - \tfrac{1}{2} x^{\tan} \cdot \nabla + 1$$

on the round cylinder. The operator $\mathcal{L}$ has four positive eigenfunctions $(1, z, \sin\theta, \cos\theta)$, three zero-eigenfunctions $(z^2 - 2, z\cos\theta, z\sin\theta)$, and countably many negative eigenfunctions. Using the ODE-lemma from Merle-Zaag [MZ98], we see that for $\tau \to -\infty$ either the plus mode is dominant or the neutral mode is dominant.

In Section 4.2, we analyze the case where the plus mode is dominant. Our key result in that section is the fine neck theorem (Theorem 4.15). It says that there exists a constant $a = a(\mathcal{M}) \neq 0$, independent of the base point $X$, such that, after re-centering, we have the asymptotic expansion

$$(25) \qquad u_X(x, \tau) = aze^{\tau/2} + o(e^{\tau/2})$$

whenever $|x| \leq 100$ and $\tau \ll \log Z(X)$. Assuming $a = 1/\sqrt{2}$ without loss of generality, the estimate (25) says that the area of the cross sections of the fine necks increases at unit rate if one moves in positive $z$-direction.[13] To prove the fine neck theorem, we first use barrier arguments and the results from Section 3 to show that the cylindrical radius $\rho(\tau)$ grows exponentially as $\tau \to -\infty$. We then project onto the plus mode (after multiplying by a

---

[13]It is an instructive exercise to check that the translating bowl soliton indeed satisfies the fine neck estimate, where $a$ is proportional to the reciprocal of the speed of the translating soliton.

cutoff function) and consider the evolution equations for the coefficients of

$$(26) \qquad\qquad P_+ \hat{u}_X = a_X z + b_X \cos\theta + c_X \sin\theta + d_X.$$

Carefully analyzing these evolution equations and estimating all the error terms, after re-centering, we obtain the estimate (25).

In Section 4.3, we analyze the case where the neutral mode is dominant. Our main result in that section is the inwards quadratic neck theorem (Theorem 4.22), which gives an inwards quadratic bending of the central neck. The result is related to the main result of Angenent-Daskalopoulos-Sesum [ADS15], but we assume neither convexity nor rotational symmetry. Roughly speaking, the method of our proof of the inwards quadratic neck theorem is as follows. We consider the expansion

$$(27) \qquad\qquad \hat{u}_X = \alpha_1(z^2 - 2) + \alpha_2 z \cos\theta + \alpha_3 z \sin\theta + o(|\vec{\alpha}|),$$

and derive some differential inequalities for the coefficients $\alpha_1, \alpha_2, \alpha_3$. Analyzing these differential inequalities carefully, we show that there is some positive constant $A > 0$ such that for $\tau \to -\infty$ we have

$$(28) \qquad\qquad \alpha_1(\tau) = \frac{-A + o(1)}{|\tau|}, \qquad |\alpha_2| + |\alpha_3| = o(|\alpha_1|).$$

The proof is quite technical, since we have to estimate the optimal graphical radius, estimate the error terms, and analyze the differential inequalities, and all those steps are interrelated. Finally, combining the inwards quadratic neck theorem with a barrier argument, we see that in the neutral mode case the solution must be compact (Corollary 4.25).

In Section 5, we assume that the plus mode is dominant.

In Section 5.1, we prove a global curvature estimate (Theorem 5.6), which says that the flow is eternal with uniformly bounded curvature. In particular, this rules out potential pathologies caused by spatial infinity. Roughly speaking, the idea is that by the fine neck estimate (25), the necks open up at spatial infinity, and thus the curvature decays to zero at spatial infinity. The actual proof is somewhat more involved, since we have to relate the cylindrical scale and the regularity scale. This gives a curvature bound on compact time intervals. Together with a local type I estimate, which follows from a simple blowdown argument, this can be upgraded to a global curvature bound. In particular, the solution is eternal.

In Section 5.2, we estimate the cap size and analyze the asymptotics. To begin with, we consider the "height of the tip" function

$$(29) \qquad\qquad \psi(t) := \inf_{x \in M_t} x_3.$$

This is a strictly increasing function and the infimum is attained at some point $p_t \in M_t$. In the cap size control theorem (Theorem 5.9), we prove that there is a uniform constant $C = C(\mathcal{M}) < \infty$ such that every point in

$M_t \setminus B_C(p_t)$ lies on a fine neck. We also show that $M_t \setminus B_C(p_t)$ is the graph of a function $r$ in cylindrical coordinates around the $z$-axis satisfying

(30) $$r(t, z, \theta) = \sqrt{2(z - \psi(t))} + o(\sqrt{z - \psi(t)}),$$

and that the height of the tip function $\psi$ satisfies

(31) $$\psi(t) = t + o(|t|).$$

Roughly speaking, the idea is that the fine neck estimate (25) can be integrated to get the shape of a parabola. In order to do this, we first prove, via a blowdown argument, that the tip does not slow down too much, which gives us enough "fast points" in space-time around which the fine neck estimate can be applied.

In Section 6, we still assume that the plus mode is dominant.

First, in the fine asymptotics theorem (Theorem 6.6) we prove that, after shifting $\mathcal{M}$ in space-time, the function $r$ from (30) becomes rotationally symmetric at a very fast rate, namely

(32) $$|\partial_\theta r| = O(1/r^{100})$$

uniformly in time. The proof is based on the Brendle-Choi neck improvement theorem [BC17] and some careful estimates controlling how the fine necks align with each other. In particular, the estimate uniformly bounds the motion of the tip in the $xy$-plane. Combining the fine asymptotics theorem with a parabolic version of the moving plane method, we show that the solution is rotationally symmetric (Theorem 6.14).

In Section 7, we complete the classification of ancient low entropy flows.

In Theorem 7.1, we show that if the plus mode is dominant, then the solution must be a bowl soliton. To this end, we analyze the rotationally symmetric solution from Section 6. Using elementary geometric arguments we show that the height function $z : M_t \to \mathbb{R}$ does not have local maxima, and similarly that the radius function $r = r(z, t)$ does not have local maxima. Thus, the function $f = \langle \nu, e_3 \rangle$ is positive. Together with the fine asymptotic theorem and the maximum principle, we conclude that $H = f$. Hence, the solution is a mean convex noncollapsed translating soliton, which by the uniqueness result for translators from Haslhofer [Has15] yields that the solution is a bowl soliton.

Finally, in Theorem 7.3 we show that if the neutral mode is dominant, then the solution must be an ancient oval. The idea is that via blowup around the tips we can find eternal low entropy flows, which by the above must be bowl solitons. We also have a central cylinder. Using the maximum principle, we show that the solution is mean convex and noncollapsed also in the region in between. Using the result of Angenent-Daskalopoulos-Sesum [ADS18], we can then conclude that the solution is an ancient oval.

Finally, let us outline the proof of the mean convex neighborhood conjecture (Theorem 1.7). Under the assumptions of Theorem 1.7, we first construct a unit-regular, cyclic, integral Brakke flow $\mathcal{M} = \{\mu_t\}_{t \geq t_0 - \delta}$ whose support is equal to the outer flow $\{M_t\}_{t \geq t_0 - \delta}$. Moreover, we can arrange that $\mathcal{M}$ has a multiplicity one cylindrical tangent flow at $X_0$, and that $\mathcal{M}$ is smooth in $P(X_0, \delta)$ at almost every time, with only spherical and cylindrical singularities. The latter properties follow from upper semicontinuity of Huisken's density and the basic regularity theory from Section 3.

We then show that, possibly after decreasing $\delta$, the mean curvature does not vanish at any regular point in $P(X_0, \delta)$. If this did not hold, then we could pass to a limit flow which on the one hand had a point with vanishing mean curvature, but on the other hand had strictly positive mean curvature by our classification result of ancient low entropy flows (Theorem 1.2).

Next, we show that, after decreasing $\delta$ again, the space-time connected component that reaches $X_0$ has positive mean curvature at all regular points. The proof is based on a continuity argument in time and uses that the singular set is small.

Finally, arguing as in [HW18] we show that the mean curvature is positive also on the nonsmooth points in the sense that the surface moves everywhere strictly inwards. This concludes the outline of the proof.

## 2. Preliminaries

2.1. **Brakke flows.** As in Ilmanen [Ilm94, Def. 6.2, 6.3] a two-dimensional *Brakke flow* in $\mathbb{R}^3$ is a family of Radon measures $\mathcal{M} = \{\mu_t\}_{t \in I}$ in $\mathbb{R}^3$ that is two-rectifiable for almost all times and satisfies

$$(33) \qquad \frac{d}{dt} \int \varphi \, d\mu_t \leq \int \left( -\varphi \mathbf{H}^2 + (\nabla \varphi)^\perp \cdot \mathbf{H} \right) d\mu_t$$

for all test functions $\varphi \in C_c^1(\mathbb{R}^3, \mathbb{R}_+)$. Here, $\frac{d}{dt}$ denotes the limsup of difference quotients, $\perp$ denotes the normal projection, and $\mathbf{H}$ denotes the mean curvature vector of the associated varifold $V_{\mu_t}$, which is defined via the first variation formula and exists almost everywhere at almost all times. The integral on the right hand side is interpreted as $-\infty$ whenever it does not make sense literally.

All Brakke flows $\mathcal{M} = \{\mu_t\}_{t \in I}$ that we encounter in the present paper have the following three additional good properties of being:

- *integral* (c.f. [Bra78, Ilm94]): $\mu_t$ is integer two-rectifiable for almost all $t$.
- *unit-regular* (c.f. [Whi05, SW16]): Every spacetime point of Gaussian density one is a regular point, i.e. for all $X = (x,t) \in \mathcal{M}$ with $\Theta_X = 1$, there exists an $\varepsilon = \varepsilon(X) > 0$ such that $t' \mapsto \mathrm{spt}(\mu_{t'}) \cap B(x,\varepsilon)$ is a smooth mean curvature flow for $t' \in [t - \varepsilon^2, t + \varepsilon^2]$.
- *cyclic* (c.f. [Whi09]): For almost all $t$ the associated $\mathbb{Z}_2$ flat chain $[V_{\mu_t}]$ is cyclic. Here, the $\mathbb{Z}_2$ flat chain $[V]$ associated to an integral varifold $V$ is called cyclic if $\partial[V] = 0$.

By Allard's closure theorem [All72, Thm. 6.4] and a result of White [Whi09, Thm. 3.3], respectively, being integral and cyclic is preserved under varifold convergence with locally bounded first variation. Limits of sequences of Brakke flows can be taken via Ilmanen's compactness theorem [Ilm94, Sec. 7]. As a consequence of the above quoted results (see also White [Whi09, Thm. 4.2, Rmk. 4.4]), being integral and cyclic is preserved under limits of Brakke flows. By the local regularity theorem [Whi05] (see also Schulze-White [SW16, Thm. 4.2]), being unit-regular is also preserved under limits of Brakke flows. In particular, Brakke flows starting at any closed embedded surface $M \subset \mathbb{R}^3$ that are constructed via Ilmanen's elliptic regularization [Ilm94] are integral, unit-regular, and cyclic. We also recall from Ilmanen [Ilm94, Sec. 10] that $M_t := \mathrm{spt}(\mu_t)$ satisfies the avoidance principle.

## 2.2. Monotonicity formula and tangent flows.

Let $\mathcal{M} = \{\mu_t\}_{t \in I}$ be a two-dimensional unit-regular, integral Brakke flow in $\mathbb{R}^3$, say with bounded area ratios.[14] Given a space-time point $X_0 = (x_0, t_0) \in \mathcal{M}$, let

$$(34) \qquad \rho_{X_0}(x,t) = \frac{1}{4\pi(t_0 - t)} e^{-\frac{|x - x_0|^2}{4(t_0 - t)}} \qquad (t < t_0).$$

By Huisken's monotonicity formula [Hui90] (see also [Ilm95]) we have

$$(35) \qquad \frac{d}{dt} \int \rho_{X_0} \, d\mu_t \leq - \int \left| \mathbf{H} - \frac{(x - x_0)^\perp}{2(t - t_0)} \right|^2 \rho_{X_0} \, d\mu_t.$$

The Gaussian density of $\mathcal{M}$ at $X_0$ is defined by

$$(36) \qquad \Theta_{X_0}(\mathcal{M}) = \lim_{t \nearrow t_0} \int \rho_{X_0}(x,t) \, d\mu_t(x).$$

It follows from the monotonicity formula that the Gaussian density is a well defined real number $\geq 1$ and that the map

$$(37) \qquad (x,t) \to \Theta_{(x,t)}(\mathcal{M})$$

is upper-semicontinous. Moreover, by the local regularity theorem for the mean curvature flow [Bra78, Whi05] there exists a universal constant $\varepsilon_0 > 0$ such that any $X \in \mathcal{M}$ with $\Theta_X \leq 1 + \varepsilon_0$ is a regular point.

---

[14]In particular, any flow with finite entropy has bounded area ratios.

Given $X \in \mathcal{M}$ and $\lambda_i \to 0$, let $\mathcal{M}^i$ be the Brakke flow which is obtained from $\mathcal{M}$ by translating $X$ to the space-time origin and parabolically rescaling by $\lambda_i^{-1}$. By the compactness theorem for Brakke flows [Ilm94] one can pass to a subsequential limit $\hat{\mathcal{M}}_X$, which is called a tangent flow at $X$. By the monotonicity formula every tangent flow is backwardly selfsimilar, i.e. $\hat{\mathcal{M}}_X \cap \{t \leq 0\}$ is invariant under parabolic dilation $\mathcal{D}_\lambda(x, t) = (\lambda x, \lambda^2 t)$. If $\mathcal{M}$ is ancient, then for any $\lambda_i \to \infty$ one can also pass along a subsequence to a backwardly selfsimilar limit $\check{\mathcal{M}}$, which is called a tangent flow at infinity.

## 3. Coarse properties of ancient low entropy flows

3.1. **Partial regularity.** The goal of this section is to prove a partial regularity result for ancient low entropy flows. To this end, we start with the following lemma.

**Lemma 3.1** (low entropy cyclic minimal cones). *Let $\mu$ be an integer two-rectifiable Radon measure in $\mathbb{R}^3$ with $Ent[\mu] < 2$. If the associated varifold $V_\mu$ is a cyclic minimal cone, then $\mu = \mathcal{H}^2 \lfloor P$ for some flat plane $P$.*

*Proof.* The proof is along the lines of [BW17, Lem. 4.1]. If $x \in \mathrm{spt}(\mu) - \{o\}$, since being cyclic is preserved under weak limits, any tangent cone $C$ at $x$ is a cyclic minimal cone which splits off a line. Such a minimal cone consists of a static configuration of even number of half-planes meeting along a common line. The entropy assumption implies that there are only two of those half-planes, and since the configuration is static, it follows that $C$ is a multiplicity one plane. Hence, by Allard's regularity theorem [All72], $x$ is a regular point, and so the only potential singularity is at the vertex $o$. As the link of $\mathrm{spt}(\mu)$ is a smooth multiplicity one geodesic in $S^2$, it is a great circle, and so $\mu$ is indeed the Hausdorff measure of a flat multiplicity one plane.                          $\square$

**Theorem 3.2** (partial regularity). *Let $\mathcal{M}$ be an ancient low entropy flow. Then there is either (1) a cylindrical singularity in which case $\mathcal{M}$ must be a round shrinking cylinder, or (2) no singularity, or (3) an at most countable number of spherical singularities.*[15]

*Proof.* Let $\check{\mathcal{M}}$ be a tangent flow at infinity (see Section 2.2). By preservation under weak limits, $\check{\mathcal{M}}$ is integral, unit regular, and cyclic. Using Lemma 3.1 it follows that any tangent cone to $\check{M}_{-1}$ is a flat multiplicity one plane. Thus, by Allard's regularity theorem [All72], $\check{M}_{-1}$ is smooth. Hence, by the classification of low entropy shrinkers by Bernstein-Wang [BW17], $\check{\mathcal{M}}$ is either a flat plane, a round shrinking sphere, or a round shrinking cylinder (all of multiplicity one). The same reasoning implies that any tangent flow $\hat{\mathcal{M}}_X$ at any space time point $X \in \mathcal{M}$ is either a flat plane, a round shrinking

---

[15]At this point of the paper we cannot exclude the possibility of more than one spherical singularity, since we cannot exclude yet the potential scenario that several compact connected components form via "neckpinches at infinity" and "contracting cusps at infinity".

sphere, or a round shrinking cylinder (all of multiplicity one).

If a tangent flow at $X$ is a flat multiplicity one plane, then $X$ is a smooth point by the local regularity theorem [Bra78, Whi05]. If there is some $X \in \mathcal{M}$ with a cylindrical tangent flow, then $\mathcal{M}$ is a round shrinking cylinder by the equality case of Huisken's monotonicity formula (see Section 2.2). Finally, spherical tangent flows can happen at most countably many times since the set of strict local maxima of the map $X \to \Theta_X(\mathcal{M})$ is countable, see e.g. [Whi97]. This proves the theorem. $\square$

By Theorem 3.2 any ancient low entropy flow $\mathcal{M} = \{\mu_t\}_{t \in (-\infty, T_E(\mathcal{M})]}$ is smooth except for at most countably many spherical singularities until it becomes extinct. Hence, recalling also that higher multiplicities are ruled out by the low entropy assumption, it is safe from now on to conflate the compact sets $M_t = \mathrm{spt}(\mu_t)$ and the Radon-measures $\mu_t$ in the notation.

**Corollary 3.3** (extinction time). *If $\mathcal{M}$ is an ancient low entropy flow then exactly one of the following happens:*[16]

  (i) $T_E(\mathcal{M}) = \infty$, *i.e. the flow $\mathcal{M}$ is eternal, or*
 (ii) $T_E(\mathcal{M}) < \infty$ *and $\mathcal{M}$ is a round shrinking cylinder, or*
(iii) $T_E(\mathcal{M}) < \infty$ *and $\mathcal{M}$ becomes extinct in at most countably many round points, or*
 (iv) $T_E(\mathcal{M}) < \infty$ *and for every $R < \infty$ there exist $T(R) < T_E(\mathcal{M})$ such that $B(0, R) \cap M_t = \emptyset$ for every $t \in (T(R), T_E(\mathcal{M})]$.*

*Proof.* Assume $T_E(\mathcal{M}) < \infty$, and suppose (iv) does not hold. Then by upper semi-continuity of Huisken's density (see Section 2.2), there exist a point $X = (x_0, T_E(\mathcal{M}))$ which $\Theta_X \geq 1$. If the tangent flow at $X$ is a plane, then this contradicts the definition of $T_E(\mathcal{M})$ by unit regularity. If the tangent flow at $X$ is a cylinder, then we are in case (ii). If the tangent flow at $X$ is a round shrinking sphere, then we are in case (iii). This proves the corollary. $\square$

3.2. **Finding necks back in time.** Let $\mathcal{M}$ be an ancient low entropy flow. Given a point $X = (x, t) \in \mathcal{M}$ and a scale $r > 0$, we consider the flow

$$(38) \qquad \mathcal{M}_{X,r} = \mathcal{D}_{1/r}(\mathcal{M} - X),$$

which is obtained from $\mathcal{M}$ by translating $X$ to the space-time origin and parabolically rescaling by $1/r$. Here, $\mathcal{D}_\lambda(x, t) = (\lambda x, \lambda^2 t)$.

**Definition 3.4.** Fix $\varepsilon > 0$. We say that $\mathcal{M}$ is $\varepsilon$-*cylindrical around $X$ at scale $r$,* if $\mathcal{M}_{X,r}$ is $\varepsilon$-close in $C^{\lfloor 1/\varepsilon \rfloor}$ in $B(0, 1/\varepsilon) \times [-1, -2]$ to the evolution of a round shrinking cylinder with radius $r(t) = \sqrt{-2t}$ and axis through the origin. The notions $\varepsilon$-*planar* and $\varepsilon$-*spherical* are defined similarly.

---

[16]At this stage of the paper we cannot yet exclude the scenarios that the solution "becomes extinct in more than one point" or that the solution "escapes to infinity".

Given any point $X = (x, t) \in \mathcal{M}$, we analyze the solution around $X$ at the diadic scales $r_j = 2^j$, where $j \in \mathbb{Z}$.

**Theorem 3.5.** *For any small enough $\varepsilon > 0$, there is a positive integer $N = N(\varepsilon) < \infty$ with the following significance. If $\mathcal{M}$ is an ancient low entropy flow, which is not a round shrinking sphere, round shrinking cylinder or a flat plane, then for every $X \in \mathcal{M}$ there exists an integer $J(X) \in \mathbb{Z}$ such that*

(39)      *$\mathcal{M}$ is not $\varepsilon$-cylindrical around $X$ at scale $r_j$ for all $j < J(X)$,*

*and*

(40)      *$\mathcal{M}$ is $\frac{\varepsilon}{2}$-cylindrical around $X$ at scale $r_j$ for all $j \geq J(X) + N$.*

*Proof.* In the following argument we will frequently use the local regularity theorem for the mean curvature flow (see [Bra78, Whi05]) without explicitly mentioning it. Given $\varepsilon > 0$, for any $X = (x, t) \in \mathcal{M}$ we define

(41)      $J(X) := \inf\{j \in \mathbb{Z} \,|\, \mathcal{M}$ is $\varepsilon$-cylindrical around $X$ at scale $r_j\}$.

Recall that by Huisken's monotonicity formula [Hui90, Ilm95] the quantity

$$\Theta_{(x,t)}(r) = \int_{M_{t-r^2}} \frac{1}{4\pi r^2} e^{-\frac{|x-y|^2}{4r^2}} \, dA(y) \tag{42}$$

is monotone, with equality only in the self-similarly shrinking case. Recall also that by Bernstein-Wang [BW17], the only self-similarly shrinking solutions with low entropy are the flat plane, the round shrinking sphere and the round shrinking cylinder.

Since $\mathcal{M}$ is non-flat and not a round shrinking sphere, its tangent flow at infinity $\check{\mathcal{M}}$ (see Section 2.2) must be a round shrinking cylinder. Hence, $J(X) < \infty$. Similarly, if $\mathcal{M}$ had a cylindrical tangent flow $\hat{\mathcal{M}}_X$, then by the equality case of Huisken's monotonicity formula $\mathcal{M}$ would be a round shrinking cylinder. Hence, $J(X) > -\infty$. Therefore, $J(X) \in \mathbb{Z}$ and the statement (39) holds true by the definition from equation (41).

To prove (40), note first that (recalling that our fixed $\varepsilon$ is small) we have

$$\Theta_X\left(r_{J(X)}\right) \geq \text{Ent}[S^1] - \tfrac{1}{100}, \tag{43}$$

and

$$\lim_{j \to \infty} \Theta_X(r_j) = \text{Ent}[S^1]. \tag{44}$$

Next, observe that by the equality case of Huisken's monotonicity formula, if $\Theta_X(r_{j+1}) - \Theta_X(r_{j-1}) = 0$ then $\mathcal{M}$ is 0-selfsimilar around $X$ at scale $r_j$. This can be made quantitative (c.f. Cheeger-Haslhofer-Naber [CHN13]). Namely, there exists a $\delta = \delta(\varepsilon) > 0$ such that if

$$\Theta_X(r_{j+1}) - \Theta_X(r_{j-1}) \leq \delta, \tag{45}$$

then

(46)      $\mathcal{M}$ is $\frac{\varepsilon}{2}$-selfsimilar around $X$ at scale $r_j$.

For $j \geq J(X)$ in our context $\frac{\varepsilon}{2}$-selfsimilar simply means $\frac{\varepsilon}{2}$-cylindrical.

Finally, using again monotonicity and quantitative rigidity we see that after going from scale $J(X)$ to scale $J(X) + N$, where $N = N(\varepsilon) < \infty$, we have

$$(47) \qquad\qquad \Theta_X \left( r_{J(X)+N(\varepsilon)} \right) \geq \mathrm{Ent}[S^1] - \delta.$$

Combining the above facts, we conclude that (40) holds. This finishes the proof of the theorem. $\qquad\square$

We fix a small enough parameter $\varepsilon > 0$ quantifying the quality of the necks for the rest of the paper.

**Definition 3.6.** The *cylindrical scale* of $X \in \mathcal{M}$ is defined by

$$(48) \qquad\qquad Z(X) = 2^{J(X)}.$$

**Remark 3.7.** *Recall that the regularity scale $R(X)$ is defined as the maximal radius $r$ such that $|A| \leq 1/r$ in the parabolic ball $P(X, r)$. It follows from the definition of regularity scale and the low entropy assumption that there exists a positive integer $N = N(\varepsilon) < \infty$ such that*

$$(49) \qquad \mathcal{M} \text{ is } \varepsilon\text{-planar around } X \text{ at scale } r_j \text{ for all } j \leq \log_2 R(X) - N.$$

3.3. **Asymptotic slope.** Let $\mathcal{M}$ be an ancient low entropy flow, which is not a round shrinking sphere, round shrinking cylinder or a flat plane. Recall from above that $\mathcal{M}$ has an asymptotic cylinder $\check{\mathcal{M}}$ for $t \to -\infty$. By Colding-Minicozzi [CM15] the axis is unique. We can assume without loss of generality that the axis is in $x_3$-direction.

**Proposition 3.8** (asymptotic cylindrical scale)**.** *For every $\delta > 0$ there exists $\Lambda = \Lambda(\delta) < \infty$ such that if $(p_0, t_0) \in \mathcal{M}$ then*

$$(50) \qquad\qquad Z(p, t_0) \leq \delta |p - p_0|$$

*for all $p \in M_{t_0}$ with $|p - p_0| \geq \Lambda Z(p_0, t_0)$.*

*Proof.* Assume without loss of generality that $(p_0, t_0) = (0, 0)$ and $Z(0, 0) = 1$. If the assertion fails, then there is a sequence $p_i \in M_0$ with $|p_i| \geq i$, but

$$(51) \qquad\qquad Z(p_i, 0) \geq \delta |p_i|$$

for some $\delta > 0$. Let $\mathcal{M}^i$ be the flow which is obtained by parabolically rescaling by $1/|p_i|$ around $(0, 0)$ and pass to a limit $\mathcal{M}^\infty$. The limit $\mathcal{M}^\infty$ is an ancient low entropy flow, which has a cylindrical singularity at $(0, 0)$. Hence, by the equality case of Huisken's monotonicity formula $\mathcal{M}^\infty$ is a round shrinking cylinder that becomes extinct at time $T = 0$, in contradiction with (51). This proves the proposition. $\qquad\square$

We normalize $\mathcal{M}$ such that $X_0 = (0, 0) \in \mathcal{M}$ and $Z(X_0) \leq 1$.

**Theorem 3.9** (asymptotic slope)**.** *For every $\delta > 0$ there exists $\Lambda = \Lambda(\delta) < \infty$ such that*

$$(52) \qquad \frac{\sup\{\sqrt{x_1^2 + x_2^2}|(x_1, x_2, x_3) \in M_t\}}{|x_3|} < \delta$$

*whenever $t \leq -10$ and $|(x_1, x_2, x_3)| \geq \Lambda\sqrt{-t}$.*

*Proof.* We first claim that for every $\delta > 0$ there exist $\Lambda_1 = \Lambda_1(\delta) < \infty$ and $T_1 = T_1(\delta) > -\infty$ such that every $(p, t) \in \mathcal{M}$ with $t \leq T_1$ and $\frac{|p|}{\sqrt{-t}} \geq \Lambda_1$ satisfies

$$(53) \qquad x_1^2 + x_2^2 \leq \frac{\delta^2}{2} x_3^2.$$

Indeed, if $(p, t) \in \mathcal{M}$ is any point with $t \leq -10$ and $\frac{|p|}{\sqrt{-t}} \gg 1$, then by Proposition 3.8 it has cylindrical scale $Z(p, t) \ll |p|$. If (53) was violated, then going back further in time this neck would intersect the central neck of our asymptotic cylinder; a contradiction.

Consequently, for every $\delta > 0$ there exist $T_2 = T_2(\delta) > -\infty$ and $L(\delta) < \infty$ such that

$$(54) \qquad \frac{1}{\sqrt{-t}} M_t \subset \left\{|z| \leq L, x_1^2 + x_2^2 \leq 3\right\} \cup \left\{|x_3| \geq L, x_1^2 + x_2^2 \leq \frac{\delta^2}{2} x_3^2\right\}$$

for all $t \leq T_2$. Using suitable large spheres as barriers, this implies the assertion. $\qquad\square$

**Corollary 3.10** (barrier for the rescaled flow)**.** *There exists an even smooth function $\varphi : \mathbb{R} \to \mathbb{R}_+$ with $\lim_{z \to \pm\infty} \varphi'(z) = 0$ such that the rescaled mean curvature flow $\bar{M}_\tau^X = e^{\frac{\tau}{2}}(M_{-e^{-\tau}} - x_0)$, where $\tau = -\log(t_0 - t)$, satisfies*

$$(55) \qquad \bar{M}_\tau^X \subset \left\{x_1^2 + x_2^2 \leq \varphi(x_3)^2\right\}$$

*for $\tau \leq \mathcal{T}(Z(X))$.*

By the corollary, any potential ends must be in direction $x_3 \to \pm\infty$.

## 4. Fine neck analysis

4.1. **Setting up the fine neck analysis.** Let $\mathcal{M}$ be a noncompact ancient low entropy flow in $\mathbb{R}^3$, which is not a round shrinking sphere, round shrinking cylinder or a flat plane. Given any point $X_0 = (x_0, t_0) \in \mathcal{M}$, we consider the rescaled flow

$$(56) \qquad \bar{M}_\tau^{X_0} = e^{\frac{\tau}{2}} \left(M_{-e^{-\tau}} - x_0\right),$$

where $\tau = -\log(t_0 - t)$. By Theorem 3.5 and Colding-Minicozzi [CM15], the rescaled flow converges for $\tau \to -\infty$ to the cylinder $\Sigma = \{x_1^2 + x_2^2 = 2\}$. Moreover, the convergence is uniform in $X_0$ once we normalize such that

$Z(X_0) \leq 1$. Hence, we can find universal functions $\sigma(\tau) > 0$ and $\rho(\tau) > 0$ with

(57) $$\lim_{\tau \to -\infty} \sigma(\tau) = 0, \quad \lim_{\tau \to -\infty} \rho(\tau) = \infty, \quad \text{and} - \rho(\tau) \leq \rho'(\tau) \leq 0,$$

such that $\bar{M}_\tau^{X_0}$ is the graph of a function $u(\cdot, \tau)$ over $\Sigma \cap B_{2\rho(\tau)}(0)$, namely

(58) $$\{x + u(x, \tau)\nu_\Sigma(x) : x \in \Sigma \cap B_{2\rho(\tau)}(0)\} \subset \bar{M}_\tau^{X_0},$$

where $\nu_\Sigma$ denotes the outward pointing unit normal to $\Sigma$, and

(59) $$\|u(\cdot, \tau)\|_{C^4(\Sigma \cap B_{2\rho(\tau)}(0))} \leq \sigma(\tau) \rho(\tau)^{-1}.$$

We will now set up a fine neck analysis following [ADS15] and [BC17].

We denote by $C < \infty$ and $\mathcal{T} > -\infty$ constants that can change from line to line and can depend on various other quantities (such as the function $\rho$ from above, the neck parameter $\varepsilon$ from Section 3.2, etc), but are independent of the point $X_0$ with $Z(X_0) \leq 1$. We also fix a nonnegative smooth cutoff function $\varphi$ satisfying $\varphi(z) = 1$ for $|z| \leq \frac{1}{2}$ and $\varphi(z) = 0$ for $|z| \geq 1$, and set

(60) $$\hat{u}(x, \tau) = u(x, \tau)\varphi\left(\frac{x_3}{\rho(\tau)}\right).$$

We recall from Angenent-Daskalopoulos-Sesum that there are shrinkers

(61) $$\Sigma_a = \{\text{surface of revolution with profile } r = u_a(z), 0 \leq z \leq a\},$$
$$\tilde{\Sigma}_b = \{\text{surface of revolution with profile } r = \tilde{u}_b(z), 0 \leq z < \infty\}$$

as illustrated in [ADS15, Fig. 1], see also [KM14]. We will refer to these shrinkers as ADS-shrinkers and KM-shrinkers, respectively. The parameter $a$ captures where the curves $u_a$ meet the $z$-axis, namely $u_a(a) = 0$, and the parameter $b$ is the asymptotic slope of the curves $u_b$, namely $\lim_{z \to \infty} u_b'(z) = b$. A detailed description of these shrinkers can be found in [ADS15, Sec. 8]. In particular, the shrinkers can be used for barrier arguments as well as for calibration arguments. To describe the latter, we fix suitable parameters $a_0, b_0, L_0 < \infty$ and consider the region $F^+ \subset \mathbb{R}^3$ bounded by $\Sigma_{a_0}, \Sigma_{b_0}$ and $\{z = L_0\}$. The region $F^+$ is foliated by $\{\Sigma_a\}_{a \geq a_0}, \Sigma$ and $\{\tilde{\Sigma}_b\}_{b \leq b_0}$. Denoting by $\nu_{\text{fol}}$ the outward unit normal of this family, by [ADS15, Lemma 4.10] we have that

(62) $$\text{div}(e^{-x^2/4}\nu_{\text{fol}}) = 0,$$

i.e. the shrinker family forms a calibration for the Gaussian area. Similarly, the region $F^- = \{(x_1, x_2, -x_3) \,|\, x \in F^+\}$ is calibrated. Let $F = F^- \cup F^+$.

**Proposition 4.1.** *There exists $\mathcal{T} > -\infty$ independent of $X_0$ such that*[17]

(63) $$\bar{M}_\tau \cap \{|x_3| \geq L_0\} \subset F$$

*for all $\tau \leq \mathcal{T}$.*

---

[17]Recall that we assume that $Z(X_0) \leq 1$.

*Proof.* By Corollary 3.10 (barrier for the rescaled flow), $\bar{M}_\tau \cap \{x_3 \geq L_0\}$ is contained in the region bounded by $\tilde{\Sigma}_{b_0}$ and $\{x_3 = L_0\}$ for sufficiently negative $\tau$. Moreover, since $\bar{M}_\tau$ is a graph with small norm over a long cylinder, it can not cross $\Sigma_{a_0} \cap \{x_3 \geq L_0\}$. This implies the assertion. $\qquad\square$

Denote by $\Delta_\tau$ the region bounded by $\bar{M}_\tau$ and $\Sigma$.

**Proposition 4.2** (c.f. [BC17, Prop. 2.2]). *For all $L \in [L_0, \rho(\tau)]$ and $\tau \leq \mathcal{T}$ we have the Gaussian area estimate*

$$(64) \qquad \int_{\bar{M}_\tau \cap \{|x_3| \geq L\}} e^{-\frac{|x|^2}{4}} - \int_{\Sigma \cap \{|x_3| \geq L\}} e^{-\frac{|x|^2}{4}}$$
$$\geq -\int_{\Delta_\tau \cap \{|x_3| = L\}} e^{-\frac{|x|^2}{4}} |N \cdot \nu_{fol}|.$$

*Proof.* Consider the region $\Delta_\tau \cap \{|x_3| \geq L\} \cap \{|x| \leq R\}$. Integrating (62) over this region, and using the divergence theorem and Proposition 4.1, we obtain

$$(65) \qquad \int_{\bar{M}_\tau \cap \{|x_3| \geq L\}} e^{-\frac{|x|^2}{4}} - \int_{\Sigma \cap \{|x_3| \geq L\}} e^{-\frac{|x|^2}{4}}$$
$$\geq -\int_{\Delta_\tau \cap \{|x_3| = L\}} e^{-\frac{|x|^2}{4}} |N \cdot \nu_{\text{fol}}| - \int_{\Delta_\tau \cap \{|x| = R\}} e^{-\frac{|x|^2}{4}} |N \cdot \nu_{\text{fol}}|,$$

where $N$ denotes the unit normal of the boundary. By the entropy bound we have that

$$(66) \qquad \int_{\Delta_\tau \cap \{|x| = R\}} e^{-\frac{|x|^2}{4}} |N \cdot \nu_{\text{fol}}| \leq CR^2 e^{-\frac{R^2}{4}},$$

and passing $R \to \infty$ the assertion follows. $\qquad\square$

The next proposition shows that closeness to the cylinder in the region $\{|x_3| \leq \frac{L}{2}\}$ implies closeness to the cylinder in the larger region $\{|x_3| \leq L\}$:

**Proposition 4.3** (c.f. [BC17, Prop. 2.3], [ADS15, Lem. 4.7]). *The graph function $u$ satisfies the integral estimates*

$$(67) \qquad \int_{\Sigma \cap \{|x_3| \leq L\}} e^{-\frac{|x|^2}{4}} |\nabla u(x, \tau)|^2 \leq C \int_{\Sigma \cap \{|x_3| \leq \frac{L}{2}\}} e^{-\frac{|x|^2}{4}} u(x, \tau)^2$$

*and*

$$(68) \qquad \int_{\Sigma \cap \{\frac{L}{2} \leq |x_3| \leq L\}} e^{-\frac{|x|^2}{4}} u(x, \tau)^2 \leq CL^{-2} \int_{\Sigma \cap \{|x_3| \leq \frac{L}{2}\}} e^{-\frac{|x|^2}{4}} u(x, \tau)^2$$

*for all $L \in [L_0, \rho(\tau)]$ and $\tau \leq \mathcal{T}$, where $C < \infty$ is a numerical constant.*

*Proof.* By the low entropy assumption we have

$$(69) \qquad \int_{\bar{M}_\tau} e^{-\frac{|x|^2}{4}} \leq \int_{\Sigma} e^{-\frac{|x|^2}{4}}.$$

Using this and Proposition 4.2, the rest of the proof is as in [BC17, proof of Prop. 2.3]. $\qquad\square$

Recall that $\bar{M}_\tau$ is expressed as graph of a function $u(x,\tau)$ over $\Sigma \cap B_{2\rho(\tau)}(0)$ satisfying the estimate (59). Using that $\bar{M}_\tau$ moves by rescaled mean curvature flow one obtains:

**Lemma 4.4** (c.f. [BC17, Lem. 2.4]). *The function $u(x,\tau)$ satisfies*

$$\partial_\tau u = \mathcal{L}u + E, \tag{70}$$

*where $\mathcal{L}$ is the linear operator on $\Sigma$ defined by*

$$\mathcal{L}f = \Delta f - \frac{1}{2}\langle x^{\tan}, \nabla f\rangle + f, \tag{71}$$

*and where the error term can be estimated by*

$$|E| \leq C\sigma(\tau)\rho^{-1}(\tau)(|u| + |\nabla u|) \tag{72}$$

*for $\tau \leq \mathcal{T}$.*

*Proof.* The proof is similar to [BC17, proof of Lem. 2.4]. □

Denote by $\mathcal{H}$ the Hilbert space of all functions $f$ on $\Sigma$ such that

$$\|f\|_{\mathcal{H}}^2 = \int_\Sigma \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} f^2 < \infty. \tag{73}$$

**Lemma 4.5** (c.f. [BC17, Lem. 2.5]). *The function $\hat{u}(x,\tau) = u(x,\tau)\,\varphi\big(\frac{x_3}{\rho(\tau)}\big)$ satisfies*

$$\partial_\tau \hat{u} = \mathcal{L}\hat{u} + \hat{E}, \tag{74}$$

*where*

$$\|\hat{E}\|_{\mathcal{H}} \leq C\rho^{-1}\|\hat{u}\|_{\mathcal{H}} \tag{75}$$

*for $\tau \leq \mathcal{T}$.*

*Proof.* As in [BC17, page 8] we compute

$$\hat{E} = E\,\varphi\Big(\frac{x_3}{\rho(\tau)}\Big) - \frac{2}{\rho(\tau)}\frac{\partial u}{\partial z}\,\varphi'\Big(\frac{x_3}{\rho(\tau)}\Big) - \frac{1}{\rho(\tau)^2}\,u\,\varphi''\Big(\frac{x_3}{\rho(\tau)}\Big) \tag{76}$$
$$+ \frac{x_3}{2\rho(\tau)}\,u\,\varphi'\Big(\frac{x_3}{\rho(\tau)}\Big) - \frac{x_3\rho'(\tau)}{\rho(\tau)^2}\,u\,\varphi'\Big(\frac{x_3}{\rho(\tau)}\Big).$$

If $|x_3| \leq \frac{\rho(\tau)}{2}$, then Lemma 4.4 gives

$$|\hat{E}| = |E| \leq C\sigma(\tau)\rho^{-1}(\tau)(|u| + |\nabla u|). \tag{77}$$

If $\frac{\rho(\tau)}{2} \leq |x_3| \leq \rho(\tau)$, then using equation (57) we obtain

$$|\hat{E}| \leq C|u| + C\rho^{-1}(\tau)|\nabla u|. \tag{78}$$

Using also Proposition 4.3 and equation (57), we infer that

$$
\int_\Sigma e^{-\frac{|x|^2}{4}} |\hat{E}|^2 \leq \frac{C\sigma^2}{\rho^2} \int_{\Sigma \cap \{|x_3| \leq \frac{\rho(\tau)}{2}\}} e^{-\frac{|x|^2}{4}} u^2 + C \int_{\Sigma \cap \{\frac{\rho(\tau)}{2} \leq |x_3| \leq \rho(\tau)\}} e^{-\frac{|x|^2}{4}} u^2
$$

$$
+ \frac{C}{\rho^2} \int_{\Sigma \cap \{|x_3| \leq \rho(\tau)\}} e^{-\frac{|x|^2}{4}} |\nabla u|^2
$$

$$
\leq \frac{C}{\rho^2} \int_{\Sigma \cap \{|x_3| \leq \frac{\rho(\tau)}{2}\}} e^{-\frac{|x|^2}{4}} u^2
$$

$$
(79) \qquad \leq \frac{C}{\rho^2} \int_\Sigma e^{-\frac{|x|^2}{4}} \hat{u}^2
$$

for $\tau \leq \mathcal{T}$. Thus, we obtain the desired result. $\qquad \square$

Let us recall some facts from [BC17] about the operator $\mathcal{L}$ defined in (71). In cylindrical coordinates this operator takes the form

$$
(80) \qquad \mathcal{L} = \frac{\partial^2}{\partial z^2} f + \frac{1}{2} \frac{\partial^2}{\partial \theta^2} f - \frac{1}{2} z \frac{\partial}{\partial z} f + f.
$$

Analysing the spectrum of $\mathcal{L}$, the Hilbert space $\mathcal{H}$ from (73) can be decomposed as

$$
(81) \qquad \mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_0 \oplus \mathcal{H}_-,
$$

where $\mathcal{H}_+$ is spanned by the four positive eigenmodes $1, z, \sin\theta, \cos\theta$, and $\mathcal{H}_0$ is spanned by the three zero-modes $z^2 - 2, z\cos\theta, z\sin\theta$. We have

$$
(82) \qquad
\begin{aligned}
\langle \mathcal{L}f, f \rangle_\mathcal{H} &\geq \frac{1}{2} \|f\|_\mathcal{H}^2 & &\text{for } f \in \mathcal{H}_+, \\
\langle \mathcal{L}f, f \rangle_\mathcal{H} &= 0 & &\text{for } f \in \mathcal{H}_0, \\
\langle \mathcal{L}f, f \rangle_\mathcal{H} &\leq -\frac{1}{2} \|f\|_\mathcal{H}^2 & &\text{for } f \in \mathcal{H}_-.
\end{aligned}
$$

Consider the functions

$$
(83) \qquad
\begin{aligned}
U_+(\tau) &:= \|P_+ \hat{u}(\cdot, \tau)\|_\mathcal{H}^2, \\
U_0(\tau) &:= \|P_0 \hat{u}(\cdot, \tau)\|_\mathcal{H}^2, \\
U_-(\tau) &:= \|P_- \hat{u}(\cdot, \tau)\|_\mathcal{H}^2,
\end{aligned}
$$

where $P_+, P_0, P_-$ denote the orthogonal projections to $\mathcal{H}_+, \mathcal{H}_0, \mathcal{H}_-$, respectively. Using Lemma 4.5 we obtain

$$
(84) \qquad
\begin{aligned}
\frac{d}{d\tau} U_+(\tau) &\geq U_+(\tau) - C\rho^{-1} (U_+(\tau) + U_0(\tau) + U_-(\tau)), \\
\left| \frac{d}{d\tau} U_0(\tau) \right| &\leq C\rho^{-1} (U_+(\tau) + U_0(\tau) + U_-(\tau)), \\
\frac{d}{d\tau} U_-(\tau) &\leq -U_-(\tau) + C\rho^{-1} (U_+(\tau) + U_0(\tau) + U_-(\tau)).
\end{aligned}
$$

To proceed, we need the following ODE-lemma.

**Lemma 4.6** (Merle-Zaag [MZ98])**.** *Let $x(\tau), y(\tau), z(\tau)$ be nonnegative functions with $(x + y + z)(\tau) > 0$ that converge to zero as $\tau \to -\infty$. Moreover, suppose for each $\varepsilon > 0$ there exists a $\tau_0 = \tau_0(\varepsilon)$ such that for $\tau \leq \tau_0$ the following holds*

$$
\begin{aligned}
|x_\tau| &\leq \varepsilon(x + y + z), \\
y_\tau &\leq -c_0 y + \varepsilon(x + z), \\
z_\tau &\geq c_0 z - \varepsilon(x + y).
\end{aligned}
$$
(85)

*Then we have $y \leq 2\varepsilon(x + z)$ for $\tau \leq \tau_0$. Moreover, either $x$ is dominant in the sense that*

$$
y + z = o(x)
$$
(86)

*as $\tau \to -\infty$, or $z$ is dominant in the more precise sense that*

$$
x + y \leq 100\varepsilon z
$$
(87)

*for $\tau \leq \tau_0(\varepsilon)$ for each $\varepsilon \leq \frac{1}{100}$.*

**Remark 4.7.** *The statement in* [MZ98] *is slightly different, but a careful inspection of their proof gives the variant of the lemma that we stated.*

If $(U_+ + U_0 + U_-)(\hat{\tau}) = 0$ for some $\hat{\tau}$, then Lemma 4.5 implies that $(U_+ + U_0 + U_-)(\tau) = 0$ for all $\tau \leq \hat{\tau}$, and by analytic continuation it follows that $\mathcal{M}$ is a round shrinking cylinder; a contradiction. We can thus apply the Merle-Zaag lemma (Lemma 4.6) to conclude that either the neutral mode is dominant, i.e.

$$
U_+ + U_- = o(U_0)
$$
(88)

or the plus mode is dominant, i.e.

$$
U_- + U_0 \leq C\rho^{-1} U_+.
$$
(89)

We will analyze these two cases in turn in the following two sections. Recall also that our analysis above depends on a choice of $X_0$ and $\rho$. However, the following proposition shows that either the neutral mode is always dominant or the plus mode is always dominant.

**Proposition 4.8.** *If a neck centered at some point $X_0$ satisfies $U_+ + U_- = o(U_0)$ for some admissible choice of $\rho_0$, then any other neck centered at any point $X$ also satisfy $U_+ + U_- = o(U_0)$ for any admissible choice of $\rho$.*

*Proof.* Any two necks converge to the same neck after rescaling as $\tau \to -\infty$. Thus, the above statement is obvious. $\qquad\square$

4.2. **Fine analysis in the plus mode.** In this section, we assume that the plus mode is dominant. We recall from above this means that after fixing a center $X_0 \in \mathcal{M}$ with $Z(X_0) \leq 1$, and a graphical scale function $\rho$, we have that

$$(90) \qquad\qquad U_- + U_0 \leq C\rho^{-1}U_+$$

for all $\tau \leq \mathcal{T}$. As before, $C < \infty$ and $\mathcal{T} > -\infty$ denote constants that can change from line to line and are independent of the point $X_0 \in \mathcal{M}$ with $Z(X_0) \leq 1$. The main goal of this section is to prove Theorem 4.15.

4.2.1. *Graphical radius.* Using (84) and (90) we obtain

$$(91) \qquad\qquad \frac{d}{d\tau}U_+ \geq U_+ - C\rho^{-1}U_+.$$

By integrating this differential inequality, for every $\mu > 0$ we can find a constant $\mathcal{T}(\mu) > -\infty$ such that

$$(92) \qquad\qquad U_+(\tau) \leq Ce^{(1-\mu)\tau}.$$

for all $\tau \leq \mathcal{T}(\mu)$. Recalling that $U_+ = ||P_+\hat{u}||_{\mathcal{H}}^2$ and using (90) we infer that

$$(93) \qquad\qquad ||\hat{u}||_{\mathcal{H}} \leq Ce^{\frac{(1-\mu)\tau}{2}}.$$

By standard interpolation inequalities this implies

$$(94) \qquad\qquad ||u(\cdot,\tau)||_{C^4([-10L_0,10L_0]\times[0,2\pi])} \leq Ce^{\frac{(1-\mu)\tau}{2}}$$

for all $\tau \leq \mathcal{T}(\mu)$. The estimate (94) is not sharp. To improve it to a sharp estimates (i.e. to remove the $\mu$) we start with the following $C^0$-estimate:

**Proposition 4.9.** *The rescaled mean curvature flow $\bar{M}_\tau$ satisfies the estimate*

$$(95) \qquad\qquad \sup_{\bar{M}_\tau \cap \{|x_3| \leq e^{-\frac{\tau}{10}}\}} \left|x_1^2 + x_2^2 - 2\right| \leq e^{\frac{\tau}{10}}$$

*for all $\tau \leq \mathcal{T}$.*

*Proof.* Fix $\mu$ sufficiently small, e.g. $\mu = \frac{1}{100}$. Recall from Section 4.1 that there are shrinkers $\Sigma_a$ and $\tilde{\Sigma}_b$ with profile functions $u_a$ and $\tilde{u}_b$.

As preparation for the barrier argument, observe that (94) implies

$$(96) \qquad\qquad \sup_{\bar{M}_\tau \cap \{|x_3| \leq L_0\}} (x_1^2 + x_2^2) \leq 2 + Ce^{\frac{(1-\mu)\tau}{2}},$$

and

$$(97) \qquad\qquad \inf_{\bar{M}_\tau \cap \{|x_3| \leq L_0\}} (x_1^2 + x_2^2) \geq 2 - Ce^{\frac{(1-\mu)\tau}{2}}.$$

To prove an upper bound, we will use the shrinkers $\tilde{\Sigma}_b$ as outer barriers. To this end, we first observe that for $z \geq 2\sqrt{2}$ the profile function $\tilde{u}_b$ satisfies the estimate

$$(98) \qquad\qquad 2 + e^{-1}b^2z^2 \leq \tilde{u}_b(z)^2 \leq 2 + b^2z^2.$$

Indeed, [ADS15, Lem. 4.11] and [ADS15, Sec. 8.12] tell us that $\tilde{u}_b$ satisfies the differential equation

$$(99) \qquad \tilde{u}_b' = \frac{w}{2z\tilde{u}_b}(\tilde{u}_b^2 - 2),$$

for $z \geq 2\sqrt{2}$, where $w = w(z)$ is a function that satisfies

$$(100) \qquad 2 \leq w(z) \leq 2 + \frac{16}{z^2}.$$

Combining (99) and (100) we see that

$$(101) \qquad 0 \leq \frac{d}{dz}\log\left(\frac{\tilde{u}_b(z)^2 - 2}{z^2}\right) \leq \frac{16}{z^3}.$$

Integrating (101) from $z$ to $\infty$ and using that the asymptotic slope of $\tilde{u}_b$ equals $b$ we infer that

$$(102) \qquad 0 \leq \log b^2 - \log\left(\frac{\tilde{u}_b(z)^2 - 2}{z^2}\right) \leq \int_{2\sqrt{2}}^{\infty} \frac{16}{x^3}dx = 1,$$

which proves (98).

Now, fixing $\hat{\tau} \leq \mathcal{T}$ we consider the shrinker $\tilde{\Sigma}_b$ with parameter

$$(103) \qquad b = 2\sqrt{eC}L_0^{-1}e^{\frac{(1-\mu)\hat{\tau}}{4}}.$$

Using (96) and (98) we see that $\tilde{\Sigma}_b \cap \{x_3 = L_0\}$ is outside of $\bar{M}_\tau \cap \{x_3 = L_0\}$ for all $\tau \leq \hat{\tau}$. Moreover, recalling that $\tilde{u}_b$ has asymptotic slope $b$ and that $\bar{M}_\tau$ has vanishing asymptotic slope (see Corollary 3.10) we also see that $\tilde{\Sigma}_b \cap \{x_3 = h\}$ is outside of $\bar{M}_\tau \cap \{x_3 = h\}$ for $h \geq h_0(\hat{\tau})$ large enough, and for all $\tau \leq \hat{\tau}$. Finally, $\tilde{\Sigma}_b \cap \{L_0 \leq x_3 \leq h\}$ is outside of $\bar{M}_\tau \cap \{L_0 \leq x_3 \leq h\}$ for $-\tau$ large enough. Hence, by the avoidance principle, $\tilde{\Sigma}_{b_1} \cap \{x_3 \geq L_0\}$ is outside of $\bar{M}_\tau \cap \{x_3 \geq L_0\}$ for all $\tau \leq \hat{\tau}$. Repeating the same argument for $x_3$ replaced by $-x_3$, and using (98), we conclude that

$$(104) \qquad \sup_{\bar{M}_\tau \cap \{|x_3| \leq e^{-\frac{\tau}{10}}\}} (x_1^2 + x_2^2) \leq 2 + e^{\frac{\tau}{10}}$$

for $\tau \leq \mathcal{T}$.

Similarly, considering the shrinkers $\Sigma_a$ as inner barriers as in [BC17, p. 10–11] and using (97) we obtain

$$(105) \qquad \inf_{\bar{M}_\tau \cap \{|x_3| \leq e^{-\frac{\tau}{10}}\}} (x_1^2 + x_2^2) \geq 2 - e^{\frac{\tau}{10}}$$

for $\tau \leq \mathcal{T}$. Together with (104) this proves the assertion. $\qquad \square$

We will now use the $C^0$-estimate to prove that the rescaled flow $\bar{M}_\tau$ is cylindrical over an exponentially expanding region for $\tau \to -\infty$.

**Proposition 4.10.** *For $\tau \leq \mathcal{T}$ the rescaled mean curvature flow $\bar{M}_\tau$ can be written as graph of a function $v(\cdot, \tau)$ over the cylinder $\Sigma \cap \{|x_3| \leq \frac{1}{2}e^{-\tau/10}\}$ with the estimate*

$$(106) \qquad \|v\|_{C^6(\Sigma \cap \{|x_3| \leq \frac{1}{2}e^{-\tau/10}\})} \leq Ce^{\tau/10}.$$

*Proof.* We first observe that for $\tau \leq \mathcal{T}$, every point $\bar{X} = (\bar{x}, \tau)$, where $\bar{x} \in \bar{M}_\tau \cap \{|x_3| \leq \frac{3}{4}e^{-\tau/10}\}$, has regularity scale comparable to 1, namely

$$(107) \qquad\qquad C^{-1} \leq R(\bar{X}) \leq C$$

for some uniform $C < \infty$, where $R(\bar{X})$ denotes the regularity scale of the rescaled flow.

Indeed, if $R(\bar{X}) \gg 1$ then we could write $\bar{M}_\tau$ is graph with small norm over a disk of radius 2 centered at $\bar{X}$, contradicting Proposition 4.9. If on the other hand $R(\bar{X}) \ll 1$ then by the argument from Section 3.2 for some $\tau' < \tau$ the rescaled flow would be $\varepsilon$-close to a cylinder or sphere with center $\bar{X}$ and radius $\frac{1}{100}$, contradicting again Proposition 4.9.

To continue, note that the geometric meaning of Proposition 4.9 is that our surface $\bar{M}_\tau$ is trapped between two cylinders of almost equal radii. Combining this with the regularity scale bound (107) we see that at every point $\bar{x} \in \bar{M}_\tau \cap \{|x_3| \leq \frac{3}{4}e^{-\tau/10}\}$ the tangent plane must be almost parallel to the one of the cylinder $\Sigma$ at $\pi(\bar{X})$, where $\pi$ denotes the nearest neighbor projection. It follows that $\bar{M}_\tau \cap \{|x_3| \leq \frac{3}{4}e^{-\tau/10}\}$ can be written as graph of a function $v(\cdot, \tau)$ over the cylinder $\Sigma$ with small $C^1$-norm. It is clear by looking at the middle region, that there is one single layer. Moreover, the $C^1$-norm bound and the regularity scale bound (107) imply

$$(108) \qquad \|\nabla^2 v\|_{C^2(|x_3| \leq \frac{3}{4}e^{-\tau/10})} \leq C,$$

and thus $v$ is the solution of a linear uniformly parabolic equation[18] with coefficients of uniformly small $C^1$-norm. Hence, using standard interior estimates the assertion follows. $\qquad\square$

We now repeat the process from Section 4.1 with improved functions $\rho$ and $\sigma$. Namely, by Proposition 4.10 we can now choose

$$(109) \qquad \rho(\tau) = e^{-\tau/20}, \text{ and } \sigma(\tau) = Ce^{\tau/20}.$$

With this new choice of $\rho$ and $\sigma$, we can write $\bar{M}_\tau$ as graph of a function $u(\cdot, \tau)$ defined over the exponentially large domain $\Sigma \cap B_{2\rho(\tau)}$ such that it satisfies the estimate (59) for $\tau \leq \mathcal{T}$.

**Proposition 4.11.** *For $\tau \leq \mathcal{T}$ the function $\hat{u}(x, \tau) = u(x, \tau)\varphi\left(\frac{x_3}{e^{-\tau/20}}\right)$ satisfies the estimate*

$$(110) \qquad\qquad \|\hat{u}\|_{\mathcal{H}} \leq Ce^{\frac{\tau}{2}}.$$

---

[18]$v_t = a_{ij}v_{ij} + b_i v_i + cv$ where $a = a(\nabla v, v)$, $b = b(\nabla v, v)$, $c = c(\nabla v, v)$. c.f. (206).

*In particular, we have*

$$(111) \qquad \sup_{\bar{M}_\tau \cap B_{10L_0}(0)} |x_1^2 + x_2^2 - 2| \leq Ce^{\frac{\tau}{2}}.$$

*Proof.* We define $U_+, U_0, U_-$ by the formulas from (83), where now $\hat{u}(x,\tau) = u(x,\tau)\varphi(\frac{x_3}{e^{-\tau/20}})$. Using Lemma 4.5 and Proposition 4.8 we then get the inequalities

$$(112) \qquad U_0 + U_- \leq Ce^{\frac{\tau}{20}}U_+$$

and

$$(113) \qquad \frac{d}{d\tau}U_+ \geq U_+ - Ce^{\frac{\tau}{20}}U_+.$$

Rewriting the latter inequality as $\frac{d}{d\tau}\log(e^{-\tau}U_+) \geq -Ce^{\frac{\tau}{20}}$ and integrating from $\tau$ to $\mathcal{T}$ yields the estimate

$$(114) \qquad U_+(\tau) \leq Ce^\tau,$$

which together with (112) implies that

$$(115) \qquad \|\hat{u}\|_{\mathcal{H}} \leq Ce^{\frac{\tau}{2}}.$$

Using standard interior estimates we conclude that

$$(116) \qquad \sup_{\bar{M}_\tau \cap B_{10L_0}(0)} |x_1^2 + x_2^2 - 2| \leq Ce^{\frac{\tau}{2}}.$$

This proves the proposition. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.2.2. *Constant functions cannot be dominant.* To obtain refined information it is useful to decompose

$$(117) \qquad P_+ = P_{1/2} + P_1,$$

where $P_{1/2}$ is the projection to the span of $z, \cos\theta, \sin\theta$, and $P_1$ is the projection to the span of the constant function 1. Accordingly, we can decompose

$$(118) \qquad U_+ := \|P_+\hat{u}\|_{\mathcal{H}}^2 = \|P_{1/2}\hat{u}\|_{\mathcal{H}}^2 + \|P_1\hat{u}\|_{\mathcal{H}}^2 =: U_{1/2} + U_1.$$

Using this decomposition, Lemma 4.5, and the assumption (90) that the plus mode is dominant[19] we obtain

$$(119) \qquad \left|\frac{d}{d\tau}U_{1/2} - U_{1/2}\right| \leq C\rho^{-1}(U_{1/2} + U_1),$$

$$(120) \qquad \left|\frac{d}{d\tau}U_1 - 2U_1\right| \leq C\rho^{-1}(U_{1/2} + U_1).$$

Applying the Merle-Zaag lemma (Lemma 4.6) with $x = e^{-\tau}U_{1/2}$, $y = 0$, $z = e^{-\tau}U_1$ we infer that either the $U_{1/2}$ is dominant, i.e.

$$(121) \qquad U_1 = o(U_{1/2})$$

---

[19] By Proposition 4.8 we can indeed use this inequality with $\rho(\tau) = e^{-\tau/20}$.

or the constant function 1 is dominant, i.e.

$$(122) \qquad U_{1/2} \leq C\rho^{-1}U_1.$$

The next proposition shows that $U_1$ cannot be dominant.

**Proposition 4.12.** *It must be the case that* $U_1 = o(U_{1/2})$.

*Proof.* Assume towards a contradiction that the estimate $U_1 = o(U_{1/2})$ does not hold. By the above discussion, we then have

$$(123) \qquad U_- + U_0 + U_{1/2} \leq Ce^{\frac{\tau}{20}}U_1$$

for all $\tau \leq \mathcal{T}$. Using Lemma 4.5 it follows that

$$(124) \qquad \left| \frac{d}{d\tau}U_1 - 2U_1 \right| \leq Ce^{\frac{\tau}{20}}U_1.$$

These two inequalities imply that

$$(125) \qquad ||\hat{u}||_{\mathcal{H}} \leq Ce^{\tau},$$

(see the proof of Proposition 4.11), and thus in particular

$$(126) \qquad ||u(\cdot,\tau)||_{C^4([-10L_0,10L_0]\times[0,2\pi])} \leq Ce^{\tau}$$

for all $\tau \leq \mathcal{T}$.

Moreover, integrating (124) and using (123) yields that $e^{-2\tau}U_1$ converges to some positive limit $K_0 > 0$ for $\tau \to -\infty$ with the estimate

$$(127) \qquad |e^{-2\tau}U_1 - K_0^2| \leq Ce^{\frac{\tau}{20}}$$

for $\tau \leq \mathcal{T}$.

We will now argue similarly as in [ADS15, proof of Lem. 5.11]: By the above estimates, there exists a constant $K$ such that the rescaled flow $\bar{M}_\tau$ satisfies

$$(128) \qquad x_1^2 + x_2^2 = 2(1 + Ke^\tau) + O(e^{\frac{21}{20}\tau})$$

uniformly on the bounded interval $[-100, 100]$, and more precisely

$$(129) \qquad K_0 = |K| \int_\Sigma (\tfrac{e}{8\pi})^{\frac{1}{4}} \tfrac{1}{4\pi} e^{-\frac{|x|^2}{4}} = |K|(\tfrac{\pi}{2e})^{\frac{1}{4}}.$$

Hence, recalling that $\tau = -\log(-t)$, the original flow satisfies

$$(130) \qquad x_1^2 + x_2^2 = (-2t)(1 + K(-t)^{-1}) + O(|t|^{-\frac{1}{20}})$$

$$(131) \qquad \qquad = 2(K - t) + O(|t|^{-\frac{1}{20}})$$

for $|x_3| \leq 100\sqrt{-t}$, where we assume for ease of notation that $X_0 = (0,0)$. Now, if we instead rescale with respect to the new center $\tilde{X}_0 = (0, K)$, i.e. if we set $\tau = -\log(K - t)$, then the corresponding rescaled flow $\tilde{M}_\tau$ satisfies

$$(132) \qquad x_1^2 + x_2^2 = 2 + O(e^{\frac{21}{20}\tau}).$$

Write $\tilde{M}_\tau$ as the graph of a function $\tilde{u}(\cdot,\tau)$ over $\Sigma \cap \{-\rho(\tau) \leq x_3 \leq \rho(\tau)\}$ with $\rho(\tau) = e^{-\frac{\tau}{20}}$ and consider the spectrum. Arguing as in the proof of Proposition 4.8 we see that $\tilde{U}_1$ is dominant. Hence,

$$(133) \qquad |e^{-2\tau}\tilde{U}_1 - \tilde{K}| \leq Ce^{\frac{\tau}{20}}$$

for some $\tilde{K} > 0$. However, using (132) we can directly compute

$$(134) \qquad \tilde{U}_1 = \left| \int \varphi\tilde{u} \cdot (\tfrac{e}{2\pi})^{\frac{1}{4}} \tfrac{1}{4\pi} e^{-\frac{|x|^2}{4}} \right| \leq Ce^{\frac{21}{10}\tau},$$

This contradicts (133), and finishes the proof of the proposition. $\qquad\square$

4.2.3. *The fine neck theorem.* By Proposition 4.11 and Proposition 4.12 we can now assume that

$$(135) \qquad U_- + U_0 \leq C\rho^{-1}U_+,$$

and

$$(136) \qquad U_1 = o(U_{1/2}),$$

where $\rho(\tau) = e^{-\tau/20}$. Recall in particular that Proposition 4.11 gives

$$(137) \qquad \|\hat{u}\|_{\mathcal{H}} \leq Ce^{\frac{\tau}{2}}.$$

Moreover, using in addition equation (136) and the assumption that our solution is not the round shrinking cylinder we see that

$$(138) \qquad \lim_{\tau \to -\infty} e^{-\tau}U_{1/2} > 0.$$

Furthermore, by using (137), we establish the following coarse estimate in a compact region.

**Lemma 4.13.** *For sufficiently large $-\tau$, we have*

$$(139) \qquad \|u(\cdot,\tau)\|_{C^2(\Sigma \cap B_{10L_0}(0))} \leq Ce^{\frac{40}{81}\tau}.$$

*Proof.* Given a fixed sufficiently large $-\tau_0$, we define parabolic regions $Q_r$ in $\Sigma \times \mathbb{R}$ by

$$(140) \quad Q_r = \{(\theta, z, \tau) : \theta \in (0, 2\pi], |z| \leq 100L_0r, -(100L_0r)^2 \leq \tau_0 - \tau \leq 0\}.$$

Let $\delta = (\tfrac{1}{10})^{1/8}$, and consider a smooth nonnegative cutoff function $\eta_1(\theta, z, \tau)$ satisfying $\eta_1 = 0$ on $\partial Q_1$ and $\eta_1 = 1$ in $Q_\delta$. Then, the function $\bar{u}_1 = u\eta_1$ satisfies

$$(141) \qquad \partial_\tau \bar{u}_1 = \mathcal{L}\bar{u}_1 + g_1,$$

where $|g_1| \leq C(|u| + |\nabla u|)$ due to Lemma 4.4. Using Proposition 4.3 this implies

$$(142) \qquad \int_{Q_1} |g_1|^3 \leq C\int_{Q_1} |u|^3 + |\nabla u|^3 \leq C\|u\|^2_{L^2(Q_1)}\|u\|_{C^1(Q_1)} \leq Ce^{\tau_0}.$$

Therefore, [Wan92, Thm. 3.14] yields

$$(143) \qquad \|u\|_{L^\infty(Q_\delta)} \leq \|\bar{u}_1\|_{L^\infty(Q_1)} \leq Ce^{\frac{\tau_0}{3}}.$$

Hence, the standard interior regularity implies $\|u\|_{C^2(Q_{\delta^2})} \leq Ce^{\frac{\tau_0}{3}}$.

Next, we consider a new cutoff $\eta_2$ satisfying $\eta_2 = 0$ on $\partial Q_{\delta^2}$ and $\eta_2 = 1$ in $Q_{\delta^3}$. Then, $\bar{u}_2 = u\eta_2$ satisfies $\partial_\tau \bar{u}_2 = \mathcal{L}\bar{u}_2 + g_2$, where

$$(144) \quad \int_{Q_{\delta^2}} |g_2|^3 \leq C \int_{Q_{\delta^2}} |u|^3 + |\nabla u|^3 \leq C\|u\|_{L^2(Q_{\delta^2})}^2 \|u\|_{C^1(Q_{\delta^2})} \leq Ce^{\frac{4}{3}\tau_0}.$$

Hence, we have $\|u\|_{C^2(Q_{\delta^4})} \leq Ce^{\frac{4}{9}\tau_0}$. We repeat this process twice more so that we can obtain the desired result. $\qquad\square$

We will now express $P_+\hat{u} \in \mathcal{H}_+$ as linear combination of the four eigenfunctions $z, \cos\theta, \sin\theta$ and $1$. Namely, let

$$(145) \qquad a^X(\tau) = \left(\frac{e}{8\pi}\right)^{\frac{1}{4}} \int z\hat{u}^X(x,\tau)\frac{1}{4\pi}e^{-\frac{|x|^2}{4}},$$

$$(146) \qquad b^X(\tau) = \left(\frac{2e}{\pi}\right)^{\frac{1}{4}} \int \cos\theta\hat{u}^X(x,\tau)\frac{1}{4\pi}e^{-\frac{|x|^2}{4}},$$

$$(147) \qquad c^X(\tau) = \left(\frac{2e}{\pi}\right)^{\frac{1}{4}} \int \sin\theta\hat{u}^X(x,\tau)\frac{1}{4\pi}e^{-\frac{|x|^2}{4}},$$

$$(148) \qquad d^X(\tau) = \left(\frac{e}{2\pi}\right)^{\frac{1}{4}} \int \hat{u}^X(x,\tau)\frac{1}{4\pi}e^{-\frac{|x|^2}{4}},$$

where the superscript $X$ is to remind us that all these coefficients depend (a priori) on the choice of base point $X$. Then, we have

$$(149) \qquad P_+\hat{u}^X = a^X z + b^X \cos\theta + c^X \sin\theta + d^X.$$

Moreover, $U_+^X = \|P_+\hat{u}^X\|_{\mathcal{H}}^2$ is given by a sum of coefficients squared:

$$(150) \qquad U_+^X = 2^{-\frac{1}{2}}\pi^{\frac{1}{2}}e^{-\frac{1}{2}}\left(4|a^X|^2 + |b^X|^2 + |c^X|^2 + 2|d^X|^2\right).$$

**Proposition 4.14.** *There coefficients defined in* (145) *satisfy the estimates*

$$(151) \qquad\qquad |d^X(\tau)| \leq Ce^{\frac{11}{20}\tau},$$

*and*

$$(152) \quad |e^{-\frac{\tau}{2}}a^X(\tau) - \bar{a}^X| + |e^{-\frac{\tau}{2}}b^X(\tau) - \bar{b}^X| + |e^{-\frac{\tau}{2}}c^X(\tau) - \bar{c}^X| \leq Ce^{\frac{\tau}{20}}$$

*for some numbers* $\bar{a}^X, \bar{b}^X, \bar{c}^X$ *that might depend on* $X$.

*Proof.* Using Lemma 4.5 and $\mathcal{L}1 = 1$ we compute

$$(153) \quad \frac{d}{d\tau}d^X(\tau) = \left(\frac{e}{2\pi}\right)^{\frac{1}{4}} \int (\mathcal{L}\hat{u} + \hat{E})\frac{1}{4\pi}e^{-\frac{|x|^2}{4}}$$

$$(154) \qquad = \left(\frac{e}{2\pi}\right)^{\frac{1}{4}} \int (\hat{u} + \hat{E})\frac{1}{4\pi}e^{-\frac{|x|^2}{4}} = d^X(\tau) + \left(\frac{e}{2\pi}\right)^{\frac{1}{4}} \int \frac{\hat{E}}{4\pi}e^{-\frac{|x|^2}{4}}.$$

Hence,

$$(155) \qquad \left| \frac{d}{d\tau} \left( e^{-\tau} d^X(\tau) \right) \right| \le C e^{-\tau} \|\hat{E}\|_{\mathcal{H}} \le C e^{-\tau} \rho^{-1} \|\hat{u}\|_{\mathcal{H}} \le C e^{-\frac{9}{20}\tau}.$$

Integrating from $\tau$ to $\mathcal{T}$ this implies (151).

In a similar manner, we compute

$$(156) \quad \left| \frac{d}{d\tau} \left( e^{-\frac{\tau}{2}} a^X(\tau) \right) \right| + \left| \frac{d}{d\tau} \left( e^{-\frac{\tau}{2}} b^X(\tau) \right) \right| + \left| \frac{d}{d\tau} \left( e^{-\frac{\tau}{2}} c^X(\tau) \right) \right|$$
$$\le C e^{-\frac{\tau}{2}} \rho^{-1} \|\hat{u}\|_{\mathcal{H}} \le C e^{\frac{\tau}{20}}.$$

Integrating from $-\infty$ to $\tau$ this implies (152). $\qquad\square$

We are now ready to prove the main theorem of this section:

**Theorem 4.15.** *Let $\mathcal{M}$ be an ancient low entropy flow which is not the round shrinking sphere, round shrinking cylinder or flat plane. If the plus mode is dominant, then there are constants $\bar{a} = \bar{a}(\mathcal{M}) \ne 0$, $C = C(\mathcal{M}) < \infty$ and a decreasing function $\mathcal{T} : \mathbb{R}_+ \to \mathbb{R}_-$ (depending on $\mathcal{M}$) with the following significance.*

*For every $X \in \mathcal{M}$ the graph function $u^X(\cdot, \tau)$ of the rescaled flow $\bar{M}_\tau^X$ satisfies the estimates*

$$(157) \qquad \|e^{-\frac{\tau}{2}} \hat{u}^X(x, \tau) - \bar{a}z - \bar{b}^X \cos\theta - \bar{c}^X \sin\theta\|_{\mathcal{H}} \le C e^{\frac{\tau}{40}},$$

*and*

$$(158) \qquad \sup_{|x_3| \le 10 L_0} \left| e^{-\frac{\tau}{2}} u^X(x, \tau) - \bar{a}z - \bar{b}^X \cos\theta - \bar{c}^X \sin\theta \right| \le C e^{\frac{\tau}{160}}$$

*for $\tau \le \mathcal{T}(Z(X))$. Here, the constant $\bar{a}$ is independent of $X$, and $\bar{b}^X$ and $\bar{c}^X$ are numbers that may depend on $X$ and satisfy*

$$(159) \qquad |\bar{b}^X| + |\bar{c}^X| \le C.$$

*Proof.* By scaling we can pretend without essential loss of generality that we only work with center points $X$ satisfying $Z(X) \le 1$ (this condition is only used to figure out more easily which constants are uniform in $X$).

Consider the difference

$$(160) \qquad D^X := \hat{u}^X - e^{\tau/2} \left( \bar{a}^X z + \bar{b}^X \cos\theta + \bar{c}^X \sin\theta \right).$$

Using equation (149) and Proposition 4.14 we see that

$$(161) \qquad |D^X| \le |\hat{u}^X - P_+ \hat{u}^X| + C(1 + |z|) e^{\frac{11}{20}\tau}.$$

Since by (135) and (137) we have

$$(162) \qquad U_- + U_0 \le C e^{\frac{21}{20}\tau},$$

it follows that

$$(163) \qquad \|D^X\|_{\mathcal{H}} \le C e^{\frac{21}{40}\tau},$$

which proves (157) modulo the claim about the coefficients.

Combining the equations (137) and (150) and Proposition 4.14 we see that

$$|\bar{b}^X| + |\bar{c}^X| \le C, \tag{164}$$

which proves (159).

We recall that $e^{-\frac{\tau}{2}} u$ corresponds to the original scale. Hence, if instead of $X = (x_1, x_2, x_3, t)$ we consider the new origin

$$X' = (x_1 - \bar{b}^X \cos\theta, x_2 - \bar{c}^X \sin\theta, x_3, t) \tag{165}$$

then the estimate (163) simplifies to

$$\|\hat{u}^{X'}(x, \tau) - e^{\tau/2} \bar{a}^X z\|_{\mathcal{H}} \le C e^{\frac{21}{40}\tau}, \tag{166}$$

i.e. the estimate (157) holds with $\bar{a}^{X'} = \bar{a}^X$, $\bar{b}^{X'} = 0$, and $\bar{c}^{X'} = 0$. If $\bar{a}^X = 0$, then (166) implies $\|\hat{u}^{X'}\|_{\mathcal{H}}^2 \le C e^{\frac{21}{20}\tau}$, contradicting (138). Here, we have used Proposition 4.8, as well as Proposition 4.12 to show that, even after re-centering, the $\frac{1}{2}$ mode dominates. Hence, $\bar{a}^X \ne 0$. Since the estimate (166) holds for any $X$ and since $\bar{a}^X$ does not vanish for any $X$, we see that $\bar{a}^X =: \bar{a}$ is independent of $X$.

It remains to prove the pointwise estimate (158). To this end, we start with

$$\|D^X\|_{L^2(\Sigma \cap \{|z| \le 10L_0\})} \le C\|D^X\|_{\mathcal{H}} \le C e^{\frac{21}{40}\tau}. \tag{167}$$

Next, using the inequality (164) and Lemma 4.13 we estimate

$$\begin{align} (168) \quad \|\nabla D^X\|_{L^2(\Sigma \cap \{|z| \le 10L_0\})} &+ \|\nabla^2 D^X\|_{L^2(\Sigma \cap \{|z| \le 10L_0\})} \\ &\le C\|\hat{u}^X\|_{C^2(\Sigma \cap \{|z| \le 10L_0\})} + C e^{\tau/2} \le C e^{\frac{40}{81}\tau}. \end{align}$$

Applying Agmon's inequality we conclude that

$$\sup_{|x_3| \le 10L_0} |D^X| \le C e^{\frac{81}{160}\tau}. \tag{169}$$

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

After a change of coordinates we can assume without loss of generality that our ancient low entropy flow $\mathcal{M}$ satisfies $\bar{a} = \bar{a}(\mathcal{M}) = 1/\sqrt{2}$. Then, after recentering as in (165), Theorem 4.15 tells us that the graph $u^X(\cdot, \tau)$ of the rescaled flow $\bar{M}_\tau^X$ satisfies

$$\sup_{|x_3| \le 10L_0} \left| e^{-\frac{\tau}{2}} u^X(x, \tau) - \frac{1}{\sqrt{2}} z \right| \le C e^{\frac{\tau}{160}} \tag{170}$$

for $\tau \le \mathcal{T}(Z(X))$.

**Corollary 4.16.** *If the rescaled flow $\bar{M}_\tau^X$ satisfies (170), then we have the estimates*

$$(171) \qquad\qquad \sup_{\{x_3 \leq -L_0\}} (x_1^2 + x_2^2) \leq 2,$$

$$(172) \qquad\qquad \inf_{\{x_3 \geq L_0\}} (x_1^2 + x_2^2) \geq 2,$$

*and*

$$(173) \qquad\qquad \inf_{\bar{M}_\tau^X} x_3 \leq -\mu e^{-\tau/2}$$

*for $\tau \leq \mathcal{T}(Z(X))$, where $\mu > 0$ is a numeric constant. In particular, the unrescaled mean curvature flow $\mathcal{M} = \{M_t\}$ satisfies*

$$(174) \qquad \inf_{p \in M_t} x_3(p) > -\infty, \qquad and \qquad \sup_{p \in M_t} x_3(p) = \infty.$$

*Proof.* The first two estimates follow easily by using the KM-shrinkers $\tilde{\Sigma}_b$ as outer barriers (using also Corollary 3.10) and the ADS-shrinkers $\Sigma_a$ as inner barriers; since $X_0$ is arbitrary, the latter also implies that $\sup_{p \in M_t} x_3(p) = \infty$.

The third estimate follows from the improved barrier argument (where the surfaces $\Sigma_a$ are shifted along the $z$-axis) from [BC17, Sec. 3].

Finally, for any $X_0 = (x_0, t_0) \in \mathcal{M}$, after recentering in space, the estimates tells us that $\bar{M}_\tau^{X_0} \cap \{x_3 \leq -L_0\}$ is contained inside the cylinder $\Sigma$ of radius $\sqrt{2}$ for $\tau \leq \mathcal{T}$. Hence, by comparison $M_{t_0} \cap \{x_3 \leq -L_0\}$ is compact (see also Corollary 4.25 and its proof). Since $X_0$ is arbitrary, this yields $\inf_{p \in M_t} x_3(p) > -\infty$. $\qquad\square$

4.3. **Fine analysis in the neutral mode.** In this section, we assume that the neutral mode is dominant. The main goal is to prove Theorem 4.22 and Corollary 4.25, which show that the solution is compact with a precise inwards quadratic expansion.

Given any center $X_0 \in \mathcal{M}$, there exists some functions $\sigma$ and $\rho$ satisfying (57) such that (59) holds, and we have

$$(175) \qquad U_- + U_+ = o(U_0), \qquad\qquad \left| \partial_\tau U_0 \right| \leq o(U_0)$$

for $\tau \leq \mathcal{T}$. In this section, $C < \infty$ and $\mathcal{T} > -\infty$ denote constants that can change from line to line and may also depend on $X_0 \in \mathcal{M}$. To distinguish the initial choice of $\rho$, we set

$$(176) \qquad\qquad \rho_0(z) = \rho(z).$$

We will later use improved scale functions, but $\rho_0$ will never change.

4.3.1. *Graphical radius.* The first goal is to prove a lower bound for the optimal graphical radius. To begin with, we consider the positive function

$$(177) \qquad \alpha(\tau) = \left( \int_{|x_3| \leq L_0} u^2(x,s) \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \right)^{1/2}.$$

**Lemma 4.17.** *For $L \in [L_0, \rho(\tau)]$, we have the estimate*

$$(178) \qquad \alpha(\tau)^2 \leq \int_{\Sigma \cap \{|x_3| \leq L\}} u^2(x,\tau) \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \leq C\alpha(\tau)^2.$$

*Proof.* By Proposition 4.3 we have

$$(179) \qquad \int_{\Sigma \cap \{2^{k-1}L_0 \leq |x_3| \leq 2^k L_0\}} u^2 e^{-\frac{|x|^2}{4}} \leq \frac{C_0}{4^k L_0^2} \int_{\Sigma \cap \{|x_3| \leq 2^{k-1}L_0\}} u^2 e^{-\frac{|x|^2}{4}}$$

where $C_0$ is a constant. Hence,

$$(180) \qquad \int_{\Sigma \cap \{|x_3| \leq 2L_0\}} u^2 \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \leq \left( 1 + \frac{C_0}{4L_0^2} \right) \alpha(\tau)^2 \leq \exp\left( \frac{C_0}{4L_0^2} \right) \alpha(\tau)^2,$$

and

$$(181) \qquad \int_{\Sigma \cap \{|x_3| \leq 2^2 L_0\}} u^2 \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \leq \left( 1 + \frac{C_0}{4^2 L_0^2} \right) \exp\left( \frac{C_0}{4L_0^2} \right) \alpha(\tau)^2$$

$$\leq \exp\left( \left( \tfrac{1}{4} + \tfrac{1}{4^2} \right) \frac{C_0}{L_0^2} \right) \alpha(\tau)^2.$$

Thus, if $2^{I-1}L_0 \leq \rho(\tau) \leq 2^I L_0$ then we inductively obtain

$$(182)$$

$$\int_{\Sigma \cap \{|x_3| \leq \rho(\tau)\}} u^2 \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \leq \exp\left( \sum_{i=1}^{I} \frac{1}{4^i} \frac{C_0}{L_0^2} \right) \alpha(\tau)^2 \leq \exp\left( \frac{2C_0}{L_0^2} \right) \alpha(\tau)^2.$$

This completes the proof.                                      $\square$

Now, define an increasing continuous function by

$$(183) \qquad \beta_1(\tau) = \sup_{\sigma \leq \tau} \alpha(\sigma).$$

By standard interior estimates, we have

$$(184) \qquad |u|(x,\tau) \leq C\beta_1(\tau)$$

for $|x_3| \leq L_0$ and $\tau \leq \mathcal{T}$.

For technical reasons, it will be best to work with a monotone function $\beta$, which simultaneously has controlled derivatives. To this end, we define

$$(185) \qquad \beta(\tau) = \sup_{\sigma \leq \tau} \left( \int_{\Sigma} u^2(x,\sigma) \varphi^2\left( \tfrac{x_3}{\rho_0(\sigma)} \right) \frac{1}{4\pi} e^{-\frac{|x|^2}{4}} \right)^{1/2},$$

where we recall that $\rho_0$ is defined in (176) to be the original graphical scale function, which is the input to this section. Clearly, $\beta$ is a locally Lipschitz,

increasing function. By equation (175), we have $\beta' = o(\beta)$ at almost every time,[20] so in particular

(186) $$0 \leq \beta'(\tau) \leq \tfrac{1}{5}\beta(\tau).$$

Moreover, by Lemma 4.17 we have
(187)
$$\beta_1(\tau) \leq \beta(\tau) = \left( \int_\Sigma u^2(x,\sigma)\varphi^2\left(\tfrac{x_3}{\rho_0(\sigma)}\right)\tfrac{1}{4\pi}e^{-\tfrac{|x|^2}{4}} \right)^{1/2} \leq C\alpha(\sigma) \leq C\beta_1(\tau),$$

where $\sigma$ is chosen such that the second equality holds. To recapitulate, we have obtained

(188) $$\beta_1(\tau) \leq \beta(\tau) \leq C\beta_1(\tau).$$

We now prove a $C^0$-estimate in terms of $\beta(\tau)$.

**Proposition 4.18.** *There are constants $c > 0$ and $C < \infty$ such that*

(189) $$|u|(x,\tau) \leq C\beta(\tau)^{\frac{1}{2}}$$

*holds whenever $|x_3| \leq c\beta(\tau)^{-\frac{1}{4}}$ and $\tau \leq \mathcal{T}$.*

*Proof.* By the estimates (184) and (188) there is a constant $K < \infty$ such that

(190) $$|u|(x,\tau) \leq K\beta(\tau)$$

whenever $|x_3| \leq L_0$ and $\tau \leq \mathcal{T}$.

Fixing $\hat{\tau} \leq \mathcal{T}$, we consider the ADS-shrinker $\Sigma_a$ with parameter

(191) $$a = c(K\beta(\hat{\tau}))^{-\frac{1}{2}},$$

where $c > 0$ is a numerical constant depending on $L_0$. Then, by [ADS15, Lemma 4.4] the profile function $u_a$ satisfies

(192) $$u_a(L_0) \leq \sqrt{2} - K\beta(\hat{\tau}).$$

Comparing (190) and (192) we infer that $\bar{M}_\tau \cap \{x_3 = L_0\}$ lies outside of $\Sigma_a \cap \{x_3 = L_0\}$ for $\tau \leq \hat{\tau}$.

On the other hand, similarly as in (101) we have

(193) $$0 \leq \frac{d}{dz} \log\left( \frac{2 - u_a(z)^2}{z^2} \right) \leq \frac{16}{z^3}$$

Integration from $\sqrt{a}$ to $a$ yields

(194) $$0 \leq \log\left( \frac{2}{a^2} \right) - \log\left( \frac{2 - u_a(\sqrt{a})^2}{a} \right) \leq 1,$$

hence,

(195) $$u_a(\sqrt{a})^2 \geq 2 - \tfrac{2}{a}.$$

---

[20]If one prefers thinking about smooth objects, one could further regularize $\beta$ to be $C^1$, maintaining similar properties.

Since $\bar{M}_{\hat{\tau}} \cap \{x_3 \geq L_0\}$ lies outside of $\Sigma_a \cap \{x_3 \geq L_0\}$, comparing the inequalities (191) and (195) yields

$$(196) \qquad u(x, \hat{\tau}) \geq -C\beta(\hat{\tau})^{\frac{1}{2}}$$

for $\{L_0 \leq x_3 \leq \sqrt{a}\}$ (and similarly for negative $x_3$). Hence, (190) implies that (196) holds for $|x_3| \leq \sqrt{a} = C_0^{\frac{1}{2}}(K\beta(\hat{\tau}))^{-\frac{1}{4}}$.

Next, we establish the upper bound. By using (101), we have

$$(197) \qquad e^{-1}b^2 L_0^2 \leq \tilde{u}_b^2(L_0) - 2 \leq 4(\tilde{u}_b(L_0) - \sqrt{2}),$$

for sufficiently small $b$. Therefore, if we choose $b^2 = 4eL_0^{-2}K\beta(\hat{\tau})$, then we have

$$(198) \qquad K\beta(\hat{\tau}) = \tfrac{1}{4e}L_0^2 b^2 \leq \tilde{u}_b(L_0) - \sqrt{2}.$$

Hence, (190) implies that $\bar{M}_\tau \cap \{x_3 \geq L_0\}$ lies inside of $\tilde{\Sigma}_b \cap \{x_3 \geq L_0\}$ for $\tau \leq \hat{\tau}$. Then, (101) implies

$$(199) \qquad \tilde{u}_b^2(1/\sqrt{b}) \leq 2 + b.$$

Thus, we can complete the proof by arguing similarly with the KM-shrinkers $\tilde{\Sigma}_b$ as outer barriers. $\qquad \square$

Similarly as in Proposition 4.10 the $C^0$-estimate from Proposition 4.18 can be upgraded to a $C^4$-estimate. Hence, we can now repeat the process from Section 4.1 with better functions $\rho$ and $\sigma$. Namely, by we can now choose

$$(200) \qquad \rho(\tau) = \beta(\tau)^{-\frac{1}{5}}, \text{ and } \sigma(\tau) = \beta(\tau)^{\frac{1}{5}}.$$

and write $\bar{M}_\tau$ as a graph of a function $u$ over $\Sigma \cap B_{2\rho(\tau)}$ such that

$$(201) \qquad \|u(\cdot, \tau)\|_{C^4(\Sigma \cap B_{2\rho(\tau)}(0))} \leq \rho(\tau)^{-2}$$

for $\tau \leq \mathcal{T}$. Note that by equation (186) the derivative $\rho'$ indeed satisfies

$$(202) \qquad -\rho(\tau) \leq \rho'(\tau) \leq 0,$$

as required by condition (57). From now on we work with the function

$$(203) \qquad \hat{u}(x, \tau) = u(x, \tau)\varphi\left(\frac{x_3}{\rho(\tau)}\right),$$

where $\rho$ is the improved graphical radius from (200).

4.3.2. *Estimates for the error terms.* In the following lemma we Taylor expand the rescaled mean curvature flow to second order:

**Lemma 4.19.** *The function $\hat{u}(x,\tau) = u(x,\tau)\varphi\left(\frac{x_3}{\rho(\tau)}\right)$ satisfies*

$$(204) \qquad \partial_t \hat{u} = \mathcal{L}\hat{u} - \tfrac{1}{2\sqrt{2}}\hat{u}^2 - \tfrac{1}{4\sqrt{2}}\partial_\theta^2\hat{u}^2 + E,$$

*where the error term can be estimated by*

$$|E| \leq C\varphi|z||\nabla u|^3 + C\varphi(|u| + |\nabla u|)^2(|u| + |\nabla u| + |\nabla^2 u|)$$
$$+ C|\varphi'|\rho^{-1}\big(|\nabla u| + |zu|\big)$$
$$+ C|\varphi''|\rho^{-2}|u|$$
$$(205) \qquad + C\varphi(1 - \varphi)\big(|u|^2 + |\nabla^2 u|^2\big).$$

*Proof.* Consider the rescaled mean curvature flow operator for graphs over the cylinder

$$\mathcal{N}(w) := \partial_t w$$

$$- \frac{[1 + (\frac{\partial_\theta w}{w})^2]\partial_z^2 w + \frac{1+(\partial_z w)^2}{w^2}\partial_\theta^2 w - 2\frac{(\partial_z w)(\partial_\theta w)^2}{w^3}\partial_z\partial_\theta w - \frac{(\partial_\theta w)^2}{w^3}}{1 + (\partial_z w)^2 + (\frac{\partial_\theta w}{w})^2}$$

$$(206) \qquad + \frac{1}{w} - \frac{w - z\partial_z w}{2\sqrt{1 + (\partial_z w)^2 + \frac{(\partial_\theta w)^2}{w^2}}}$$

c.f. [GKS18, eqn. (2.1)]. Now set $w^s = \sqrt{2} + su$ and let $f(s) = \mathcal{N}(w^s)$. Then Taylor's theorem implies that

$$(207) \qquad f(1) = f(0) + f'(0) + \frac{1}{2}f''(0) + \frac{1}{6}f'''(s_0)s_0^3$$

for some $s_0 \in [0, 1]$. Since the constant function $\sqrt{2}$ and the function $\sqrt{2}+u$ both solve the rescaled mean curvature flow equation we have

$$(208) \qquad f(0) = 0 \text{ and } f(1) = 0.$$

Furthermore, one easily computes that

$$(209) \qquad \frac{d}{ds}|_{s=0}\mathcal{N}(w^s) = \partial_t u - \big(\partial_z^2 u + \tfrac{1}{2}\partial_\theta^2 u - \tfrac{1}{2}z\partial_z u + u\big) = \partial_t u - \mathcal{L}u,$$

and

$$(210) \qquad \frac{d^2}{ds^2}|_{s=0}\mathcal{N}(w^s) = \frac{1}{\sqrt{2}}u^2 + \tfrac{1}{2\sqrt{2}}\partial_\theta^2 u^2 =: -\mathcal{Q}(u).$$

Observe that when computing $\mathcal{N}$, every term involving $s$ is multiplied by either $u$ or its derivatives. Thus, the numerators of all the third derivative terms will include a product of at least three $u$ factors and its derivatives. Observe further that the expression for $\mathcal{N}$ is linear in the second derivatives of $w$, and, in particular, no second order derivatives appear in the denominator. Note that only the last summand defining $\mathcal{N}(w)$ involves a $z$ as a prefactor, and that this term contains no second order terms. Finally, for

$|z| \leq 2\rho(\tau)$, the norm $||u||_{C^2}$ is very small, so all the terms in which a product of four or more $u$ factors and its derivatives appear are dominated by corresponding terms of third degree. This information implies that we can safely estimate

$$(211) \quad \left|\frac{d^3}{ds^3}|_{s=s_0}\mathcal{N}(w^s)\right| \leq C_0|z||\nabla u|^3 + E_0(|u|, |\nabla u|) + Q_0(|u|, |\nabla u|)|\nabla^2 u|,$$

where $C_0$ is a constant, $E_0$ is a cubic plynomial and $Q_0$ is a quadratic polynomial. Putting things together we infer that

$$(212) \quad \left|\partial_t u - \left(\mathcal{L}u + \tfrac{1}{2}\mathcal{Q}(u)\right)\right|$$
$$\leq C_0|z||\nabla u|^3 + E_0(|u|, |\nabla u|) + Q_0(|u|, |\nabla u|)|\nabla^2 u|.$$

We now want to see what this computation translates to in terms of $\hat{u}$. For the linear term, using (202) we get

$$(213) \quad \left|\partial_t \check{u} - \mathcal{L}\check{u} - \varphi(\partial_t u - \mathcal{L}u)\right| \leq \rho^{-2}|\varphi''||u| + 2\rho^{-1}|\varphi'|(|\nabla u| + |zu|)$$

For the quadratic term $\mathcal{Q}$, using that $\rho$ is $\theta$-independent we get

$$(214) \quad |\mathcal{Q}(\hat{u}) - \varphi\mathcal{Q}(u)| = |\varphi^2\mathcal{Q}(u) - \varphi\mathcal{Q}(u)| = \varphi(1 - \varphi)|\mathcal{Q}(u)|.$$

Putting everything together, this implies the assertion. $\qquad\square$

For our function

$$(215) \qquad\qquad \hat{u}(x, \tau) = u(x, \tau)\varphi\left(\tfrac{x_3}{\rho(\tau)}\right),$$

where $\rho$ is the improved graphical scale from (200), we let

$$(216) \qquad U_+ = \|P_+\hat{u}\|_{\mathcal{H}}^2, \qquad U_0 = \|P_0\hat{u}\|_{\mathcal{H}}^2, \qquad U_- = \|P_-\hat{u}\|_{\mathcal{H}}^2.$$

By assumption (175) and Proposition 4.8 we still have

$$(217) \qquad\qquad U_+ + U_- = o(U_0).$$

Therefore, we can now expand

$$(218) \qquad\qquad \hat{u} = \alpha_1\psi_1 + \alpha_2\psi_2 + \alpha_3\psi_3 + o(|\vec{\alpha}|),$$

where

$$(219) \qquad\qquad \vec{\alpha}(\tau) = (\alpha_1(\tau), \alpha_2(\tau), \alpha_3(\tau))$$

are time dependent coefficients, and where $\psi_1, \psi_2$ and $\psi_3$ are the three zero eigenfunctions satisfying $\mathcal{L}\psi_i = 0$, explicitly

$$(220) \qquad\qquad \psi_1 = 2^{-\frac{3}{2}}\left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}}(z^2 - 2),$$

$$(221) \qquad\qquad \psi_2 = \left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}}z\cos\theta,$$

$$(222) \qquad\qquad \psi_3 = \left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}}z\sin\theta.$$

Moreover, Lemma 4.17 and equation (217) yield that

$$(223) \qquad\qquad C^{-1}\alpha(\tau) \leq |\vec{\alpha}|(\tau) \leq C\alpha(\tau)$$

for $\tau \leq \mathcal{T}$, where $\alpha(\tau)$ is the function defined in equation (177).

Our next goal is to derive a system of ODEs for the coefficients $\alpha_1, \alpha_2, \alpha_3$ from the expansion (218). To this end, we start with the following error estimate.

**Lemma 4.20.** *The error term $E$ from Lemma 4.19 satisfies the estimate*

$$(224) \qquad |\langle E, \psi_i \rangle| \leq C\beta(\tau)^{2+\frac{1}{5}}$$

*for $\tau \leq \mathcal{T}$.*

*Proof.* Using equation (201) and Lemma 4.19 we see that

$$(225) \qquad |E| \leq C\rho(\tau)^{-2}.$$

Together with $|\psi_i| \leq C\rho(\tau)^2$ on the support of $E$ this yields the coarse estimate

$$(226) \qquad |E||\psi_i| \leq C.$$

Using this, we compute

$$
\left| \int_{\Sigma \cap \left\{ |x_3| \geq \rho(\tau)^{\frac{1}{10}} \right\}} E\psi_i e^{-\frac{|x|^2}{4}} \right| \leq C \int_{\Sigma \cap \left\{ \rho(\tau)^{\frac{1}{10}} \leq |x_3| \leq \rho(\tau) \right\}} \exp\left( -\tfrac{1}{4}\rho(\tau)^{\frac{1}{5}} \right)
$$
$$(227) \qquad\qquad\qquad\qquad \leq C\beta(\tau)^3$$

for $\tau \leq \mathcal{T}$.

On the other hand, if $|x_3| \leq \frac{1}{2}\rho(\tau)$ we have $\varphi' = \varphi'' = 1 - \varphi = 0$. Therefore, for $|x_3| \leq \rho(\tau)^{\frac{1}{10}}$ we can estimate

$$(228) \quad |E||\psi_i| \leq C(z^2 + 2)\big(|z||\nabla u|^3 + (|u| + |\nabla u|)^2(|u| + |\nabla u| + |\nabla^2 u|)\big)$$

$$(229) \qquad\qquad \leq C\rho^{\frac{3}{10}}\rho^{-2}(|u|^2 + |\nabla u|^2)$$

$$(230) \qquad\qquad \leq C\rho^{-1}(|u|^2 + |\nabla u|^2).$$

Thus, using also Proposition 4.3 and Lemma 4.17, we infer that

$$(231) \quad \left| \int_{\Sigma \cap \left\{ |x_3| \leq \rho(\tau)^{\frac{1}{10}} \right\}} E\psi_i e^{-\frac{|x|^2}{4}} \right| \leq \frac{C}{\rho(\tau)} \int_{\Sigma \cap \{|x_3| \leq \rho(\tau)\}} (u^2 + |\nabla u|^2)e^{-\frac{|x|^2}{4}}$$

$$(232) \qquad\qquad\qquad\qquad \leq C\rho(\tau)^{-1}\alpha(\tau)^2$$

$$(233) \qquad\qquad\qquad\qquad \leq C\beta(\tau)^{2+\frac{1}{5}}.$$

This proves the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.3.3. *The inwards quadratic neck theorem.* The next proposition gives a system of ODEs for the coefficients from the expansion (218).

**Proposition 4.21.** *The coefficients $\alpha_1, \alpha_2, \alpha_3$ from the expansion (218) satisfy*

(234)                    $$\tfrac{d}{d\tau}\alpha_1 = -2A\alpha_1^2 - A(\alpha_2^2 + \alpha_3^2) + o(\beta^2),$$

(235)                    $$\tfrac{d}{d\tau}\alpha_2 = -A\alpha_1\alpha_2 + o(\beta^2),$$

(236)                    $$\tfrac{d}{d\tau}\alpha_3 = -A\alpha_1\alpha_3 + o(\beta^2),$$

*where $A = \tfrac{1}{2}\left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}}$.*

*Proof.* Using the lemmas above and $\partial_\theta\psi_1 = 0$, we compute

(237)        $$\tfrac{d}{d\tau}\alpha_1 = \langle \partial_t\hat{u}, \psi_1\rangle$$

(238)        $$= \langle \mathcal{L}\hat{u} - \tfrac{1}{2\sqrt{2}}\hat{u}^2 - \tfrac{1}{4\sqrt{2}}\partial_\theta^2\hat{u}^2 + E, \psi_1\rangle$$

(239)        $$= \langle -\tfrac{1}{2\sqrt{2}}\hat{u}^2 + E, \psi_1\rangle$$

(240)        $$= -\tfrac{1}{2\sqrt{2}}\left(c_{111}\alpha_1^2 + c_{122}\alpha_2^2 + c_{133}\alpha_3^2\right)$$
$$- \tfrac{1}{\sqrt{2}}(c_{112}\alpha_1\alpha_2 + c_{113}\alpha_1\alpha_3 + c_{123}\alpha_2\alpha_3) + o(\beta^2),$$

where

(241)                    $$c_{ijk} = \int \psi_i\psi_j\psi_k \tfrac{1}{4\pi}e^{-|x|^2/4}.$$

For $\zeta = 2, 3$, we have $\partial_\theta^2\psi_\zeta = -\psi_\zeta$. Therefore,

(242)        $$\tfrac{d}{d\tau}\alpha_\zeta = \langle \partial_t\hat{u}, \psi_\zeta\rangle$$

(243)        $$= \langle \mathcal{L}\hat{u} - \tfrac{1}{2\sqrt{2}}\hat{u}^2 - \tfrac{1}{4\sqrt{2}}\partial_\theta^2\hat{u}^2 + E, \psi_\zeta\rangle$$

(244)        $$= \langle -\tfrac{1}{4\sqrt{2}}\hat{u}^2 + E, \psi_\zeta\rangle$$

(245)        $$= -\tfrac{1}{4\sqrt{2}}\left(c_{\zeta11}\alpha_1^2 + c_{\zeta22}\alpha_2^2 + c_{\zeta33}\alpha_3^2\right)$$

(246)        $$- \tfrac{1}{2\sqrt{2}}(c_{\zeta12}\alpha_1\alpha_2 + c_{\zeta13}\alpha_1\alpha_3 + c_{\zeta23}\alpha_2\alpha_3) + o(\beta^2).$$

The only nonvanishing coefficients are

(247)   $$c_{111} = 2^{-\frac{9}{2}}\left(\tfrac{e}{2\pi}\right)^{\frac{3}{4}}\tfrac{1}{4\pi}\int_0^{2\pi} e^{-\frac{1}{2}}\sqrt{2}\, d\theta \int_{-\infty}^{\infty}(z^2 - 2)^3 e^{-\frac{z^2}{4}}\, dz = 2^{\frac{3}{2}}\left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}},$$

and

(248)

$$c_{122} = 2^{-\frac{3}{2}}\left(\tfrac{e}{2\pi}\right)^{\frac{3}{4}}\tfrac{1}{4\pi}\int_0^{2\pi}(\cos\theta)^2 e^{-\frac{1}{2}}\sqrt{2}\, d\theta \int_{-\infty}^{\infty}(z^2 - 2)z^2 e^{-\frac{z^2}{4}}\, dz = 2^{\frac{1}{2}}\left(\tfrac{e}{2\pi}\right)^{\frac{1}{4}},$$

and $c_{122} = c_{133}$. Setting

(249)                    $$A = \frac{1}{2\sqrt{2}}c_{122} = \frac{1}{4\sqrt{2}}c_{111},$$

and putting things together, this implies the assertion.                    $\square$

**Theorem 4.22.** *The coefficients* $\alpha_1, \alpha_2, \alpha_3$ *from the expansion* (218) *satisfy*

$$(250) \qquad \alpha_1(\tau) = \frac{-2^{\frac{5}{4}}\pi^{\frac{1}{4}}e^{-\frac{1}{4}} + o(1)}{|\tau|},$$

*and*

$$(251) \qquad |\alpha_2| + |\alpha_3| = o(|\alpha_1|)$$

*for* $\tau \leq \mathcal{T}$.

*Proof.* To analyze the system of ODEs from Proposition 4.21, we consider the function

$$(252) \qquad \alpha_4 = (\alpha_2^2 + \alpha_3^2)^{\frac{1}{2}}.$$

Then, we have

$$(253) \qquad \tfrac{d}{d\tau}\alpha_1 = -2A\alpha_1^2 - A\alpha_4^2 + o(\beta^2),$$

and

$$(254) \qquad \tfrac{d}{d\tau}\alpha_4 = -A\alpha_1\alpha_4 + o(\beta^2),$$

provided $\alpha_4(\tau) \neq 0$.

We also define

$$(255) \qquad \alpha_5(\tau) = |\alpha_1|(\tau) + \alpha_4(\tau).$$

**Claim 4.23.** *For* $\tau \leq \mathcal{T}$ *we have*

$$(256) \qquad \alpha_1(\tau) < 0,$$

*and*

$$(257) \qquad \alpha_5(\tau) = \sup_{\sigma \leq \tau} \alpha_5(\sigma) > 0.$$

*Proof of Claim 4.23.* By (223) and (255) the quantities $\alpha$, $|\vec{\alpha}|$, and $\alpha_5$ are all comparable. In particular, they are strictly greater than zero for $\tau \leq \mathcal{T}$.

We recall that $\alpha(\tau) \to 0$ as $\tau \to -\infty$. Hence, there exists a sequence $\tau_i \to -\infty$ such that

$$(258) \qquad \alpha_5(\tau_i) = \sup_{\sigma \leq \tau_i} \alpha_5(\sigma).$$

Recalling also that $\beta$ and $\beta_1$ are comparable and monotone, and that $\beta_1(\tau) = \sup_{\sigma \leq \tau} \alpha(\sigma)$, we can estimate

$$(259) \qquad \beta(\tau) \leq C\beta_1(\tau_i) \leq C \sup_{\sigma \leq \tau_i} \alpha_5(\sigma) = C\alpha_5(\tau_i)$$

for $\tau \leq \tau_i$.

Suppose towards a contradiction that $\alpha_1(\tau_i) \geq 0$ for some $i$.

If $\alpha_4(\tau_i) = 0$, we have $\alpha_1(\tau_i) > 0$, and (253) and (259) imply that $\alpha_1(\tau) > \alpha_1(\tau_i)$ for some $\tau < \tau_i$. Together with $\alpha_4(\tau) \geq 0 = \alpha_4(\tau_i)$ this contradicts (258). If $\alpha_4(\tau_i) > 0$, then adding (253) and (254), and using (259) yields

$$(260) \qquad \frac{d}{d\tau}|_{\tau=\tau_i}\alpha_5(\tau) \leq -\tfrac{1}{2}A\alpha_5(\tau_i)^2 + o(\beta(\tau_i)^2) < 0;$$

this contradicts again (258). Hence, $\alpha_1(\tau_i) < 0$.

By the above, to finish the proof of the claim it is enough to show that whenever $\hat{\tau} \leq \mathcal{T}$ is a time such that

$$(261) \qquad -\alpha_1(\hat{\tau}) > 0,$$

and

$$(262) \qquad \alpha_5(\hat{\tau}) = \sup_{\sigma \leq \hat{\tau}} \alpha_5(\sigma),$$

then there is some $\delta > 0$ such that the functions $-\alpha_1(\tau)$ and $\alpha_5(\tau)$ are increasing for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$.

We first consider the case $\alpha_4(\hat{\tau}) = 0$. Then, by (259) there exists some $\delta > 0$ such that

$$(263) \qquad \left| A\alpha_1\alpha_4(\tau) + \tfrac{d}{d\tau}\alpha_4(\tau) \right| \leq \tfrac{A}{100}\alpha_1^2(\tau), \qquad \alpha_4(\tau) \leq \tfrac{1}{100}|\alpha_1|(\tau)$$

holds for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$ if $\alpha_4(\tau) \neq 0$, and also

$$(264) \qquad \left| 2A\alpha_1^2(\tau) + A\alpha_4^2(\tau) + \tfrac{d}{d\tau}\alpha_1(\tau) \right| \leq \tfrac{A}{100}\alpha_1^2(\tau)$$

holds for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$. These inequalities imply that $-\alpha_1(\tau)$ is increasing, in particular $\alpha_1(\tau) < 0$ and $\alpha_5(\tau) = -\alpha_1(\tau) + \alpha_4(\tau)$ for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$. Analyzing (263) and (264) again with this additional input we conclude that $\alpha_5(\tau)$ is increasing for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$.

Finally, we consider the case $\alpha_4(\hat{\tau}) > 0$. Then, there exists some $\delta > 0$ such that $\alpha_4(\tau) > 0$ for all $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$. Hence, adding (253) and (254), and using (259) yields

$$(265) \qquad \frac{d}{d\tau}\alpha_5(\tau) \geq \tfrac{1}{4}A\alpha_5^2(\tau)$$

for sufficiently small $\delta$. Therefore, $\alpha_5(\tau)$ is increasing for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$. Also, it is obvious that $-\alpha_1(\tau)$ is increasing for $\tau \in [\hat{\tau}, \hat{\tau} + \delta)$ provided $\delta$ is small enough. This proves the claim. $\qquad\square$

Continuing the proof of the theorem, by the claim we have $\alpha_1(\tau) < 0$ and

$$(266) \qquad \frac{d}{d\tau}\alpha_1 = -A\alpha_1^2 - A|\vec{\alpha}|^2 + o(|\vec{\alpha}|^2),$$

$$(267) \qquad \frac{d}{d\tau}\alpha_2 = -A\alpha_1\alpha_2 + o(|\vec{\alpha}|^2),$$

$$(268) \qquad \frac{d}{d\tau}\alpha_3 = -A\alpha_1\alpha_3 + o(|\vec{\alpha}|^2).$$

Hence,

$$(269) \qquad \frac{d}{d\tau}(|\alpha_1| + |\vec{\alpha}|) = A(|\alpha_1| + |\vec{\alpha}|)^2 + o(|\vec{\alpha}|^2),$$

which yields

$$(270) \qquad |\alpha_1|(\tau) + |\vec{\alpha}|(\tau) = \frac{1 + o(1)}{A|\tau|}.$$

Plugging this into (266) shows

$$(271) \qquad \tfrac{d}{d\tau}(-\alpha_1) \geq \frac{1 + o(1)}{2A|\tau|^2}.$$

Since $\lim_{\tau \to -\infty} \alpha_1(\tau) = 0$, integration from $-\infty$ to $\tau$ yields

$$(272) \qquad -\alpha_1(\tau) \geq \frac{1 + o(1)}{2A|\tau|}.$$

Combining (270) and (272) we conclude that

$$(273) \qquad \alpha_1 = \frac{-1 + o(1)}{A|\tau|} \text{ and } |\alpha_2| + |\alpha_3| = o(|\alpha_1|).$$

This proves the theorem. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

**Corollary 4.24.** *For $\tau \leq \mathcal{T}$ the rescaled flow $\bar{M}_\tau^{X_0} \cap \{|z| \geq L_0\}$ is contained inside the cylinder $\Sigma$ of radius $\sqrt{2}$.*

*Proof.* Theorem 4.22 implies that $\bar{M}_\tau^{X_0}$ satisfies

$$(274) \qquad\qquad u(L_0, \tau) < 0$$

for $\tau \leq \mathcal{T}$. Hence, given arbitrarily small $b > 0$, the KM-barrier $\tilde{\Sigma}_b$ lies outside of $\bar{M}_\tau^{X_0} \cap \{x_3 = L_0\}$. Using also Corollary 3.10 and the avoidance principle we infer that $\tilde{\Sigma}_b$ lies outside of $\bar{M}_\tau^{X_0}$ on $\{x_3 \geq L_0\}$. Namely, $\bar{M}_\tau^{X_0}$ satisfies $x_1^2 + x_2^2 \leq 2$ for $x_3 \geq L_0$. In the same manner, $\bar{M}_\tau^{X_0}$ satisfies $x_1^2 + x_2^2 \leq 2$ for $x_3 \leq L_0$. This proves the corollary. $\qquad\square$

**Corollary 4.25.** *For an ancient low entropy solution, if the neutral mode is dominant, then the solution is compact.*

*Proof.* Denote the rescaled flow by $\bar{M}_\tau = \bar{M}_\tau^{X_0}$, and the original flow by $\mathcal{M} = \{M_t\}$. Then, by the above corollary, there exists some $t_1 \in (-\infty, t_0)$ and $L < \infty$ such that $M_{t_1} \cap \{|z| \geq L\}$ is contained inside the cylinder of radius $\sqrt{2(t_0 - t_1)}$. The lemma below implies that $M_t$ is compact for all $t > t_0$. Since $X_0$ was arbitrary, this proves the corollary. $\qquad\square$

**Lemma 4.26.** *Let $M_0 \subset \mathbb{R}^3$ be a complete embedded surface such that*

$$(275) \qquad\qquad M_0 \setminus B(0, L) = \Sigma \setminus B(0, L)$$

*for some $L < \infty$, where $\Sigma$ denotes the cylinder of radius $\sqrt{2}$. Then the mean curvature flow $\{M_t\}$ of $M_0$ is compact for all $t > 1$.*

*Proof.* Let $\mathcal{M}$ be the mean curvature flow with initial condition $M_0$. For $z_i \to \pm\infty$ consider the shifted flows $\mathcal{M}^i = \mathcal{M} - (z_i, 0)$. For $i \to \infty$ the flows $\mathcal{M}^i$ converge to a limit $\mathcal{M}^\infty$, whose initial time slice is $\Sigma$. By uniqueness, $\mathcal{M}^\infty$ must be a round shrinking cylinder that becomes extinct at time 1. We conclude that $M_t$ is compact (this includes the possibility that $M_0$ is noncompact, but $M_t$ becomes compact via a contracting cusp singularity at spatial infinity) for all $t > 1$. $\qquad\square$

## 5. Curvature bound and cap size control

Throughout this section, $\mathcal{M}$ will always be an ancient low entropy flow where the plus mode is dominant. We will frequently use the fine neck theorem (Theorem 4.15) and its corollary (Corollary 4.16). By an affine change of coordinates, we can assume without loss of generality that

$$(276) \qquad \bar{a} = 1/\sqrt{2},$$

that the axis of the asymptotic cylinder is the $x_3$ axis, and that

$$(277) \qquad \min_{p \in M_0} x_3(p) = 0 \text{ is attained at the origin.}$$

5.1. **Global curvature bound.** We will prove a global in space-time curvature bound. To this end, we start with the following proposition.

**Proposition 5.1.** *For every $\Lambda < \infty$ there exists a $\rho = \rho(\mathcal{M}, \Lambda) < \infty$ with the following significance. If $p_0, p \in M_{t_0}$ are points such that $Z(p_0, t_0) \leq \Lambda$ and*

$$(278) \qquad d := x_3(p) - x_3(p_0) \geq \rho,$$

*then*

$$(279) \qquad Z(p, t_0) > \Lambda.$$

*Proof.* Let $C = C(\mathcal{M}) < \infty$ and $\mathcal{T}(\Lambda) > -\infty$ be the constants from the fine neck theorem (Theorem 4.15) applied at any point with cylindrical scale at most $\Lambda$. We will show that if $p_0, p \in M_{t_0}$ are points with $Z(p_0, t_0) \leq \Lambda$ and $Z(p, t_0) \leq \Lambda$, then $d := x_3(p) - x_3(p_0)$ is bounded from above.

Let $t_1 = t_0 - d^2$, and let $\tau$ be the corresponding rescaled time, namely

$$(280) \qquad \tau = -\log(t_0 - t_1) = -2\log d.$$

Observe that if $d$ is sufficiently large, at time $\tau$ we will see necks around $p_0$ and $p$. Those two necks have to align with each other, else looking further back in time they would intersect. Moreover, if $d$ is sufficiently large, then $\tau \leq \mathcal{T}(\Lambda)$, as required for applying the fine neck theorem.

Letting $r_0$ and $r$ be (representitives of) the radii of the fine necks around $p_0$ and $p$ at time $t_1$, applying the fine neck theorem (Theorem 4.15) at time

$\tau$ with $z = 0$ twice, once centered at $(p_0, t_0)$ and once centered at $(p, t_0)$, we obtain

$$(281) \qquad\qquad |r - r_0| \leq 2Ce^{\frac{\tau}{160}}.$$

On the other hand, applying the fine neck theorem (Theorem 4.15) centered at $(p_0, t_0)$ at time $\tau$ with $z = 1$, we obtain

$$(282) \qquad\qquad |r - r_0| \geq \frac{1}{\sqrt{2}} - Ce^{\frac{\tau}{160}}.$$

Comparing (281) and (282) we conclude that $-\tau$ is bounded; hence $d$ is bounded above. This finishes the proof of the proposition. $\qquad\square$

**Corollary 5.2.** *For every compact interval $I \subseteq (-\infty, T_E(\mathcal{M}))$ there exist constants $\delta > 0$, $\Lambda < \infty$ and $C < \infty$ (possibly depending on $\mathcal{M}$ and $I$) such that for every $p \in M_t$ with $t \in I$ and $x_3(p) \geq \Lambda$ either*

$$(283) \qquad\qquad R(p, t) \geq \delta,$$

*or*

$$(284) \qquad\qquad (p, t) \text{ is } \varepsilon\text{-spherical at scale } CR(p, t).$$

*Proof.* Suppose that there exists a sequence of points $p_k \in M_{t_k}$ with $t_k \in I$ such that $x_3(p_k) \to \infty$ and $R(p_k, t_k) \to 0$. Recall that by Section 3.2, there are constants $0 \ll C \ll C' < \infty$ such that if $(p_k, t_k)$ is not $\varepsilon$-spherical at scale $CR(p_k, t_k)$, then $Z(p_k, t_k) \leq C'R(p_k, t_k)$. For $k$ large, the latter option is impossible by Proposition 5.1, and so the assertion follows. $\qquad\square$

**Lemma 5.3.** *The surface $M_{t_0}$ has no compact connected components for any $t_0 \in (-\infty, T_E(\mathcal{M}))$.*

*Proof.* Suppose towards a contradiction that there is a compact connected component $N_{t_0}$ at some time $t_0$. Let $\mathcal{N}$ be the space-time connected component of $N_{t_0}$, and denote its time $t$-slice by $N_t$. By Corollary 4.16 there is some $t_1 < t_0$ such that $N_{t_1}$ is non-compact.

Let $\delta, \Lambda$, and $C$ be the constants from Corollary 5.2, corresponding to the interval $I = [t_1, t_0]$. Using Corollary 4.16 and spherical barriers we see that there is a constant $\Lambda' \in (\Lambda, \infty)$ such that

$$(285) \qquad\qquad N_t \subset B(0, \Lambda') \cup \{x_3 \geq \Lambda'\}.$$

for all $t \in I$. After possibly increasing the $\Lambda'$ we can assume that

$$(286) \qquad\qquad \Lambda' \geq \sup_{p \in N_{t_0}} |p| + 10\sqrt{t_1 - t_0} + 10C\delta.$$

Now if $N_t$ is compact at any given $t \in [t_1, t_0]$, then it must be the case that $R(p, t) \geq \delta$ for every $p \in N_t$ with $x_3(p) \geq \Lambda'$. Otherwise, Corollary 5.2 would imply that $(p, t)$ is $\varepsilon$-spherical at a scale $\leq C\delta$, and $N_t$ could not reach $N_{t_0}$ at time $t_0$. Hence,

$$(287) \qquad\qquad -\frac{d}{dt} \sup_{p \in N_t} x_3(p) \leq 2\delta^{-1}$$

whenever $N_t$ is compact and $\sup_{p \in N_t} x_3(p) \geq \Lambda'$. Combining (285) and (287) we conclude that $N_t$ is compact for all $t \in [t_1, t_0]$; a contradiction.          $\square$

**Corollary 5.4.** *There exist a constant $C = C(\mathcal{M}) < \infty$ such that*

$$(288) \qquad\qquad Z(p,t) \leq CR(p,t)$$

*for all $(p,t) \in \mathcal{M}$. In particular, $\lim_{x_3(p) \to \infty} R(p,t) = \infty$.*

*Proof.* By Lemma 5.3 the flow cannot be $\varepsilon$-spherical at any scale. Thus, by Section 3.2 the regularity scale and the cylindrical scale must be comparable. Together with Proposition 5.1 this yields $\lim_{x_3(p) \to \infty} R(p,t) = \infty$.          $\square$

Lemma 5.3 also quickly implies a local type I curvature estimate:

**Proposition 5.5.** *For all $\Lambda < \infty$ there is a $C(\Lambda) < \infty$ with the following significance. If $\mathcal{M}$ reaches a point $p_0$ at time $t_0$, then we have the curvature bound*

$$(289) \qquad\qquad |A(p,t)| \leq \frac{C}{\sqrt{t_0 - t}}$$

*for all $p \in B(p_0, \Lambda\sqrt{t_0 - t})$ at all $t < t_0$.*

*Proof.* Suppose towards a contradiction that for some $\Lambda < \infty$ there is a sequence of points $(p_i, t_i) \in \mathcal{M}$ such that the estimate fails for $C_i = i$.

After shifting $(p_i, t_i)$ to $(0,0)$ and rescaling, we get a sequence of flows $\mathcal{M}^i$ which reach the point $0$ at time $0$, such that there are points $q_i \in M^i_{-1} \cap B(0, \Lambda)$ with $|A|(q_i, -1) > i$. Moreover, the time slices of $\mathcal{M}^i$ have no compact connected components by Lemma 5.3.

By the entropy assumption we can pass to a subsequential limit $\mathcal{M}^i \to \mathcal{M}$, which is an ancient low entropy flow. By Brakke's clearing out lemma [Bra78] the limit flow $\mathcal{M}$ reaches the origin at time $0$. Moreover, the limit $\mathcal{M}$ cannot have any compact connected component at any $t < 0$.

Thus, $\mathcal{M}$ has no spherical singularities when $t < 0$ and so $\mathcal{M} \cap \{t < 0\}$ is smooth. In particular, $\mathcal{M}$ satisfies $|A| \leq C$ on $B(0, 2\Lambda) \times [-2, -1/2]$ for some $C < \infty$; for $i$ large enough this contradicts $|A|(q_i, -1) > i$.          $\square$

Combining the above results we now get a global curvature bound:

**Theorem 5.6** (Global curvature bound)**.** *Let $\mathcal{M}$ be an ancient low entropy flow such that the plus mode is dominant. Then $\mathcal{M}$ is eternal with globally bounded curvature, namely*

$$(290) \qquad\qquad T_E(\mathcal{M}) = \infty,$$

*and there is a constant $C = C(\mathcal{M}) < \infty$ such that*

$$(291) \qquad\qquad \sup_{(p,t) \in \mathcal{M}} |A|(p,t) \leq C.$$

*Proof.* By Corollary 5.4 the flow is eternal, i.e. $T_E(\mathcal{M}) = \infty$, and by Lemma 5.3 the time slices have no compact connected components. If $(p,t) \in \mathcal{M}$ is any point with $R(p,t) < \frac{1}{10}$, then by Proposition 5.1 and Corollary 5.4

there exists a point $(q, t)$ with $0 < x_3(q) - x_3(p) < C\rho$ such that $R(q, t) = 1$. But then, by unit regularity there is a point $(q', t + \delta)$ with $|q' - q| \leq 1$ and $R(q', t + \delta) \leq 2$, where $\delta > 0$ is a uniform constant. Using the fine neck theorem (Theorem 4.15) we see that $|x_1(p) - x_1(q)| + |x_2(p) - x_2(q)|$ is bounded. Applying the local type I curvature estimate (Proposition 5.5) centered at $(q', t + \delta)$ we conclude that $|A(p, t)| \leq C$. $\qquad \square$

5.2. **Cap size control and asymptotics.** The goal of this section is to prove Theorem 5.9, which states that our surfaces $M_t$ have a cap of uniformly controlled size and open up like a parabola. To begin with, let us consider the "height of the tip" function

$$(292) \qquad \psi(t) := \inf_{x \in M_t} x_3.$$

**Lemma 5.7.** *The function $\psi$ is strictly increasing and satisfies*

$$(293) \qquad \lim_{t \to \pm\infty} \psi(t) = \pm\infty.$$

*Proof.* Applying Corollary 4.16 we see that $\psi(t) > -\infty$ for every $t \in \mathbb{R}$ and that the infimum is actually a minimum (using also Corollary 3.10). By the comparison with planes / the strong maximum principle, the function $\psi$ is strictly increasing. Applying Corollary 4.16 again we infer that $\lim_{t \to -\infty} \psi(t) = -\infty$. Finally, by comparison with conical expanders (using also Corollary 3.10) we have $\lim_{t \to \infty} \psi(t) = \infty$. $\qquad \square$

**Proposition 5.8** (Fast tip points)**.** *There exists a constant $Q = Q(\mathcal{M}) < \infty$ such that every $p \in M_t$ with $x_3(p) = \psi(t)$ satisfies $R(p, t) \leq Q$.*

*Proof.* If the assertion fails we can find a sequence $(p_j, t_j) \in \mathcal{M}$ with $x_3(p_j) = \psi(t_j)$ such that $Q_j := R(p_j, t_j) \to \infty$. Let $\mathcal{M}^j$ be the flow that is obtained by parabolically rescaling by $Q_j^{-1}$ around $X_j = (p_j, t_j)$. Note that $\mathcal{M}^j$ has expansion parameter $\bar{a}^j = Q_j^{-1}\bar{a} \to 0$.

Up to a subsequence, we can pass to a limit $\mathcal{M}^\infty$. The limit $\mathcal{M}^\infty$ is itself a noncompact ancient low entropy flow. Since its zero time slice is contained in a half space, $\mathcal{M}^\infty$ cannot be the cylinder. Hence, by the fine neck theorem (Theorem 4.15) it has an expansion parameter $a^\infty \neq 0$. However, it follows from the fine neck theorem and its proof that $\bar{a}^j \to a^\infty$, contradicting $\bar{a}^j \to 0$. This finishes the proof of the proposition. $\qquad \square$

For each time, select a point $p_t \in M_t$ such that $x_3(p_t) = \psi(t)$.

**Theorem 5.9** (Cap size control)**.** *Let $\mathcal{M}$ be an ancient low entropy flow such that the plus mode is dominant. Then there exists a $C = C(\mathcal{M}) < \infty$ such that for $t \in \mathbb{R}$ every point in $M_t \setminus B_C(p_t)$ lies on a fine neck. In particular, the surface $M_t$ has exactly one end.*

*Moreover, $M_t \setminus B_C(p_t)$ is the graph of a function $r$ in cylindrical coordinates around the z-axis satisfying*

(294)
$$r(t, z, \theta) = \sqrt{2(z - \psi(t))} + o(\sqrt{z - \psi(t)})$$

*for $z \geq \psi(t) + C$, and the height of the tip function $\psi$ satisfies*

(295)
$$\psi(t) = t + o(|t|)$$

*Proof.* Let $p \in M_t$, and let $t_*$ be the time in which $\psi(t_*) = x_3(p)$. By Proposition 5.8 (fast tip points) we have $R(p_{t_*}) \leq Q$, an by the global curvature bound (Theorem 5.6) we have $t_* - t \geq c(x_3(p) - \psi(t))$. Thus, applying the fine neck theorem (Theorem 4.15) and Proposition 4.9 at $(p_{t_*}, t_*)$ we see that if $x_3(p) - \psi(t)$ is sufficiently large, then $p$ lies on a fine neck. Together with Corollary 3.10 and Corollary 4.16 this proves that we have a cap of controlled size $C = C(\mathcal{M}) < \infty$, and exactly one end.

Fix $t$. For $z \geq \psi(t) + C$, let $r(z)$ be (a representitive of) the radius of the fine neck of $M_t$ at $x_3(x) = z$. By the above and the fine neck theorem (Theorem 4.15) we have

(296)
$$r(z + r(z)) = r(z) + 1 + o((z - \psi(t))^{-1/160})).$$

This can be iterated to

(297)
$$r\left(z + kr(z) + \left[1 + o((z - \psi(t))^{-1/160}))\right] \frac{k(k-1)}{2}\right)$$
$$= r(z) + k\left[1 + o((z - \psi(t))^{-1/160}))\right].$$

Setting $h = z - \psi(t)$ we get

(298)
$$r\left(z + kr(z) + \left[1 + o(h^{-1/160})\right] \frac{k(k-1)}{2}\right) = r(z) + k\left[1 + o(h^{-1/160})\right].$$

The argument of $r$ in the left hand side of (298) is

(299)
$$s := z + kr(z) + [1 + o(h^{-1/160})]\frac{k(k-1)}{2},$$

and solving for $k$, we get

(300)
$$k = \frac{-(r(z) - \frac{1}{2}) + \sqrt{[2 + o(h^{-1/160})](s - z)}}{1 + o(h^{-1/160})}.$$

Combining (298) and (300), and taking $h$ very large, we obtain

(301)
$$r(h) = \sqrt{2h} + o(\sqrt{h}).$$

This proves (294).

Finally, in light of (301), one can put a a bowl with tip curvature $(1 + \delta)$ inside the domain bounded by $M_{t_0}$, and a bowl with tip curvature $(1 - \delta)$ outside that domain. By the avoidance principle, this implies $(1 - 2\delta)t \leq \psi(t) \leq (1 + 2\delta)t$ for large enough $t$, hence

(302)
$$\psi(t) = t + o(t)$$

for positive $t$. In fact, the above steps are uniform in the original time $t_0$: For every $\delta$ there exist a uniform $s$ such that

$$(303) \qquad (1 - 2\delta)(t - t_0) \leq \psi(t) - \psi(t_0) \leq (1 + 2\delta)(t - t_0),$$

whenever $t - t_0 \geq s$. Thus, for every $t < 0$ we have

$$(304) \qquad (1 - 2\delta)|t| \leq -\psi(t) \leq (1 + 2\delta)|t|,$$

and so

$$(305) \qquad \psi(t) = t + o(|t|)$$

holds for all $t$. This finishes the proof of the theorem. □

## 6. Rotational symmetry

6.1. **Fine expansion away from the cap.** Let $\mathcal{M}$ be as in the previous section. The goal of this section is to prove Theorem 6.6, which shows that the cylindrical end becomes rotationally symmetric at very fast rate, and also controls the position of the cap in the $xy$-plane uniformly in time.

Given a point $q \in \mathbb{R}^3$ and a direction $w \in S^2$, we denote by $K$ the *normalized rotation vector field* that corresponds to a rotation by $\pi/2$ around the axis $W = \{q + wt : t \in \mathbb{R}\}$, namely

$$(306) \qquad K(x) = SJS^{-1}(x - q), \qquad \text{where } J = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and $S \in \mathrm{SO}_3$ is any rotation matrix with $Se_3 = w$.[21]

**Definition 6.1** (c.f. [BC17, Def. 4.3]). A point $X \in \mathcal{M}$ with $H(X) > 0$ is called *$\delta$-symmetric* if there exists a normalized rotation vector field $K$ such that

$$(307) \qquad |\langle K, \nu \rangle H| \leq \delta \text{ in the parabolic ball } P(X, 10H^{-1}(X)).$$

The following proposition shows that $\mathcal{M}$ becomes *$\delta$-symmetric* at a very fast rate if one moves away from the cap.

**Proposition 6.2.** *There exist a constant $C = C(\mathcal{M}) < \infty$ with the following significance. If $X = (x, t) \in \mathcal{M}$ is any point with*

$$(308) \qquad x_3 - \psi(t) \geq C,$$

*then*

$$(309) \qquad X \text{ is } (x_3 - \psi(t))^{-300}\text{-symmetric.}$$

---

[21]$S$ is only determined up to a one-parameter choice, but $K$ is well-defined.

*Proof.* We recall from Theorem 5.9 (cap size control) that there exists a constant $C_0 < \infty$ such that every point $x \in M_t \setminus B_{C_0}(p_t)$ lies on the center of a fine neck. In particular, every point $X = (x, t)$ with

$$(310) \qquad\qquad x_3 - \psi(t) \geq C_0$$

satisfies the a-priori assumptions of the neck-improvement theorem from Brendle-Choi [BC17, Thm. 4.4]. Hence, if $X = (x, t) \in \mathcal{M}$ is a point with

$$(311) \qquad\qquad x_3 - \psi(t) \geq 2^{\frac{j}{400}} C_0$$

then we can iteratively apply the neck-improvement theorem [BC17, Thm. 4.4] to conclude that $X$ is $2^{-j}\varepsilon$-symmetric. This implies the assertion.  $\square$

**Corollary 6.3** (strong symmetry). *There exist a constant $C = C(\mathcal{M}) < \infty$ with the following significance. If $X = (x, t) \in \mathcal{M}$ is any point with $x_3 - \psi(t) \geq C$, then there exist a direction $w_X \in S^2$ and a point $q_X \in \mathbb{R}^3$ with*

$$(312) \qquad\qquad |w_X - e_3| \leq \tfrac{1}{100}, \qquad \langle q_X, e_3 \rangle = x_3,$$

*such that the normalized rotation vector field $K_X(y) = S_X J S_X^{-1}(y - q_X)$, where $S_X \in SO_3$ with $S_X e_3 = w_X$, satisfies the estimate*

$$(313) \qquad\qquad \sup_{P(X, 10H^{-1}(X))} |\langle K_X, \nu \rangle H| \leq (x_3 - \psi(t))^{-300}.$$

*Proof.* Since fine necks are very close to the asymptotic cylinder, we always have $|w_X - e_3| \leq \tfrac{1}{100}$. In addition, by moving $q_X$ along the axis $W$ we can always arrange that $\langle q_X, e_3 \rangle = x_3$. Hence, the corollary follows from the proposition.  $\square$

**Definition 6.4.** We call any triple $(X, w_X, q_X)$ that satisfies the conclusion of Corollary 6.3 a *strongly symmetric triple*.

The following lemma shows that nearby strongly symmetric triples at the same time align well with each other.

**Lemma 6.5** (alignment). *There exists a constant $C = C(\mathcal{M}) < \infty$ with the following significance. If $(X, w_X, q_X)$ and $(Y, w_Y, q_Y)$ are strongly symmetric triples with $X = (x, t)$, $Y = (y, t)$ and $|x - y| H(X) \leq 1$, then*

$$(314) \qquad\qquad |w_X - w_Y| \leq C(x_3 - \psi(t))^{-300},$$

*and*

$$(315) \qquad |\langle q_X - q_Y, e_1 \rangle| + |\langle q_X - q_Y, e_2 \rangle| \leq C(x_3 - \psi(t))^{-\frac{599}{2}}.$$

*Proof.* Without loss of generality, after suitable rotations and translations, we can assume that $t = 0$, $x_3 = 0$, $w_X = e_3$, $q_X = 0$, and

$$(316) \qquad\qquad S_Y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix},$$

where $\varphi$ is a fixed angle with $0 \leq \varphi \leq \tfrac{1}{10}$.

We express $\mathcal{M} \cap P(X, 10H^{-1}(X))$ in cylindrical coordinates over the $z$-axis, namely we parametrize as

(317) $$(\theta, z, t) \mapsto \big(r(\theta, z, t) \cos \theta, r(\theta, z, t) \sin \theta, z\big).$$

In these coordinates, one can directly compute that

(318) $$\nu = \frac{\big(\cos \theta + r^{-1} r_\theta \sin \theta, \sin \theta - r^{-1} r_\theta \cos \theta, -r_z\big)}{\sqrt{1 + r^{-2} |r_\theta|^2 + |r_z|^2}},$$

and

(319) $$\big|\langle K_X, \nu \rangle\big| = \frac{|r_\theta|}{\sqrt{1 + r^{-2} |r_\theta|^2 + |r_z|^2}}.$$

Since $X$ is the center of a (fine) neck, we have

(320) $$r^{-2} |r_\theta|^2 + |r_z|^2 \leq 1 + 10\varepsilon,$$

and

(321) $$\frac{1 - \varepsilon}{r} \leq H(X) \leq \frac{1 + \varepsilon}{r}.$$

Combining these equations with Corollary 6.3 (strong symmetry) we infer that

(322) $$\frac{|r_\theta|}{r} \leq 2(-\psi(0))^{-300}.$$

Together with Theorem 5.9 this yields the estimate

(323) $$|r_\theta| \leq C r^{-599}$$

in the parabolic ball $P(X, 10H^{-1}(X))$.

Now, for the normalized rotation vector with center $q_Y = (q_1, q_2, q_3)$ and axis $w_Y = (0, -\sin \varphi, \cos \varphi)$, we compute

(324) $$K_Y(x) = S_Y J S_Y^{-1}(x - q_Y) = \begin{bmatrix} (x_2 - q_2) \cos \varphi + (x_3 - q_3) \sin \varphi \\ -(x_1 - q_1) \cos \varphi \\ -(x_1 - q_1) \sin \varphi \end{bmatrix},$$

and

$$\langle K_Y, \nu \rangle \sqrt{1 + r^{-2} |r_\theta|^2 + |r_z|^2}$$
$$= r^{-1} (q_1 x_2 - x_1 q_2) \cos \varphi + r^{-1} x_1 (x_3 - q_3) \sin \varphi + r_z (x_1 - q_1) \sin \varphi$$
(325) $$+ r^{-2} r_\theta \big[ x_2 (x_2 - q_2) \cos \varphi + x_2 (x_3 - q_3) \sin \varphi + x_1 (x_1 - q_1) \cos \varphi \big].$$

Arguing as above, using Theorem 5.9 and Corollary 6.3 we obtain

(326) $$|\langle K_Y, \nu \rangle| \leq C r^{-599}$$

in the parabolic ball $P(X, 8H^{-1}(X))$. In addition, we have the rough estimate

(327) $$|q_1| + |q_2| + |q_3| \leq 10r.$$

Now, from equation (325), using the estimates (320), (323), (326) and (327), we infer that

$$(328) \quad \frac{q_1 x_2 - x_1 q_2}{r} \cos \varphi + \frac{x_1}{r}(x_3 - q_3) \sin \varphi + r_z(x_1 - q_1) \sin \varphi \leq C r^{-599}.$$

At time $t = 0$, we consider the points with $\theta = 0$. Then, $(x_1, x_2) = (r, 0)$ and equation (328) yields

$$(329) \quad - q_2 \cos \varphi + (x_3 - q_3 + r_z(r - q_1)) \sin \varphi \leq C r^{-599}.$$

In the case $q_2 \leq 0$, we consider the points with $x_3 = 20r$. Then, using also $|q_3| \leq 10r$ and $\cos \varphi \geq \frac{1}{2}$, we obtain

$$(330) \quad \tfrac{1}{2}|q_2| + (10r + r_z(r - q_1)) \sin \varphi \leq C r^{-599}.$$

Moreover, since $X$ lies on a (fine) neck, we have $|r_z| \leq \varepsilon$. Hence,

$$(331) \quad \tfrac{1}{2}|q_2| + 5r \sin \varphi \leq C r^{-599}.$$

Since $\sin \varphi \geq 0$, we infer that

$$(332) \quad |q_2| \leq C r^{-599},$$

and

$$(333) \quad |\sin \varphi| \leq C r^{-600}.$$

In the case $q_2 \geq 0$, we obtain the same estimates by considering points with $x_3 = -20r$. Similarly, considering points with $\theta = \pi/2$ we obtain

$$(334) \quad |q_1| \leq C r^{-599}.$$

Since $|w_X - w_Y| \leq C|\sin \varphi|$, these inequalities prove the lemma.      □

We are now ready to prove the main theorem of this section:

**Theorem 6.6** (fine asymptotics)**.** *There exist a fixed vector $q = (q_1, q_2, 0) \in \mathbb{R}^3$ and a large constant $C < \infty$ (both depending on $\mathcal{M}$) such that for all $t \in \mathbb{R}$ the surface $(M_t - q) \cap \{x_3 - \psi(t) \geq C\}$ can be expressed in cylindrical coordinates over the z-axis with the estimate*

$$(335) \quad |\partial_\theta r|(\theta, z, t) \leq r(\theta, z, t)^{-100}.$$

*Proof.* Given any time $t \in \mathbb{R}$, we choose a sequence of points $y_j \in M_t$ satisfying

$$(336) \quad \langle y_j, e_3 \rangle = \psi(t) + j.$$

By Corollary 6.3 (strong symmetry), for $j$ large enough (which we will always assume from now on), these points are part of strongly symmetric triples $(Y_j, w_j, q_j)$, where $Y_j = (y_j, t)$. Lemma 6.5 gives the estimates

$$(337) \quad |w_{j+1} - w_j| \leq C j^{-300},$$

and

$$(338) \quad |\langle q_{i+1} - q_i, e_1 \rangle| + |\langle q_{i+1} - q_i, e_2 \rangle| \leq C j^{-299}$$

Since $w_j$ converges to $e_3$, using (337) we get

$$(339) \qquad |w_j - e_3| \leq C \sum_{k=j}^{\infty} k^{-300} \leq C j^{-299}.$$

Similarly, equation (338) implies that $\langle q_j, e_1 \rangle e_1 + \langle q_j, e_2 \rangle e_2$ converges to a limit $q \in \mathbb{R}^2 \times \{0\}$ with the estimate

$$(340) \qquad |\langle q_j - q, e_1 \rangle| + |\langle q_j - q, e_2 \rangle| \leq C \sum_{k=j}^{\infty} k^{-299} \leq C j^{-298}.$$

By translating $\mathcal{M}$, we may assume $q = 0$. Then, we have

$$(341) \qquad |w_j - e_3| \leq C j^{-299},$$

$$(342) \qquad |\langle q_j, e_1 \rangle| + |\langle q_j, e_2 \rangle| \leq C j^{-298}.$$

Together with Theorem 5.9 and Corollary 6.3 (strong symmetry), this yields the assertion. $\qquad \square$

### 6.2. Moving plane method.

The goal of this section is to prove Theorem 6.14, which says that $\mathcal{M}$ is rotationally symmetric.

Let us start with some preliminaries about the Hopf lemma and the strong maximum principle for graphical mean curvature flow. Suppose $u$ and $v$ are graphical solutions of the mean curvature flow in the parabolic ball $P(0, \delta)$ around the origin of size $\delta > 0$, namely

$$(343) \qquad u_t = \sqrt{1 + |Du|^2} \operatorname{div} \left( \frac{Du}{\sqrt{1 + |Du|^2}} \right),$$

$$(344) \qquad v_t = \sqrt{1 + |Dv|^2} \operatorname{div} \left( \frac{Dv}{\sqrt{1 + |Dv|^2}} \right).$$

Then, their difference $w = u - v$ satisfies the linear equation

$$(345) \qquad w_t = a_{ij} w_{ij} + b_i w_i,$$

where

$$(346) \qquad a_{ij}(x, t) = \delta_{ij} - \frac{u_i u_j}{1 + |Du|^2}(x, t),$$

and

$$(347) \qquad b_i(x, t) = -\frac{v_{ij}(u_j + v_j)}{1 + |Du|^2}(x, t) + \frac{v_{kl} v_k v_l (u_i + v_i)}{(1 + |Du|^2)(1 + |Dv|^2)}(x, t).$$

Hence, standard results for linear parabolic equations (see e.g. [Lie96]) yield:

**Lemma 6.7** (Hopf lemma)**.** *Let $u$ and $v$ be graphical solutions of the mean curvature flow in $P(0, \delta) \cap \{x_1 \leq 0\}$. Suppose that $u(0, 0) = v(0, 0)$, and that $u < v$ in $P(0, \delta) \cap \{x_1 < 0\}$. Then, $\frac{\partial u}{\partial x_1}(0, 0) > \frac{\partial v}{\partial x_1}(0, 0)$.*

**Proposition 6.8** (strong maximum principle). *Let $u$ and $v$ be graphical solutions of the mean curvature flow in $P(0,\delta)$. Suppose that $u(0,0) = v(0,0)$, and that $u \leq v$ in $P(0,\delta)$. Then, $u = v$ in $P(0,\delta)$.*

Now, as in the previous section, let $\mathcal{M} = \{M_t\}_{t \in \mathbb{R}}$ be an ancient low entropy flow in $\mathbb{R}^3$, where the plus mode is dominant. By the global curvature bound (Theorem 5.6) and the entropy assumption, there exists some uniform graphical scale $\delta_0 > 0$ at which the Hopf lemma (Lemma 6.7) and the strong maximum principle (Proposition 6.8) can be applied.

To set up the (parabolic variant of the) moving plane method, given a constant $\mu \geq 0$ we consider the surfaces

$$(348) \qquad\qquad M_t^{\mu -} = M_t \cap \{x_1 < \mu\},$$

$$(349) \qquad\qquad M_t^{\mu +} = M_t \cap \{x_1 > \mu\}.$$

Moreover, we set $M_t^{\mu} = M_t \cap \{x_1 = \mu\}$, and denote by $M_t^{\mu <}$ the surface that is obtained from $M_t^{\mu +}$ by reflection about the plane $\{x_1 = \mu\}$, namely

$$(350) \qquad\qquad M_t^{\mu <} = \left\{ (2\mu - x_1, x_2, x_3) : x \in M_t^{\mu +} \right\}.$$

Similarly, denoting by $K_t \subset \mathbb{R}^3$ the closed domain bounded by $M_t$ we consider the regions

$$(351) \qquad\qquad K_t^{\mu -} = K_t \cap \{x_1 < \mu\},$$

$$(352) \qquad\qquad K_t^{\mu +} = K_t \cap \{x_1 > \mu\},$$

and

$$(353) \qquad\qquad K_t^{\mu <} = \left\{ (2\mu - x_1, x_2, x_3) : x \in K_t^{\mu +} \right\}.$$

**Definition 6.9.** We say *the moving plane can reach $\mu$* if for all $\tilde{\mu} \geq \mu$ we have the inclusion $K_t^{\tilde{\mu} <} \subseteq K_t^{\tilde{\mu} -}$ for all $t \in \mathbb{R}$.

The following proposition shows that the reflected domain cannot touch at spatial infinity.

**Proposition 6.10** (no contact at infinity). *For every $\mu > 0$, there exists a constant $h_\mu < \infty$ such that*

$$(354) \qquad\qquad K_t^{\mu <} \cap \{x_3 \geq \psi(t) + h_\mu\} \subseteq Int(K_t^{\mu -})$$

*for every $t \in \mathbb{R}$.*

*Proof.* Denote the position vector of $M_t \cap \{x_3 \geq \psi(t) + C\}$ by

$$(355) \qquad\qquad X(\theta, z, t) = (r(\theta, z, t) \cos\theta, r(\theta, z, t) \sin\theta, z).$$

Given $(z_0, t_0)$ with $z_0 - \psi(t_0) \geq C$, using Theorem 6.6 (fine asymptotics), we obtain

(356)
$$|r(z_0, \theta, t_0) - r(z_0, \theta_0, t_0)| \leq \frac{2\pi}{r(z_0, \theta_0, t_0)^{100}},$$

from which we directly infer that

(357)
$$K_{t_0}^{\mu<} \cap \{x_3 = z_0\} \cap \{x_1 \leq \tfrac{1}{2}\mu\} \subseteq \operatorname{Int}(K_{t_0}^{\mu-})$$

for sufficiently large $z_0 - \psi(t_0)$.

Let $\tilde{\mu} \geq \mu$. Using Theorem 6.6 (fine asymptotics), we see that if $\theta \in (-\frac{\pi}{2}, 0)$ is such that

(358)
$$r(\theta, z_0, t_0) \cos \theta \in [\tfrac{1}{2}\tilde{\mu}, \tfrac{3}{2}\tilde{\mu}],$$

then for sufficiently large $z_0 - \psi(t_0)$ we have

(359)
$$\langle X_\theta, e_2 \rangle = r_\theta \sin \theta + r \cos \theta \geq -r^{-100}|\sin \theta| + r \cos \theta > 0.$$

Hence, we have a graph with positive slope that can be reflected. Repeating the same argument for $\theta \in (0, \frac{\pi}{2})$ we conclude that

(360)
$$K_{t_0}^{\mu<} \cap \{x_3 = z_0\} \cap \{x_1 \geq \tfrac{1}{2}\mu\} \subseteq \operatorname{Int}(K_{t_0}^{\mu-})$$

for sufficiently large $z_0 - \psi(t_0)$. This proves the proposition. $\qquad \square$

**Corollary 6.11** (start plane). *There exists some $\mu < \infty$, such that the moving plane can reach $\mu$.*

*Proof.* By Theorem 5.9 (cap size control) and Theorem 6.6 (fine asymptotics) there exists a $C = C(\mathcal{M}) < \infty$ such that whenever $x_3 - \psi(t) \leq C$ we have

(361)
$$x_1^2 + x_2^2 \leq 10C,$$

namely we uniform control for the position of the cap. Together with Proposition 6.10 (no contact at infinity) this implies the assertion. $\qquad \square$

Now, for $h_\mu$ as in the proposition and $\delta > 0$ we define

(362)
$$E_t^\mu = \{x_3 \leq \psi(t) + h_{\mu/2}\}, \qquad E_t^{\mu,\delta} = \{x \in E_t^\mu : d(x, M_t^\mu) \geq \delta\}.$$

**Lemma 6.12** (distance gap). *Suppose the moving plane can reach $\mu > 0$. Then, there exists a positive increasing function $\alpha : (0, \delta_0) \to \mathbb{R}_+$ such that*

(363)
$$d(M_t^{\mu-}, K_t^{\mu<} \cap E_t^{\mu,\delta}) \geq \alpha(\delta) > 0$$

*for all $t \in \mathbb{R}$.*

*Proof.* We will first show that

(364) $$K_t^{\mu<} \subseteq \mathrm{Int}(K_t^{\mu-})$$

for all $t \in \mathbb{R}$. Indeed, by definition, we have $K_t^{\mu<} \subseteq K_t^{\mu-}$. If (364) fails, then there must be some $t \in \mathbb{R}$ and some $p \in M_t^{\mu<} \cap M_t^{\mu-}$; this contradicts the strong maximum principle (Proposition 6.8). This proves (364).

   Now, suppose towards a contradiction that for some $\delta > 0$ we have

(365) $$\inf_{t\in\mathbb{R}} d(M_t^{\mu-}, K_t^{\mu<} \cap E_t^{\mu,\delta}) = 0.$$

Choose a sequence of space-time points $(x_i, t_i) \in \mathcal{M}$ such that $x_i \in M_{t_i}^{\mu<} \cap E_{t_i}^{\mu,\delta}$ and $\lim_{i\to\infty} d(x_i, K_{t_i}^{\mu-}) = 0$. By Proposition 6.10 (no contact at infinity), Theorem 6.6 (fine asymptotics), and the uniform cap position control (361), the distance between $x_i$ and the point $\psi(t_i)e_3$ is uniformly bounded. Hence, by Theorem 5.6 (global curvature bound) and Theorem 5.9 (cap size control), we can take subsequential limits $\overline{\mathcal{M}}$ and $\bar{x}$ of the flows $\mathcal{M} - (\psi(t_i)e_3, t_i)$ and points $x_i - \psi(t_i)e_3$. Applying the strong maximum principle (Proposition 6.8) for $\overline{\mathcal{M}}$ at the spacetime point $(\bar{x}, 0)$ gives a contradiction. This proves the lemma. $\square$

**Lemma 6.13** (angle gap). *Suppose that the moving plane can reach $\mu > 0$. Then, there exists a positive constant $\theta_\mu > 0$ such that $|\langle \nu(x,t), e_1 \rangle| \geq \theta_\mu$ holds on $M_t^\mu \cap E_t^\mu$ for all $t \in \mathbb{R}$.*

*Proof.* First, the Hopf lemma (Lemma 6.7) shows that $|\langle \nu(x,t), e_1 \rangle| \neq 0$.

   Now, suppose towards a contradiction there is a sequence $(x_i, t_i) \in \mathcal{M}$ such that $x_i \in M_{t_i}^\mu \cap E_{t_i}^\mu$ and $\lim_{i\to\infty} |\langle \nu(x_i, t_i), e_1 \rangle| = 0$. Then, as in the proof of Lemma 6.12, we can take subsequential limits $\overline{\mathcal{M}}$ and $\bar{x}$ of the flows $\mathcal{M} - (\psi(t_i)e_3, t_i)$ and points $x_i - \psi(t_i)e_3$. Applying the Hopf lemma (Lemma 6.7) for $\overline{\mathcal{M}}$ at the spacetime point $(\bar{x}, 0)$ gives a contradiction. This proves the lemma. $\square$

**Theorem 6.14.** $\mathcal{M}$ *is rotationally symmetric.*

*Proof.* It is enough to show that the moving plane can reach $\mu = 0$. Consider the interval

(366) $$I = \{\mu \geq 0 : \text{the moving plane can reach } \mu\}$$

Note that $I \neq \emptyset$ by Corollary 6.11 (start plane). Let $\mu := \inf I$, and observe that $\mu \in I$. Suppose towards a contradiction that $\mu > 0$.

   First, by Proposition 6.10 (no contact at infinity) we have

(367) $$K_t^{\frac{\mu}{2}<} \cap (E_t^\mu)^c \subseteq K_t^{\frac{\mu}{2}-}$$

for all $t \in \mathbb{R}$.

Next, by Lemma 6.13 (angle gap) there exists a $\delta_1 \in (0, \min\{\delta_0, \frac{\mu}{2}\})$ such that for $\delta \in (0, \delta_1)$ we have

$$K_t^{(\mu-\delta)<} \cap E_t^{\mu} \cap \{x_1 \geq \mu - 2\delta_1\} \subseteq K_t^{(\mu-\delta)-} \tag{368}$$

for all $t \in \mathbb{R}$.

Finally, combining the above with Lemma 6.12 (distance gap) we conclude that every $\delta \in (0, \min\{\delta_1, \alpha(\delta_1)\})$ we have

$$K_t^{(\mu-\delta)<} \subseteq K_t^{(\mu-\delta)-}. \tag{369}$$

for all $t \in \mathbb{R}$. Hence, the moving plane can reach $\mu - \delta$; a contradiction. This proves the theorem. $\qquad\square$

## 7. Classification of ancient low entropy flows

**7.1. The noncompact case.** Let $\mathcal{M}$ be an ancient noncompact low entropy flow that is not a round shrinking sphere, round shrinking cylinder or a flat plane. By Corollary 4.25, the plus mode is dominant. As before, we normalize such that $a = a(\mathcal{M}) = 1/\sqrt{2}$.

**Theorem 7.1.** $\mathcal{M}$ *is the bowl soliton.*

*Proof.* From Section 5 we know that $\mathcal{M}$ has a cap of controlled size and opens up like a parabola. From Section 6 we know that $\mathcal{M}$ is rotationally symmetric. Thus, at each time $M_t$ can be described by rotating a curve around the $z$-axis. The curve $\gamma_t(s)$ ($0 \leq s < \infty$) is in the $xz$-plane and satisfies $x(\gamma(0)) = 0$ and $\gamma'(0) = (1, 0, 0)$.

We will first show that the height function $z : M_t \to \mathbb{R}$ does not have any local maxima. Assume towards a contradiction that at some time $t_0$ there is a local maximum of the height function. Denote its value by $h(t_0)$. Since the mean curvature vector points downward, $h(t)$ is decreasing. On the other hand, the minimum of the height function $\psi(t)$ is an inreasing function. In particular, $h - \psi$ is decreasing. Since local maxima of $z$ cannot disapear as we go back in time (c.f. [Ang88]), it follows that the height function has a local maximum for all $t \leq t_0$. Using Proposition 5.8 we see that

$$\lim_{t \to -\infty} (h - \psi)(t) = \infty; \tag{370}$$

this contradicts the fact that the cap size is bounded. Thus, the height function $z$ does not have any local maxima, and the $z$-componenent of the vector $\gamma_t'(s)$ is nonzero for all $s \neq 0$.

We can now describe $M_t$ by a function $r(z, t)$, where $\psi(t) \leq z < \infty$. To proceed we need the following lemma.

**Lemma 7.2.** *The radius function $r = r(z, t)$ satisfies*

$$(371) \qquad\qquad rr_z = 1 + o(1)$$

*uniformly as $z - \psi(t) \to \infty$. In particular, the radius function does not have critical points outside a cap of controlled size.*

*Proof of Lemma 7.2.* We recall from [BC17, Lem. 6.5] that on $\varepsilon$-necks one has the estimate

$$(372) \qquad\qquad |rr_z| + |r^2 r_{zz}| \leq C,$$

in particular

$$(373) \qquad\qquad |\partial_z^2 r^2| \leq C_0/r.$$

Now, we consider a point $(z_0, t_0)$ in a neck satisfying $\partial_z(r^2)(z_0, t_0) \neq 2$. Without loss of generality, we assume $\partial_z(r^2)(z_0, t_0) - 2 = 2\delta > 0$. Then, we have $\partial_z(r^2)(z, t_0) \geq 2 + \delta$ for $|z - z_0| \leq \tilde{\delta} r_0$ where $r_0 = r(z_0, t_0)$ and $\tilde{\delta} = \min\{1, \frac{1}{2}C_0^{-1}\delta\}$. Therefore, we obtain

$$(374) \qquad r^2(z_0 + \tilde{\delta} r_0, t_0) - r^2(z_0, t_0) \geq (2 + \delta)\tilde{\delta} r_0.$$

Since $((0, 0, z_0), t_0)$ is the center of a fine neck with rescaled time parameter $\tau$ satisfying $e^{-\frac{\tau}{2}} = r_0/\sqrt{2}$, using Theorem 5.9 we see that

$$(375) \qquad \left| r\left(z_0 + \tilde{\delta} r_0, \, t_0\right) - r(z_0, t_0) - \tilde{\delta} \right| \leq C r_0^{-\frac{1}{40}}.$$

Combining the above inequalities yields

$$(376) \qquad\qquad C r_0^{\frac{39}{40}} + O(1) \geq \tilde{\delta} \delta r_0.$$

Therefore, $\tilde{\delta} = \frac{1}{2}C_0^{-1}\delta$ for sufficiently large $r_0$, and thus $\delta \leq C r_0^{-\frac{1}{80}}$. Namely,

$$(377) \qquad\qquad \partial_z(r^2) - 2 = O(r^{-\frac{1}{80}}).$$

This proves the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Continuing the proof of the theorem, suppose towards a contradiction that $r$ has a local maximum at some time $t_0$. The mean curvature vector at such a maximum points towards the $z$-axis and has size at least $1/r(z, t)$, hence

$$(378) \qquad\qquad \frac{d}{dt}r(z_{max}(t), t) \leq -1/r(z_{max}(t), t).$$

Integrating this differential inequality from $t_0$ back to time $-\infty$ (again, using [Ang88] to conclude that the maximum point persists), we obtain that $r^2(z_{\max}(t), t)$ grows at least linearly in $-t$. In particular, $z_{\max}(t) - \psi(t) \to \infty$ as $t \to -\infty$, which, as $r_z(z_{\max}(t), t) = 0$, contradicts Lemma 7.2.

By the above, the (upwards) unit normal vector $\nu$ of our surface can never be horizontal or vertical (except at the tip). Thus

$$(379) \qquad\qquad f := \langle \nu, e_3 \rangle > 0.$$

By Lemma 7.2 we have the asymptotics

$$(380) \qquad f(z,t) = \frac{1 + o(1)}{r}$$

uniformly as $z - \psi(t) \to \infty$. In particular, there exists some universal $h_0 < \infty$ such that $f(\psi(t) + h, t) \geq \frac{1}{10\sqrt{h}}$ for all $h \geq h_0$. We claim that there exists some universal $\delta > 0$, such that $f(z,t) \geq \delta$ whenever $\psi(t) \leq z \leq \psi(t) + h_0$. Otherwise, taking a sequence of $(t_i, z_i)$ with $f(t_i, z_i) \to 0$ and passing to a subsequential limit, we obtain a flow (low entropy, with dominant plus mode, rotationally symmetric, with $a = \frac{1}{\sqrt{2}}$) with $r_z(z_0, 0) = 0$ for some $z_0$. By [Ang88], for all $t < 0$, the radius function $r$ would have a local maximum, which was previously seen to be impossible. Hence, we have established the existence of a positive function $g : [0, \infty) \to \mathbb{R}_+$ such that

$$(381) \qquad f(z,t) \geq g(z - \psi(t)).$$

Moreover, by Theorem 5.9 we have the asymptotics

$$(382) \qquad H(z,t) = \frac{1 + o(1)}{r}$$

uniformly as $z - \psi(t) \to \infty$. Combining this with (380) and (381), we see that $H/f$ converges to 1 at spatial infinity uniformly for all $t$. In particular, there is some $C < \infty$ such that

$$(383) \qquad \left| \frac{H}{f} \right| \leq C.$$

Suppose towards a contradiction that

$$(384) \qquad c := \inf \frac{H}{f} < 1.$$

Select $X_i = (x_i, t_i) \in \mathcal{M}$ such that $\frac{H}{f}(X_i) \to c$. Let $\mathcal{M}^i$ be the flow that is obtained from $\mathcal{M}$ by shifting $X_i$ to the space-time origin, and pass to a limit $\mathcal{M}^\infty$. For the limit $\mathcal{M}^\infty$ the function $\frac{H}{f}$ attains a minimum $c < 1$ at the space-time origin. Together with $\frac{H}{f} \to 1$ at spatial infinity, this contradicts the strong maximum principle for the evolution equation

$$(385) \qquad \partial_t \frac{H}{f} = \Delta \frac{H}{f} + 2\langle \nabla \log f, \nabla \frac{H}{f} \rangle.$$

Hence,

$$(386) \qquad \inf \frac{H}{f} \geq 1.$$

A similar argument shows that

$$(387) \qquad \sup \frac{H}{f} \leq 1.$$

Thus $H = f$. Hence, $M_t$ is a mean convex and noncollapsed (since the cap is compact) translating soliton, and thus by [Has15] must be the bowl soliton. $\qquad \square$

7.2. **The compact case.** In this section, we treat the case where the neutral mode is dominant.

**Theorem 7.3.** *If the neutral mode is dominant, then $\mathcal{M}$ is an ancient oval.*

*Proof.* From Section 4.3 we know that $\mathcal{M} = \{M_t\}_{t\in(-\infty, T_E(\mathcal{M}))}$ is compact. Hence, the flow becomes extinct in a point at $T_E(\mathcal{M}) < \infty$, and each time-slice has the topology of a sphere.

Since the blowdown for $t \to -\infty$ is a cylinder, for $t \leq \mathcal{T}$ there is a central neck that divides $M_t$ into two halves $M_t^\pm$. We will show that $M_t^+$ is mean convex and $\alpha$-noncollapsed for $t \leq \mathcal{T}$.

Consider the "height of the tip" function

(388) $$\psi_+(t) = \max_{p\in M_t} x_3(p).$$

Since the blowdown for $t \to -\infty$ is a cylinder, we have

(389) $$\lim_{t\to-\infty} \frac{\psi_+(t)}{\sqrt{|t|}} = \infty.$$

Let $t_j \to -\infty$ be a sequence such that $\psi'_+(t_j)\sqrt{|t_j|} \to \infty$. Then we can find points $p_j \in M_{t_j}^+$ with $\psi_+(t_j) = x_3(p_j)$ and

(390) $$\lim_{j\to\infty} \frac{R(p_j, t_j)}{\sqrt{|t_j|}} = 0.$$

Let $\mathcal{M}^j$ be the sequence of flows that is obtained from $\mathcal{M}$ by shifting $(p_j, t_j)$ to the origin and parabolically rescaling by $R(p_j, t_j)^{-1}$, and pass to a limit $\mathcal{M}^\infty$. The limit $\mathcal{M}^\infty$ is an eternal low entropy flow, which is non-flat. Hence, by Theorem 7.1 it is a translating bowl soliton.

Therefore, for each $j$ large enough $M_{t_j}^+$ is an embedded disk which has the central neck $Z_{t_j}$ as a collar and also another neck $Z_{t_j}^+$ bounding a convex cap $C_{t_j}^+$ that is $\alpha$-noncollapsed, say for $\alpha = \frac{1}{100}$. Let $N_{t_j}^+$ be the topological annulus with collars $Z_{t_j}$ and $Z_{t_j}^+$. Let $\{N_t\}_{t\leq t_j}$ be the time-dependent topological annulus which is obtained from $N_{t_j}^+$ by following the points backwards in time along the mean curvature evolution. By the preservation of necks backwards in time (see Section 3.2), the annulus $N_t$ has two necks as collars for all $t \leq t_j$. Moreover, for $t \ll t_j$ the entire domain $N_t$ constitutes a single neck. Hence, by the parabolic maximum principle (for $H$ and $\frac{|A|}{H}$) applied with two collar boundaries, the annulus $N_t$ is mean convex and satisfies $\frac{|A|}{H} \leq 2$ for all $t \leq t_j$.

Now, adding the cap $C_{t_j}^+$, we infer that $M_{t_j}^+$ is mean convex and satisfies $|A|/H \leq 100$. Since $t_j \to -\infty$, by applying the parabolic maximum principle again, but this time for the disk $M_t^+$ which has only one collar boundary,

the central neck $Z_t$, we conclude that $M_t^+$ is mean convex and satisfies $|A|/H \leq 100$ for $t \leq \mathcal{T}$.

The same argument applies to $M_t^-$. Hence, $\mathcal{M} = \{M_t\}_{t\in(-\infty,T_E(\mathcal{M}))}$ is mean convex and satisfies $|A|/H \leq 100$. Together with the entropy assumption, this implies that $\mathcal{M} = \{M_t\}_{t\in(-\infty,T_E(\mathcal{M}))}$ is $\alpha$-noncollapsed for some $\alpha > 0$. Thus, by [ADS18] it is an ancient oval. $\qquad\square$

## 8. Applications

8.1. **Proof of the mean convex neighborhood conjecture.** The purpose of this section is to prove Theorem 1.7 (mean convex neighborhoods at all times). Since spherical singularities, by compactness, clearly have mean convex neighborhoods, it suffices to prove:

**Theorem 8.1** (mean convex neighborhood theorem). *Let $X_0 = (x_0, t_0)$ be a (backwardly) singular point of $\{M_t\}_{t\geq 0}$ and suppose that*

$$(391) \qquad \lim_{\lambda\to\infty} \lambda(K_{t_0+\lambda^{-2}t} - x_0) = \bar{B}^2(\sqrt{2(t_0 - t)}) \times \mathbb{R},$$

*smoothly with multiplicity one. Then there exist an $\varepsilon = \varepsilon(X_0) > 0$ such that whenever $t_0 - \varepsilon < t_1 < t_2 < t_0 + \varepsilon$ then*

$$(392) \qquad\qquad K_{t_2} \cap B(x_0, \varepsilon) \subseteq K_{t_1} \setminus M_{t_1}.$$

*The same result holds with $M$ and $K$ replaced with $M'$ and $K'$.*

We recall that given a closed embedded surface $M \subset \mathbb{R}^3$, we denote by $\{M_t\}_{t\geq 0}$ the outer flow. Observe that if $x \in M_t$ is a regular point, then there exists a $\delta > 0$ such that $M_t \cap B(x, \delta)$ splits $B(x, \delta)$ into two connected components: one in $\text{Int}(K_t)$ and the other in $\mathbb{R}^3 \setminus K_t$.

Theorem 8.1 will be proved after the following two auxiliary results.

**Proposition 8.2.** *Under the assumptions of Theorem 8.1, there exists a constant $\delta = \delta(X_0) > 0$ and a unit-regular, cyclic, integral Brakke flow $\mathcal{M} = \{\mu_t\}_{t\geq t_0-\delta}$ whose support is $\{M_t\}_{t\geq t_0-\delta}$ such that*
   (i) *The tangent flow to $\mathcal{M}$ at $X_0$ is a multiplicity one cylinder.*
   (ii) *The flow $\mathcal{M}$ has only multiplicity one cylindrical and spherical singularities in $\bar{B}(x_0, 2\delta) \times [t_0 - \delta, t_0 + \delta]$.*
   (iii) *$\mathcal{M}$ is smooth in $\bar{B}(x_0, 2\delta)$ for a.e. $t \in [t_0 - \delta, t_0 + \delta]$ and is smooth outside of a set of Hausdorff dimension 1 for every $t \in [t_0 - \delta, t_0 + \delta]$.*
   (iv) *$H \neq 0$ for every regular point of $\mathcal{M}$ in $\bar{B}(x_0, 2\delta) \times [t_0 - \delta, t_0 + \delta]$.*
   (v) *There exist $A < \infty$ and $c > 0$ such that if $X = (x, t)$ is a point of $\mathcal{M}$ in $\bar{B}(x_0, 2\delta) \times [t_0 - \delta, t_0 + \delta]$ with $R(X) \leq c$ then $\mathcal{M}$ is smooth and connected in $P(X, AR(X))$ and there is a point $X' = (x', t') \in \mathcal{M} \cap P(X, AR(x))$ with $R(X') \geq 2R(X)$ and with $|x' - x_0| \leq \max\{|x - x_0| - cR(X'), \delta/2\}$.*

*Proof.* It follows from [HW18, Thm. B3] that there exists an outer Brakke flow starting from $M$, whose support is $\{M_t\}_{t\geq 0}$. In particular, this, together with monotonicity, implies that $\mathrm{Ent}[\mathcal{H}^n\lfloor M_t]$ is uniformly bounded. Thus, (391) implies that there exists $t_* < t_0$ such that

$$(393) \qquad \frac{1}{(4\pi(t_0 - t_*))^{n/2}} \int \exp\left(\frac{-|x - x_0|^2}{4(t_0 - t_*)}\right) d\mathcal{H}^n\lfloor M_{t_*} < 2.$$

Therefore, applying [HW18, Thm. B3] once more, this time with the initial time $t_*$, we get a unit-regular, cyclic, integral Brakke flow $\mathcal{M} := \{\mu_t\}_{t\geq t_*}$ whose support is $\{M_t\}_{t\geq t_*}$. Together with (391) and (393) it follows that $\mathcal{M}$ has a multiplicity 1 tangent cylinder at $(x_0, t_0)$. This proves (i).

(ii) now follows from the upper semi-continuity of the density (see Section 2.2), the proof of Theorem 3.2 (partial regularity) and the gap result from Bernstein-Wang [BW17].

(iii) follows from (ii) and standard stratification, see e.g. [Whi97].

Suppose towards a contradiction that (iv) does not hold. Then there exists a sequence of smooth point $X_i = (x_i, t_i) \to X_0 = (x_0, t_0)$ with $H(X_i) = 0$. Denoting by $r_i = R(X_i)$ the regularity scale, since $X_0$ is singular and $X_i \to X_0$, we have $r_i \to 0$.

Let $\mathcal{M}^i$ be the flow that is obtained from $\mathcal{M}$ by shifting $X_i$ to the origin and parabolically rescaling by $1/r_i$, and pass (using Ilmanen's compactness theorem [Ilm94]) to a Brakke flow limit $\mathcal{M}^\infty$.

By construction as a blowup limit, $\mathcal{M}^\infty$ is an ancient, integral Brakke flow. By the local regularity theorem for the mean curvature flow and [HW18, Thm. B5], $\mathcal{M}^\infty$ is unit regular. By a result of White [Whi09], the limit $\mathcal{M}^\infty$ is cyclic. Furthermore, we have

$$(394) \qquad\qquad \mathrm{Ent}[\mathcal{M}^\infty] \leq \mathrm{Ent}[S^1 \times \mathbb{R}].$$

This entropy bound follows from the same reasoning as in the proof of [Her18, Thm. 1]. For the sake of completeness, we include it here as well. Since the $\mathcal{M}^i$ have uniformly bounded entropy (by monotonicity), and as $\mathcal{M}^i \to \mathcal{M}^\infty$, it follows from the definition of the entropy of $\mathcal{M}^\infty$ that there exist a sequences of points $(x'_k, t'_k)$ in the original space-time with $(x'_k, t'_k) \to (x_0, t_0)$ and scales $r'_k \to 0$, such that for every $\varepsilon > 0$, for $k$ large enough,

$$(395) \qquad \frac{1}{4\pi r_k'^2} \int \exp(-|x - x'_k|^2/4r_k'^2) d\mu_{t'_k - r_k'^2} \geq \mathrm{Ent}[\mathcal{M}^\infty] - \varepsilon.$$

On the other hand, since $(x_0, t_0)$ is a cylindrical singularity, there exist $r > 0$ such that for every point in $(p, t) \in B(x_0, r) \times [t_0 - r, t_0 + r]$ we have

$$(396) \qquad \frac{1}{4\pi r^2} \int \exp(-|x - p|^2/4r^2) d\mu_{t - r^2} \leq \mathrm{Ent}[\mathbb{S}^1 \times \mathbb{R}] + \varepsilon.$$

By Huisken's monotonicity formula and by the arbitrariness of $\varepsilon$, this implies (394). Hence, $\mathcal{M}^\infty$ is an ancient low entropy flow.

Now, by Theorem 1.2 (classification of ancient low entropy flows) the limit $\mathcal{M}^\infty$ is either a flat plane, round shrinking sphere, round shrinking cylinder, a bowl soliton, or an ancient oval. If the limit is non-flat, then for $i$ large enough that contradicts $H(X_i) = 0$. If the limit is a flat plane, then we obtain a contradiction with the fact $X_i$ has regularity scale $r_i \to 0$, by the local regularity theorem. This proves (iv).

For (v), observe by inspection of the four non-planar ancient low entropy flows that there exist $A, C < \infty$ such that if $X = (x, t)$ is a point on such a flow $\mathcal{M}^\infty$, then for every unit vector $v$, there exists a point in $X' \in P(X, AR(X)) \cap \{y : (y - x) \cdot v \geq R(X)\} \cap \mathcal{M}^\infty$ with

$$(397) \qquad 4R(X) \leq R(X') \leq CR(X).$$

Using this, (v) follows by a contradiction argument similarly as in the proof of (iv). This finishes the proof of the proposition. $\qquad\square$

By wiggling with the constant $\delta$ we can assume that, in addition to (i)-(v) of Proposition 8.2 we also have:

(vi) $t_0 - \delta$ is a smooth time for the flow in $\bar{B}(x_0, 2\delta)$, and $M_{t_0-\delta} \cap \bar{B}(x_0, 2\delta)$ is $\varepsilon_0$-cylindrical, where $\varepsilon_0 > 0$ is fixed.

Let $\mathcal{N}$ the unique (space-time) connected component of $\mathcal{M} \cap (\bar{B}(x_0, \delta) \times [t_0 - \delta, t_0 + \delta])$ which contains the point $X_0 = (x_0, t_0)$.[22] In particular, there exist $\varepsilon > 0$ such that

$$(398) \qquad \mathcal{M} \cap \Big( B(x_0, 2\varepsilon) \times [t_0 - \varepsilon, t_0 + \varepsilon] \Big) \subseteq \mathcal{N}.$$

We will show that $B(x_0, \varepsilon) \times [t_0 - \varepsilon, t_0 + \varepsilon]$ is a mean convex neighborhood.

For any smooth point $X = (x, t) \in \mathcal{M} \cap (\bar{B}(x_0, 2\delta) \times [t_0 - \delta, t_0 + \delta])$, if $\mathbf{H}(X)$ points into $K_t$ we say that $H(X) > 0$, and if $\mathbf{H}(X)$ points outside of $K_t$, we say that $H(X) < 0$.

**Lemma 8.3.** $H > 0$ *for all regular points in* $\mathcal{N}$.

*Proof.* For every time $s \in [t_0 - \delta, t_0 + \delta]$, let $\mathcal{A}^s$ be the set of (space-time) connected components of $\mathcal{N} \cap (\mathbb{R}^3 \times [t_0 - \delta, s])$, and let $\mathcal{A}^s_-$ be the set of space time connected components of $\mathcal{N} \cap (\mathbb{R}^3 \times [t_0 - \delta, s))$. Set

$$(399) \qquad \begin{aligned} I := \{ s \in [t_0 - \delta, t_0 + \delta] \mid &\text{For every } \mathcal{F} \in \mathcal{A}^s \text{ and } X_1, X_2 \in \mathrm{Reg} \cap \mathcal{F}, \\ &\mathrm{Sign}(H(X_1)) = \mathrm{Sign}(H(X_2)) \}. \end{aligned}$$

Note that $t_0 - \delta \in I$ by property (vi). Also note that if $s \notin I$, then $s' \notin I$ for every $s' > s$. Thus $I$ is an interval containing $t_0 - \delta$. The assertion of

---

[22]In this section $B(x, r)$ always denotes an open ball, and $\bar{B}(x, r)$ a closed ball.

the lemma is therefore equivalent to the assertion that $t_0 + \delta \in I$. Suppose towards a contradiction that $t_0 + \delta \notin I$, and set

$$(400) \qquad\qquad\qquad s_0 := \sup I.$$

**Claim 8.4.** $s_0 \notin I$.

*Proof of Claim 8.4.* If $s_0 = t_0 + \delta$ this is clear by assumption. Assume that $s_0 < t_0 + \delta$ were in $I$. Since $\mathcal{N} \cap \bar{B}(x_0, \delta)$ is compact, each two connected components of $\mathcal{F}_1 \neq \mathcal{F}_2$ in $\mathcal{A}^{s_0}$ are a positive distance apart. Moreover, there will be no new connected components entering $\bar{B}(x_0, \delta)$ for some time (again, since supports of Brakke flows are closed). Thus, there exists an $s > s_0$ such that each connected components $\mathcal{F}^s \in A^s$ has a unique connected component $\mathcal{F}^{s_0} \in A^{s_0}$ such that $\mathcal{F}^{s_0} \subseteq \mathcal{F}^s$.

We claim that if $s - s_0$ is sufficiently small, every smooth point in $\mathcal{F}^s$ can be connected by a path of smooth points in $\bar{B}(x_0, \delta)$ to a point in $\mathcal{F}^{s_0}$. Indeed, by repeated use of Lemma 8.2 (v), any regular point $X = (x, t) \in \mathcal{F}^s$, can be connected via regular points to some $(x', t') = X' \in \mathcal{F}^s$ with $R(X') \geq c$ and with $|x' - x_0| \leq \delta - c^2$. Thus, is $s - s_0$ is sufficiently small, $X'$ can be connected via smooth points to a point in $\mathcal{F}^{s_0}$ (this argument is similar to [KL17, Propsition 3.5]).

As $H \neq 0$, this implies that $H$ does not change sign on $\mathcal{F}^s$. This proves the claim. □

Thus, if $t_0 + \delta \notin I$ then $I^c = [s_0, t_0 + \delta]$ for some $s_0 \geq t_0 - \delta$. As $s_0 \in I^c$, there exists some connected component $\mathcal{F} \in \mathcal{A}^{s_0}$ such that $H$ does not have a fixed sign over regular points in $\mathcal{F}$. Since for each $s < s_0$, we have $s \in I$, there are three potential ways in which this can happen for the first time at time $s_0$. We will rule them out.

1. *Appearance of a new connected component*: $\mathcal{F} \subseteq \partial B(x_0, \delta) \times \{s_0\}$, i.e. a new connected component enters through the boundary. In this case, all points of $\mathcal{F}$ are regular, as the tangent flows are contained in half-spaces. Since $\mathcal{F}$ is connected (by definition), and as $H \neq 0$, this implies that the sign of $H$ is constant along $\mathcal{F}$. Thus, this case is excluded.

2. *Merger of two connected components*: There exist two distinct space-time connected components $\mathcal{F}_\pm \in \mathcal{A}^{s_0}$ such that $H > 0$ (resp. $H < 0$) on $\mathcal{F}_+$ (resp. $\mathcal{F}_-$) and such that $\overline{\mathcal{F}_+} \cap \overline{\mathcal{F}_-} \neq \emptyset$. Since $H \neq 0$ at regular points, the points of $\overline{\mathcal{F}_+} \cap \overline{\mathcal{F}_-}$ are all singular (in particular, $s_0$ is a singular time). Moreover, by avoidance, $\overline{\mathcal{F}_+} \cap \overline{\mathcal{F}_-} \subseteq \partial B(x_0, \delta) \times \{s_0\}$.

**Claim 8.5.** There exists a time $t < s_0$ such that $M_t \cap B(x_0, 2\delta)$ is smooth and some *spatial* connected component of $M_t \cap \bar{B}(x_0, 2\delta)$ contains points both from $\mathcal{F}_+$ and from $\mathcal{F}_-$.

*Proof of Claim 8.5.* Consider the spatial connected components of such $\mathcal{F}_\pm$, $\mathcal{F}_\pm(t)$, and observe how they extend to $\bar{B}(x_0, 2\delta)$. If they never intersect on smooth times, then via a density argument (or, indeed, by the maximum

principle), we see that, at singular times as well, if $(x', t')$ is a point reached by both smooth flows then $x' \in \partial B(x_0, 2\delta)$. We thus obtain two disjoint unit multiplicity Brakke flows in $B(x_0, 3\delta/2)$, which have a point in $\overline{\mathcal{F}_+} \cap \overline{\mathcal{F}_-}$ as a joint point they reach at time $s_0$, but this is again impossible since the Huisken density near $(x_0, t_0)$ is less than 2. $\qquad\square$

As $H \neq 0$ on $M_t \cap \bar{B}(x_0, 2\delta)$, this gives a contradiction to the definitions of $\mathcal{F}_\pm$. Indeed, at time $t$, the signs of $H$ on $\mathcal{F}_\pm$ have to be the same. Thus, this case is excluded as well.

*3. Merger of an existing component with a new boundary component*: In this case, we may assume without loss of generality that $\mathcal{F}$ contains $\mathcal{F}_+ \cup \mathcal{F}_\partial$, where on $\mathcal{F}_+ \in \mathcal{A}_-^{s_0}$ we have $H > 0$ on regular points and where $\mathcal{F}_\partial \subseteq \mathcal{F} \cap (\partial B(x_0, \delta) \times \{s_0\})$ is connected, intersects $\overline{\mathcal{F}_+}$, and contains at least one point which is not in

$$(401) \qquad \overline{\mathcal{M} \cap (\bar{B}(x_0, \delta) \times [t_0 - \delta, s_0))},$$

and all such points have $H < 0$ (note that, as in the first case, such points are always regular). This case is excluded similarly to Case 2 above: Taking such a point, and looking at the connected component of the point reaching it at a slightly earlier smooth time $t$ in a slightly bigger ball, we see it intersects with the continuation of $\mathcal{F}_+(t)$ in that ball, which cannot be.

We have thus excluded the existence of such an $\mathcal{F}$, so $I^c = \emptyset$. This concludes the proof of the lemma. $\qquad\square$

*Proof of Theorem 8.1.* Having established Proposition 8.2 and Lemma 8.3, the final step of the proof of the mean convex neighborhood theorem is now similar to the final step of the proof of Theorem 3.5 in [HW18]. For convenience of the reader, we include the argument here as well.

**Claim 8.6.** If $t_0 - \varepsilon \leq t_1 < t_2 \leq t_0 + \varepsilon$, then

$$(402) \qquad K_{t_2} \cap B(x_0, \varepsilon) \subseteq K_{t_1}.$$

*Proof of Claim 8.6.* Suppose to the contrary that there is point $y \in K_{t_2} \cap B(x_0, \varepsilon)$ that is not in $K_{t_1}$ for some $t_1 < t_2$ with $t_1, t_2 \in [t_0 - \varepsilon, t_0 + \varepsilon]$. Let

$$(403) \qquad 0 < \rho < \text{dist}(y, K_{t_1} \cup \partial B(x_0, \varepsilon)).$$

Let $t \in (t_1, t_0)$ be the first time $\geq t_1$ such that

$$(404) \qquad \text{dist}(y, K_t) = \rho.$$

Let $p \in K_t$ be a point such that $\text{dist}(y, p) = \text{dist}(y, K_t)$. Note that the tangent flow at $(p, t)$ lies in a half-space (more specifically the half-space $\{x : x \cdot (y - p) \leq 0\}$). Hence $(p, t)$ is a regular point of the flow. By the previous lemma, the mean curvature at $(p, t)$ is positive and so for $s < t$ very close to $t$, we have $\text{dist}(y, K_s) < \rho$; a contradiction. This proves the claim. $\qquad\square$

Continuing the proof of the theorem, it only remains to show that for $t_1 < t_2$ with $t_1, t_2 \in [t_0 - \varepsilon, t_0 + \varepsilon]$, if

$$(405) \qquad\qquad\qquad x \in B(x_0, \varepsilon) \cap K_{t_2},$$

then $x$ is in the interior of $K_{t_1}$. If $x$ is the interior of $K_{t_2}$, then by Claim 8.6 it is in the interior of $K_{t_1}$. Thus we may assume that $x$ is in the boundary of $K_{t_2}$. For $s \in [t_1, t_2)$ sufficiently close to $t_2$, the point $x$ is in the interior of $K_s$. If $(x, t_2)$ is a regular point, this is because the mean curvature is positive. If $(x, t_2)$ is a singular point, this is true by (ii) of Proposition 8.2. Since $x$ is in the interior of $K_s$, and since $K_s \subseteq K_{t_1}$, we conclude that $x$ is in the interior of $K_{t_1}$. This completes the proof of the mean convex neighborhood theorem. $\qquad\qquad\square$

## 8.2. **Proof of the uniqueness conjectures for weak flows.**
In this final section, we prove the uniqueness results for mean curvature flow through singularities, namely Theorem 1.9 (nonfattening) and Theorem 1.12 (evolution of two-spheres).

*Proof of Theorem 1.9.* Observe that if $T \leq T_{\mathrm{disc}}$ then, by definition, the outer flow $\{M_t\}_{t \in [0,T]}$ agrees with the level set flow of $M$ and the inner flow of $M$. In particular, if $(x_0, T)$ is a multiplicity one cylindrical (resp. spherical) singularity of $\{M_t\}$ then *either* $\mathcal{K}_{X,\lambda}$, *or* $\mathcal{K}'_{X,\lambda}$ converges for $\lambda \to 0$ smoothly with multiplicity one to a round shrinking solid cylinder or a round shrinking solid ball (see also item (i) of Proposition 8.2 and its proof).

The result now follows from combining the main theorem of Hershkovits-White [HW18], which establishes that $T < T_{\mathrm{disc}}$ assuming the existence of mean convex neighborhoods a priori, and Theorem 1.7, which proves the existence of mean convex neighborhoods. $\qquad\qquad\square$

*Proof of Theorem 1.12.* Let $M \subset \mathbb{R}^3$ be an embedded two-sphere, and let $\mathcal{M} = \{M_t\}$ be the outer flow starting at $M$. By a result of White [Whi95], whenever $M_t$ is smooth, then it has genus zero.

Suppose towards a contradiction that there is some singular point with a tangent flow that is not spherical or cylindrical. Let $T$ be the first time when this happens (more precisely, the infimum of times where this happens). Then $T \leq T_{\mathrm{disc}}$ (by Theorem 1.9) and for a.e. $t \in [0, T]$ the flow is smooth. Hence, we can run the argument from Ilmanen [Ilm95], and get, assuming the multiplicity one conjecture, that at time $T$ at every singular point there exists a tangent flow that is smooth with multiplicity one.

Such a tangent flow, by smoothness and unit-regularity, cannot be a nontrivial static or quasistatic cone. Hence it is a properly embedded shrinker. Suppose towards a contradiction there are two loops in the shrinker $N$ that have nonvanishing intersection number modulo two. Then for some smooth time $t' < T$ we would see two loops in $M_{t'}$ with nonvanishing intersection number modulo two; a contradiction. Hence, by a result of Brendle [Bre16]

the shrinker $N$ is a sphere, cylinder or a plane. Thus, arguing as in the first paragraph of the proof of Theorem 1.9, by the mean convex neighborhood theorem (Theorem 1.7) and compactness, there is some $\delta > 0$ such that the flow has only cylindrical and spherical singularities on $[0, T + \delta)$; this contradicts the definition of $T$. Hence, the flow $\mathcal{M}$ has only spherical and cylindrical singularities. Thus, by Theorem 1.9 (nonfattening) the solution is unique. $\qquad\square$

## References

[ADS15]    S. Angenent, P. Daskalopoulos, and N. Sesum. Unique asymptotics of ancient convex mean curvature flow solutions. *arXiv:1503.01178*, 2015.

[ADS18]    S. Angenent, P. Daskalopoulos, and N. Sesum. Uniqueness of two-convex closed ancient solutions to the mean curvature flow. *arXiv:1804.07230*, 2018.

[AIC95]    S. Angenent, T. Ilmanen, and D. Chopp. A computed example of nonuniqueness of mean curvature flow in $\mathbf{R}^3$. *Comm. Partial Differential Equations*, 20(11-12):1937–1958, 1995.

[All72]    W. Allard. On the first variation of a varifold. *Ann. of Math. (2)*, 95:417–491, 1972.

[Ang88]    Sigurd Angenent. The zero set of a solution of a parabolic equation. *J. Reine Angew. Math.*, 390:79–96, 1988.

[Ang92]    S. Angenent. Shrinking doughnuts. In *Nonlinear diffusion equations and their equilibrium states, 3 (Gregynog, 1989)*, volume 7 of *Progr. Nonlinear Differential Equations Appl.*, pages 21–38. Birkhäuser Boston, Boston, MA, 1992.

[AV97]    S. Angenent and J. Velazquez. Degenerate neckpinches in mean curvature flow. *J. Reine Angew. Math.*, 482:15–66, 1997.

[AW94]    S. Altschuler and L. Wu. Translating surfaces of the non-parametric mean curvature flow with prescribed contact angle. *Calc. Var. Partial Differential Equations*, 2(1):101–111, 1994.

[BC17]    S. Brendle and K. Choi. Uniqueness of convex ancient solutions to mean curvature flow in $\mathbb{R}^3$. *arXiv:1711.00823v2*, 2017.

[BH16]    S. Brendle and G. Huisken. Mean curvature flow with surgery of mean convex surfaces in $\mathbb{R}^3$. *Invent. Math.*, 203(2):615–654, 2016.

[BHH16]    R. Buzano, R. Haslhofer, and O. Hershkovits. The moduli space of two-convex embedded spheres. *arXiv:1607.05604*, 2016.

[BK17]    R. Bamler and B. Kleiner. Uniqueness and stability of Ricci flow through singularities. *arXiv:1709.04122*, 2017.

[Bra78]    K. Brakke. *The motion of a surface by its mean curvature*, volume 20 of *Mathematical Notes*. Princeton University Press, Princeton, N.J., 1978.

[Bre13]    S. Brendle. Rotational symmetry of self-similar solutions to the Ricci flow. *Invent. Math.*, 194(3):731–764, 2013.

[Bre16]    S. Brendle. Embedded self-similar shrinkers of genus 0. *Ann. of Math. (2)*, 183(2):715–728, 2016.

[BW17]    J. Bernstein and L. Wang. A topological property of asymptotically conical self-shrinkers of small entropy. *Duke Math. J.*, 166(3):403–435, 2017.

[CGG91]    Y.G. Chen, Y. Giga, and S. Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.*, 33(3):749–786, 1991.

[CHN13]    J. Cheeger, R. Haslhofer, and A. Naber. Quantitative stratification and the regularity of mean curvature flow. *Geom. Funct. Anal.*, 23(3):828–847, 2013.

[CIM15]    T. Colding, T. Ilmanen, and W. Minicozzi. Rigidity of generic singularities of mean curvature flow. *Publ. Math. Inst. Hautes Études Sci.*, 121:363–382, 2015.

[CM12]   T. Colding and W. Minicozzi. Generic mean curvature flow I; generic singularities. *Ann. of Math. (2)*, 175(2):755–833, 2012.

[CM15]   T. Colding and W. Minicozzi. Uniqueness of blowups and Lojasiewicz inequalities. *Ann. of Math. (2)*, 182(1):221–285, 2015.

[CM16]   T. Colding and W. Minicozzi. The singular set of mean curvature flow with generic singularities. *Invent. Math.*, 204(2):443–471, 2016.

[CM17]   T. Colding and W. Minicozzi. Arnold-Thom gradient conjecture for the arrival time. *arXiv:1712.05381*, 2017.

[CMP15]  T. Colding, W. Minicozzi, and E. Pedersen. Mean curvature flow. *Bull. Amer. Math. Soc. (N.S.)*, 52(2):297–333, 2015.

[Das14]  P. Daskalopoulos. Ancient solutions to geometric flows. In *Proceedings of the International Congress of Mathematicians, Vol. III (Seoul, 2014)*, pages 773–788, Seoul, 2014.

[Eck07]  K. Ecker. A formula relating entropy monotonicity to Harnack inequalities. *Comm. Anal. Geom.*, 15(5):1025–1061, 2007.

[ES91]   L. Evans and J. Spruck. Motion of level sets by mean curvature. I. *J. Differential Geom.*, 33(3):635–681, 1991.

[Gan18]  Z. Gang. A description of a space-and-time neighborhood of generic singularities formed by mean curvature flow. *arXiv:1803.10903v2*, 2018.

[GKS18]  Z. Gang, D. Knopf, and M. Sigal. Neckpinch dynamics for asymmetric surfaces evolving by mean curvature flow. *Mem. Amer. Math. Soc.*, 253(1210):v+78, 2018.

[Ham95]  R. Hamilton. Harnack estimate for the mean curvature flow. *J. Differential Geom.*, 41(1):215–226, 1995.

[Has15]  R. Haslhofer. Uniqueness of the bowl soliton. *Geom. Topol.*, 19(4):2393–2406, 2015.

[Her18]  O. Hershkovits. Translators asymptotic to cylinders. *arXiv:1805.10553*, 2018.

[HH16]   R. Haslhofer and O. Hershkovits. Ancient solutions of the mean curvature flow. *Comm. Anal. Geom.*, 24(3):593–604, 2016.

[HK17a]  R. Haslhofer and B. Kleiner. Mean curvature flow of mean convex hypersurfaces. *Comm. Pure Appl. Math.*, 70(3):511–546, 2017.

[HK17b]  R. Haslhofer and B. Kleiner. Mean curvature flow with surgery. *Duke Math. J.*, 166(9):1591–1626, 2017.

[HK18]   R. Haslhofer and D. Ketover. Minimal two-spheres in three-spheres. *arXiv:1708.06567*, 2018.

[HM11]   R. Haslhofer and R. Müller. A compactness theorem for complete Ricci shrinkers. *Geom. Funct. Anal.*, 21(5):1091–1116, 2011.

[HP99]   G. Huisken and A. Polden. Geometric evolution equations for hypersurfaces. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 45–84. Springer, Berlin, 1999.

[Hui84]  G. Huisken. Flow by mean curvature of convex surfaces into spheres. *J. Differential Geom.*, 20(1):237–266, 1984.

[Hui90]  G. Huisken. Asymptotic behavior for singularities of the mean curvature flow. *J. Differential Geom.*, 31(1):285–299, 1990.

[Hui93]  G. Huisken. Local and global behaviour of hypersurfaces moving by mean curvature. In *Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990)*, volume 54 of *Proc. Sympos. Pure Math.*, pages 175–191. Amer. Math. Soc., Providence, RI, 1993.

[HW18]   O. Hershkovits and B. White. Non-fattening of mean curvature flow at singularities of mean convex type. *Comm. Pure Appl. Math. (to appear)*, 2018.

[Ilm94]  T. Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Mem. Amer. Math. Soc.*, 108(520):x+90, 1994.

[Ilm95]   T. Ilmanen. Singularities of mean curvature flow of surfaces. *https://people.math.ethz.ch/ ilmanen/papers/sing.ps*, 1995.

[Ilm98]   T. Ilmanen. Lectures on mean curvature flow and related equations. *https://people.math.ethz.ch/ ilmanen/papers/notes.pdf*, 1998.

[Ilm03]   T. Ilmanen. Problems in mean curvature flow. *https://people.math.ethz.ch/ ilmanen/classes/eil03/problems03.ps*, 2003.

[IW16]    J. Isenberg and H. Wu. Mean curvature flow of noncompact hypersurfaces with type-II curvature blow-up. *arXiv:1603.01664*, 2016.

[KKM18]   N. Kapouleas, S. Kleene, and N. Moller. Mean curvature self-shrinkers of high genus: non-compact examples. *J. Reine Angew. Math.*, 739:1–39, 2018.

[KL17]    Bruce Kleiner and John Lott. Singular Ricci flows I. *Acta Math.*, 219(1):65–134, 2017.

[KM14]    S. Kleene and N. Moller. Self-shrinkers with a rotational symmetry. *Trans. Amer. Math. Soc.*, 366(8):3943–3963, 2014.

[Lie96]   G. Lieberman. *Second order parabolic differential equations*. World Scientific Publishing Co., Inc., River Edge, NJ, 1996.

[MS08]    J. Metzger and F. Schulze. No mass drop for mean curvature flow of mean convex hypersurfaces. *Duke Math. J.*, 142(2):283–312, 2008.

[MSHS15]  F. Martin, A. Savas-Halilaj, and K. Smoczyk. On the topology of translating solitons of the mean curvature flow. *Calc. Var. Partial Differential Equations*, 54(3):2853–2882, 2015.

[MZ98]    F. Merle and H. Zaag. Optimal estimates for blowup rate and behavior for nonlinear heat equations. *Comm. Pure Appl. Math.*, 51(2):139–196, 1998.

[Ngu14]   X. Nguyen. Construction of complete embedded self-similar surfaces under mean curvature flow, Part III. *Duke Math. J.*, 163(11):2023–2056, 2014.

[Per02]   G. Perelman. The entropy formula for the Ricci flow and its geometric applications. *arXiv:math/0211159*, 2002.

[Per03]   G. Perelman. Ricci flow with surgery on three-manifolds. *arXiv:math/0303109*, 2003.

[Rub13]   H. Rubinstein. Some of Hyam's favourite problems. *Contemp. Math.*, 597:165–175, 2013.

[Sim83]   L. Simon. *Lectures on geometric measure theory*, volume 3 of *Proceedings of the Centre for Mathematical Analysis, Australian National University*. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[Smi82]   F. Smith. On the existence of embedded minimal 2-spheres in the 3-sphere, endowed with an arbitrary Riemannian metric. *Phd thesis, Supervisor: Leon Simon, University of Melbourne*, 1982.

[SW16]    F. Schulze and B. White. A local regularity theorem for mean curvature flow with triple edges. *arXiv:1605.06592*, 2016.

[SX17]    J. Spruck and L. Xiao. Complete translating solitons to the mean curvature flow in $\mathbb{R}^3$ with nonnegative mean curvature. *arXiv:1703.01003*, 2017.

[Top11]   P. Topping. Reverse bubbling in geometric flows. In *Surveys in geometric analysis and relativity*, volume 20 of *Adv. Lect. Math. (ALM)*, pages 491–508. Int. Press, Somerville, MA, 2011.

[Wan92]   L. Wang. On the regularity theory of fully nonlinear parabolic equations: I. *Comm. Pure Appl. Math.*, 45(1):27–76, 1992.

[Wan11]   X. Wang. Convex solutions to the mean curvature flow. *Ann. of Math. (2)*, 173(3):1185–1239, 2011.

[Wan16]   L. Wang. Asymptotic structure of self-shrinkers. *arXiv:1610.04904*, 2016.

[Whi95]   B. White. The topology of hypersurfaces moving by mean curvature. *Comm. Anal. Geom.*, 3(1-2):317–333, 1995.

[Whi97]   B. White. Stratification of minimal surfaces, mean curvature flows, and harmonic maps. *J. Reine Angew. Math.*, 488:1–35, 1997.

[Whi00]   B. White. The size of the singular set in mean curvature flow of mean convex sets. *J. Amer. Math. Soc.*, 13(3):665–695, 2000.
[Whi02]   B. White. Evolution of curves and surfaces by mean curvature. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pages 525–538, Beijing, 2002. Higher Ed. Press.
[Whi03]   B. White. The nature of singularities in mean curvature flow of mean-convex sets. *J. Amer. Math. Soc.*, 16(1):123–138, 2003.
[Whi05]   B. White. A local regularity theorem for mean curvature flow. *Ann. of Math. (2)*, 161(3):1487–1519, 2005.
[Whi09]   B. White. Currents and flat chains associated to varifolds, with an application to mean curvature flow. *Duke Math. J.*, 148(1):41–62, 2009.

KYEONGSU CHOI, DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVE, CAMBRIDGE, MA 02139, USA

ROBERT HASLHOFER, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 40 ST GEORGE STREET, TORONTO, ON M5S 2E4, CANADA

OR HERSHKOVITS, DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, 450 SERRA MALL, STANFORD, CA 94305, USA

*E-mail:* choiks@mit.edu, roberth@math.toronto.edu, orher@stanford.edu