# Planarity and the mean curvature flow of pinched submanifolds in higher codimension

Keaton Naff

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

# Abstract

Planarity and the mean curvature flow of pinched submanifolds in higher codimension

Keaton Naff

In this thesis, we explore the role of planarity in mean curvature flow in higher codimension and investigate its implications for singularity formation in a certain class of flows. In Chapter 1, we show that the blow-ups of compact $n$-dimensional solutions to mean curvature flow in $\mathbb{R}^N$ initially satisfying the pinching condition $|A|^2 < c|H|^2$ for a suitable constant $c = c(n)$ must be codimension one. We do this by establishing a new a priori estimate via a maximum principle argument.

In Chapter 2, we consider ancient solutions to the mean curvature flow in $\mathbb{R}^{n+1}$ ($n \geq 3$) that are weakly convex, uniformly two-convex, and satisfy derivative estimates $|\nabla A| \leq \gamma_1 |H|^2$, $|\nabla^2 A| \leq \gamma_2 |H|^3$. We show that such solutions are noncollapsed. The proof is an adaptation of the foundational work of Huisken and Sinestrari on the flow of two-convex hypersurfaces. As an application, in arbitrary codimension, we classify the singularity models of compact $n$-dimensional ($n \geq 5$) solutions to the mean curvature flow in $\mathbb{R}^N$ that satisfy the pinching condition $|A|^2 < c|H|^2$ for $c = \min\{\frac{1}{n-2}, \frac{3(n+1)}{2n(n+2)}\}$. Using recent work of Brendle and Choi, together with the estimate of Chapter 1, we conclude that any blow-up model at the first singular time must be a codimension one shrinking sphere, shrinking cylinder, or translating bowl soliton.

Finally, in Chapters 3 and 4, we prove a canonical neighborhood theorem for the mean curvature flow of compact $n$-dimensional submanifolds in $\mathbb{R}^N$ ($n \geq 5$) satisfying a pinching condition $|A|^2 < c|H|^2$ for $c = \min\{\frac{1}{n-2}, \frac{3(n+1)}{2n(n+2)}\}$. We first discuss, in some detail, a well-known

compactness theorem of the mean curvature flow. Then, adapting an argument of Perelman and using the conclusions of Chapter 2, we characterize regions of high curvature in the pinched solutions of the mean curvature flow under consideration.

# Table of Contents

# Acknowledgements

I would like to begin by thanking my advisor, Simon Brendle, for his steadfast guidance and support during my time at Columbia. From our first discussions until the completion of this thesis, he has been incredibly generous with his encouragement, patience, and advice. I am constantly inspired by his determination and example; humbled by his wisdom and brilliance. Thank you.

I am very grateful for the teachings, discussions, and mentorship I have had during the past five years. I would like to thank Professors Toti Daskalopolous, Gerhard Huisken, Johan De Jong, Chiu-Chu Melissa Liu, Duong Phong, Ovidiu Savin, Natasa Šešum, Michael Thaddeus, and Mu-Tao Wang. I will fondly remember the time I spent as a teaching assistant to Prof. Richard Hamilton – I greatly enjoyed our conversations and his stories.

At each stage of graduate school, I have benefited from the guidance and friendship of my exceptional colleagues. I am grateful to Beomjun Choi, Tsz-Kiu Aaron Chow, Clara Dolfen, Alexis Drouot, Elena Giorgi, Mitchell Faulk, Teng Fei, Bin Guo, Zhijie Huang, Florian Johne, Monica Marinescu, Donovan McFeron, Renata Picciotto, Guillaume Remy, Henri Roesch, Kevin Smith, Xuan Wu, Hui Yu, and Jingze Zhu for our joyful time together the past few years. I especially thank Jordan Keller, Sebastien Picard, Freid Tong, and Pei-Ken Hung for welcoming me into the geometry and PDE community at Columbia and guiding me during my second year.

I appreciate the hard work of our department's administrative staff: Shatieyce Bailey, Abby Dove, Deniz Macleod, Crispina Pincus, Alenia Reynoso, and Nathan Schweer.

Long before I was a student of math, I was, and still am, a student of life. In that regard, I am eternally indebted to my wonderful family and friends for their unwavering support and affection. I am particularly grateful to my good friend Chase for the years we spent learning math together. Finally, I thank Enika for her enduring love, patience, and encouragement. You have filled my years as a graduate student with abundant joy.

# Chapter 1

# The Planarity Estimate for Pinched Solutions of the MCF

In this chapter, we will begin our investigation into singularity formation in the mean curvature flow in higher codimension. The work discussed here was completed in [38]. There is little hope to classify singularity models in general, even in codimension one. However, for suitable assumptions on initial data, this problem becomes more feasible. For hypersurfaces, assumptions like mean-convexity, convexity, and two-convexity ([31, 32, 33]) strongly restrict the types of singularity models that can appear. In higher codimension, the second fundamental form is much more complicated and preserved curvature conditions for the mean curvature flow are, so far, relatively rare. Here, we are interested in a preserved curvature pinching condition discovered by Andrews and Baker in [2]. Research in this setting thus far [2, 41] has suggested that singularity models for these pinched flows might always be codimension one. This behavior is quite interesting, as the original flow may be of arbitrarily high codimension. Colding and Minicozzi [21] have recently shown that if the asymptotic shrinker of a singularity model is a multiplicity-one cylinder, then the solution must be codimension one. The multiplicity-one assumption is difficult to verify in practice because embedded initial data need not remain embedded in higher codimension, and so higher multiplicity can occur. Here we take an alternative approach and we prove that blow-ups are codimension one directly.

The condition introduced by Andrews-Baker [2] is a curvature pinching inequality $|A|^2 < c|H|^2$, where $c$ is a constant. Importantly, they proved this condition is preserved by the mean curvature flow if $c \leq \frac{4}{3n}$. Let us call a solution of the mean curvature flow satisfying $|A|^2 < c|H|^2$ a *$c$-pinched flow*. The dimensional bound $\frac{4}{3n}$ is a technical constraint imposed by the proof that

$c$-pinching is preserved. It is more natural to consider $c = \frac{1}{n-k}$ for $k \in \{1, \dots, n-1\}$. In codimension one, the condition $|H| > 0$ is equivalent to mean-convexity and the condition $|A|^2 < \frac{1}{n-k}|H|^2$ implies $k$-convexity. Note that $\frac{1}{n-k} \leq \frac{4}{3n}$ if $n \geq 4k$. The mean curvature flow of convex and two-convex hypersurfaces has been studied in the foundational works of Huisken [31] and Huisken-Sinestrari [32, 33]. The results of [2] and [41] are extensions of these results established by Huisken and Sinestrari to higher codimension, assuming the stronger pinching condition. To contextualize our result, we must first discuss these.

In the seminal paper [31] - which draws upon Hamilton's foundational work on the Ricci flow [23] - Huisken proved the mean curvature flow evolves closed convex hypersurfaces into spherical singularities. Using the techniques developed there - in particular, the delicate Stampacchia iteration - Andrews and Baker proved that the mean curvature flow in $\mathbb{R}^N$ will deform closed $n$-dimensional initial data satisfying

$$|A|^2 < c_1|H|^2, \quad c_1 := \min\left\{\frac{4}{3n}, \frac{1}{n-1}\right\} = \begin{cases} \frac{4}{3n} & \text{if } n = 2, 3 \\ \frac{1}{n-1} & \text{if } n \geq 4 \end{cases}$$

to a point in finite time. Moreover, the flow is asymptotic to a family of shrinking round spheres contained in some $(n+1)$-dimensional affine subspace of $\mathbb{R}^N$.

Because $\frac{4}{3n} \leq \frac{1}{n-1}$ for $n \leq 4$ and $c$-pinching is only preserved for $c \leq \frac{4}{3n}$, the previous result is the best currently possible if $n \leq 4$. Suppose now we have a closed initial submanifold satisfying the weaker pinching condition studied by Nguyen [41]:

$$|A|^2 < c_2|H|^2, \quad c_2 := \min\left\{\frac{4}{3n}, \frac{1}{n-2}\right\} = \begin{cases} \frac{4}{3n} & \text{if } n = 5, 6, 7 \\ \frac{1}{n-2} & \text{if } n \geq 8 \end{cases}.$$

In [32] and [33], Huisken and Sinestrari established several very important a priori estimates for the study of flows of two-convex hypersurfaces. In particular, they establish a *convexity estimate* for mean-convex flows, which shows blow-ups must be weakly convex, and a *cylindrical estimate* for two-convex flows, which shows blow-ups must satisfy $|A|^2 \leq \frac{1}{n-1}H^2$. Also very important in their work is a *pointwise derivative estimate* for the second fundamental form. In higher codimension, we no longer have a notion of convexity, but the derivative and cylindrical estimates still make sense. In [41], Nguyen first proves a pointwise derivative estimate using the pinching condition, and then uses a blow-up argument to establish the cylindrical estimate for $c_2$-pinched flows. Specifically, his quantitative cylindrical estimate shows the following alternative: for a $c_2$-pinched flow, either the flow becomes everywhere $c_1$-pinched, hence spherical, or there are regions of the manifold which are becoming arbitrarily close to a codimension one round cylinder $S^{n-1} \times \mathbb{R}$ contained in $\mathbb{R}^N$ prior to the first singular time.

The result of [41] leaves open the possibility that a potential "cap" to a forming cylindrical singularity may not lie in the same $(n+1)$-dimensional subspace as the "neck". Presently, we will show that this does not occur and more generally that all high curvature regions prior to singularity formation must become nearly planar. Results of this type have been obtained before in the curve shortening flow. In [1], Altschuler proved singularities of the curve shortening flow in $\mathbb{R}^3$ are always planar.

Here is our setting. We suppose $n \geq 5$ and our initial data $M_0 = F_0(M) \subset \mathbb{R}^N$ is an $n$-dimensional, closed, immersed submanifold satisfying

$$|A|^2 < c_n|H|^2, \quad c_n := \min\left\{\frac{4}{3n}, \frac{3(n+1)}{2n(n+2)}\right\} = \begin{cases} \frac{3(n+1)}{2n(n+2)} & \text{if } \ n = 5, 6, \text{ or } 7 \\ \frac{4}{3n} & \text{if } \ n \geq 8 \end{cases}.$$

For $n = 5$ and $n = 6$, the constant $c_n$ is strictly between $\frac{1}{n-1}$ and $\frac{4}{3n}$. This value of $c_n$ in these dimensions is the largest we can allow in our new estimates in the proof of our main theorem below. For $n \geq 7$, we have $\frac{3(n+1)}{2n(n+2)} \geq \frac{4}{3n}$, with equality when $n = 7$. The value of $c_n$ in higher dimensions is the largest allowed by estimates used in [2] to prove that $c$-pinching is preserved. We will use these estimates as well. Under these assumptions, we consider a maximal solution $M_t = F(M,t) \subset \mathbb{R}^N$, $t \in [0,T)$, to the mean curvature flow, where $T$ is the first singular time.

For the purpose of studying planarity, we define a tensor $\hat{A}$ by

$$\hat{A}(X,Y) := A(X,Y) - \frac{\langle A(X,Y), H\rangle}{|H|^2}H, \tag{1.1}$$

for vector fields $X$ and $Y$ tangent to $M_t$. Since $c_n|H|^2 > |A|^2 \implies |H| > 0$ for $t \in [0,T)$, this tensor is well-defined along the flow. Under suitable assumptions, $\hat{A}$ vanishes identically if and only if our submanifold is a hypersurface inside an $(n+1)$-dimensional affine subspace of $\mathbb{R}^N$. See Proposition 1.7 in Section 4.1. Here is the main theorem of this chapter:

**Theorem 1.1.** *Suppose $n \geq 5$ and $N > n+1$. Let $c_n = \min\{\frac{4}{3n}, \frac{3(n+1)}{2n(n+2)}\}$. Suppose $M_t = F(M,t) \subset \mathbb{R}^N$, $t \in [0,T)$, is a smooth family of $n$-dimensional, closed, immersed submanifolds evolving by mean curvature flow, which initially satisfies $|A|^2 < c_n|H|^2$. Then there are constants $\sigma > 0$ and $C < \infty$, depending only upon the initial submanifold $M_0$, such that*

$$|\hat{A}|^2 \leq C|H|^{2-\sigma}$$

*holds pointwise on $M \times [0,T)$.*

Together with Proposition 1.7 below, this result shows that at the first singular time, blow-ups must be codimension one. We show this in Corollary 1.8. In particular, an estimate for $|\hat{A}|/|H|$ is a measure of how far our submanifold is from being planar. The result above shows that the mean curvature flow preserves near-planarity under the $c_n$-pinching assumption. Since $\frac{1}{n-2} \leq \frac{4}{3n}$ for $n \geq 8$, the $c_2$-pinching condition considered in [41] is included in the theorem above, at least when $n \geq 8$. Our result also applies for weaker pinching constants of the form $c = \frac{1}{n-k}$ if $n \geq 4k$. These weaker pinching constraints will allow a wider range of singularities models and our result shows these must also be codimension one.

Let $\tilde{c}_2 := \min\{\frac{3(n+1)}{2n(n+2)}, \frac{1}{n-2}\} = \min\{c_2, \frac{3(n+1)}{2n(n+2)}\}$. For $\tilde{c}_2$-pinched flows, the planarity estimate and Nguyen's cylindrical and derivative estimates all hold. Consequently, for $\tilde{c}_2$-pinched flows, using a recent classification result of Brendle-Choi [12, 13], it is possible to give a complete classification of singularity models at the first singular time. The pinching constant and, consequently, the classification are sharp when $n \geq 8$. We can prove the following corollary by establishing the noncollapsing of singularity models. The proof is an adaptation of the work of Huisken-Sinestrari [33]. We will discuss this in detail in Chapter 2.

**Corollary 1.2.** *Suppose $n \geq 5$ and $N > n+1$. Let $\tilde{c}_2 = \min\{\frac{3(n+1)}{2n(n+2)}, \frac{1}{n-2}\}$. Consider a family of closed $n$-dimensional submanifolds in $\mathbb{R}^N$ evolving by mean curvature flow which initially satisfy $|A|^2 < \tilde{c}_2|H|^2$. At the first singular time, the only possible blow-up limits are codimension one shrinking round spheres, shrinking round cylinders, and translating bowl solitons.*

The structure of this chapter is as follows. In Section 4.1, we record useful notation and standard identities for the higher codimension mean curvature flow. We also show that if $\hat{A}$ vanishes, then the submanifold is codimension one. In Section 1.2, we derive the evolution equation for $|\hat{A}|^2$. In Section 1.3, we prove Theorem 1.1 via the maximum principle.

There is a connection between the mean curvature flow of convex and two-convex hypersurfaces and the Ricci flow of initial data with positive isotropic curvature (PIC). Positive isotropic curvature was introduced by Micallef and Moore [37] for the study of minimal two-spheres and has been studied in the Ricci flow since Hamilton's fundamental paper [27]. If a submanifold $M \subset \mathbb{R}^N$ satisfies $|A|^2 < \frac{1}{n-2}|H|^2$, or if $M$ is a two-convex hypersurface, then the induced metric on $M$ has positive isotropic curvature. Consequently, if $M$ satisfies $|A|^2 < \frac{1}{n-1}|H|^2$, or if $M$ is a convex hypersurface, then the induced metric on $M \times \mathbb{R}$ has positive isotropic curvature (this property is called PIC1). In [6], Brendle showed that the Ricci flow evolves PIC1 initial data into round spheres. As for PIC initial data, Hamilton's breakthrough in [27] was to show that in dimension four the Ricci flow of PIC initial data only develops neck-pinch singularities. The study of PIC initial data for the Ricci flow in higher dimensions $(n \geq 12)$ has recently been addressed by Brendle in [10]. Although the results for the Ricci flow and the mean curvature flow in these contexts are very similar, it is interesting that the proofs have tended to be somewhat different.

## 1.1. Notation and Preliminaries on the MCF in Higher Codimension

In this section we will record notation, identities, and results that we will use in the proof of our theorem. We suppose we are given a solution of mean curvature flow $M_t = F(M,t) \subset \mathbb{R}^N$, where $M$ is an abstract manifold and $F : M \times [0,T) \to \mathbb{R}^N$ is a smooth familiy of parametrizations satisfying $\partial_t F = H$. Here $H$ denotes the mean curvature vector, while $g$ and $A$ denote the metric and the second fundamental form. We let $TM$ and $NM$ denote the (time-dependent) tangent and normal bundles of $M$ and recall both bundles are subbundles of $M \times \mathbb{R}^N \cong F^*T\mathbb{R}^N$.

We let $D$ denote the Euclidean derivative on $F^*T\mathbb{R}^N$ and let $\nabla_X Y = (D_X Y)^\top$ and $\nabla^\perp_X \nu = (D_X \nu)^\perp$ denote the induced connections on $TM$ and $NM$. It is possible to view the second fundamental form as either a section of $T^*M \otimes T^*M \otimes NM$ or a section of $T^*M \otimes T^*M \otimes F^*T\mathbb{R}^N$. By a minor abuse of the notation, we let $\nabla^\perp$ denote the induced connection on the former bundle and $\nabla$ denote the induced connection on the latter bundle. We can similarly view $H$ in either $NM$ or $F^*T\mathbb{R}^N$, and so we distinguish $\nabla^\perp H$ and $DH$. We will do our computations in local coordinates on $M$. For a fixed point $p_0 \in M$ and a fixed time $t_0 \in [0,T)$, we consider normal coordinates $(x_1, \dots, x_n)$ around $p$ such that $e_i = \partial_{x_i}|_{p_0}$ is an orthonormal basis of $T_{p_0}M$ and $\nabla_i \, \partial_{x_j}|_{p_0} = \nabla_{e_i} \partial_{x_j}|_{p_0} = 0$ at time $t_0$. We will use latin indices $i, j, k, \dots$ to indicate tangential components of tensors. We will not make use of the natural space-time connection for bundles over $M \times [0,T)$, although one certainly could.

We use Einstein summation notation: repeated Latin indices in multiplied tensor components will by default indicate summation from $1$ to $n$. Sometimes we will include the summation symbol to emphasize its presence. Since we work with an orthonormal basis we will raise and lower indices freely (except those of the metric tensor). For example, for $(0,2)$-tensors $T$ and $S$, we allow ourselves to write

$$T_{ik} S_{jk} = T_{ik} {S_j}^k = g^{kl} T_{ik} S_{jl} = \sum_{k=1}^{n} T_{ik} S_{jk}.$$

Since we do not use a covariant time derivative, it is important to keep track of the metric when differentiating in time. We recall that in higher codimension the evolution equations for the metric and its inverse are

$$\frac{\partial}{\partial t} g_{ij} = -2\langle A_{ij}, H\rangle, \tag{1.2}$$

$$\frac{\partial}{\partial t} g^{ij} = 2\langle A^{ij}, H\rangle = 2\langle A_{ij}, H\rangle. \tag{1.3}$$

Sometimes, such as above, we will not use indices for the components of tensors taking values in the normal bundle. We will use the inner product $\langle \cdot\,, \cdot\rangle$ to indicate summation over normal directions. Other times, we will use Greek indices $\alpha, \beta, \gamma, \dots$ to indicate normal components of tensors. In these instances, repeated Greek indices will *usually* indicate summation from $1$ to $N-n$. However, often we will only sum from $2$ to $N-n$ and in these cases, we will include a summation symbol to emphasize that. To illustrate our convention, suppose $\nu_1, \dots, \nu_{N-n}$ is local orthonormal frame for the normal bundle. Then $A_{ij\alpha} = \langle A(e_i, e_j), \nu_\alpha\rangle$ and

$$|\langle A_{ij}, A_{kl}\rangle|^2 = \left|\sum_{\alpha=1}^{N-n} A_{ij\alpha}A_{kl\alpha}\right|^2 = \langle A_{ij}, A_{kl}\rangle\langle A_{ij}, A_{kl}\rangle = \sum_{i,j,k,l=1}^{n}\sum_{\alpha,\beta=1}^{N-n} A_{ij\alpha}A_{kl\alpha}A_{ij\beta}A_{kl\beta}.$$

While the meaning of the left-hand side is perhaps a bit less obvious, it has the advantage making our computation much more succinct. For the norm of traced tensors, summation will always take place inside the norm. For example,

$$|\langle A_{ik}, A_{jk}\rangle|^2 = \left|\sum_{k=1}^{n}\langle A_{ik}, A_{jk}\rangle\right|^2 = \langle A_{ik}, A_{jk}\rangle\langle A_{il}, A_{jl}\rangle = \sum_{i,j,k,l=1}^{n}\langle A_{ik}, A_{jk}\rangle\langle A_{il}, A_{jl}\rangle.$$

We call $|H|$ the *scalar mean curvature*. Since we assume $c_n|H|^2 > |A|^2 \implies |H| > 0$, we can define

$$\nu_1 := \frac{H}{|H|} \tag{1.4}$$

to be the *principal normal direction*. Of course, $|\nu_1| = 1$. From now on, $\nu_1$ will always denote the vector defined in (1.4) and $\nu_2, \dots, \nu_{N-n}$ will denote a local frame of normal vectors orthogonal to

$\nu_1$. Having defined the principal direction, the $(0, 2)$-tensor

$$h_{ij} := \langle A_{ij}, \nu_1 \rangle = A_{ij1} \tag{1.5}$$

is the *principal component of the second fundamental form*. This is the only nonzero component of the second fundamental form if our submanifold happens to be codimension one. With this notation, we can express $\hat{A}$, defined in (1.1), as

$$\hat{A}_{ij} = \sum_{\alpha=2}^{N-n} A_{ij\alpha} \nu_\alpha. \tag{1.6}$$

Note $\hat{A}_{ij\alpha} = A_{ij\alpha}$ for $\alpha \geq 2$. Moreover, we have the identities

$$A_{ij} = \hat{A}_{ij} + h_{ij}\nu_1 = \hat{A}_{ij} + \mathring{h}_{ij}\nu_1 + \frac{1}{n}|H| g_{ij}\nu_1, \tag{1.7}$$

$$H = |H|\nu_1, \tag{1.8}$$

$$|A|^2 = |\hat{A}|^2 + |\mathring{h}|^2 + \frac{1}{n}|H|^2. \tag{1.9}$$

We will often use that $g^{ij}\hat{A}_{ij} = 0$ as well as the obvious orthogonality relations

$$\langle \hat{A}_{ij}, \nu_1 \rangle = \langle \nabla_k^\perp \nu_1, \nu_1 \rangle = 0. \tag{1.10}$$

**Remark 1.3.** Unfortunately, the notation used here is slightly different from the notation used in [2] and in [41]. For the reader who has [2] or [41] on hand, here is how to translate: the full second fundamental form, denoted by $A$ here, is denoted by $h$ in [2] and $A$ in [41]. The principal component of the second fundamental form, denoted by $h$ here, is denoted by $h_1$ in [2] and $A_1$ in [41]. Finally, the tensor $\hat{A}$ here is denoted by $h_-$ in [2] and $A_-$ in [41].

The curvature and normal curvature are denoted by $R$ and $R^\perp$ respectively, and our sign convention is that

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z - \nabla_{[Y,X]} Z,$$
$$R^\perp(X,Y)\nu = \nabla^\perp_Y \nabla^\perp_X \nu - \nabla^\perp_X \nabla^\perp_Y \nu - \nabla^\perp_{[Y,X]} \nu.$$

In higher codimension, the fundamental Gauss, Codazzi, and Ricci equations in Euclidean space in a local frame take the form

$$R_{ijkl} = \langle A_{ik}, A_{jl}\rangle - \langle A_{il}, A_{jk}\rangle, \tag{1.11}$$
$$\nabla^\perp_i A_{jk} = \nabla^\perp_j A_{ik}, \tag{1.12}$$
$$R^\perp_{ij\alpha\beta} = \sum_{k=1}^{n} (A_{ik\alpha}A_{jk\beta} - A_{jk\alpha}A_{ik\beta}). \tag{1.13}$$

We also define a vector-valued version of the normal curvature by

$$R^\perp_{ij}(\nu_\alpha) = R^\perp_{ij\alpha\beta}\nu_\beta = A_{ik\alpha}A_{jk} - A_{jk\alpha}A_{ik}.$$

In particular, we note that $R^\perp_{ij}(\nu_1) = h_{ik}A_{jk} - h_{jk}A_{ik}$, which in view of (1.7) gives

$$R^\perp_{ij}(\nu_1) = \mathring{h}_{ik}\hat{A}_{jk} - \mathring{h}_{jk}\hat{A}_{ik}. \tag{1.14}$$

For $\alpha \geq 2$, we have

$$\begin{aligned} R^\perp_{ij}(\nu_\alpha) &= \hat{A}_{ik\alpha}(\mathring{h}_{jk}\nu_1 + \hat{A}_{jk}) - \hat{A}_{jk\alpha}(\mathring{h}_{ik}\nu_1 + \hat{A}_{jk}) \\ &= -\langle R^\perp_{ij}(\nu_1), \nu_\alpha\rangle \nu_1 + \hat{A}_{ik\alpha}\hat{A}_{jk} - \hat{A}_{jk\alpha}\hat{A}_{ik}. \end{aligned}$$

To summarize, for $\alpha, \beta \in \{1, \dots, N-n\}$, we have

$$R^{\perp}_{ij\alpha\beta} = \begin{cases} 0 & \alpha = \beta \\ \mathring{h}_{ik}\hat{A}_{jk\beta} - \mathring{h}_{jk}\hat{A}_{ik\beta} & \alpha = 1,\ \beta \geq 2 \\ \mathring{h}_{jk}\hat{A}_{ik\alpha} - \mathring{h}_{ik}\hat{A}_{jk\alpha} & \alpha \geq 2,\ \beta = 1 \\ \hat{A}_{ik\alpha}\hat{A}_{jk\beta} - \hat{A}_{jk\alpha}\hat{A}_{ik\beta} & \alpha,\ \beta \geq 2 \end{cases}. \tag{1.15}$$

We define a new connection for the orthogonal decomposition of $NM = E_1 \oplus \hat{E}$ where $\hat{E}$ consists of normal vectors $\hat{\nu}$ which are everywhere orthogonal to $\nu_1$, $\langle \hat{\nu}, \nu_1 \rangle = 0$, and $E_1 = C^{\infty}(M)\nu_1$. Define $\hat{\nabla}^{\perp}$ on $\hat{E}$ by

$$\hat{\nabla}^{\perp}_i \hat{\nu} := \nabla^{\perp}_i \hat{\nu} - \langle \nabla^{\perp}_i \hat{\nu}, \nu_1 \rangle \nu_1. \tag{1.16}$$

Since, by definition, $\hat{A}$ is a section of $T^*M \otimes T^*M \otimes \hat{E}$, it is natural to define the connection $\hat{\nabla}^{\perp}$ on $\hat{A}$ by

$$\hat{\nabla}^{\perp}_i \hat{A}_{jk} := \nabla^{\perp}_i \hat{A}_{jk} - \langle \nabla^{\perp}_i \hat{A}_{jk}, \nu_1 \rangle \nu_1. \tag{1.17}$$

For clarity, let us summarize the relationship between the derivatives $D, \nabla, \nabla^{\perp}$, and $\hat{\nabla}^{\perp}$ acting on $A, H$, and $\hat{A}$. If we view $A_{jk}$ and $H$ as taking values in $F^*T\mathbb{R}^N$, then we can decompose $\nabla_i A_{jk}$ and $D_i H$ into the tangential and normal components to get

$$\begin{aligned} \nabla_i A_{jk} &= (\nabla_i A_{jk})^{\perp} + (\nabla_i A_{jk})^{\top} \\ &= \nabla^{\perp}_i A_{jk} - \sum_{l=1}^{n} \langle A_{jk}, A_{il} \rangle e_l, \end{aligned} \tag{1.18}$$

$$\begin{aligned} D_i H &= (D_i H)^{\perp} + (D_i H)^{\top} \\ &= \nabla^{\perp}_i H - \sum_{j=1}^{n} \langle H, A_{ij} \rangle e_j. \end{aligned} \tag{1.19}$$

Similarly, and more relevant to the coming computations, if we view $A_{jk}$, $H$, and $\hat{A}_{jk}$ as taking values in the normal bundle, then we can decompose $\nabla_i^\perp A_{jk}$, $\nabla_i^\perp H$, and $\nabla_i^\perp \hat{A}_{jk}$ using $NM = E_1 \oplus \hat{E}$ to get

$$\nabla_i^\perp A_{jk} = \big(\hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp \nu_1\big) + \big(\langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle + \nabla_i h_{jk}\big)\nu_1, \tag{1.20}$$

$$\nabla_i^\perp H = |H|\nabla_i^\perp \nu_1 + (\nabla_i |H|)\nu_1, \tag{1.21}$$

$$\nabla_i^\perp \hat{A}_{jk} = \hat{\nabla}_i^\perp \hat{A}_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle \nu_1. \tag{1.22}$$

Note that $\nabla_i^\perp A_{jk}$ is shorthand for $(\nabla_i^\perp A_{jk\alpha})\nu_\alpha$. As a direct consequence of the orthogonality relations, we have

**Proposition 1.4** (Decomposition of derivatives)**.**

$$|\nabla^\perp A|^2 = |\hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp \nu_1|^2 + |\langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle + \nabla_i h_{jk}|^2. \tag{1.23}$$

$$|\nabla^\perp H|^2 = |H|^2 |\nabla^\perp \nu_1|^2 + |\nabla |H||^2. \tag{1.24}$$

$$|\nabla^\perp \hat{A}|^2 = |\hat{\nabla}^\perp \hat{A}|^2 + |\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2. \tag{1.25}$$

We will use these identities in Sections 1.2 and 1.3.

It is very useful to consider the implications of the Codazzi equation for the decomposition of $\nabla_i^\perp A_{jk}$ above. Projecting the Codazzi equation (1.12) onto $E_1$ and $\hat{E}$ implies the both of the tensors

$$\nabla_i h_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle \quad \text{and} \quad \hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp \nu_1$$

are symmetric in $i, j, k$. Consequently, it is equivalent to trace over $j, k$ or trace over $i, k$, and this implies

$$\sum_{k=1}^n \nabla_k h_{ik} + \langle \nabla_k^\perp \hat{A}_{ik}, \nu_1\rangle = \nabla_i |H|, \tag{1.26}$$

$$\sum_{k=1}^n \hat{\nabla}_k^\perp \hat{A}_{ik} + h_{ik}\nabla_k^\perp \nu_1 = |H|\nabla_i^\perp \nu_1. \tag{1.27}$$

Next, we review the evolution equations for $A$ and $H$ in higher codimension. We let $\frac{\partial}{\partial t}^{\perp}$ denote the projection of the time derivative onto the normal bundle and $\Delta^{\perp}$ denote the Laplacian with respect to the connection $\nabla^{\perp}$.

**Proposition 1.5** (Evolution of $A$ and $H$)**.** *With the summation convention, the evolution equations of $A_{ij}$ and $H$ are*

$$\Big(\frac{\partial}{\partial t}^{\perp} - \Delta^{\perp}\Big)A_{ij} = -\langle H, A_{ik}\rangle A_{jk} - \langle H, A_{jk}\rangle A_{ik} + \langle A_{ij}, A_{kl}\rangle A_{kl} \tag{1.28}$$
$$- 2\langle A_{ik}, A_{jl}\rangle A_{kl} + \langle A_{ik}, A_{kl}\rangle A_{jl} + \langle A_{jl}, A_{kl}\rangle A_{ik},$$
$$\Big(\frac{\partial}{\partial t}^{\perp} - \Delta^{\perp}\Big)H = \langle H, A_{kl}\rangle A_{kl}. \tag{1.29}$$

*The evolution equations of $|A|^2$ and $|H|^2$ are*

$$\frac{\partial}{\partial t}|A|^2 = \Delta|A|^2 - 2|\nabla^{\perp}A|^2 + 2|\langle A_{ij}, A_{kl}\rangle|^2 + 2|R^{\perp}|^2, \tag{1.30}$$
$$\frac{\partial}{\partial t}|H|^2 = \Delta|H|^2 - 2|\nabla^{\perp}H|^2 + 2|\langle A_{ij}, H\rangle|^2. \tag{1.31}$$

*Proof.* The equations above have been derived in both [46] and [2], but under slightly different notation and conventions (see also Remark 1.3). For the convenience of the reader, we will give a derivation here.

The first step is to prove the important Simons' identity, by commuting second derivatives of the second fundamental form. To that end, by standard commutation identities we have

$$\nabla_i^{\perp}\nabla_k^{\perp}A_{jl\alpha} = \nabla_k^{\perp}\nabla_i^{\perp}A_{jl\alpha} + R_{ikjp}A_{pl\alpha} + R_{iklp}A_{jp\alpha} + R^{\perp}_{ik\alpha\beta}A_{jl\beta}.$$

Plugging in the Gauss equation (1.11) and Ricci equation (1.13), we obtain

$$\nabla_i^{\perp}\nabla_k^{\perp}A_{jl\alpha} = \nabla_k^{\perp}\nabla_i^{\perp}A_{jl\alpha} + \big(\langle A_{ij}, A_{kp}\rangle - \langle A_{ip}, A_{jk}\rangle\big)A_{pl\alpha}$$
$$+ \big(\langle A_{il}, A_{kp}\rangle - \langle A_{ip}, A_{kl}\rangle\big)A_{jp\alpha} + \big(A_{ip\alpha}A_{kp\beta} - A_{ip\beta}A_{kp\alpha}\big)A_{jl\beta}.$$

Note that $(A_{ip\alpha}A_{kp\beta} - A_{ip\beta}A_{kp\alpha})A_{jl\beta} = \langle A_{kp}, A_{jl}\rangle A_{ip\alpha} - \langle A_{ip}, A_{jl}\rangle A_{kp\alpha}$, so we may write

$$\begin{aligned}\nabla_i^\perp \nabla_k^\perp A_{jl} = \nabla_k^\perp \nabla_i^\perp A_{jl} &+ \big(\langle A_{ij}, A_{kp}\rangle - \langle A_{ip}, A_{jk}\rangle\big)A_{pl} \\ &+ \big(\langle A_{il}, A_{kp}\rangle - \langle A_{ip}, A_{kl}\rangle\big)A_{jp} + \langle A_{kp}, A_{jl}\rangle A_{ip} - \langle A_{ip}, A_{jl}\rangle A_{kp}.\end{aligned}$$

Using the Codazzi equation (1.12) twice gives

$$\begin{aligned}\nabla_i^\perp \nabla_j^\perp A_{kl} &= \nabla_i^\perp \nabla_k^\perp A_{jl} \\ &= \nabla_k^\perp \nabla_l^\perp A_{ij} + \big(\langle A_{ij}, A_{kp}\rangle - \langle A_{ip}, A_{jk}\rangle\big)A_{pl} \\ &\quad + \big(\langle A_{il}, A_{kp}\rangle - \langle A_{ip}, A_{kl}\rangle\big)A_{jp} + \langle A_{kp}, A_{jl}\rangle A_{ip} - \langle A_{ip}, A_{jl}\rangle A_{kp}.\end{aligned}$$

Finally, if we trace over the indices $k$ and $l$, gather terms, and rearrange a bit, we obtain

$$\begin{aligned}\nabla_i^\perp \nabla_j^\perp H = \Delta^\perp A_{ij} &- \langle H, A_{ip}\rangle A_{jp} + \langle A_{ij}, A_{pq}\rangle A_{pq} \\ &- 2\langle A_{ip}, A_{jq}\rangle A_{pq} + \langle A_{iq}, A_{pq}\rangle A_{jp} + \langle A_{jq}, A_{pq}\rangle A_{ip}.\end{aligned} \tag{1.32}$$

Next, we compute $\partial_t^\perp A_{ij}$. To do so, let us denote $F_i = dF(\partial_{x_i})$ and recall that $\partial_t F = H$. In our notation we have $\partial_t F_i = D_i H$ and $[\partial_t, D_i] = 0$. Then

$$\frac{\partial}{\partial t}^\perp A_{ij} = \frac{\partial}{\partial t}^\perp (D_i F_j)^\perp = \left(\frac{\partial}{\partial t}\Big(D_i F_j - g^{pq}\langle D_i F_j, F_p\rangle F_q\Big)\right)^\perp = \big(D_i D_j H\big)^\perp,$$

where we have used that $D_i F_j = 0$ at the origin in our coordinates. Now by (1.19) we have $D_j H = \nabla_j^\perp H - \langle H, A_{jp}\rangle F_p$. Differentiating and taking the normal projection, we obtain

$$\frac{\partial}{\partial t}^\perp A_{ij} = \nabla_i^\perp \nabla_j^\perp H - \langle H, A_{jp}\rangle A_{ip}. \tag{1.33}$$

Combining (1.32) and (1.33) proves the claimed evolution equation for $A$. The evolution equation for $H$ is a straightforward consequence of the evolution equation for $A_{ij}$ and the identity

$$\Big(\frac{\partial}{\partial t}^{\perp} - \Delta^{\perp}\Big)H = \Big(\frac{\partial}{\partial t}g^{ij}\Big)H + g^{ij}\Big(\frac{\partial}{\partial t}^{\perp} - \Delta^{\perp}\Big)A_{ij},$$

in view of (1.3).

To obtain the evolution equation for $|A|^2 = g^{ij}g^{kl}\langle A_{ij}, A_{kl}\rangle$, we use (1.3) once more to get

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)|A|^2 + 2|\nabla^{\perp}A|^2 = 2\Big\langle\Big(\frac{\partial}{\partial t}^{\perp} - \Delta^{\perp}\Big)A_{ij}, A_{ij}\Big\rangle + 8\langle H, A_{ij}\rangle\langle H, A_{ij}\rangle.$$

Plugging in the evolution equation for $A_{ij}$, the terms of the form $|\langle H, A\rangle|^2$ cancel. Relabeling and organizing the indices a bit, we are left with

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)|A|^2 + 2|\nabla^{\perp}A|^2 = 2|\langle A_{ij}, A_{kl}\rangle|^2 + 4\langle A_{ik}, A_{kl}\rangle\langle A_{ij}, A_{jl}\rangle - 4\langle A_{ik}, A_{jl}\rangle\langle A_{ij}, A_{kl}\rangle.$$

By the Ricci equation (1.13), we have

$$\begin{aligned}
|R^{\perp}|^2 &= R^{\perp}_{ij\alpha\beta}R^{\perp}_{ij\alpha\beta} \\
&= \big(A_{ik\alpha}A_{jk\beta} - A_{jk\alpha}A_{ik\beta}\big)\big(A_{il\alpha}A_{jl\beta} - A_{jl\alpha}A_{il\beta}\big) \\
&= \langle A_{ik}, A_{il}\rangle\langle A_{jk}, A_{jl}\rangle + \langle A_{jk}, A_{jl}\rangle\langle A_{ik}, A_{il}\rangle \\
&\quad - \langle A_{jk}, A_{il}\rangle\langle A_{ik}, A_{jl}\rangle - \langle A_{ik}, A_{jl}\rangle\langle A_{jk}, A_{il}\rangle.
\end{aligned}$$

After another relabeling of indices, we see that

$$2|R^{\perp}|^2 = 4\langle A_{ik}, A_{kl}\rangle\langle A_{ij}, A_{jl}\rangle - 4\langle A_{ik}, A_{jl}\rangle\langle A_{ij}, A_{kl}\rangle,$$

which establishes the evolution equation for $|A|^2$. Finally, the evolution equation of $|H|^2$ follows easily from the evolution equation of $H$. □

It will be useful to expand each of the reaction terms

$$|\langle A_{ij}, H\rangle|^2 = \langle A_{ij}, H\rangle\langle A_{ij}, H\rangle,$$
$$|\langle A_{ij}, A_{kl}\rangle|^2 = \langle A_{ij}, A_{kl}\rangle\langle A_{ij}, A_{kl}\rangle,$$
$$|R^{\perp}_{ij}|^2 = R^{\perp}_{ij\alpha\beta}R^{\perp}_{ij\alpha\beta},$$

on the right-hand side of the evolution equations (1.30) and (1.31), using the formula (1.7). It is straightforward to see that

$$\langle A_{ij}, H\rangle = |H|h_{ij} = \frac{1}{n}|H|^2 g_{ij} + |H|\mathring{h}_{ij}. \tag{1.34}$$

Similarly,

$$\begin{aligned}\langle A_{ij}, A_{kl}\rangle &= h_{ij}h_{kl} + \langle \hat{A}_{ij}, \hat{A}_{kl}\rangle \\ &= \frac{1}{n^2}|H|^2 g_{ij}g_{kl} + \frac{1}{n}|H|(g_{ij}\mathring{h}_{kl} + \mathring{h}_{ij}g_{kl}) + \mathring{h}_{ij}\mathring{h}_{kl} + \langle \hat{A}_{ij}, \hat{A}_{kl}\rangle.\end{aligned} \tag{1.35}$$

As for the third reaction term, recalling (1.15), we have

$$\begin{aligned}|R^{\perp}|^2 &= \sum_{\beta=1}^{N-n} R^{\perp}_{ij1\beta}R^{\perp}_{ij1\beta} + \sum_{\alpha=1}^{N-n} R^{\perp}_{ij\alpha1}R^{\perp}_{ij\alpha1} + \sum_{\alpha,\beta=2}^{N-n} R^{\perp}_{ij\alpha\beta}R^{\perp}_{ij\alpha\beta} \\ &= 2|R^{\perp}_{ij}(\nu_1)|^2 + \sum_{\alpha,\beta=2}^{N-n}\Big|(\hat{A}_{ik\alpha}\hat{A}_{jk\beta} - \hat{A}_{jk\alpha}\hat{A}_{ik\beta})\Big|^2.\end{aligned} \tag{1.36}$$

Using (1.14), (1.34), (1.35), (1.36), we obtain the following proposition for use in later sections.

**Proposition 1.6** (Decomposition of reaction terms)**.**

$$|h|^2 = |\mathring{h}|^2 + \frac{1}{n}|H|^2. \tag{1.37}$$

$$|\langle A_{ij}, H\rangle|^2 = |H|^2|h|^2. \tag{1.38}$$

$$|\langle A_{ij}, A_{kl}\rangle|^2 = |h|^4 + 2\Big|\sum_{i,j=1}^{n} \mathring{h}_{ij}\hat{A}_{ij}\Big|^2 + |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2. \tag{1.39}$$

$$|R^\perp|^2 = |\hat{R}^\perp|^2 + 2|R^\perp_{ij}(\nu_1)|^2, \tag{1.40}$$

*where*

$$|R^\perp_{ij}(\nu_1)|^2 = \Big|\sum_{k=1}^{n}(\mathring{h}_{ik}\hat{A}_{jk} - \mathring{h}_{jk}\hat{A}_{ik})\Big|^2, \tag{1.41}$$

$$|\hat{R}^\perp|^2 := \sum_{\alpha,\beta=2}^{N-n}\Big|\sum_{k=1}^{n}(\hat{A}_{ik\alpha}\hat{A}_{jk\beta} - \hat{A}_{jk\alpha}\hat{A}_{ik\beta})\Big|^2. \tag{1.42}$$

The expression $|\hat{R}^\perp|^2$ will appear in our evolution equation for $|\hat{A}|^2$. For brevity, from now on we will write

$$|\mathring{h}_{ij}\hat{A}_{ij}|^2 = \Big|\sum_{i,j=1}^{n}\mathring{h}_{ij}\hat{A}_{ij}\Big|^2.$$

We now give a proof that the vanishing of $\hat{A}$ implies codimension one. In application, if the initial flow is $c$-pinched, then the blow up will satisfy $|A|^2 \le c|H|^2$. The argument below works as long as the tensor $|H|g_{ij} - h_{ij}$ is positive definite (which is equivalent to $h_{ij}$ be $(n-1)$-convex).

**Proposition 1.7.** *Let $n \ge 2$ and $N > n+1$. Suppose $F : M \to \mathbb{R}^N$ is an immersion of a connected, $n$-dimensional manifold satisfying $|H| > 0$. Assume $|H|g_{ij} - h_{ij}$ is positive definite and $\hat{A} \equiv 0$. Then $F(M)$ is an immersed hypersurface in an $(n+1)$-dimensional affine subspace of $\mathbb{R}^N$.*

*Proof.* Because $|H| > 0$, the principal normal $\nu_1$ is well-defined. The vanishing of $\hat{A}$ in addition to our pinching assumption implies $\nu_1$ is parallel with respect to $\nabla^\perp$. Specifically, recall that by projecting the Codazzi equation, we have (1.27):

$$|H|\nabla_i^{\perp}\nu_1 = \sum_{k=1}^{n} \hat{\nabla}_k^{\perp}\hat{A}_{ik} + h_{ik}\nabla_k^{\perp}\nu_1 = h_{ik}\nabla_k^{\perp}\nu_1.$$

Since the tensor $|H|g_{ik} - h_{ik}$ is positive definite, we must have $\nabla^{\perp}\nu_1 = 0$.

Now let $\gamma : [0,1] \to M$ be a smooth path connecting any pair of distinct points $p_0 = \gamma(0)$ and $p_1 = \gamma(1)$ in $M$. Define $\nu_2(0), \dots, \nu_{N-n}(0)$ to be the completion of $\nu_1(p_0)$ to an orthonormal basis of $N_{p_0}M$. For $\beta \in \{2, \dots, N-n\}$ and $s \in [0,1]$, let $\nu_\beta(s) \in N_{\gamma(s)}M$ be the parallel transport of $\nu_\beta(0)$ with respect to $\nabla^{\perp}$. Because $\nu_1$ is parallel with respect to $\nabla^{\perp}$ and $\langle \nu_\beta(0), \nu_1(p_0)\rangle = 0$, we have $\big\langle \nu_\beta(s), \nu_1\big(\gamma(s)\big)\big\rangle = 0$ for all $s \in [0,1]$. If we let $e_1, \dots, e_n$ denote a parallel orthonormal basis of $T_{\gamma(s)}M$ along $\gamma$, then

$$(D_{\gamma'(s)}\nu_\beta(s))^{\top} = \sum_{i=1}^{n} \langle D_{\gamma'(s)}\nu_\beta(s), e_i\rangle e_i = -\sum_{i=1}^{n} \langle \nu_\beta(s), A(\gamma'(s), e_i)\rangle e_i = 0,$$

since $\nu_\beta(s)$ is orthogonal to $A = h\nu_1$. It follows that

$$D_{\gamma'(s)}\nu_\beta(s) = \nabla^{\perp}_{\gamma'(s)}\nu_\beta(s) + (D_{\gamma'(s)}\nu_\beta(s))^{\top} = 0,$$

which shows $\nu_\beta(s)$ is parallel along $\gamma$ with respect to the ambient connection $D$ as well. On the other hand, the constant unit vector field $\omega_\beta$ in $\mathbb{R}^N$ defined by the condition $\omega_\beta(F(p_0)) = \nu_\beta(0)$, is also parallel along $F(\gamma(s))$ with respect to $D$. By uniqueness of parallel transport, this implies $\nu_\beta(1)$ agrees with the restriction of $\omega_\beta$ to $F(M)$. Since $p_1$ was arbitrary, we see that the restriction of the vector fields $\omega_2, \dots, \omega_{N-n}$ form a parallel orthonormal basis of the complement of $\nu_1$ in $NM$ at every point on $M$. This implies the ambient coordinate functions $y_\beta : \mathbb{R}^N \to \mathbb{R}$ given by $y_\beta(x) = \langle x, \omega_\beta\rangle$ are constant on $F(M)$. It follows that $F(M)$ must lie in a translation of the $(n+1)$-dimensional subspace of $\mathbb{R}^N$ orthogonal to $\omega_2, \dots, \omega_{N-n}$. $\square$

Finally, as the main application of Theorem 1.1, we deduce that singularity models must be codimension one.

**Corollary 1.8.** *Suppose $n \geq 5$ and $N > n+1$. Let $c_n = \min\{\frac{4}{3n}, \frac{3(n+1)}{2n(n+2)}\}$. Suppose $M_t \subset \mathbb{R}^N$, $t \in [0,T)$, is a smooth family of $n$-dimensional closed submanifolds evolving by mean curvature flow which initially satisfies $|A|^2 < c_n|H|^2$. Then at the first singular time every blow-up limit must be codimension one.*

*Proof.* Since we have strict inequality $|A|^2 < c_n|H|^2$, we can find $\delta > 0$, depending on $M_0$, such that $f := (c_n - \delta)|H|^2 - |A|^2$ satisfies $f \geq \delta|H|^2$ on $M_0$. Then, by Theorem 2 in [2] (note Remark 1.3), $f \geq \delta|H|^2$ for $t \in [0,T)$. Now, suppose $\tilde{M}_t$ is a smooth blow-up limit. On the blow-up limit, we must also have $\tilde{f} := (c_n - \delta)|\tilde{H}|^2 - |\tilde{A}|^2 \geq \delta|\tilde{H}|^2 \geq 0$. Since $c_n - \delta < \frac{4}{3n}$, it follows by work in Section 3 of [2] that $(\partial_t - \Delta)\tilde{f} \geq 0$. En route to proving Theorem 1.1, we will also establish this inequality in Lemma 1.15 below. Hence, by the strong maximum principle, either $\tilde{f} \equiv 0$ or $\tilde{f} > 0$ on the blow-up limit. If $\tilde{f} \equiv 0$, then $\tilde{H} \equiv 0$, and consequently $\tilde{A} \equiv 0$. In this case, the blow-up must be a codimension one hyperplane. Otherwise, $\tilde{f} > 0$, and thus $|\tilde{H}| > 0$ everywhere on the blow-up. Theorem 1.1 implies that $|A|^2|H|^2 - |\langle A, H\rangle|^2 \leq C|H|^{4-\sigma}$. In the limit, we deduce $\hat{\tilde{A}} = 0$. Since $|\tilde{A}|^2 \leq c_n|\tilde{H}|^2$ easily implies $|\tilde{H}|\tilde{g}_{ij} - \tilde{h}_{ij}$ is positive definite, Proposition 1.7 implies the blow-up limit must be codimension one. □

## 1.2. Evolution of $|\hat{A}|^2$

In this section, we compute the evolution equation of $|\hat{A}|^2$. We do this by using the formulas stated in Section 4.1. To begin, we recall the useful standard identity

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)\Big(\frac{u}{v}\Big) = \frac{1}{v}\Big(\frac{\partial}{\partial t} - \Delta\Big)u - \frac{u}{v^2}\Big(\frac{\partial}{\partial t} - \Delta\Big)v + \frac{2}{v}\nabla_k v\, \nabla_k\Big(\frac{u}{v}\Big). \tag{1.43}$$

Now it follows from (1.1) that

$$|\hat{A}|^2 = |A|^2 - |\langle A_{ij}, H\rangle|^2|H|^{-2}.$$

So we will need the evolution equations of $|A|^2$ and $|\langle A_{ij}, H\rangle|^2|H|^{-2}$. We have already recorded, in (1.30) and (1.31), the evolution equations of $|A|^2$ and $|H|^2$. The latter of these equations combined with (1.43) (set $u = |\langle A_{ij}, H\rangle|^2$ and $v = |H|^2$) implies

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)\frac{|\langle A_{ij}, H\rangle|^2}{|H|^2} = |H|^{-2}\Big(\frac{\partial}{\partial t} - \Delta\Big)|\langle A_{ij}, H\rangle|^2 \\ - |H|^{-4}|\langle A_{ij}, H\rangle|^2\big( - 2|\nabla^\perp H|^2 + 2|\langle A_{kl}, H\rangle|^2\big) \\ + 2|H|^{-2}\Big\langle \nabla_k |H|^2, \nabla_k \frac{|\langle A_{ij}, H\rangle|^2}{|H|^2}\Big\rangle. \tag{1.44}$$

Before computing the evolution of $|\langle A_{ij}, H\rangle|^2$, we simplify the terms on the second and third lines using Propositions 1.4 and 1.6. In particular, using $|\langle A_{ij}, H\rangle|^2 = |H|^2|h|^2$ and (1.24), we rewrite the two terms on the second line of (1.44) as

$$2|H|^{-4}|\langle A_{ij}, H\rangle|^2|\nabla^\perp H|^2 = 2|h|^2|\nabla^\perp \nu_1|^2 + 2|H|^{-2}|h|^2|\nabla |H||^2, \tag{1.45}$$

$$-2|H|^{-4}|\langle A_{ij}, H\rangle|^4 = -2|h|^4. \tag{1.46}$$

As for the gradient term on the third line of (1.44), we compute that $\nabla_k|H|^2 = 2|H|\nabla_k|H|$ and $\nabla_k(|H|^{-2}|\langle A_{ij}, H\rangle|^2) = \nabla_k|h|^2 = 2h_{ij}\nabla_k h_{ij}$. Therefore,

$$2|H|^{-2}\Big\langle \nabla_k |H|^2, \nabla_k \frac{|\langle A_{ij}, H\rangle|^2}{|H|^2}\Big\rangle = 8|H|^{-1}h_{ij}\nabla_k|H|\nabla_k h_{ij}. \tag{1.47}$$

To summarize (1.45), (1.46), and (1.47), we have shown so far that

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)\frac{|\langle A_{ij}, H\rangle|^2}{|H|^2} = |H|^{-2}\Big(\frac{\partial}{\partial t} - \Delta\Big)|\langle A_{ij}, H\rangle|^2 \\ - 2|h|^4 + 2|h|^2|\nabla^\perp \nu_1|^2 + 2|H|^{-2}|h|^2|\nabla|H||^2 \\ + 8|H|^{-1}h_{ij}\nabla_k|H|\nabla_k h_{ij}. \tag{1.48}$$

For the evolution of $|\langle A_{ij}, H\rangle|^2$, we have the following lemma.

**Lemma 1.9.**

$$
\begin{aligned}
&|H|^{-2}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\langle A_{ij},H\rangle|^2\\
&=4|\mathring{h}_{ij}\hat{A}_{ij}|^2+2|R^{\perp}_{ij}(\nu_1)|^2+4|h|^4 \qquad (1.49)\\
&\quad-4|H|^{-1}\mathring{h}_{ij}\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle-4\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle\\
&\quad-4|h|^2|\nabla^{\perp}\nu_1|^2-2|H|^{-2}|h|^2|\nabla|H||^2-8|H|^{-1}h_{ij}\nabla_k|H|\nabla_k h_{ij}-2|\nabla h|^2.
\end{aligned}
$$

*Proof.* Recall that any time $h$ is traced with $\hat{A}$, we may replace $h$ with $\mathring{h}$ because $\hat{A}$ is traceless. To begin, we use (1.28) and (1.29) to obtain

$$
\begin{aligned}
\Big\langle\Big(\frac{\partial}{\partial t}^{\perp}-\Delta^{\perp}\Big)A_{ij},H\Big\rangle&=-\langle H,A_{ik}\rangle\langle A_{jk},H\rangle-\langle H,A_{jk}\rangle\langle A_{ik},H\rangle+\langle A_{ij},A_{kl}\rangle\langle A_{kl},H\rangle\\
&\quad-2\langle A_{ik},A_{jl}\rangle\langle A_{kl},H\rangle+\langle A_{ik},A_{kl}\rangle\langle A_{jl},H\rangle+\langle A_{jl},A_{kl}\rangle\langle A_{ik},H\rangle,\\
\Big\langle A_{ij},\Big(\frac{\partial}{\partial t}^{\perp}-\Delta^{\perp}\Big)H\Big\rangle&=\langle A_{kl},H\rangle\langle A_{ij},A_{kl}\rangle.
\end{aligned}
$$

Tracing each of these equations with a copy of $\langle A_{ij},H\rangle$, we get

$$
\begin{aligned}
\Big\langle\Big(\frac{\partial}{\partial t}^{\perp}-\Delta^{\perp}\Big)A_{ij},H\rangle\langle A_{ij},H\rangle&=-2\langle A_{ik},H\rangle\langle A_{jk},H\rangle\langle A_{ij},H\rangle+\langle A_{ij},A_{kl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle\\
&\quad-2\langle A_{ik},A_{jl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle+2\langle A_{ik},A_{kl}\rangle\langle A_{jl},H\rangle\langle A_{ij},H\rangle,\\
\Big\langle A_{ij},\Big(\frac{\partial}{\partial t}^{\perp}-\Delta^{\perp}\Big)H\Big\rangle\langle A_{ij},H\rangle&=\langle A_{ij},A_{kl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle.
\end{aligned}
$$

Combining these latter formulas together gives

$$
\begin{aligned}
\Big(\Big(\frac{\partial}{\partial t}-\Delta\Big)\langle A_{ij},H\rangle\Big)\langle A_{ij},H\rangle&=-2\langle A_{ik},H\rangle\langle A_{jk},H\rangle\langle A_{ij},H\rangle+2\langle A_{ij},A_{kl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle\\
&\quad-2\langle A_{ik},A_{jl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle+2\langle A_{ik},A_{kl}\rangle\langle A_{jl},H\rangle\langle A_{ij},H\rangle\\
&\quad-2\langle\nabla^{\perp}_k A_{ij},\nabla^{\perp}_k H\rangle\langle A_{ij},H\rangle.
\end{aligned}
$$

Therefore, recalling (1.3), we have

$$
\begin{aligned}
&\Big(\frac{\partial}{\partial t} - \Delta\Big)|\langle A_{ij}, H\rangle|^2 \\
&= 2\Big(\frac{\partial}{\partial t} g^{ij}\Big) g^{kl}\langle A_{ik}, H\rangle\langle A_{jl}, H\rangle + 2\Big(\Big(\frac{\partial}{\partial t} - \Delta\Big)\langle A_{ij}, H\rangle\Big)\langle A_{ij}, H\rangle - 2|\nabla\langle A_{ij}, H\rangle|^2 \\
&= 4\langle A_{ij}, A_{kl}\rangle\langle A_{kl}, H\rangle\langle A_{ij}, H\rangle \qquad (1.50)\\
&\quad - 4\langle A_{ik}, A_{jl}\rangle\langle A_{kl}, H\rangle\langle A_{ij}, H\rangle + 4\langle A_{ik}, A_{kl}\rangle\langle A_{jl}, H\rangle\langle A_{ij}, H\rangle \\
&\quad - 4\langle \nabla_k^\perp A_{ij}, \nabla_k^\perp H\rangle\langle A_{ij}, H\rangle - 2|\nabla\langle A_{ij}, H\rangle|^2.
\end{aligned}
$$

To finish the proof, we multiply by $|H|^{-2}$ and then rewrite each of the remaining terms using $A = \hat{A} + h\nu$. For the term in the first line of (1.50), we have

$$
\begin{aligned}
4|H|^{-2}\langle A_{ij}, A_{kl}\rangle\langle A_{kl}, H\rangle\langle A_{ij}, H\rangle &= 4|H|^{-2}|H|^2 h_{ij}h_{kl}\langle A_{ij}, A_{kl}\rangle \\
&= 4|h|^4 + 4h_{ij}h_{kl}\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle \\
&= 4|h|^4 + 4\mathring{h}_{ij}\mathring{h}_{kl}\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle \\
&= 4|h|^4 + 4|\mathring{h}_{ij}\hat{A}_{ij}|^2. \qquad (1.51)
\end{aligned}
$$

For the terms in the second line of (1.50), recalling (1.41), we have

$$
\begin{aligned}
|R_{ij}^\perp(\nu_1)|^2 &= |\mathring{h}_{ik}\hat{A}_{jk} - \mathring{h}_{jk}\hat{A}_{ik}|^2 \\
&= |h_{ik}A_{jk} - h_{jk}A_{ik}|^2 \\
&= \langle h_{ik}A_{jk} - h_{jk}A_{ik}, h_{il}A_{jl} - h_{jl}A_{il}\rangle \\
&= 2h_{ik}h_{il}\langle A_{jk}, A_{jl}\rangle - 2h_{ik}h_{jl}\langle A_{jk}, A_{il}\rangle \\
&= 2\langle A_{jk}, A_{jl}\rangle\langle A_{ik}, \frac{H}{|H|}\rangle\langle A_{il}, \frac{H}{|H|}\rangle - 2\langle A_{jk}, A_{il}\rangle\langle A_{ik}, \frac{H}{|H|}\rangle\langle A_{jl}, \frac{H}{|H|}\rangle.
\end{aligned}
$$

After reindexing (e.g. $j \to k \to l \to i \to j$ on the first term and $j \to i \to l \to j$, $k \to k$ on the second term), this gives

$$4|H|^{-2}\big(\langle A_{ik},A_{kl}\rangle\langle A_{jl},H\rangle\langle A_{ij},H\rangle-\langle A_{ik},A_{jl}\rangle\langle A_{kl},H\rangle\langle A_{ij},H\rangle\big)=2|R^{\perp}_{ij}(\nu_1)|^2. \tag{1.52}$$

Together (1.51) and (1.52) verify the claimed reaction terms on the first line (1.49).

It remains to analyze terms in third line of (1.50). By (1.20) and (1.21)

$$\begin{aligned}\langle\nabla^{\perp}_k A_{ij},\nabla^{\perp}_k H\rangle &= |H|\langle\nabla^{\perp}_k A_{ij},\nabla^{\perp}_k\nu_1\rangle+\nabla_k|H|\langle\nabla^{\perp}_k A_{ij},\nu_1\rangle\\ &= |H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle+|H|h_{ij}|\nabla^{\perp}\nu_1|^2+\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle+\nabla_k|H|\nabla_k h_{ij}.\end{aligned}$$

Consequently, the first term on the third line of (1.50) is

$$\begin{aligned}-4|H|^{-2}\langle\nabla^{\perp}_k A_{ij},\nabla^{\perp}_k H\rangle\langle A_{ij},H\rangle &= -4|H|^{-1}h_{ij}\langle\nabla^{\perp}_k A_{ij},\nabla^{\perp}_k H\rangle\\ &= -4|H|^{-1}h_{ij}\Big(|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle+|H|h_{ij}|\nabla^{\perp}\nu_1|^2\\ &\qquad+\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle+\nabla_k|H|\nabla_k h_{ij}\Big)\\ &= -4|H|^{-1}\mathring{h}_{ij}\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle-4|H|^{-1}h_{ij}\nabla_k|H|\nabla_k h_{ij}\\ &\qquad-4\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle-4|h|^2|\nabla^{\perp}\nu_1|^2.\end{aligned} \tag{1.53}$$

As for the second term on the third line of (1.50), we have

$$\begin{aligned}-2|H|^{-2}|\nabla\langle A_{ij},H\rangle|^2 &= -2|H|^{-2}|\nabla(|H|h_{ij})|^2\\ &= -2|H|^{-2}|h|^2|\nabla|H||^2-2|\nabla h|^2-4|H|^{-1}h_{ij}\nabla_k|H|\nabla_k h_{ij}.\end{aligned} \tag{1.54}$$

Equations (1.53) and (1.54) give the claimed six gradient terms on the right-hand side of equation (1.49). This completes the proof of the lemma. □

Substituting the lemma statement (1.49) into (1.48) and combining like terms yields

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\langle A_{ij},H\rangle|^2}{|H|^2} &= 4|\mathring{h}_{ij}\hat{A}_{ij}|^2+2|R^{\perp}_{ij}(\nu_1)|^2+2|h|^4\\ &\quad -4|H|^{-1}\mathring{h}_{ij}\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle-4\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle\\ &\quad -2|\nabla h|^2-2|h|^2|\nabla^{\perp}\nu_1|^2.\end{aligned}$$

We negate the expression above and add (1.30), the evolution equation of $|A|^2$, to get

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 &= -2|\nabla^{\perp}A|^2+2|\langle A_{ij},A_{kl}\rangle|^2+2|R^{\perp}|^2\\ &\quad -4|\mathring{h}_{ij}\hat{A}_{ij}|^2-2|R^{\perp}_{ij}(\nu_1)|^2-2|h|^4\\ &\quad +4|H|^{-1}\mathring{h}_{ij}\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle+4\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle\\ &\quad +2|\nabla h|^2+2|h|^2|\nabla^{\perp}\nu_1|^2.\end{aligned} \tag{1.55}$$

By (1.39) and (1.40) in Proposition 1.6, the reaction terms above satisfy

$$2|\langle A_{ij},A_{kl}\rangle|^2-4|\mathring{h}_{ij}\hat{A}_{ij}|^2-2|h|^4=2|\langle\hat{A}_{ij},\hat{A}_{kl}\rangle|^2, \tag{1.56}$$

$$2|R^{\perp}_{ij}|^2-2|R^{\perp}_{ij}(\nu_1)|^2=2|\hat{R}^{\perp}|^2+2|R^{\perp}_{ij}(\nu_1)|^2. \tag{1.57}$$

As for the gradient terms, using (1.20), we obtain

$$|\nabla^{\perp}A|^2=|\nabla^{\perp}\hat{A}|^2+|\nabla h|^2+|h|^2|\nabla^{\perp}\nu_1|^2+2\mathring{h}_{ij}\langle\nabla^{\perp}\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle+2\nabla_k\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle.$$

Rearranging this gives

$$-2|\nabla^{\perp}A|^2+2|\nabla h|^2+2|h|^2|\nabla^{\perp}\nu_1|^2+4\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle=-2|\nabla^{\perp}\hat{A}|^2-4\nabla_k\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle. \tag{1.58}$$

Substituting (1.56), (1.57), and (1.58) into equation (1.55), we finally get

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 &= 2|\langle \hat{A}_{ij},\hat{A}_{kl}\rangle|^2+2|\hat{R}^{\perp}|^2+2|R^{\perp}_{ij}(\nu_1)|^2 \\ &\quad -2|\nabla^{\perp}\hat{A}|^2+4|H|^{-1}\mathring{h}_{ij}\nabla_k|H|\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle-4\nabla_k\mathring{h}_{ij}\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle.\end{aligned} \tag{1.59}$$

Note that, by the orthogonality relations (1.10) and the identity $\mathring{A}=\hat{A}+\mathring{h}\nu_1$,

$$\begin{aligned}\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle &= -\langle\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle=-\langle\mathring{A}_{ij},\nabla^{\perp}_k\nu_1\rangle,\\ \nabla_k\mathring{h}_{ij} &= \langle\nabla^{\perp}_k\mathring{A}_{ij},\nu_1\rangle-\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle.\end{aligned}$$

From these and (1.59), we may now conclude the following proposition.

**Proposition 1.10** (Evolution of $|\hat{A}|^2$)**.**

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 &= 2|\langle \hat{A}_{ij},\hat{A}_{kl}\rangle|^2+2|\hat{R}^{\perp}|^2+2|R^{\perp}_{ij}(\nu_1)|^2 \\ &\quad -2|\nabla^{\perp}\hat{A}|^2+4\sum_{i,j,k=1}^{n}Q_{ijk}\langle\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle,\end{aligned} \tag{1.60}$$

*where*

$$Q_{ijk}:=\langle\nabla^{\perp}_k\mathring{A}_{ij},\nu_1\rangle-\langle\nabla^{\perp}_k\hat{A}_{ij},\nu_1\rangle-|H|^{-1}\mathring{h}_{ij}\nabla_k|H|. \tag{1.61}$$

## 1.3. Proof of Theorem 1.1

In a similar fashion to the proof of the pointwise derivative estimate in [41], will compare the evolution of $|\hat{A}|^2$ to the evolution of $c|H|^2-|A|^2$ along the mean curvature flow. Recall the setup of Theorem 1.1: we are given an $n$-dimensional mean curvature flow $M_t=F(M,t)\subset\mathbb{R}^N$ which satisfies $c_n|H|^2>|A|^2$ initially, where

$$c_n:=\begin{cases}\frac{3(n+1)}{2n(n+2)} & \text{if } n=5,6, \text{ or } 7\\ \frac{4}{3n} & \text{if } n\geq 8\end{cases}.$$

Note $c_n = \min\{\frac{3(n+1)}{2n(n+2)}, \frac{4}{3n}\}$. As $M$ is compact, we can find small real numbers $\varepsilon_0, \varepsilon_1 > 0$ depending on $M_0$ such that $(c_n - \varepsilon_0)|H|^2 - |A|^2 > \varepsilon_1$ holds everywhere on $M$ initially. Define a new constant

$$c_0 := \begin{cases} \frac{3(n+1)}{2n(n+2)} - \varepsilon_0 & \text{if } n = 5, 6, \text{ or } 7 \\ \frac{4}{3n} & \text{if } n \geq 8 \end{cases}$$

and

$$f := c_0|H|^2 - |A|^2. \tag{1.62}$$

Note that $c_0 \leq c_n$ only depends upon initial data if $n = 5, 6$, or $7$. The subtraction of $\varepsilon_0$ in these lower dimensions is because we need a bit more breathing room in our estimates than the critical dimensional constant $\frac{3(n+1)}{2n(n+2)}$ allows. By Theorem 2 in [2] (note Remark 1.3), we have $f > \varepsilon_1$ on $M \times [0, T)$. In particular, $|H| > 0$ on $M \times [0, T)$.

Let $\delta > 0$ be a small constant, which we will determine towards the end of the proof. We computed the evolution equation of $|\hat{A}|^2$ in the previous section. By (1.30) and (1.31) in Proposition 1.5, the evolution equation for $f$ is given by

$$\Big(\frac{\partial}{\partial t} - \Delta\Big) f = 2(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2) + 2(c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R^\perp_{ij}|^2). \tag{1.63}$$

The pinching condition implies that both terms on the right-hand side of (1.63) are nonnegative on $M \times [0, T)$. This is proven in [2] and stated in [41], but we will establish this ourselves in Lemma 1.15 below.

The first step of the proof of Theorem 1.1, and our main task, is to analyze the evolution equation of the scale-invariant quantity $|\hat{A}|^2/f$. We will show this ratio satisfies a favorable evolution equation with a right-hand side that has a nonpositive term. Specifically, we will show that

$$\Big(\frac{\partial}{\partial t} - \Delta\Big)\frac{|\hat{A}|^2}{f} \leq 2\Big\langle \nabla \frac{|\hat{A}|^2}{f}, \nabla \log f\Big\rangle - \delta \frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t} - \Delta\Big) f. \tag{1.64}$$

Then we will analyze the evolution of the non-scale-invariant quantity $|\hat{A}|^2/f^{1-\sigma}$. We will show for $\sigma > 0$ sufficiently small, the nonpositive term above can be used to control the nonnegative terms introduced by the additional factor of $f^\sigma$. Theorem 1.1 will then follow from the maximum principle.

By equation (1.63), Proposition 1.10, and the useful identity (1.43), the evolution equation of $|\hat{A}|^2/f$ is given by

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f} &= \frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - |\hat{A}|^2\frac{1}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle \\ &= \frac{1}{f}\Big(2|\langle\hat{A}_{ij},\hat{A}_{kl}\rangle|^2 + 2|\hat{R}^\perp|^2 + 2|R^\perp_{ij}(\nu_1)|^2\Big) + \frac{1}{f}\Big(-2|\nabla^\perp\hat{A}|^2 + 4Q_{ijk}\langle\hat{A}_{ij},\nabla^\perp_k\nu_1\rangle\Big) \\ &\quad - |\hat{A}|^2\frac{1}{f^2}\Big(2(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2)\Big) - |\hat{A}|^2\frac{1}{f^2}\Big(2(c_0|\langle A_{ij},H\rangle|^2 - |\langle A_{ij},A_{kl}\rangle|^2 - |R^\perp_{ij}|^2)\Big) \\ &\quad + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle.\end{aligned}$$

Rearranging these terms, we have

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f} &= \frac{1}{f}\Big(2|\langle\hat{A}_{ij},\hat{A}_{kl}\rangle|^2 + 2|\hat{R}^\perp|^2 + 2|R^\perp_{ij}(\nu_1)|^2\Big) \\ &\quad - \frac{1}{f}\Big(2\frac{|\hat{A}|^2}{f}(c_0|\langle A_{ij},H\rangle|^2 - |\langle A_{ij},A_{kl}\rangle|^2 - |R^\perp_{ij}|^2)\Big) \\ &\quad + \frac{1}{f}\Big(4Q_{ijk}\langle\hat{A}_{ij},\nabla^\perp_k\nu_1\rangle - 2|\nabla^\perp\hat{A}|^2 - 2\frac{|\hat{A}|^2}{f}(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2)\Big) \\ &\quad + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle.\end{aligned} \tag{1.65}$$

We analyze the right-hand side in two steps. We must estimate the reaction terms on the first line by the reaction terms on the second line and the gradient term $4Q_{ijk}\langle\hat{A}_{ij},\nabla^\perp_k\nu_1\rangle$ by the good Bochner terms coming from the evolution of $|\hat{A}|^2$ and $f$.

### 1.3.1. Reaction Term Estimates

We begin by estimating the reaction terms. We will make use of the following two estimates. The first estimate is proven on page 372 in [2] Section 3. The second estimate is a matrix inequality which is Theorem 1 in [36].

**Lemma 1.11.**

$$\Big|\sum_{i,j=1}^{n}\mathring{h}_{ij}\hat{A}_{ij}\Big|^2+|R_{ij}^{\perp}(\nu_1)|^2\leq 2|\mathring{h}|^2|\hat{A}|^2, \tag{1.66}$$

$$|\langle\hat{A}_{ij},\hat{A}_{kl}\rangle|^2+|\hat{R}^{\perp}|^2\leq\frac{3}{2}|\hat{A}|^4. \tag{1.67}$$

*Proof.* The arguments given in [2] to prove inequality (1.66) are simple and short, so we will repeat them in our notation here. We will express inequality (1.67) so that it is an immediate consequence of Theorem 1 in [36].

Fix any point $p\in M$ and time $t\in[0,T)$. Let $e_1,\dots,e_n$ be an orthonormal basis which identifies $T_pM\cong\mathbb{R}^n$ at time $t$ and then choose $\nu_2,\dots,\nu_{N-n}$ to be a basis of the orthogonal complement of principal normal $\nu_1$ in $N_pM$ at time $t$. For each $\beta\in\{2,\dots,N-n\}$, define a matrix $A_\beta=\langle A,\nu_\beta\rangle$ whose components are given by $(A_\beta)_{ij}=A_{ij\beta}$. Note, by definition, $\hat{A}_{ij\beta}=A_{ij\beta}$ when $\beta\geq 2$, but we want to match the notation of [36] for the moment. Then $\hat{A}=\sum_{\beta=2}^{N-n}A_\beta\nu_\beta$. We also have $\mathring{h}=\langle\mathring{A},\nu_1\rangle$.

To prove (1.66), let $\lambda_1,\dots,\lambda_n$ denote the eigenvalues of $\mathring{h}$. Assume the orthonormal basis is an eigenbasis of $\mathring{h}$. Now

$$\begin{aligned}\Big|\sum_{i,j=1}^{n}\mathring{h}_{ij}\hat{A}_{ij}\Big|^2&=\sum_{\beta=2}^{N-n}\sum_{i,j,k,l=1}^{n}\mathring{h}_{ij}\mathring{h}_{kl}A_{ij\beta}A_{kl\beta}\\&=\sum_{\beta=2}^{N-n}\Big(\sum_{i,j=1}^{n}\mathring{h}_{ij}A_{ij\beta}\Big)^2\\&=\sum_{\beta=2}^{N-n}\Big(\sum_{i=1}^{n}\lambda_iA_{ii\beta}\Big)^2.\end{aligned}$$

By Cauchy-Schwarz,

$$\Big|\sum_{i,j=1}^{n}\mathring{h}_{ij}\hat{A}_{ij}\Big|^2\leq\sum_{\beta=2}^{N-n}\Big(\sum_{i=1}^{n}\lambda_j^2\Big)\Big(\sum_{i=1}^{n}A_{ii\beta}^2\Big)=|\mathring{h}|^2\sum_{\beta=2}^{N-n}\sum_{i=1}^{n}A_{ii\beta}^2. \tag{1.68}$$

By (1.41),

$$|R^{\perp}_{ij}(\nu_1)|^2 = \sum_{\beta=2}^{N-2}\sum_{i,j=1}^{n}\Big(\sum_{k=1}^{n}(\mathring{h}_{ik}A_{jk\beta} - \mathring{h}_{jk}A_{ik\beta})\Big)^2$$
$$= \sum_{\beta=2}^{N-2}\sum_{i,j=1}^{n}(\lambda_i - \lambda_j)^2 A^2_{ij\beta}$$
$$= \sum_{\beta=2}^{N-2}\sum_{i\neq j}(\lambda_i - \lambda_j)^2 A^2_{ij\beta}.$$

Since $(\lambda_i - \lambda_j)^2 \le 2(\lambda_i^2 + \lambda_j^2) \le 2|\mathring{h}|^2$, we have

$$|R^{\perp}_{ij}(\nu_1)|^2 \le 2|\mathring{h}|^2 \sum_{\beta=2}^{N-2}\sum_{i\neq j} A^2_{ij\beta}. \tag{1.69}$$

Summing (1.68) and (1.69), we obtain

$$\Big|\sum_{i,j=1}^{n}\mathring{h}_{ij}\hat{A}_{ij}\Big|^2 + |R^{\perp}_{ij}(\nu_1)|^2 \le |\mathring{h}|^2\sum_{\beta=2}^{N-n}\sum_{i=1}^{n}A^2_{ii\beta} + 2|\mathring{h}|^2\sum_{\beta=2}^{N-2}\sum_{i\neq j}A^2_{ij\beta} \le 2|\mathring{h}|^2|\hat{A}|^2,$$

which is (1.66)

To establish (1.67), for $\alpha, \beta \in \{2, \dots, N-n\}$ define

$$S_{\alpha\beta} := \mathrm{tr}(A_\alpha A_\beta) = \sum_{i,j=1}^{n} A_{ij\alpha}A_{ij\beta} \qquad \text{and} \qquad S_\alpha := |A_\alpha|^2 = \sum_{i,j=1}^{n} A_{ij\alpha}A_{ij\alpha}.$$

Let $S := S_2 + \dots + S_{N-n} = |\hat{A}|^2$.

Now

$$|\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 = \sum_{i,j,k,l=1}^{n} \sum_{\alpha,\beta=2}^{N-n} A_{ij\alpha}A_{kl\alpha}A_{ij\beta}A_{kl\beta}$$
$$= \sum_{\alpha\beta=2}^{N-n} \Big( \sum_{i,j=1}^{n} A_{ij\alpha}A_{ij\beta}\Big)\Big( \sum_{k,l=1}^{n} A_{kl\alpha}A_{kl\beta}\Big)$$
$$= \sum_{\alpha,\beta=2}^{N-n} S_{\alpha\beta}^2.$$

In addition, recalling the definition (1.42), we may write

$$|\hat{R}^\perp|^2 = \sum_{\alpha,\beta=2}^{N-n} \big|A_\alpha A_\beta - A_\beta A_\alpha\big|^2,$$

where $(A_\alpha A_\beta)_{ij} = (A_\alpha)_{ik}(A_\beta)_{kj} = (A_\alpha)_{ik}(A_\beta)_{jk}$ denotes standard matrix multiplication and $|\cdot|$ is the usual square norm of the matrix. We see that inequality (1.67) is equivalent to

$$\sum_{\alpha,\beta=2}^{N-n} \big|A_\alpha A_\beta - A_\beta A_\alpha\big|^2 + \sum_{\alpha,\beta=2}^{N-n} S_{\alpha\beta}^2 \leq \frac{3}{2}S^2. \tag{1.70}$$

Now if $N-n=2$, inequality (1.67) is trivial since $|R^\perp|^2 = 0$ and $|\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 = |\hat{A}|^4$. Otherwise, if $N-n \geq 3$, inequality (1.70) is precisely Theorem 1 in [36]. This completes the proof. □

As an immediate consequence of the previous lemma, we have the following estimate for the reaction terms coming from the evolution of $|\hat{A}|^2$.

**Lemma 1.12** (Upper bound for the reaction terms of $(\partial_t - \Delta)|\hat{A}|^2$)**.**

$$|\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 + |\hat{R}^\perp|^2 + |R_{ij}^\perp(\nu_1)|^2 \leq \frac{3}{2}|\hat{A}|^4 + 2|\mathring{h}|^2|\hat{A}|^2. \tag{1.71}$$

Next we express the reaction term in the evolution of $f$ in terms of $\hat{A}$, $\mathring{h}$, and $|H|$. In view of (1.62), observe that

$$\frac{nc_0 - 1}{n}|H|^2 = |\hat{A}|^2 + |\mathring{h}|^2 + f. \tag{1.72}$$

We have following lower bound for the reaction terms in the evolution of $f$.

**Lemma 1.13** (Lower bound for the reaction terms of $(\partial_t - \Delta)f$)**.** *If* $\frac{1}{n} < c_0 \leq \frac{4}{3n}$, *then*

$$\frac{|\hat{A}|^2}{f}(c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R_{ij}^{\perp}|^2) \geq \frac{2}{nc_0 - 1}|\hat{A}|^4 + \frac{nc_0}{nc_0 - 1}|\mathring{h}|^2|\hat{A}|^2. \tag{1.73}$$

*Proof.* We do a computation that is similar to a computation in [2], except we do not throw away the pinching term $f$. By (1.37), (1.38), (1.39), (1.40), we have

$$\begin{aligned}
&c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R_{ij}^{\perp}|^2 \\
&\quad= \frac{1}{n}c_0|H|^4 + c_0|\mathring{h}|^2|H|^2 \\
&\qquad - |\mathring{h}|^4 - \frac{2}{n}|\mathring{h}|^2|H|^2 - \frac{1}{n^2}|H|^4 - 2|\mathring{h}^{ij}\hat{A}_{ij}|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 \\
&\qquad - |\hat{R}^{\perp}|^2 - 2|R_{ij}^{\perp}(\nu_1)|^2 \\
&\quad= \frac{1}{n}(c_0 - \frac{1}{n})|H|^4 + (c_0 - \frac{1}{n})|\mathring{h}|^2|H|^2 - \frac{1}{n}|\mathring{h}|^2|H|^2 - |\mathring{h}|^4 \\
&\qquad - 2|\mathring{h}_{ij}\hat{A}_{ij}|^2 - 2|R_{ij}^{\perp}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2.
\end{aligned}$$

Use (1.72) and cancel terms to get

$$\begin{aligned}
&c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R_{ij}^{\perp}|^2 \\
&\quad= \frac{1}{n}(f + |\hat{A}|^2 + |\mathring{h}|^2)|H|^2 + (f + |\hat{A}|^2 + |\mathring{h}|^2)|\mathring{h}|^2 - \frac{1}{n}|\mathring{h}|^2|H|^2 - |\mathring{h}|^4 \\
&\qquad - 2|\mathring{h}_{ij}\hat{A}_{ij}|^2 - 2|R_{ij}^{\perp}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2 \\
&\quad= \frac{1}{n}(f + |\hat{A}|^2)|H|^2 + (f + |\hat{A}|^2)|\mathring{h}|^2 \\
&\qquad - 2|\mathring{h}_{ij}\hat{A}_{ij}|^2 - 2|R_{ij}^{\perp}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2.
\end{aligned}$$

Using (1.72) once more for the remaining factor of $|H|^2$ gives

$$\begin{aligned}
&c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R^{\perp}_{ij}|^2 \\
&\quad = \frac{1}{n}(f + |\hat{A}|^2)\Big(c_0 - \frac{1}{n}\Big)^{-1}(f + |\hat{A}|^2 + |\mathring{h}|^2) + (f + |\hat{A}|^2)|\mathring{h}|^2 \\
&\quad\quad - 2|\mathring{h}_{ij}\hat{A}_{ij}|^2 - 2|R^{\perp}_{ij}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2 \\
&\quad = \frac{1}{nc_0 - 1}f(f + 2|\hat{A}|^2 + |\mathring{h}|^2) + f|\mathring{h}|^2 + \frac{1}{nc_0 - 1}|\hat{A}|^4 + \frac{nc_0}{nc_0 - 1}|\hat{A}|^2|\mathring{h}|^2 \\
&\quad\quad - 2|\mathring{h}_{ij}\hat{A}_{ij}|^2 - 2|R^{\perp}_{ij}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2.
\end{aligned}$$

Now by the two estimates in Lemma 1.11

$$2|\mathring{h}_{ij}\hat{A}_{ij}|^2 + 2|R^{\perp}_{ij}(\nu_1)|^2 + |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 + |\hat{R}^{\perp}|^2 \le 4|\mathring{h}|^2|\hat{A}|^2 + \frac{3}{2}|\hat{A}|^4.$$

Therefore

$$\begin{aligned}
&\frac{1}{nc_0 - 1}|\hat{A}|^4 + \frac{nc_0}{nc_0 - 1}|\hat{A}|^2|\mathring{h}|^2 - 2|\mathring{h}^{ij}\hat{A}_{ij}|^2 - 2|R^{\perp}_{ij}(\nu_1)|^2 - |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 - |\hat{R}^{\perp}|^2 \\
&\quad\quad \ge \Big(\frac{1}{nc_0 - 1} - \frac{3}{2}\Big)|\hat{A}|^4 + \Big(\frac{nc_0}{nc_0 - 1} - 4\Big)|\mathring{h}|^2|\hat{A}|^2.
\end{aligned}$$

Since $c_0 \le \frac{4}{3n}$, we have

$$\frac{1}{nc_0 - 1} - \frac{3}{2} \ge \frac{3}{2}, \qquad \frac{nc_0}{nc_0 - 1} - 4 \ge 0.$$

Consequently, we have

$$\begin{aligned}
c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R^{\perp}_{ij}|^2 &\ge \frac{2}{nc_0 - 1}f|\hat{A}|^2 + \frac{nc_0}{nc_0 - 1}f|\mathring{h}|^2 + \frac{1}{nc_0 - 1}f^2 \qquad (1.74) \\
&\ge \frac{2}{nc_0 - 1}f|\hat{A}|^2 + \frac{nc_0}{nc_0 - 1}f|\mathring{h}|^2.
\end{aligned}$$

Multiplying both sides by $\frac{|\hat{A}|^2}{f}$ completes the proof of the lemma. $\square$

Putting Lemmas 1.12 and 1.13 together, we have

**Lemma 1.14** (Reaction term estimate)**.** *If* $0 < \delta \le \frac{1}{2}$ *and* $\frac{1}{n} < c_0 \le \frac{4}{3n}$*, then*

$$|\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 + |\hat{R}^\perp|^2 + |R^\perp_{ij}(\nu_1)|^2 \le (1-\delta)\frac{|\hat{A}|^2}{f}(c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R^\perp_{ij}|^2). \quad (1.75)$$

*Proof.* In view of (1.71) and (1.73), we have

$$\begin{aligned} |\langle \hat{A}_{ij}, \hat{A}_{kl}\rangle|^2 + |\hat{R}^\perp|^2 + |R^\perp_{ij}(\nu_1)|^2 - (1-\delta)\frac{|\hat{A}|^2}{f}(c_0|\langle A_{ij}, H\rangle|^2 - |\langle A_{ij}, A_{kl}\rangle|^2 - |R^\perp_{ij}|^2) \\ \le \frac{3}{2}|\hat{A}|^4 + 2|\mathring{h}|^2|\hat{A}|^2 - \frac{2(1-\delta)}{nc_0 - 1}|\hat{A}|^4 - \frac{nc_0(1-\delta)}{nc_0-1}|\mathring{h}|^2|\hat{A}|^2 \\ = \Big(\frac{3}{2} - \frac{2(1-\delta)}{nc_0-1}\Big)|\hat{A}|^4 + \Big(2 - \frac{nc_0(1-\delta)}{nc_0-1}\Big)|\mathring{h}|^2|\hat{A}|^2. \end{aligned}$$

If $c_0 \le \frac{4}{3n}$, then

$$\frac{1}{nc_0 - 1} \ge 3, \quad \text{and} \quad \frac{nc_0}{nc_0-1} \ge 4.$$

Therefore if $\delta \le \frac{1}{2}$

$$\begin{aligned} \frac{3}{2} - \frac{2(1-\delta)}{nc_0-1} \le \frac{3}{2} - 6(1-\delta) \le 0, \\ 2 - \frac{nc_0(1-\delta)}{nc_0-1} \le 2 - 4(1-\delta) \le 0, \end{aligned}$$

which gives (1.75). ∎

### 1.3.2. Gradient Term Estimates

Having analyzed the reaction terms, we turn our attention to the gradient terms. For this, we will use equation (1.23). Recalling that $\hat{A}_{jk}$ is traceless, it is straightforward to verify that

$$|\nabla_i h_{jk} + \langle \nabla^\perp_i \hat{A}_{jk}, \nu_1\rangle|^2 = |\nabla_i \mathring{h}_{jk} + \langle \nabla^\perp_i \hat{A}_{jk}, \nu_1\rangle|^2 + \frac{1}{n}|\nabla|H||^2, \quad (1.76)$$

$$|\hat{\nabla}^\perp_i \hat{A}_{jk} + h_{jk}\nabla^\perp_i \nu_1|^2 = |\hat{\nabla}^\perp_i \hat{A}_{jk} + \mathring{h}_{jk}\nabla^\perp_i \nu_1|^2 + \frac{1}{n}|H|^2|\nabla^\perp \nu_1|^2. \quad (1.77)$$

Observe that the first term in (1.76) is just

$$|\langle \nabla_i^\perp \mathring{A}_{jk}, \nu_1 \rangle|^2 = |\nabla_i \mathring{h}_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1 \rangle|^2, \tag{1.78}$$

which will be useful later on. Now as observed in [31] (cf. [23] for the Ricci flow), it follows from the Codazzi identity (1.12) that the tensor

$$E_{ijk} = \frac{1}{n+2}\Big(g_{ij}\nabla_k^\perp H + g_{jk}\nabla_i^\perp H + g_{ki}\nabla_j^\perp H\Big)$$

is an irreducible component of $\nabla_i^\perp A_{jk}$ consisting of its various traces. In other words, $E_{ijk}\nabla_i^\perp A_{jk} = |E|^2$. This allows one to get an improved estimate over the trivial one. Namely,

$$|E|^2 = \frac{3}{n+2}|\nabla^\perp H|^2 \le |\nabla^\perp A|^2. \tag{1.79}$$

By consequence of this estimate and (1.74), we can conclude nonnegativity of $(\partial_t - \Delta) f$:

**Lemma 1.15** (Nonnegativity of $(\partial_t - \Delta)f$)**.** *If $\frac{1}{n} < c_0 \le \frac{4}{3n}$, then*

$$\Big(\frac{\partial}{\partial t} - \Delta\Big) f \ge 2|h|^2 f + \frac{2}{nc_0 - 1}|\hat{A}|^2 f + \Big(1 - \frac{c_0(n+2)}{3}\Big)|\nabla^\perp A|^2.$$

*In particular, $(\partial_t - \Delta) f \ge 0$ if $f \ge 0$.*

*Proof.* We first observe, using (1.72), that

$$\begin{aligned}\frac{2}{nc_0-1} f|\hat{A}|^2 + \frac{nc_0}{nc_0-1} f|\mathring{h}|^2 + \frac{1}{nc_0-1} f^2 &= \frac{1}{nc_0-1}(|\mathring{h}|^2 + |\hat{A}|^2 + f)f + |\mathring{h}|^2 f + \frac{1}{nc_0-1}|\hat{A}|^2 f \\ &= \frac{1}{n}|H|^2 f + |\mathring{h}|^2 f + \frac{1}{nc_0-1}|\hat{A}|^2 f \\ &= |h|^2 f + \frac{1}{nc_0-1}|\hat{A}|^2 f.\end{aligned}$$

Consequently by (1.63) and (1.74), we have

$$\Big(\frac{\partial}{\partial t}-\Delta\Big)f \geq 2|h|^2 f + \frac{2}{nc_0-1}|\hat{A}|^2 f + 2(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2)$$

Hence, by (1.79)

$$\Big(\frac{\partial}{\partial t}-\Delta\Big)f \geq 2|h|^2 f + \frac{2}{nc_0-1}|\hat{A}|^2 f + \Big(1-\frac{c_0(n+2)}{3}\Big)|\nabla^\perp A|^2.$$

This is the same inequality obtained in [2] and [41]. □

There is an analogue of (1.79) in both the principal direction and its orthogonal complement. We observed in Section 4.1 that the projection of the Codazzi identity onto $\nu_1$ and its orthogonal complement implies the tensors $\nabla_i h_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle$ and $\hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp \nu_1$ are symmetric in $i, j, k$. Recalling (1.26) and (1.27), it follows that an irreducible component of each tensor is given by

$$\begin{aligned} E^{(1)}_{ijk} &:= \frac{1}{n+2}\Big(g_{ij}\nabla_k|H| + g_{jk}\nabla_i|H| + g_{ki}\nabla_j|H|\Big), \\ E^{(\perp)}_{ijk} &:= \frac{1}{n+2}\Big(g_{ij}|H|\nabla_k^\perp \nu_1 + g_{jk}|H|\nabla_i^\perp\nu_1 + g_{ki}|H|\nabla_j^\perp\nu_1\Big). \end{aligned}$$

You can readily confirm that $E^{(1)}_{ijk}(\nabla_i h_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle) = |E^{(1)}|^2$ and $\langle E^{(\perp)}_{ijk}, \hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp\nu_1\rangle = |E^{(\perp)}|^2$. As in (1.79), we obtain that

$$\frac{3}{n+2}|\nabla|H||^2 \leq |\nabla_i h_{jk} + \langle \nabla_i^\perp \hat{A}_{jk}, \nu_1\rangle|^2, \tag{1.80}$$

$$\frac{3}{n+2}|H|^2|\nabla^\perp \nu_1|^2 \leq |\hat{\nabla}_i^\perp \hat{A}_{jk} + h_{jk}\nabla_i^\perp \nu_1|^2. \tag{1.81}$$

Now expanding the right-hand side of both inequalities above using (1.76) and (1.77); recalling

(1.78); and noting that $\frac{3}{n+2} - \frac{1}{n} = \frac{2(n-1)}{n(n+2)}$, we arrive at the estimates

$$\frac{2(n-1)}{n(n+2)}|\nabla|H||^2 \le |\langle \nabla_i^\perp \mathring{A}_{jk}, \nu_1 \rangle|^2, \tag{1.82}$$

$$\frac{2(n-1)}{n(n+2)}|H|^2|\nabla^\perp \nu_1|^2 \le |\hat{\nabla}_i^\perp \hat{A}_{jk} + \mathring{h}_{jk}\nabla_i^\perp \nu_1|^2. \tag{1.83}$$

The second of these two estimates implies the following useful lower bound.

**Lemma 1.16** (Lower bound for Bochner term of $(\partial_t - \Delta)|\hat{A}|^2$)**.**

1. *If $\frac{1}{n} < c_0 \le \frac{4}{3n}$, then*

$$2|\hat{\nabla}^\perp \hat{A}|^2 \ge \frac{4n-10}{n+2}|\mathring{h}|^2|\nabla^\perp \nu_1|^2 + \frac{6(n-1)}{n+2}|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{6(n-1)}{n+2} f|\nabla^\perp \nu_1|^2. \tag{1.84}$$

2. *If $\frac{1}{n} < c_0 \le \frac{3(n+1)}{2n(n+2)}$, then*

$$2|\hat{\nabla}^\perp \hat{A}|^2 \ge 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2 + 4|\hat{A}|^2|\nabla^\perp \nu_1|^2 + 4f|\nabla^\perp \nu_1|^2. \tag{1.85}$$

*Proof.* We begin by applying Young's inequality

$$\begin{aligned}|\hat{\nabla}_i^\perp \hat{A}_{jk} + \mathring{h}_{jk}\nabla_i^\perp \nu_1|^2 &= |\hat{\nabla}^\perp \hat{A}|^2 + 2\langle \hat{\nabla}_i^\perp \hat{A}_{jk}, \mathring{h}_{jk}\nabla_i^\perp \nu_1 \rangle + |\mathring{h}|^2|\nabla^\perp \nu_1|^2 \\ &\le 2|\hat{\nabla}^\perp \hat{A}|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2.\end{aligned}$$

Multiplying both sides of (1.72) by $\frac{2(n-1)}{(n+2)(nc_0-1)}$ gives

$$\frac{2(n-1)}{n(n+2)}|H|^2 = \frac{2(n-1)}{(n+2)(nc_0-1)}(f + |\hat{A}|^2 + |\mathring{h}|^2).$$

In view of (1.83), our observations give us that

$$\frac{2(n-1)}{(n+2)(nc_0-1)}(f + |\hat{A}|^2 + |\mathring{h}|^2)|\nabla^\perp \nu_1|^2 \le 2|\hat{\nabla}^\perp \hat{A}|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2.$$

Subtracting the $|\mathring{h}|^2|\nabla^\perp\nu_1|^2$ term on the right-hand side gives

$$\frac{2(n-1)}{(n+2)(nc_0-1)}(f+|\hat{A}|^2)|\nabla^\perp\nu_1|^2+\Big(\frac{2(n-1)}{(n+2)(nc_0-1)}-2\Big)|\mathring{h}|^2|\nabla^\perp\nu_1|^2\le 2|\hat{\nabla}^\perp\hat{A}|^2.$$

If $c_0\le\frac{4}{3n}$, then $nc_0-1\le\frac{1}{3}$ and

$$\frac{2(n-1)}{(n+2)(nc_0-1)}\ge\frac{6(n-1)}{n+2}.$$

Plugging this in above gives the first estimate of the lemma. If $c_0\le\frac{3(n+1)}{2n(n+2)}$, then $nc_0-1\le\frac{n-1}{2(n+2)}$ and

$$\frac{2(n-1)}{(n+2)(nc_0-1)}\ge 4.$$

This establishes the second estimate in the lemma. □

Next we obtain improved lower bounds for the Bochner term in the evolution equation of $f$.

**Lemma 1.17** (Lower bound for Bochner term of $(\partial_t-\Delta)f$)**.**

1. *If* $\frac{1}{n}<c_0\le\frac{4}{3n}$*, then*

$$2\frac{|\hat{A}|^2}{f}\big(|\nabla^\perp A|^2-c_0|\nabla^\perp H|^2\big)\ge\frac{5n-8}{3(n-1)}\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2+\frac{10n-16}{n+2}|\hat{A}|^2|\nabla^\perp\nu_1|^2. \tag{1.86}$$

2. *If* $\frac{1}{n}<c_0\le\frac{3(n+1)}{2n(n+2)}$*, then*

$$2\frac{|\hat{A}|^2}{f}\big(|\nabla^\perp A|^2-c_0|\nabla^\perp H|^2\big)\ge\frac{3}{2}\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2+6|\hat{A}|^2|\nabla^\perp\nu_1|^2. \tag{1.87}$$

*Proof.* By (1.23) and (1.24), we have

$$\begin{aligned}|\nabla^\perp A|^2-c_0|\nabla^\perp H|^2&=|\langle\nabla_i^\perp\hat{A}_{jk},\nu_1\rangle+\nabla_i h_{jk}|^2-c_0|\nabla|H||^2\\&\quad+|\hat{\nabla}_i^\perp\hat{A}_{jk}+h_{jk}\nabla_i^\perp\nu_1|^2-c_0|H|^2|\nabla^\perp\nu_1|^2.\end{aligned}$$

In view of (1.76), (1.78) and (1.82), we have

$$|\langle\nabla_i^\perp\hat{A}_{jk},\nu_1\rangle+\nabla_i h_{jk}|^2-c_0|\nabla|H||^2=|\langle\nabla_i^\perp\mathring{A}_{jk},\nu_1\rangle|^2-\frac{nc_0-1}{n}|\nabla|H||^2$$
$$\geq\Big(1-\frac{(n+2)(nc_0-1)}{2(n-1)}\Big)|\langle\nabla_i^\perp\mathring{A}_{jk},\nu_1\rangle|^2.$$

In view of (1.81) and (1.72), we have

$$|\hat{\nabla}_i^\perp\hat{A}_{jk}+h_{jk}\nabla_i^\perp\nu_1|^2-c_0|H|^2|\nabla^\perp\nu_1|^2\geq\Big(\frac{3}{n+2}-c_0\Big)|H|^2|\nabla^\perp\nu_1|^2$$
$$=\frac{n}{nc_0-1}\Big(\frac{3}{n+2}-c_0\Big)(f+|\hat{A}|^2+|\mathring{h}|^2)|\nabla^\perp\nu_1|^2$$
$$\geq\frac{n}{nc_0-1}\Big(\frac{3}{n+2}-c_0\Big)f|\nabla^\perp\nu_1|^2.$$

Thus, by the three previous computations

$$2\frac{|\hat{A}|^2}{f}(|\nabla^\perp A|^2-c_0|\nabla^\perp H|^2)\geq\Big(2-\frac{(n+2)(nc_0-1)}{(n-1)}\Big)\frac{|\hat{A}|^2}{f}|\langle\nabla_i^\perp\mathring{A}_{jk},\nu_1\rangle|^2$$
$$+\frac{2n}{nc_0-1}\Big(\frac{3}{n+2}-c_0\Big)|\hat{A}|^2|\nabla^\perp\nu_1|^2.$$

If $c_0\leq\frac{4}{3n}$, then $nc_0-1\leq\frac{1}{3}$ and

$$2-\frac{(n+2)(nc_0-1)}{(n-1)}\geq 2-\frac{n+2}{3(n-1)}=\frac{5n-8}{3(n-1)},$$
$$\frac{2n}{nc_0-1}\Big(\frac{3}{n+2}-c_0\Big)\geq 6n\Big(\frac{9n-4(n+2)}{(n+2)3n}\Big)=\frac{10n-16}{n+2}.$$

This establishes the first inequality of the lemma. If $c_0\leq\frac{3(n+1)}{2n(n+2)}$, then $nc_0-1\leq\frac{n-1}{2(n+2)}$ and

$$2-\frac{(n+2)(nc_0-1)}{(n-1)}\geq 2-\frac{1}{2}=\frac{3}{2},$$
$$\frac{2n}{nc_0-1}\Big(\frac{3}{n+2}-c_0\Big)\geq\frac{4n(n+2)}{n-1}\Big(\frac{6n-3(n+1)}{2n(n+2)}\Big)=6.$$

This establishes the second inequality of the lemma. □

Finally, we must estimate the remaining gradient term that appears in the evolution equation of $|\hat{A}|^2$. The term is of the form $4Q_{ijk}\langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle$ where $Q$ is defined in (1.61). By Cauchy-Schwarz, we obtain the following useful estimate:

$$\begin{aligned} |\langle \hat{A}, \nabla^\perp \nu_1\rangle|^2 &= \sum_{i,j,k=1}^{n} \langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle^2 \\ &\leq \sum_{i,j,k=1}^{n} \Big(\sum_{\beta=2}^{N-2} \hat{A}^2_{ij\beta}\Big)\Big(\sum_{\beta=2}^{N-2} \langle \nabla^\perp_k \nu_1, \nu_\beta\rangle^2\Big) \\ &\leq |\hat{A}|^2 |\nabla^\perp \nu_1|^2. \end{aligned} \tag{1.88}$$

**Lemma 1.18** (Upper bound for gradient term of $(\partial_t - \Delta)|\hat{A}|^2$)**.**

1. *If* $\frac{1}{n} < c_0 \leq \frac{4}{3n}$, *then*

$$\begin{aligned} 4Q_{ijk}\langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle \leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + \frac{5n-9}{3(n-1)} \frac{|\hat{A}|^2}{f} |\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\ + 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{3(n-1)}{n-3} f |\nabla^\perp \nu_1|^2 + \frac{2(n+2)}{n+3}|\mathring{h}|^2|\nabla^\perp \nu_1|^2. \end{aligned} \tag{1.89}$$

2. *If* $\frac{1}{n} < c_0 \leq \frac{3(n+1)}{2n(n+2)} - \varepsilon_0$ *and* $\varepsilon = \frac{2n(n+2)}{3(n-1)}\varepsilon_0$, *then*

$$\begin{aligned} 4Q_{ijk}\langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle \leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + (1-\varepsilon)\frac{3}{2}\frac{|\hat{A}|^2}{f} |\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\ + 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + 4f|\nabla^\perp \nu_1|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2. \end{aligned} \tag{1.90}$$

*Proof.* Using the triangle inequality on (1.61), we get

$$|Q| \leq |\langle \nabla^\perp \mathring{A}, \nu_1\rangle| + |\langle \nabla^\perp \hat{A}, \nu_1\rangle| + |H|^{-1}|\mathring{h}||\nabla |H||. \tag{1.91}$$

We will first treat the case $\frac{1}{n} < c_0 \leq \frac{4}{3n}$. It easily follows from (1.62), the definition of $f$, that

$$f \leq \Big(c_0 - \frac{1}{n}\Big)|H|^2 \leq \frac{1}{3n}|H|^2.$$

Consequently, using the estimate (1.82), we obtain

$$\frac{|\hat{A}|^2}{|H|^2}|\nabla|H||^2 \leq \frac{n(n+2)}{2(n-1)}\frac{1}{3n}\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2 = \frac{n+2}{6(n-1)}\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2. \tag{1.92}$$

Then (1.88) and (1.91) give

$$\begin{aligned}4Q_{ijk}\langle\hat{A}_{ij},\nabla^\perp_k\nu_1\rangle &\leq 4|Q||\langle\hat{A},\nabla^\perp\nu_1\rangle| \\ &\leq 4\Big(|\langle\nabla^\perp\mathring{A},\nu_1\rangle| + |\langle\nabla^\perp\hat{A},\nu_1\rangle| + |H|^{-1}|\mathring{h}||\nabla|H||\Big)|\hat{A}||\nabla^\perp\nu_1|.\end{aligned}$$

Now to each of these three summed terms above we apply Young's inequality with constants $a_1, a_2, a_3 > 0$. Specifically, we have

$$\begin{aligned}4|\langle\nabla^\perp\hat{A},\nu_1\rangle||\hat{A}||\nabla^\perp\nu_1| &\leq 2a_1|\langle\nabla^\perp\hat{A},\nu_1\rangle|^2 + \frac{2}{a_1}|\hat{A}|^2|\nabla^\perp\nu_1|^2, \\ 4|\langle\nabla^\perp\mathring{A},\nu_1\rangle||\hat{A}||\nabla^\perp\nu_1| &= 4|\langle\nabla^\perp\mathring{A},\nu_1\rangle|\frac{|\hat{A}|}{f^{\frac{1}{2}}}f^{\frac{1}{2}}|\nabla^\perp\nu_1| \\ &\leq 2a_2\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2 + \frac{2}{a_2}f|\nabla^\perp\nu_1|^2, \\ 4|H|^{-1}|\mathring{h}||\nabla|H|||\hat{A}||\nabla^\perp\nu_1| &\leq 2a_3\frac{|\hat{A}|^2}{|H|^2}|\nabla|H||^2 + \frac{2}{a_3}|\mathring{h}|^2|\nabla^\perp\nu_1|^2 \\ &\leq 2a_3\frac{n+2}{6(n-1)}\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2 + \frac{2}{a_3}|\mathring{h}|^2|\nabla^\perp\nu_1|^2.\end{aligned}$$

Note we used (1.92) in the last inequality. Hence

$$\begin{aligned}4Q_{ijk}\langle\hat{A}_{ij},\nabla^\perp_k\nu_1\rangle \leq{}& 2a_1|\langle\nabla^\perp\hat{A},\nu_1\rangle|^2 + \big(2a_2 + 2a_3\frac{n+2}{6(n-1)}\big)\frac{|\hat{A}|^2}{f}|\langle\nabla^\perp\mathring{A},\nu_1\rangle|^2 \\ &+ \frac{2}{a_1}|\hat{A}|^2|\nabla^\perp\nu_1|^2 + \frac{2}{a_2}f|\nabla^\perp\nu_1|^2 + \frac{2}{a_3}|\mathring{h}|^2|\nabla^\perp\nu_1|^2.\end{aligned} \tag{1.93}$$

Now set

$$a_1 = 1, \qquad a_2 = \frac{2(n-3)}{3(n-1)}, \qquad a_3 = \frac{n+3}{n+2}.$$

In this case,

$$\begin{aligned}
2a_2 + 2a_3\frac{n+2}{6(n-1)} &= \frac{4(n-3)}{3(n-1)} + \frac{n+3}{n+2}\frac{n+2}{3(n-1)} = \frac{5n-9}{3(n-1)},\\
\frac{2}{a_2} &= \frac{3(n-1)}{n-3},\\
\frac{2}{a_3} &= \frac{2(n+2)}{n+3}.
\end{aligned}$$

Plugging these into (1.93), we conclude

$$\begin{aligned}
4Q_{ijk}\langle \hat{A}_{ij}, \nabla_k^\perp \nu_1\rangle \leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + \frac{5n-9}{3(n-1)}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2\\
+ 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{3(n-1)}{n-3}f|\nabla^\perp \nu_1|^2 + \frac{2(n+2)}{n+3}|\mathring{h}|^2|\nabla^\perp \nu_1|^2,
\end{aligned}$$

as claimed.

Now if $\frac{1}{n} < c_0 \leq \frac{3(n+1)}{2n(n+2)} - \varepsilon_0$, then $c_0 - \frac{1}{n} \leq \frac{n-1}{2n(n+2)} - \varepsilon_0$. Therefore if we take $\varepsilon = \frac{2n(n+2)}{3(n-1)}\varepsilon_0$, then

$$c_0 - \frac{1}{n} \leq (1-3\varepsilon)\frac{n-1}{2n(n+2)}.$$

In this case,

$$f \leq \Big(c_0 - \frac{1}{n}\Big)|H|^2 \leq (1-3\varepsilon)\frac{n-1}{2n(n+2)}|H|^2.$$

Again using (1.82), it follows that

$$\begin{aligned}
\frac{|\hat{A}|^2}{|H|^2}|\nabla|H||^2 &\leq (1-3\varepsilon)\frac{n(n+2)}{2(n-1)}\frac{n-1}{2n(n+2)}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2\\
&= \frac{1}{4}(1-3\varepsilon)\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2.
\end{aligned}$$

Proceeding as we did before, we obtain the inequality

$$4Q_{ijk}\langle \hat{A}_{ij}, \nabla_k^\perp \nu_1\rangle \leq 2a_1|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + \big(2a_2 + \frac{1}{2}a_3(1-3\varepsilon)\big)\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \qquad (1.94)$$
$$+ \frac{2}{a_1}|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{2}{a_2}f|\nabla^\perp \nu_1|^2 + \frac{2}{a_3}|\mathring{h}|^2|\nabla^\perp \nu_1|^2.$$

Set

$$a_1 = 1, \quad a_2 = \frac{1}{2}, \quad a_3 = 1.$$

In this case,

$$2a_2 + \frac{1}{2}a_3(1-3\varepsilon) = \frac{3}{2}(1-\varepsilon),$$
$$\frac{2}{a_2} = 4$$
$$\frac{2}{a_3} = 2.$$

Plugging these into (1.94), we get

$$4Q_{ijk}\langle \hat{A}_{ij}, \nabla_k^\perp \nu_1\rangle \leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + (1-\varepsilon)\frac{3}{2}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2$$
$$+ 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + 4f|\nabla^\perp \nu_1|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2,$$

as claimed. □

Finally, we combine the conclusions of Lemmas 1.16, 1.17, and 1.18 to get our desired estimate.

**Lemma 1.19** (Gradient term estimate)**.** *Suppose either* $n \geq 8$, $\frac{1}{n} < c_0 \leq \frac{4}{3n}$ *and* $0 < \delta \leq \frac{1}{5n-8}$*; or* $\frac{1}{n} < c_0 \leq \frac{3(n+1)}{2n(n+2)} - \varepsilon_0$*, and* $0 < \delta \leq \min\{\frac{1}{2}, \frac{2n(n+2)}{3(n-1)}\varepsilon_0\}$*. Then, in either case,*

$$4Q_{ijk}\langle \hat{A}_{ij}, \nabla_k^\perp \nu_1\rangle \leq 2|\nabla^\perp \hat{A}|^2 + 2(1-\delta)\frac{|\hat{A}|^2}{f}(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2).$$

*Proof.* First suppose $n \geq 8$, $\frac{1}{n} < c_0 \leq \frac{4}{3n}$ and $0 < \delta \leq \frac{1}{5n-8}$. Expanding $|\nabla^\perp \hat{A}|^2$ using (1.25) and using the inequality (1.84) in Lemma 1.16 gives us

$$\begin{aligned}2|\nabla^\perp \hat{A}|^2 &= 2|\hat{\nabla}^\perp \hat{A}|^2 + 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 \\ &\geq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + \frac{4n-10}{n+2}|\mathring{h}|^2|\nabla^\perp \nu_1|^2 + \frac{6(n-1)}{n+2}|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{6(n-1)}{n+2} f|\nabla^\perp \nu_1|^2.\end{aligned}$$

Multiplying inequality (1.86) in Lemma 1.17 by $(1-\delta)$ and using that $1-\delta \geq \frac{1}{2}$ on the coefficient of $|\hat{A}|^2|\nabla^\perp \nu_1|^2$ gives

$$2(1-\delta)\frac{|\hat{A}|^2}{f}\big(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2\big) \geq (1-\delta)\frac{5n-8}{3(n-1)}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 + \frac{5n-8}{n+2}|\hat{A}|^2|\nabla^\perp \nu_1|^2.$$

Putting these together, we get

$$\begin{aligned}&2|\nabla^\perp \hat{A}|^2 + 2(1-\delta)\frac{|\hat{A}|^2}{f}(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2) \\ &\quad \geq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + (1-\delta)\frac{5n-8}{3(n-1)}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\ &\qquad + \frac{11n-14}{n+2}|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{6(n-1)}{(n+2)} f|\nabla^\perp \nu_1|^2 + \frac{4n-10}{n+2}|\mathring{h}|^2|\nabla^\perp \nu_1|^2.\end{aligned}$$

On the other hand, the first estimate (1.89) of Lemma 1.18 gives us that

$$\begin{aligned}4Q_{ijk}\langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle &\leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + \frac{5n-9}{3(n-1)}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\ &\quad + 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + \frac{3(n-1)}{n-3} f|\nabla^\perp \nu_1|^2 + \frac{2(n+2)}{n+3}|\mathring{h}|^2|\nabla^\perp \nu_1|^2.\end{aligned}$$

Therefore, it only remains to compare the coefficients of like terms in the two inequalities above. For the coefficients of $\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2$, we need

$$\frac{5n-9}{3(n-1)} \leq (1-\delta)\frac{5n-8}{3(n-1)} \iff \delta \leq \frac{1}{5n-8}.$$

Comparing the coefficients of the remaining terms implies we need

$$2n+4 \leq 11n-14 \iff 2 \leq n,$$
$$n+2 \leq 2(n-3) \iff 8 \leq n,$$
$$2(n+2)^2 \leq (4n-10)(n+3) \iff 19 \leq n(n-3).$$

Each of these inequalities is true if $n \geq 8$ completing the proof for the first case.

Now suppose $\frac{1}{n} < c_0 < \frac{3(n+1)}{2n(n+2)}$ and $0 < \delta < \min\{\frac{1}{2}, \frac{2n(n+2)}{3(n-1)}\varepsilon_0\}$. Arguing as before, this time using (1.85) in Lemma 1.16 and (1.87) in Lemma 1.17 yields

$$\begin{aligned}
&2|\nabla^\perp \hat{A}|^2 + 2(1-\delta)\frac{|\hat{A}|^2}{f}(|\nabla^\perp A|^2 - c_0|\nabla^\perp H|^2) \\
&\quad \geq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + (1-\delta)\frac{3}{2}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\
&\qquad + 7|\hat{A}|^2|\nabla^\perp \nu_1|^2 + 4f|\nabla^\perp \nu_1|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2.
\end{aligned}$$

Note we again used $\delta \leq \frac{1}{2}$ to simplify the coefficient of $|\hat{A}|^2|\nabla^\perp \nu_1|^2$. On the other hand, by (1.90), we have

$$\begin{aligned}
4Q_{ijk}\langle \hat{A}_{ij}, \nabla^\perp_k \nu_1\rangle &\leq 2|\langle \nabla^\perp \hat{A}, \nu_1\rangle|^2 + (1-\varepsilon)\frac{3}{2}\frac{|\hat{A}|^2}{f}|\langle \nabla^\perp \mathring{A}, \nu_1\rangle|^2 \\
&\quad + 2|\hat{A}|^2|\nabla^\perp \nu_1|^2 + 4f|\nabla^\perp \nu_1|^2 + 2|\mathring{h}|^2|\nabla^\perp \nu_1|^2,
\end{aligned}$$

where recall $\varepsilon = \frac{2n(n+2)}{3(n-1)}\varepsilon_0$. By assumption, $\delta \leq \varepsilon$, and this completes the proof of the lemma. □

### 1.3.3. Concluding Argument

We now complete the proof of Theorem 1.1. Let $\delta$ be sufficiently small so that each of our above lemmas holds. Taking $\delta = \min\{\frac{1}{5n-8}, \frac{2n(n+2)}{3(n-1)}\varepsilon_0\}$ suffices. We begin by splitting off the desired nonpositive term in the evolution equation

$$
\begin{aligned}
\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f} &= \frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - |\hat{A}|^2\frac{1}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle \\
&= 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle - \delta\frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f \\
&\qquad + \frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - (1-\delta)\frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f.
\end{aligned} \tag{1.95}
$$

By the inequalities of Lemmas 1.14 and 1.19, the sum of terms on the second line of (1.95) are nonpositive:

$$
\begin{aligned}
&\frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - (1-\delta)\frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f \\
&\qquad = \frac{1}{f}\Big(2|\langle\hat{A}_{ij},\hat{A}_{kl}\rangle|^2 + 2|\hat{R}^{\perp}|^2 + 2|R^{\perp}_{ij}(\nu_1)|^2\Big) \\
&\qquad\quad - \frac{1}{f}\Big(2(1-\delta)\frac{|\hat{A}|^2}{f}(c_0|\langle A_{ij},H\rangle|^2 - |\langle A_{ij},A_{kl}\rangle|^2 - |R^{\perp}_{ij}|^2)\Big) \\
&\qquad\quad + \frac{1}{f}\Big(4Q_{ijk}\langle\hat{A}_{ij},\nabla^{\perp}_k\nu_1\rangle - 2|\nabla^{\perp}\hat{A}|^2 - 2(1-\delta)\frac{|\hat{A}|^2}{f}(|\nabla^{\perp}A|^2 - c_0|\nabla^{\perp}H|^2)\Big) \\
&\qquad \le 0.
\end{aligned}
$$

Thus we have finally obtained our initial claim (1.64) at the beginning of this section:

$$
\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f} \le 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle - \delta\frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f. \tag{1.96}
$$

Recall that $(\frac{\partial}{\partial t}-\Delta)f$ is nonnegative at each point in space-time. Let $\sigma=\delta$. We compute that

$$
\begin{aligned}
\Big(\frac{\partial}{\partial t}-\Delta\Big)f^{1-\sigma} &= (1-\sigma)f^{-\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f + \sigma(1-\sigma)f^{-\sigma-1}|\nabla f|^2 \\
&\ge (1-\sigma)f^{-\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f.
\end{aligned}
$$

Then, making use of (1.43), we have

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f^{1-\sigma}} &= \frac{1}{f^{1-\sigma}}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - |\hat{A}|^2\frac{1}{f^{2-2\sigma}}\Big(\frac{\partial}{\partial t}-\Delta\Big)f^{1-\sigma} + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f^{1-\sigma}},\nabla\log f^{1-\sigma}\Big\rangle\\
&\le \frac{1}{f^{1-\sigma}}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - |\hat{A}|^2\frac{1}{f^{2-2\sigma}}(1-\sigma)f^{-\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f\\
&\quad + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f^{1-\sigma}},\nabla\log f^{1-\sigma}\Big\rangle\\
&= f^{\sigma}\Big(\frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - \frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f\Big) + \sigma\frac{|\hat{A}|^2}{f^2}f^{\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f\\
&\quad + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f^{1-\sigma}},\nabla\log f^{1-\sigma}\Big\rangle.\end{aligned}$$

Now again by (1.43) and (1.96)

$$\begin{aligned}\frac{1}{f}\Big(\frac{\partial}{\partial t}-\Delta\Big)|\hat{A}|^2 - \frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f &= \Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f} - 2\Big\langle\nabla\frac{|\hat{A}|^2}{f},\nabla\log f\Big\rangle\\
&\le -\delta\frac{|\hat{A}|^2}{f^2}\Big(\frac{\partial}{\partial t}-\Delta\Big)f.\end{aligned}$$

Therefore, since $\sigma=\delta$,

$$\begin{aligned}\Big(\frac{\partial}{\partial t}-\Delta\Big)\frac{|\hat{A}|^2}{f^{1-\sigma}} &\le -\delta\frac{|\hat{A}|^2}{f^2}f^{\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f + \sigma\frac{|\hat{A}|^2}{f^2}f^{\sigma}\Big(\frac{\partial}{\partial t}-\Delta\Big)f\\
&\quad + 2\Big\langle\nabla\frac{|\hat{A}|^2}{f^{1-\sigma}},\nabla\log f^{1-\sigma}\Big\rangle\\
&= 2\Big\langle\nabla\frac{|\hat{A}|^2}{f^{1-\sigma}},\nabla\log f^{1-\sigma}\Big\rangle.\end{aligned}$$

Hence by the maximum principle, there exists a constant $C$ depending only upon the initial manifold $M_0$ such that $|\hat{A}|^2 \le Cf^{1-\sigma}$ on $M\times[0,T)$. Since $f\le c_0|H|^2$, this implies $|\hat{A}|^2\le C|H|^{2-2\sigma}$. This completes the proof of Theorem 1.1.

# Chapter 2

# Noncollapsing and Classification of Singularity Models

Having established a new a priori estimate in Chapter 1, we now turn our focus to singularity models of the higher codimension mean curvature flow. In this chapter, we prove a noncollapsing result for ancient, weakly convex, uniformly two-convex solutions of the mean curvature satisfying two derivative estimates. The result will have several consequences, but our specific goal is Corollary 2.4. The work in this chapter was completed in [39].

The uniqueness of ancient solutions that are uniformly two-convex and noncollapsed (in the sense of Sheng and Wang [47]) has been studied in the noncompact case by Brendle and Choi [12, 13] and in the compact case by Angenent, Daskalopoulos, and Sesum [3, 4]. For applications to higher codimension, we replace the noncollapsed assumption with two derivative estimates. Our main theorem in this chapter is:

**Theorem 2.1.** *Suppose $n \geq 3$ and let $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$ be an $n$-dimensional, complete ancient solution to the mean curvature flow in $\mathbb{R}^{n+1}$ that is weakly convex, uniformly two-convex, and satisfies pointwise derivative estimates: $|\nabla A| \leq \gamma_1 |H|^2$ and $|\nabla^2 A| \leq \gamma_2 |H|^3$. Then the solution is noncollapsed.*

Note that by work of Haslhofer and Kleiner [29] mean-convex ancient solutions of the mean curvature flow that are noncollapsed must in fact be weakly convex. If we combine the main result above with the uniqueness results of [13] and [4], we have the following corollary.

**Corollary 2.2.** *Suppose $n \geq 3$ and let $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$, $t \in (-\infty, 0]$, be an $n$-dimensional, complete, connected ancient solution to the mean curvature flow in $\mathbb{R}^{n+1}$ that is weakly convex, uniformly two-convex, and satisfies pointwise derivative estimates: $|\nabla A| \leq \gamma_1 |H|^2$*

*and* $|\nabla^2 A| \leq \gamma_2 |H|^3$. *Then the solution is either a family of shrinking round spheres, a family of shrinking round cylinders, a translating bowl soliton, an ancient oval, or a static flat hyperplane.*

Broadly, our discussion here fits within an ongoing program aimed at characterizing self-similar and ancient solutions of the Ricci flow and the mean curvature flow. See for example [8, 11, 28, 5, 12, 13, 3, 4], to name only some of the recent related works. Specifically, however, we use the theorem above and its corollary to classify singularity models of the mean curvature flow in higher codimension.

The work of Brendle and Choi [13] shows the only possible blow-up models at the first singular time for closed, *embedded*, two-convex hypersurfaces evolving under the mean curvature flow are the spheres, cylinders, and bowls. The embeddedness assumption ensures that blow-up limits are noncollapsed (see [48, 49] or [29]), in addition to ancient, weakly convex, and uniformly two-convex. Noncollapsing of blow-ups is an incredibly useful assumption, as demonstrated, for example, by the efficient works of Haslhofer and Kleiner [29, 30]. The work of Huisken and Sinestrari [32, 33], however, shows that in higher dimensions noncollapsing is not necessary for the analysis of solutions if one has pointwise derivative estimates and a pinching estimate. For instance, see [5], where, under such assumptions Bourni and Langford proved a uniqueness theorem for translators of the mean curvature flow.

By replacing the noncollapsed assumption in the main result of [13] by an assumption of derivative estimates, we can classify blow-ups models for pinched solutions of the mean curvature flow in higher codimension, where embeddedness is no longer preserved. Our work also applies to immersed solutions in codimension one. Indeed, the blow-ups of closed, *immersed*, two-convex solutions of the mean curvature flow are still ancient, weakly convex, and uniformly two-convex. In fact, by Theorem 5.3 in [33], blow-ups satisfy the cylindrical estimate $|A|^2 \leq \frac{1}{n-1}|H|^2$, which implies weak convexity and uniform two-convexity. By Theorems 6.1 and 6.3 in [33], blow-ups also satisfy $|\nabla A| \leq \gamma_1 |H|^2$ and $|\nabla^2 A| \leq \gamma_2 |H|^3$. Combining these results with the main theorem above allows us to drop the embeddedness assumption of Corollary 1.2 in the work of Brendle and Choi [13].

**Corollary 2.3.** *Let $n \geq 3$. Consider an arbitrary closed, immersed, two-convex hypersurface in $\mathbb{R}^{n+1}$, and evolve it by mean curvature flow. At the first singular time, the only possible blow-up limits are shrinking round spheres, shrinking round cylinders, and translating bowl solitons.*

By replacing two-convexity with a stronger pinching assumption, we can show the above classification holds in higher codimension as well. Namely, we consider closed, $n$-dimensional initial data in $\mathbb{R}^N$ that satisfies a natural curvature pinching condition $|A|^2 < c\,|H|^2$. This kind of pinching was first considered in [2], where the authors showed these inequalities are preserved by the flow if $c \leq \frac{4}{3n}$. Under these pinching conditions, Andrews and Baker [2] and Nguyen [41] have suitably extended to higher codimension many of the main ideas in the important and impactful works of Huisken [31] on the flow of convex hypersurfaces and Huisken and Sinestrari [32, 33] on the flow of two-convex hypersurfaces. For more discussion, see the introduction in Chapter 1.

At present, we are interested in the pinching condition with $c = \frac{1}{n-2}$. If $n \geq 8$, then $\frac{1}{n-2} \leq \frac{4}{3n}$. Suppose $F : M \times [0, T) \to \mathbb{R}^N$ is a smooth, closed, $n$-dimensional solution of the mean curvature flow in $\mathbb{R}^N$ initially satisfying $|A|^2 < \frac{1}{n-2}|H|^2$. In codimension one, this pinching condition implies two-convexity of the hypersurface. It was first studied by Nguyen in [41]. In direct analogy with the work of Huisken and Sinestrari, Nguyen proved the following cylindrical and derivative estimates:

- (Cylindrical Estimate.) For every constant $\eta > 0$, there exists a constant $C_\eta < \infty$, depending only upon initial data, such that the estimate

  $$|A|^2 \leq \Big(\frac{1}{n-1} + \eta\Big)|H|^2 + C_\eta$$

  holds for all $t \in [0, T)$.

- (Derivative Estimates.) There exist constants $\gamma_1, \gamma_2, C_1, C_2 < \infty$, depending only upon the initial data, such that the estimates

$$|\nabla A| \leq \gamma_1 |H|^2 + C_1,$$
$$|\nabla^2 A| \leq \gamma_2 |H|^3 + C_2$$

hold for all $t \in [0, T)$.

In addition to the estimates above, in Chapter 1 we established a new planarity estimate for pinched solutions to the mean curvature flow in higher codimension. We considered a tensor,

$$\hat{A}_{ij} := A_{ij} - \frac{\langle A_{ij}, H\rangle}{|H|^2} H,$$

which consists of the components of the second fundamental form that are orthogonal to the mean curvature vector. Under the pinching assumption, we showed that $\hat{A}$ vanishes if and only if the solution is codimension one. Then we proved the following estimate:

- (Theorem 1.1.) There exists a constant $\sigma > 0$ and a constant $C < \infty$, depending only upon the initial data, such that the estimate

$$|\hat{A}|^2 \leq C|H|^{2-\sigma}$$

holds for all $t \in [0, T)$.

We can include the dimensions $n = 5, 6$, and 7 if we strengthen our pinching assumption. Recall from the previous chapter that

$$c_1 := \min\left\{\frac{4}{3n}, \frac{1}{n-1}\right\}, \quad c_2 := \min\left\{\frac{4}{3n}, \frac{1}{n-2}\right\}, \quad c_n := \min\left\{\frac{4}{3n}, \frac{3(n+1)}{2n(n+2)}\right\}.$$

For $n \geq 5$, the cylindrical estimate holds if $c \leq c_2$ and shows that singularity models are $c_1$-pinched; for $n \geq 2$ the derivative estimates hold if $c \leq \frac{4}{3n}$; and for $n \geq 5$ the planarity estimate holds if $c \leq c_n$. The constants $\frac{4}{3n}$ and $\frac{3(n+1)}{2n(n+2)}$ are technical constants that arise in the proofs of [2] and Theorem 1.1. If $n \leq 4$, then $\frac{4}{3n} \leq c_1$; so only spherical singularities can occur (see [2]).

In particular, all three estimates hold for $n \geq 5$ if $c \leq \tilde{c}_2 := \min\{c_2, c_n\}$. Consequently, the blow-up limits of a $\tilde{c}_2$-pinched flow must be ancient, codimension one, and satisfy the estimates $|A|^2 \leq \frac{1}{n-1}|H|^2, |\nabla A| \leq \gamma_1 |H|^2, |\nabla^2 A| \leq \gamma_2 |H|^3$, precisely as for immersed solutions in codimension one. By Theorem 2.1 above, this again gives the following classification.

**Corollary 2.4.** *Let $n \geq 5$ and $N > n$. Let $\tilde{c}_2 = \min\{\frac{3(n+1)}{2n(n+2)}, \frac{1}{n-2}\}$. Consider a closed, $n$-dimensional solution to the mean curvature flow in $\mathbb{R}^N$ initially satisfying $|A|^2 < \tilde{c}_2 |H|^2$. At the first singular time, the only possible blow-up limits are codimension one shrinking round spheres, shrinking round cylinders, and translating bowl solitons.*

Let us briefly explain the arguments needed to show that Corollaries 2.3 and 2.4 follow from Theorem 2.1. For both immersed two-convex solutions in codimension one and $\tilde{c}_2$-pinched solutions in higher codimension, blow-ups satisfy the cylindrical estimate $|A|^2 \leq \frac{1}{n-1}|H|^2$. Now by the strong maximum principle, a weakly convex ancient solution that is not strictly convex must split off a line. If a blow-up splits off a line, the cylindrical estimate implies the remaining principal curvatures are all equal and hence, by the Schur lemma (since $n \geq 3$), the blow-up must be a family of shrinking cylinders. In the immersed and codimension one setting, if the blow-up is compact, then it is convex, and the original flow must become convex prior to the first singular time. The result of Huiksen [31] then shows the blow-up is a family of shrinking spheres. In the $\tilde{c}_2$-pinched and higher codimension setting, if the blow-up is compact, it is $c_1$-pinched. Then the work of Nguyen [41] shows the original flow is $c_1$-pinched prior to the first singular time, and the work of Andrews and Baker [2] implies the blow-up is a family of shrinking spheres. Note that [31] and [41] preclude the possibility of an ancient oval arising as a blow-up limit at the first singular time.

The remaining case to consider is when the blow-up is noncompact and strictly convex, which is addressed by Theorem 2.1 and the main result of [13]. For the sake of generality, we have replaced cylindrical estimate in our assumptions with an assumption of uniform two-convexity. This way the convexity assumptions in our main theorem match the weaker convexity assumptions in codimension one. To show uniform two-convexity suffices for the conclusions of the theorem, we prove, in Proposition 2.11 below, that weakly convex, uniformly two-convex, ancient solutions satisfying pointwise derivative estimates are in fact $\frac{1}{n-1}$-two-convex. That is, if $\lambda_1$ and $\lambda_2$ denote the smallest two eigenvalues of the second fundamental form and $H$ denotes the scalar mean curvature, then $\lambda_1 + \lambda_2 \geq \frac{1}{n-1}H$.

Finally, let us discuss the proof of the main theorem. Broadly, we will prove it in two steps. In the first step, we will adapt the tools and ideas developed by Huisken and Sinestrari to our ancient and convex setting. We will show that when the ancient solution is strictly convex (and not a family of shrinking round spheres), it has the structure of a long tube with either one or two convex caps attached, depending upon whether the solution is noncompact or compact. Many ideas carry over without much change; some are even a bit simpler in our (surgery-free) setting. The structure theorems give diameter, mean curvature, and uniform convexity estimates independent of time. Once this is known, we can show each time slice of the ancient solution is $\alpha$-noncollapsed for a uniform choice of $\alpha$.

The organization of this chapter is as follows. In Section 2.1, we collect the various notations, definitions, and auxiliary results that will be used in subsequent sections. Because we are working in codimension one in this chapter, the notation will differ from the previous chapter. In Section 2.2, we prove that when the ancient solution is noncompact, it has the structure of a long tube with a convex cap attached. In Section 2.3, we prove a similar structure theorem for when the solution is compact. The work in these three sections follows the pioneering work of Huisken and Sinestrari in [33]. In Section 2.4, we show the ancient solution is noncollapsed. In the Section 2.5, we include additional details for the proof of Proposition 2.11.

## 2.1. Preliminaries

In this section, we give some definitions and auxiliary results for the proof of our main theorem. Many of the statements and proofs in this section are to a certain extent standard after [33] and only require reasonable adaptations of the analogous statements and proofs.

Let us begin with some notation. Let $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$ denote a (possibly immersed) ancient solution of the mean curvature flow. Let $(p,t) \in M \times (-\infty, 0]$ be a spacetime point. Since we are working with a hypersurface, from now on we will let $H$ denote the scalar mean curvature, as opposed to the mean curvature vector. We let $\nu$ denote the outward pointing normal vector and $h = \langle A, -\nu \rangle$ denote the scalar-valued second fundamental form. Let $\lambda_1 \leq \cdots \leq \lambda_n$ denote the principal curvatures. We say the solution is weakly convex if $\lambda_1 \geq 0$; strictly convex if $\lambda_1 > 0$; uniformly convex if $\lambda_1 \geq \beta H$ for some $\beta > 0$; and analogously for the notion of two-convexity. Let $g(t)$ denote the induced metric on $M$ by the immersion $F(\cdot, t)$ and $B_{g(t)}(p, r) \subset M$ the intrinsic ball of radius $r$ centered at $p$. We will also be interested in parabolic neighborhoods. Following the notation introduced on pp.189-190 in [33], we define $P(p, t, r, \theta)$ to be the set of space time points $(q, s)$ such that $q \in B_{g(t)}(p, r)$ and $s \in [t - \theta, t]$. For the purposes of rescaling, we also define

$$\hat{P}(p, t, L, \theta) := P\Big(p, t, \frac{n-1}{H(p,t)} L, \Big(\frac{n-1}{H(p,t)}\Big)^2 \theta\Big).$$

**Lemma 2.5.** *Suppose $n \geq 3$ and let $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$ be an $n$-dimensional, complete ancient solution to the mean curvature flow in $\mathbb{R}^{n+1}$ that is weakly convex and uniformly two-convex. Then the solutions is either a family of shrinking round cylinder, a static hyperplane, or strictly convex and embedded.*

*Proof.* If the convexity is not strict, then by the strong maximum principle, the solution must split a line. It follows that the cross-section is a uniformly convex, complete, ancient solution of the mean curvature flow. Now if the cross-section is not strictly convex, the solution must be flat and hence a static hyperplane. If the cross-section is strictly and uniformly convex, then by work of Hamilton [24] and Huisken-Sinestari [34], the cross-section is a family of shrinking round spheres, and hence the ancient solution is a family of shrinking round cylinders.

In higher dimensions, strict convexity and completeness imply $F(M, t)$ is the boundary of a convex body in $\mathbb{R}^{n+1}$. This follows from an older result of Sacksteder [45], but is slightly easier to establish in our setting. The boundary of a strictly convex body in $\mathbb{R}^{n+1}$ is homeomorphic to either $S^n$ or $\mathbb{R}^n$, and in particular is simply connected. Since $F(\cdot, t)$ is a covering, this implies it is an embedding. □

Both the cylinder and the static flat hyperplane are clearly noncollapsed, so it suffices to establish the main theorem under the additional assumptions of strict convexity and embeddedness. Throughout this section and following sections, we need only consider solutions of the mean curvature flow satisfying the hypotheses.

**Definition 2.6.** An $n$-dimensional ($n \geq 3$) ancient solution $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$ satisfies $(*)$ if:

- The solution is connected, complete, strictly convex, embedded, and uniformly two-convex.
- The solution satisfies the pointwise derivative estimates $|\nabla A| \leq \gamma_1 |H|^2$ and $|\nabla^2 A| \leq \gamma_2 |H|^3$.

In our setting, it is equivalent to work with the level sets $M_t = F(M, t)$, for $t \in (-\infty, 0]$. By abuse of notation, we will sometimes identify the points $p \in M$ and $F(p, t) \in M_t$. Let $\mathcal{M} = \{M_t\}_{t \in (-\infty, 0]}$. For brevity, we will say $\mathcal{M}$ (or sometimes $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$) with these properties satisfies $(*)$.

We will use the following definition to characterize neck regions. Since we are not doing surgery, we will not need precise parametrizations of neck regions, as originally introduced by Hamilton in [27] and used extensively in [33]. One great property of Hamilton's (intrinsic) constant mean curvature foliation of neck regions is that it is canonical. However, in the extrinsic setting, the following definition is perhaps a bit simpler.

**Definition 2.7.** Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a convex hypersurface. Given constants $\varepsilon, L > 0$, we say a point $p \in M$ lies at the center of an $(\varepsilon, L)$-neck if, after rescaling so that $(n-1)H(p)^{-1} = 1$, there exist an embedded round cylinder $\Sigma := S^{n-1} \times [-L, L] \subset \mathbb{R}^{n+1}$ (of radius 1) and a function $u : \Sigma \to \mathbb{R}$, with the following properties:

- We have $\{x + u(x)\nu_\Sigma(x) : x \in \Sigma\} \subset F(M)$ and $||u||_{C^{10}(\Sigma)} \leq \varepsilon$.
- $F(p) \in \{x + u(x)\nu_\Sigma(x) : x \in S^{n-1} \times \{0\}\}$, i.e. the point $p$ lies on the central sphere.

Let $N := F^{-1}(\{x + u(x)\nu_\Sigma(x) : x \in \Sigma\})$. We will say $N$ is an $(\varepsilon, L)$-neck and has length $2L$. Up to a choice of orientation, there exists a unit vector $\omega \in S^n$ which is tangent to the axis of $\Sigma$. This defines a height function $y : M \to \mathbb{R}$ by $y(q) = \langle F(q) - F(p), \omega\rangle$. The sets $S_y := \{q \in N : y(q) = y\}$ for $y \in [-L, L]$ are the cross-sectional spheres of $N$. The axis of the neck $N$ refers to either of the two unit vectors parallel to the axis of $\Sigma$.

In the terminology introduced in Section 3 of [33], Definition 2.7 defines a "geometric neck"; i.e. a neck that is parametrized by a cylinder (in this case, as a graph). This ensures the restriction of the embedding $F$ to the region $N \subset M$ is close to the standard embedding of a cylinder into $\mathbb{R}^{n+1}$. For the detection of necks however, it is much simpler to check if the curvature is close to that of a cylinder. These ideas go back to Hamilton's work on necks in [27]. We will use the following proposition, which is a direct consequence of Proposition 3.5 in [33].

**Proposition 2.8.** *Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a convex hypersurface. Given constants $\varepsilon_0 \in (0, \frac{1}{n})$ and $L \geq 100$, there exists $\varepsilon_1 \in (0, \varepsilon_0)$, depending only upon $n$, $L$, and $\varepsilon_0$, such that the following holds. Suppose $p \in M$ satisfies:*

- *$\lambda_1(p) \leq \varepsilon_1 H(p)$ and $\lambda_n(p) - \lambda_2(p) \leq \varepsilon_1 H(p)$;*
- *For every $q \in B_g(p, (L+10)\frac{n-1}{H(p)})$, $\sum_{k=1}^{8} H(p)^{-k-1}|\nabla^k h(q)| \leq \varepsilon_1$.*

*Then $p$ lies at the center of an $(\varepsilon_0, L)$-neck $N$.*

In the language of [33], a point $p \in M$ that satisfies the two properties of the above proposition is said to lie at the center of an $(\varepsilon_1, L+10)$-extrinsic "curvature neck". The above proposition shows there is little difference between curvature necks and the geometric necks defined above.

Next, we give two standard lemmas concerning the control of curvature. Because these results are standard consequences of the pointwise derivative estimates, we omit the proofs. The first shows that the curvature at one point controls the curvature of all points in a suitable parabolic neighborhood (cf. Lemma 6.6 in [33]).

**Lemma 2.9.** *Suppose $\mathcal{M}$ satisfies $(*)$. Let $(p_0, t_0)$ be a spacetime point. Then there exists a constant $\hat{r} := \hat{r}(n, \gamma_1, \gamma_2) > 0$ such that for every $(p, t) \in \hat{P}(p_0, t_0, \hat{r}, \hat{r}^2)$ we have*

$$\frac{1}{4}H(p_0, t_0) \leq H(p, t) \leq 4H(p_0, t_0).$$

Interior estimates now give pointwise estimates for all higher order derivatives of the second fundamental form (cf. Corollary 6.4 in [33]).

**Lemma 2.10.** *Suppose $\mathcal{M}$ satisfies $(*)$. For all nonnegative integers $k, l$ there exist constants $\gamma_{k,l} := \gamma_{k,l}(n, \gamma_1, \gamma_2) < \infty$ such that the pointwise estimates,*

$$\Big|\frac{\partial}{\partial t}^l \nabla^k h(p_0, t_0)\Big| \leq \gamma_{k,l} H^{2l+k+1}(p_0, t_0),$$

*hold for every $(p_0, t_0) \in M \times (-\infty, 0]$.*

Each of our model solutions (the spheres, cylinders, bowls, and ovals) is $\frac{1}{n-1}$-two-convex. With the higher derivative estimates established, we can now show that $\mathcal{M}$ is $\frac{1}{n-1}$-two-convex. Several steps in our proof are inspired by similar arguments given in [15].

**Proposition 2.11.** *Suppose $\mathcal{M}$ satisfies $(*)$. Then $\mathcal{M}$ is $\frac{1}{n-1}$-two-convex.*

*Proof.* Recall that $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of the second fundamental form, $h$, with multiplicity. Let

$$\beta := \inf_{(p,t)\in M\times(-\infty,0]} \frac{\lambda_1(p,t) + \lambda_2(p,t)}{H(p,t)}.$$

We have assumed $\mathcal{M}$ is uniformly two-convex, so $0 < \beta \leq \frac{1}{n-1}$. Our goal is to show $\beta = \frac{1}{n-1}$.

*Step 1:* To that end, consider a sequence of points $(p_j, t_j) \in M \times (-\infty, 0]$ such that $(\lambda_1(p_j, t_j) + \lambda_2(p_j, t_j))/H(p_j, t_j) \to \beta$ as $j \to \infty$. Let $Q_j = (n-1)H(p_j, t_j)^{-1}$ and let $\hat{r}$ be the constant appearing in Lemma 2.9. Let $F : M \times (-\infty, 0] \to \mathbb{R}^{n+1}$ be a parametrization of our solution $\mathcal{M}$. For each $j$, we first consider the restriction of $F$ to the parabolic neighborhood $\hat{P}(p_j, t_j, \hat{r}, \hat{r}^2)$. Next we perform a parabolic rescaling and spacetime translation to define solutions $F_j$ by

$$F_j(p, \tau) = \frac{1}{Q_j}\big[F\big(p, Q_j^2\tau + t_j\big) - F(p_j, t_j)\big].$$

Let $\mathcal{M}^{(j)}$ denote the resulting flow. Let $g_j, h_j$, and $H_j$ denote the metric, second fundamental form, and mean curvature of $\mathcal{M}^{(j)}$. The solution $F_j$ is defined in the parabolic neighborhood $P_j(\hat{r}) := P(p_j, 0, \hat{r}, \hat{r}^2) = B_{g_j(0)}(p_j, \hat{r}) \times [-\hat{r}^2, 0]$. By construction, $H_j(p_j, 0) = n - 1$ and $F_j(p_j, 0) = 0 \in \mathbb{R}^{n+1}$. Lemmas 2.9 and 2.10 imply that

$$\sup_{P_j(\hat{r})} \Big|\frac{\partial}{\partial t}^l \nabla^k h_j\Big| \leq C(k, l, n, \gamma_1, \gamma_2).$$

This gives local uniform control in any $C^k$ norm over solutions $F_j$ in the parabolic neighborhoods $P_j(\hat{r})$. Standard compactness results (see, for example, the next chapter) then imply the solutions $F_j$ converge smoothly and uniformly on compact subsets of $P_j(\hat{r})$ to a locally defined limit flow $\hat{\mathcal{M}}$ defined in a parabolic neighborhood $\Omega \times [-T, 0]$ for some $T > 0$. We may assume $\Omega$ is a smooth bounded domain in $\mathbb{R}^n$ and that the points $p_j$ converge to $p \in \Omega$. Let us denote the limiting geometric quantities by $g$, $h$, and $H$.

*Step 2:* We now analyze the limit flow in $\Omega \times [-T, 0]$. Let $u(q, t) := (\lambda_1 + \lambda_2 - \beta H)(q, t)$. The limit is weakly convex. Moreover, $u \geq 0$ and our choice of points $(p_j, t_j)$ implies that $u(p, 0) = 0$. By the strong maximum principle, these conditions imply $u$ vanishes identically in $\Omega \times (-T, 0)$. To see this, consider the $(0, 2)$-tensor $U_{ik} := h_{ik} - \frac{\beta}{2} H g_{ik}$. The sum of the two smallest eigenvalues of $U$ is $u$. We argue by contradiction: suppose for some $\tau \in (-T, 0)$, there exists a point $q_0$ such that $u(q_0, \tau) > 0$. Recalling the evolution equations for $H$ and $h^i_k$, we have

$$\frac{\partial}{\partial t} U^i_k = \Delta U^i_k + |A|^2 U^i_k.$$

Since $u(q_0,\tau) > 0$ and $u(q,\tau) \geq 0$ for all $q \in \Omega$, we can find a smooth nonnegative function $f_0$ defined on $\Omega$ such that $f_0(q) \leq u(q,\tau)$ for all $q \in \Omega$, $f_0(q_0) \geq \frac{1}{2}u(q_0,\tau)$, and $f(q) = 0$ for all $q \in \partial\Omega$. Now let $f$ be a solution to the heat equation $\frac{\partial}{\partial t}f = \Delta f$ with initial condition $f(q,\tau) = f_0(q)$ and boundary condition $f(q,t) = 0$ for all $q \in \partial\Omega$. Since $f_0(q_0) > 0$, the strict maximum principle for scalar equations implies $f > 0$ in $\Omega \times (\tau, 0]$. Moreover, the tensor $\tilde{U}_{ik} := U_{ik} - \frac{f}{2}g_{ik}$ satisfies the equation

$$\begin{aligned}\frac{\partial}{\partial t}\tilde{U}^i_k &= \Delta\tilde{U}^i_k + |A|^2\tilde{U}^i_k + \frac{1}{2}|A|^2 f\delta^i_k \\ &\geq \Delta\tilde{U}^i_k + |A|^2\tilde{U}^i_k.\end{aligned}$$

The sum of the first two eigenvalues of $\tilde{U}$ is $u - f$. At the initial time $\tau$, $\tilde{U}$ is weakly two-convex since by construction $u(\cdot,\tau) - f_0 \geq 0$. By the weak maximum principle for tensors, weak two-convexity of $\tilde{U}$ must hold in $\Omega \times [\tau, 0]$. However, in this case we conclude $u(p,0) \geq f(p,0) > 0$, a contradiction.

*Step 3:* We have shown that $\lambda_1 + \lambda_2 \equiv \beta H$ in $\Omega \times (-T, 0)$. Fix a time $\tau \in (-T, 0)$ and let us only consider $g, h$ and $H$ at the time $\tau$. For each $q \in \Omega$, let us consider the set of orthonormal two-frames

$$E_q := \{\{e_1, e_2\} \subset T_q\Omega : |e_1| = |e_2| = 1,\ \langle e_1, e_2\rangle = 0,\ h(e_1,e_1) + h(e_2,e_2) = \lambda_1(q) + \lambda_2(q)\}.$$

The strict maximum principle for tensors implies that the set $E_q$ is invariant under parallel transport (with respect to $g(\tau)$). The proof follows directly from the results of Chapter 9 of [7]. For the convenience of the reader, we include the details of this argument in the Section 2.5. We next construct a parallel subbundle of $T\Omega$ out of the eigenspaces of $h$. To that end, for each $q \in \Omega$, consider the eigenspaces $V_{1,q} := \ker(h - \lambda_1 g)$ and $V_{2,q} := \ker(h - \lambda_2 g)$. Define a vector space $\tilde{E}_q := V_{1,q} + V_{2,q}$. Note it is possible $\lambda_1(q) = \lambda_2(q)$, in which case $\tilde{E}_q = V_{1,q}$.

**Claim:** $v \in \tilde{E}_q$ if and only if there exists $\{e_1, e_2\} \in E_q$ such that $v \in \operatorname{span}\{e_1, e_2\}$.

We will finish the proof of the proposition assuming the claim and then give a proof of the claim. An immediate consequence of the claim is that $\tilde{E}_q$ is invariant under parallel transport. In particular, the dimension of the vector spaces $\tilde{E}_q$ is constant and so $\tilde{E} := \bigcup_{q\in\Omega} \tilde{E}_q$ is a parallel vector subbundle of $T\Omega$. The classical theorem of de Rham implies $(\Omega, g(\tau))$ splits locally as an isometric product $\Omega_1^k \times \Omega_2^{n-k}$ of smaller dimensional spaces, where $k$ denotes the rank of the bundle $\tilde{E}$. Since the embedding of $(\Omega, g(\tau))$ is strictly two-convex, it has strict positive isotropic curvature as an intrinsic manifold. This implies (see [7]) the only possible splittings are $k = n-1$ or $k = n$ (no splitting).

If $k = n-1$, then the eigenvector $e_n$, corresponding to the largest eigenvalue of $h$, is a parallel vector. Because the vector is parallel, $\mathrm{Ric}(e_n, e_n) = 0$. On the other hand, the Gauss equation implies that $\mathrm{Ric}(e_n, e_n) = \lambda_n(H - \lambda_n)$, which is not zero. Thus $k = n-1$ cannot occur. If $k = n$, then $h$ has only $\lambda_1$ and $\lambda_2$ as eigenvalues (with multiplicity). Since $\beta \le 1$ and $\lambda_1 \ge 0$, we have

$$\lambda_1 + \lambda_2 = \beta H \le \beta\lambda_1 + \beta(n-1)\lambda_2 \le \lambda_1 + \beta(n-1)\lambda_2.$$

Therefore, $(1 - \beta(n-1))\lambda_2 \le 0$. Since $\lambda_2 > 0$, we conclude $\beta \ge \frac{1}{n-1}$, as was to be shown.

□

Here is the proof of the linear algebra claim necessary to complete the proof of the proposition.

*Proof of Claim.* For simplicity, since the claim does not depend upon $q$, let us suppress it in our notation. First we consider the simpler case $\lambda_1 = \lambda_2$. In this case $\dim V_1 \ge 2$. The identity defining $E$ becomes $h(e_1, e_1) + h(e_2, e_2) = 2\lambda_1$, which implies $h(e_1, e_1) = h(e_2, e_2) = \lambda_1$. This shows $\mathrm{span}\{e_1, e_2\} \subset \tilde{E}$. Since $\dim V_1 \ge 2$, the converse of the claim is also clear.

Now suppose $\lambda_2 > \lambda_1$. In this case, $\dim V_1 = 1$, the spaces $V_1$ and $V_2$ are orthogonal, and $h$ has a unique $\lambda_1$-eigenvector, which we denote by $e_1$ for the remainder of the proof. First suppose $v \in \tilde{E}$. Let $v_1 = \langle v, e_1 \rangle$ and write $v = \tilde{v} + v_1 e_1$, where $\tilde{v} \in V_2$. If $\tilde{v} = 0$, then we can take $e_2$ to be any $\lambda_2$-eigenvector and we will have $v \in \mathrm{span}\{e_1, e_2\}$. If not, then take $e_2 = |\tilde{v}|^{-1}\tilde{v}$. Again, this gives $v \in \operatorname{span}\{e_1, e_2\}$ and in either case $\{e_1, e_2\} \in E$.

Now for the other direction, suppose $\{v, w\} \in E$. Then $|v|^2 = |w|^2 = 1$, $\langle v, w\rangle = 0$, and $h(v, v) + h(w, w) = \lambda_1 + \lambda_2$. As above, let $v_1 = \langle v, e_1\rangle$ and $w_1 = \langle w, e_1\rangle$ and write $v = \tilde{v} + v_1 e_1$ and $w = \tilde{w} + w_1 e_1$. We must show $\tilde{v}, \tilde{w} \in V_2$. The case when either $\tilde{v}$ or $\tilde{w}$ is zero is clear, so we may assume both are nonzero. Our assumptions on $v$ and $w$ imply $1 = |\tilde{v}|^2 + v_1^2 = |\tilde{w}|^2 + w_1^2$ and $v_1 w_1 + \langle \tilde{v}, \tilde{w}\rangle = 0$. We compute

$$\begin{aligned} h(\tilde{v}, \tilde{v}) - \lambda_2|\tilde{v}|^2 + h(\tilde{w}, \tilde{w}) - \lambda_2|\tilde{w}|^2 &= h(v, v) + h(w, w) - \lambda_1(v_1^2 + w_1^2) - \lambda_2(|\tilde{v}|^2 + |\tilde{w}|^2) \\ &= \lambda_1 + \lambda_2 - \lambda_1(v_1^2 + w_1^2) - \lambda_2(|\tilde{v}|^2 + |\tilde{w}|^2) \\ &= \lambda_1(v_1^2 + |\tilde{v}|^2) + \lambda_2(w_1^2 + |\tilde{w}|^2) - \lambda_1(v_1^2 + w_1^2) - \lambda_2(|\tilde{v}|^2 + |\tilde{w}|^2) \\ &= (\lambda_2 - \lambda_1)(w_1^2 - |\tilde{v}|^2). \end{aligned}$$

Now $\lambda_2 = \inf\{|x|^{-2}h(x, x) : \langle x, e_1\rangle = 0,\, x \neq 0\}$. Therefore $h(\tilde{v}, \tilde{v}) \geq \lambda_2|\tilde{v}|^2$ and $h(\tilde{w}, \tilde{w}) \geq \lambda_2|\tilde{w}|^2$. If either of these inequalities is strict, then $(\lambda_2 - \lambda_1)(w_1^2 - |\tilde{v}|^2) > 0$ and hence $w_1^2 > |\tilde{v}|^2$. Similarly, by symmetry, $v_1^2 > |\tilde{w}|^2$. However, we also have $v_1 w_1 + \langle \tilde{v}, \tilde{w}\rangle = 0$, which implies $v_1^2 w_1^2 = \langle \tilde{v}, \tilde{w}\rangle^2 \leq |\tilde{v}|^2|\tilde{w}|^2$. Thus, neither inequality can be strict and so $|\tilde{v}|^{-2}h(\tilde{v}, \tilde{v}) = |\tilde{w}|^{-2}h(\tilde{w}, \tilde{w}) = \lambda_2$. We conclude that both $\tilde{v}, \tilde{w}$ are in $V_2$ and hence $\operatorname{span}\{v, w\} \subset \tilde{E}$. $\square$

The final four results of this section concern the detection of necks. The first is a rephrasing of Theorem 7.14 in [33] with a variation on its proof.

**Theorem 2.12** (Theorem 7.14 in [33])**.** *Assume $n \geq 2$. Given constants $c_1, \eta_0 > 0$, we can find constants $\hat{a}$ and $\hat{b}$ with the following property. Let $F : M \to \mathbb{R}^{n+1}$ be a complete, connected, immersed hypersurface in $\mathbb{R}^{n+1}$ with $H > 0$. Suppose that $p_0 \in M$. Moreover, suppose that $|\nabla H| \leq c_1 H^2$ and $\lambda_1 \geq \eta_0 H$ for each point in the set $U := \{p \in M : d_g(p_0, p) < \hat{a}H(p_0)^{-1}, H(p) > \hat{b}^{-1}H(p_0)\}$. Then $U = M$; in particular, M is compact.*

*Proof.* Choose $\hat{a}$ so that $\frac{\eta_0}{c_1}\log(1 + \frac{c_1\hat{a}}{100}) > 2\pi$. Let $\hat{b} := 2 + 100c_1\hat{a}$. Then $U$ is open and nonempty.

Let $f(p) := \langle F(p) - F(p_0), \omega\rangle$ where $\omega := -\nu(p_0)$. Clearly, $f$ has a strict local minimum at

$p_0$. For each $s > 0$, we denote by $U_s$ the connected component of $\{f < s\}$ that contains the point $p_0$. Let $s_*$ denote the supremum of all $s$ with the property that $U_s$ contains no critical points of $f$ other than $p_0$.

Clearly, $U_s \subset U$ if $s > 0$ is sufficiently small. Let $s_0$ denote the supremum of all $s \in (0, s_*]$ with the property that $U_s \subset U$. We claim that $U_s \subset \{p \in M : d_g(p_0, p) \leq \frac{1}{2}\hat{a}H(p_0)^{-1}, H(p) \geq 2\hat{b}^{-1}H(p_0)\}$ for $s \in (0, s_0)$. In other words, the sets $U_s$ are contained in $U$ until the next height at which $f$ has a critical point. To see this, fix a real number $s \in (0, s_0)$ and an arbitrary point $x \in U_s$. Let $\gamma(r)$ denote the integral curve of the vector field $-\frac{\nabla f}{|\nabla f|} = -\frac{\omega^\top}{|\omega^\top|}$ starting at $p$. Clearly, $\gamma$ converges to $p_0$ since $p_0$ is the only critical point of $f$ in $U_s$. Note that $\gamma$ is parametrized by arc length. We assume that $\gamma(r)$ is defined for $r \in [0, r_0)$ and satisfies $\gamma(0) = p$ and $\gamma(r) \to p_0$ as $r \to r_0$. Since the path $\gamma$ is contained in $U$, we know that $|\nabla H| \leq c_1 H^2$ and $\lambda_1 \geq \eta_0 H$ at each point on $\gamma$. Integrating the gradient estimate gives

$$H(\gamma(r)) \geq \frac{H(p_0)}{1 + c_1(r_0 - r)H(p_0)}.$$

Uniform convexity implies

$$\frac{d}{dr}\langle \nu, \omega \rangle = -\frac{h(\omega^\top, \omega^\top)}{|\omega^\top|} \leq -\eta_0 H|\omega^\top| = -\eta_0 H\sqrt{1 - \langle \omega, \nu \rangle^2}.$$

This gives

$$\frac{d}{dr}\arcsin(\langle \omega, \nu \rangle) \leq -\eta_0 H \leq -\eta_0 \frac{H(p_0)}{1 + c_1(r_0 - r)H(p_0)}.$$

Since $\arcsin(\langle \omega, \nu \rangle)$ takes values in the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$, it follows that

$$\int_0^{r_0} \eta_0 \frac{H(p_0)}{1 + c_1(r_0 - r)H(p_0)}\, dr \leq 2\pi.$$

This gives $\frac{\eta_0}{c_1}\log(1 + c_1 r_0 H(p_0)) \leq 2\pi$. In view of our choice of $\hat{a}$, we obtain $r_0 H(p_0) \leq \frac{1}{2}\hat{a}$. This implies $d_g(p_0, p) \leq r_0 \leq \frac{1}{2}\hat{a}H(p_0)^{-1}$ and, in view of our choice of $\hat{b}$, $H(p) \geq H(p_0)(1 + c_1 r_0 H(p_0))^{-1} \geq 2\hat{b}^{-1}H(p_0)$. We have shown that $U_s \subset \{p \in M : d_g(p_0, p) \leq \frac{1}{2}\hat{a}H(p_0)^{-1}, H(p) \geq$

$2\hat{b}^{-1}H(p_0)\} \subset U$ for all $s \in (0, s_0)$. Consequently, $s_0 = s_*$, as originally claimed.

For each $s \in (0, s_*)$, $U_s$ is diffeomorphic to $B^n$. Moreover, in view of the maximality of $s_*$, there exists a sequence $s_j \nearrow s_*$ and a sequence of points $q_j \in \partial U_{s_j}$ such that $|\nabla f(q_j)| \to 0$. Therefore, $q_j \to q_*$ where $q_*$ is a critical point of $f$. Clearly, $q_* \in U$ and $f(q_*) = s_*$. Since $q_* \in U$, the second fundamental form at $q_*$ is positive definite. This implies that the Hessian of $f$ at $q_*$ is either positive or negative definite and therefore $f$ has a strict local maximum or minimum at $q_*$. Since $q_j \to q_*$ and $f(q_j) \nearrow f(q_*)$, $f$ cannot have a strict local minimum at $q_*$. Consequently, $f$ must have a strict local maximum at $q_*$. For each $s \in (0, s_*)$ we denote by $V_s$ the connected component of $\{f > s\}$ which contains the point $q_*$. Since $q_j \to q_*$, we conclude that $q_j \in \partial V_{s_j}$ for $j$ large. If $j$ is sufficiently large, then $V_{s_j}$ is diffeomorphic to $B^n$.

To summarize, both $\partial U_{s_j}$ and $\partial V_{s_j}$ are connected components of $\{f = s_j\}$ (here we use that $n \geq 2$) and both $\partial U_{s_j}$ and $\partial V_{s_j}$ contain the point $q_j$. Therefore $\partial U_{s_j} = \partial V_{s_j}$. Since $M$ is connected, we conclude $\bar{U}_{s_j} \cup \bar{V}_{s_j} = M$. Since $\operatorname{diam}_g(V_{s_j}) \to 0$, it follows that $\bigcup_j U_{s_j}$ is dense in $M$. Since $\bigcup_j U_{s_j}$ is contained in the set $\{p \in M : d_g(p_0, p) \leq \frac{1}{2}\hat{a}H(p_0)^{-1}, H(p) \geq 2\hat{b}^{-1}H(p_0)\}$, we conclude $M$ itself is contained in the set $\{p \in M : d_g(p_0, p) \leq \frac{1}{2}\hat{a}H(p_0)^{-1}, H(x) \geq 2\hat{b}^{-1}H(p_0)\}$. This completes the proof. □

In our setting, we are concerned with a complete, connected, convex hypersurface that satisfies the pointwise gradient estimate $|\nabla H| \leq c_1 H^2$ everywhere (with constant $c_1 = n\gamma_1$). We will use the theorem above to find necks at a controlled distance from points which do lie on necks. Specifically, we will use the following immediate corollary.

**Corollary 2.13.** *Given constants $\eta, \gamma_1 > 0$, there exist positive constants $\hat{a}, \hat{b}$, depending only upon $n$, $\eta$, and $\gamma_1$, with the following property. Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a complete, connected, convex hypersurface satisfying the pointwise derivative estimate $|\nabla h| \leq \gamma_1 H^2$. Suppose $p_0 \in M$ is a point such that $\lambda_1(p_0) > \eta H(p_0)$. Then either, $M = B_g(p_0, \hat{a}H(p_0)^{-1})$ or there exists a point $p_1 \in B_g(p_0, \hat{a}H(p_0)^{-1})$ where $\lambda_1(p_1) \leq \eta H(p_1)$. Moreover, in either case, every point $p \in B_g(p_0, \hat{a}H(p_0)^{-1})$ satisfies $H(p) \geq \hat{b}^{-1}H(p_0)$*

*Proof.* With $c_1 := n\gamma_1$ and $\eta$ as given, let $\hat{a}$ and $\hat{b}$ be chosen as in the proof of the theorem above. Since the derivative estimate holds everywhere, the set $U = \{x \in M : d_g(p_0, x) < \hat{a}H(p_0)^{-1}, H(x) > \hat{b}^{-1}H(p_0)\}$ is the entire ball $B_g(p_0, \hat{a}H(p_0)^{-1})$. If $\lambda_1 > \eta H$ holds everywhere in $U$, the theorem implies the first alternative holds. Otherwise, clearly the second alternative holds. □

The next lemma is an auxiliary result for the proof of the Neck Detection Lemma below. Our goal is to show that whenever $\frac{\lambda_1}{H}(p_0, t_0)$ is sufficiently small then $p_0$ lies at the center of an $(\varepsilon_0, L)$-neck at time $t_0$. We begin by first showing smallness of $\frac{\lambda_1}{H}(p_0, t_0)$ implies the second fundamental form is close to the second fundamental form of a round cylinder in a small intrinsic ball around $p_0$ at time $t_0$.

**Lemma 2.14.** *Suppose $\mathcal{M}$ satisfies $(*)$. Let $\rho := \frac{1}{4}\hat{r}$, where $\hat{r}$ is the positive constant appearing in Lemma 2.9. For every $\phi \in (0, \frac{1}{n})$, we can find $\eta := \eta(\phi, n, \gamma_1, \gamma_2) \in (0, \phi)$ with the property that if $(p_0, t_0)$ is a spacetime point that satisfies $\lambda_1(p_0, t_0) \leq \eta H(p_0, t_0)$, then for every $p \in B_{g(t_0)}\big(p_0, 2\rho\frac{n-1}{H(p_0,t_0)}\big)$ there holds:*

*(1) $\lambda_1(p, t_0) \leq \phi H(p_0, t_0)$ and $\lambda_n(p, t_0) - \lambda_2(p, t_0) \leq \phi H(p_0, t_0)$;*

*(2) $\sum_{k=1}^{8} H(p_0, t_0)^{-k-1}|\nabla^k h(p, t_0)| \leq \phi$;*

*(3) $(1 - \phi)H(p_0, t_0) \leq H(p, t_0) \leq (1 + \phi)H(p_0, t_0)$.*

*Proof.* We argue by contradiction. Suppose the assertion is not true. Then for some $\phi \in (0, \frac{1}{n})$, there exists a sequence of flows $\mathcal{M}^{(j)}$ satisfying $(*)$ for some fixed values of $\gamma_1$ and $\gamma_2$ and a sequence of positive constants $\eta_j \to 0$ which are counterexamples to the assertion. Let the flow $\mathcal{M}^{(j)}$ be given by an embedding $F_j : M_j \times (-\infty, 0] \to \mathbb{R}^{n+1}$ and let $g_j$, $h_j$, $H_j$, and $\lambda_{i,j}$ denote the metric, second fundamental form, mean curvature, and principal curvatures ($i = 1, \ldots n$) on $\mathcal{M}^{(j)}$. Then for each $j$ there exist spacetime points $(p_j, t_j) \in M_j \times (-\infty, 0]$ such that $\lambda_{1,j}(p_j, t_j) \leq$

$\eta_j H_j(p_j,t_j)$, but

$$\sup_{q\in B_j}\max\left\{\frac{\lambda_{1,j}(q,t_0)}{H_j(p_j,t_j)},\quad \frac{\lambda_{n,j}(q,t_j)-\lambda_{2,j}(q,t_j)}{H_j(p_j,t_j)},\quad \sum_{k=1}^{8}\frac{|\nabla^k h_j(q,t_j)|}{H_j(p_j,t_j)^{k+1}},\quad \left|\frac{H_j(q,t_j)}{H_j(p_j,t_j)}-1\right|\right\}>\phi,$$

where $B_j := B_{g_j(t_j)}\big(p_j, 2\rho\frac{n-1}{H_j(p_j,t_j)}\big)$.

Let $Q_j = (n-1)H_j(p_j,t_j)^{-1}$. The desired inequalities are scale and translation invariant, so as in the proof of Lemma 2.11, for each $j$ we define a translated and rescaled solution $\tilde{F}_j$ by

$$\tilde{F}_j(p,\tau) = \frac{1}{Q_j}\big[F_j\big(p, Q_j^2\tau + t_j\big) - F_j(p_j,t_j)\big].$$

Let $\tilde{\mathcal{M}}^{(j)}$ denote the resulting flow. Let $\tilde{g}_j, \tilde{h}_j$, $\tilde{H}_j$, and $\tilde{\lambda}_{i,j}$ denote the geometric quantities of $\tilde{\mathcal{M}}^{(j)}$. By construction, each flow $\tilde{\mathcal{M}}^{(j)}$ satisfies $\tilde{H}_j(p_j,0) = n-1$ and $\tilde{\lambda}_{1,j}(p_j,0) \le (n-1)\eta_j$. For $r>0$, let $\tilde{P}_j(r)$ denote the parabolic neighborhood $P(p_j,0,r,r^2)$ in $\tilde{\mathcal{M}}^{(j)}$. Lemmas 2.9 and 2.10 implies uniform estimates

$$\sup_{\tilde{P}_j(4\rho)}\left|\frac{\partial}{\partial t}^{l}\nabla^k\tilde{h}_j\right| \le C(k,l,n,\gamma_1,\gamma_2).$$

As before, a subsequence of the solutions $\tilde{F}_j$ converge locally in the $C^{20}$-topology within the parabolic neighborhoods $\tilde{P}_j(4\rho)$ to a smooth limit flow $\hat{\mathcal{M}}$ defined in a small intrinsic parabolic neighborhood $P_\infty(4\rho) := P(p_\infty, 0, 4\rho, 16\rho^2)$. Let $g, h, H$ and $\lambda_i$ denote the geometric quantities on $\hat{\mathcal{M}}$.

Now we analyze the limit flow. By assumption and Proposition 2.11, $\tilde{\lambda}_{1,j}\ge 0$, $\tilde{\lambda}_{1,j}+\tilde{\lambda}_{2,j} \ge \frac{1}{n-1}\tilde{H}_j$, and $\tilde{\lambda}_{1,j}(p_j,0) \le (n-1)\eta_j$, where $\eta_j \to 0$ as $j\to\infty$. Hence in the limit, we have $\lambda_1 \ge 0$, $\lambda_1+\lambda_2 \ge \frac{1}{n-1}H$ and $\lambda_1(p_\infty,0)=0$. As in the proof of Proposition 2.11, the maximum principle implies $\lambda_1 \equiv 0$ in $P_\infty(4\rho)$. Consequently, the limit flow must split a line and in view of the estimate $\lambda_1+\lambda_2 \ge \frac{1}{n-1}H$, this gives $\lambda_2 = \cdots = \lambda_n = \frac{1}{n-1}H$. The Schur lemma implies the cross-section is a piece of an $(n-1)$-sphere and hence that $P_\infty(4\rho)$ is a small parabolic neighborhood in an evolving family of shrinking cylinders. In particular, $\lambda_1 \equiv 0$, $\lambda_n \equiv \lambda_2$, $\sum_{k=1}^{8}|\nabla^k h| \equiv 0$, and

$H \equiv n-1$ in $P_\infty(4\rho)$. By convergence of the $\tilde{F}_j$ in the $C^{20}$-topology, this implies

$$\sup_{q \in B_j} \max \left\{ \frac{\lambda_{1,j}(q,t_0)}{H_j(p_j,t_j)}, \quad \frac{\lambda_{n,j}(q,t_j) - \lambda_{2,j}(q,t_j)}{H_j(p_j,t_j)}, \quad \sum_{k=1}^{8} \frac{|\nabla^k h_j(q,t_j)|}{H_j(p_j,t_j)^{k+1}}, \quad \Big| \frac{H_j(q,t_j)}{H_j(p_j,t_j)} - 1 \Big| \right\} \to 0$$

along a subsequence as $j \to \infty$, in contradiction with our previous assumption for $j$ sufficiently large. □

By iterating the lemma above finitely many times, we can prove our version of the Neck Detection Lemma (cf. Lemma 7.4 in [33], Lemma 4.2 in [41], and Theorems 2.14 and 2.15 in [14]).

**Lemma 2.15** (Neck Detection)**.** *Suppose $\mathcal{M}$ satisfies $(*)$. Given $\varepsilon_0 \in (0, \frac{1}{n})$ and $L \geq 100$, there exists $\eta_0 := \eta_0(\varepsilon_0, L, n, \gamma_1, \gamma_2) \in (0, \varepsilon_0)$ with the property that if $(p_0, t_0)$ is a spacetime point with $\lambda_1(p_0,t_0) \leq \eta_0 H(p_0,t_0)$, then $p_0 \in M$ lies at the center of an $(\varepsilon_0, L)$-neck at time $t_0$.*

*Proof.* Let $\varepsilon_0$ and $L$ be given. After a rescaling, we can assume without loss of generality that $H(p_0,t_0) = n-1$. Choose $\varepsilon_1 \in (0, \varepsilon_0)$ so that the conclusion of Proposition 2.8 holds. Recall the choice of $\varepsilon_1$ depends only upon $n$, $\varepsilon_0$, and $L$. By Proposition 2.8, to show that $p_0$ lies at the center of a $(\varepsilon_0, L)$-neck at time times $t_0$, it suffices to show we can choose $\eta_0$ so that the following two properties hold:

(a) $\lambda_1(p_0,t_0) \leq \varepsilon_1 H(p_0,t_0)$ and $\lambda_n(p_0,t_0) - \lambda_2(p_0,t_0) \leq \varepsilon_1 H(p_0,t_0)$.

(b) For every $p \in B_{g(t_0)}(p_0, L+10)$, $\sum_{k=1}^{8} H(p_0,t_0)^{-k-1} |\nabla^k h(p,t_0)| \leq \varepsilon_1$.

Recall from the previous lemma that $\rho := \frac{1}{4}\hat{r}$. Choose $N$ sufficiently large such that

$$1 + \sum_{m=2}^{N} \frac{m-1}{m} > \frac{L+10}{\rho}.$$

We will now determine $\eta_0$. Begin by choosing $\eta_N \in (0, \varepsilon_1)$ such that the conclusions of Lemma 2.14 hold with $\phi = \varepsilon_1$. For $m = 1, \dots, N-1$, having chosen $\eta_{m+1}$, first choose $\tilde{\eta}_m \in (0, \eta_{m+1})$ so that the conclusions of Lemma 2.14 hold with $\phi = \eta_{m+1}$ and then let $\eta_m := \min\{\frac{\tilde{\eta}_m}{1+\tilde{\eta}_m}, \frac{1}{N-m}\}$. Finally, by one further application of the Lemma 2.14, choose $\tilde{\eta}_0 \in (0, \eta_1)$ so that Lemma 2.14

holds with $\phi = \eta_1$ and define $\eta_0 := \min\{\tilde{\eta}_0, \frac{\varepsilon_1}{n-1}\}$. It is clear that $\eta_0$ has been chosen in a way that depends only upon $\varepsilon_0, L, n, \gamma_1$ and $\gamma_2$.

Now let us examine the consequences of our choices. Assume $\lambda_1(p_0, t_0) \leq \eta_0 H(p_0, t_0)$. First since $\eta_0 \leq \varepsilon_1$, we have $\lambda_1(p_0, t_0) \leq \varepsilon_1 H(p_0, t_0)$. Since $\mathcal{M}$ is $\frac{1}{n-1}$-two-convex, we have $\lambda_2(p_0, t_0) \geq (\frac{1}{n-1} - \eta_0) H(p_0, t_0)$ which implies

$$\lambda_n(p_0, t_0) - \lambda_2(p_0, t_0) \leq H(p_0, t_0) - (n-1)\lambda_2(p_0, t_0) \leq (n-1)\eta_0 H(p_0, t_0) \leq \varepsilon_1 H(p_0, t_0).$$

This shows that (a) holds for our choice of $\eta_0$. As for (b): Define $\rho_1 = \rho$. For $m = 2, \ldots, N$, define

$$\rho_m := \rho\Big(1 + \sum_{j=1}^{m-1} \frac{N-j}{N+1-j}\Big) = \rho\Big(1 + \sum_{j=N+2-m}^{N} \frac{j-1}{j}\Big)$$

Note that $\rho_{m+1} = \rho_m + \rho\frac{N-m}{N+1-m}$ and $\rho_N > L + 10$ by our choice of $N$. Let us say condition $(\Gamma_m)$ holds if for all $p \in B_{g(t_0)}(p_0, \rho_m)$, we have

(i) $\lambda_1(p, t_0) \leq \eta_m H(p_0, t_0)$ and $\lambda_n(p, t_0) - \lambda_2(p, t_0) \leq \eta_m H(p_0, t_0)$;

(ii) $\sum_{k=1}^{8} H(p_0, t_0)^{-k-1} |\nabla^k h(p, t_0)| \leq \eta_m$;

(iii) $(1 - \eta_m) H(p_0, t_0) \leq H(p, t_0) \leq (1 + \eta_m) H(p_0, t_0)$.

Our choice of $\eta_0 \leq \tilde{\eta}_0$ and Lemma 2.14 implies that $(\Gamma_1)$ holds. Suppose that $1 \leq m < N$ and $(\Gamma_m)$ holds. Now consider an arbitrary point $p \in B_{g(t_0)}(p_0, \rho_m)$. First, because $\eta_m \leq \frac{\tilde{\eta}_m}{1+\tilde{\eta}_m}$, we have

$$\lambda_1(p, t_0) \leq \frac{\tilde{\eta}_m}{1+\tilde{\eta}_m} H(p_0, t_0) \leq \frac{\tilde{\eta}_m}{1+\tilde{\eta}_m} \frac{1}{1-\eta_m} H(p, t_0) \leq \tilde{\eta}_m H(p, t_0).$$

Second, because $\eta_m \leq \frac{1}{N-m}$, we have

$$\frac{n-1}{H(p, t_0)} \geq \frac{1}{1+\eta_m} \frac{n-1}{H(p_0, t_0)} = \frac{1}{1+\eta_m} \geq \frac{N-m}{N+1-m}.$$

We chose $\tilde{\eta}_m$ to satisfy the conclusions of Lemma 2.14 with $\phi = \eta_{m+1}$. Given that $\lambda_1(p, t_0) \leq \tilde{\eta}_m H(p, t_0)$ and $\frac{n-1}{H(p,t_0)} \geq \frac{N-m}{N+1-m}$, we conclude that conditions (i), (ii), and (iii) above hold in the ball $B_{g(t_0)}(p, 2\rho\frac{N-m}{N+1-m})$ with $\eta_m$ replaced by $\eta_{m+1}$. Since $p \in B_{g(t_0)}(p_0, \rho_m)$ is arbitrary and $\rho_m + \rho\frac{N-m}{N+1-m} = \rho_{m+1}$, this implies $(\Gamma_{m+1})$ holds. By finite induction $(\Gamma_N)$ holds and this implies (b) holds for $\eta_0$, completing the proof of the lemma. □

## 2.2. Tube and Cap Decomposition: Noncompact Case

In this section, we will assume $M$ is noncompact. We begin this section by recalling some useful analysis on the convexity of necks from Proposition 7.18, Lemma 7.19, and the surrounding discussion in [33]. The following lemma highlights a difference between our setting (a noncompact, convex hypersurface) and the setting of Huisken and Sinestrari (a closed, two-convex hypersurface). In the setting of Huiksen-Sinestari, the axes of different necks need not align if the (intrinsic) distance between the neck regions is large compared to the curvature scales of the necks. In our setting, we can show every neck must have approximately the same axis.

**Lemma 2.16.** *Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a noncompact, complete, convex hypersurface. There exists a unit vector $\omega \in S^n$ with the following two properties:*

*(1) $\langle \nu(q), \omega\rangle \geq 0$ for every $q \in M$.*

*(2) There exists a constant $C := C(n)$ such that if $N \subset M$ is an $(\varepsilon, L)$-neck, then $\langle \nu(q), \omega\rangle \leq C\varepsilon$ for every $q \in N$.*

*Proof.* We begin by constructing $\omega$. Let $\Omega$ denote the convex interior of $F(M)$. Choose any sequence of points $x_k \in \Omega$ such that $|x_k| \to \infty$. We can find a subsequence of these points such that $-\frac{x_k}{|x_k|}$ converges to a limit $\omega \in S^n$. Part (1) of the lemma now follows from convexity. To see this, consider any point $x \in \bar{\Omega}$ and fix some $s \geq 0$. For $k$ sufficiently large, $s_k := s|x_k|^{-1} \in [0, 1]$. By convexity, $(1-s_k)x + s_k x_k \in \bar{\Omega}$ for $k$ sufficiently large. As $k \to \infty$, $(1-s_k)x + s_k x_k$ converges to $x - s\omega$, and therefore $x - s\omega \in \bar{\Omega}$ for all $s \geq 0$. Let $q \in M$. Then $F(q) - \omega \in \bar{\Omega}$ and because $\Omega$ is convex and $\nu(q)$ is outward-pointing, $\langle x - F(q), -\nu(q)\rangle \geq 0$ for all $x \in \bar{\Omega}$. Taking $x = F(q) - \omega$ gives $\langle \nu(q), \omega\rangle \geq 0$, as claimed.

Now suppose $N \subset M$ is an $(\varepsilon, L)$-neck. By definition $N$ can be expressed as a small graph over an embedded round cylinder $\Sigma$ of length $2L$ in $\mathbb{R}^{n+1}$. Let $\omega_0$ be the unit vector parallel to the axis of $\Sigma$ such that $\langle \omega, \omega_0 \rangle \geq 0$. It follows from the definition of a neck that:

(i) $|\langle \nu(q), \omega_0 \rangle| \leq C\varepsilon$ for every $q \in N$.

(ii) If $e$ is a unit vector orthogonal to $\omega_0$, then there exists a point $q \in N$ where $|\nu(q) - e| \leq C\varepsilon$.

If $\omega = \omega_0$, then part (2) is a consequence of (i). Otherwise, consider $v := \omega - \langle \omega, \omega_0 \rangle \omega_0 \neq 0$. By (ii), we can find a point $q \in N$ where $|\nu(q) + \frac{v}{|v|}| \leq C\varepsilon$. Then at $q$ we have

$$0 \leq \langle \nu(q), \omega \rangle = -\Big\langle \frac{v}{|v|}, \omega \Big\rangle + \Big\langle \nu(q) + \frac{v}{|v|}, \omega \Big\rangle \leq -\sqrt{1 - \langle \omega, \omega_0 \rangle^2} + C\varepsilon.$$

This gives $|\omega - \omega_0| \leq C\varepsilon$ and therefore, with (i), we conclude $\langle \nu, \omega \rangle \leq C\varepsilon$ everywhere on $N$. $\square$

Now suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a noncompact, complete, convex hypersurface that satisfies the gradient estimate $|\nabla h| \leq \gamma_1 H$. Suppose $p \in M$ lies at the center of an $(\varepsilon, L)$-neck $N$. By the lemma above, we can find $\omega \in S^n$ such that $\langle \nu, \omega \rangle \geq 0$ everywhere on $M$ and $\langle \nu, \omega \rangle \leq C\varepsilon$ everywhere on $N$. Let $y$ denote the height along the axis defined by $\omega$, normalized so that $p$ is contained in the hyperplane $y = 0$. Let $\Sigma_0$ denote the intersection of $N$ with the level set $y = 0$. The definition of a neck implies $\Sigma_0$ is compact and very close to a round $(n-1)$-dimensional sphere. We will call the $y$-direction vertical and all other directions, orthogonal to $\omega$, horizontal.

By assumption $\omega$ is nearly tangent to $N$. As in [33], we consider integral curves of the height function $y$. For each $q \in \Sigma_0$, let $\gamma(\tau) := \gamma(\tau, q)$ be a solution to the ODE

$$\begin{cases} \dot{\gamma} = \frac{\omega^\top(\gamma)}{|\omega^\top(\gamma)|^2} & \tau \geq 0, \\ \gamma(0) = q, \end{cases}$$

where $\omega^\top$ denotes the projection of $\omega$ to the tangent space of $M$. The curves are defined so that $\frac{d}{d\tau} y(\gamma(\tau)) = \langle \gamma'(\tau), \omega \rangle = 1$. So by our normalization $y(\Sigma_0) = 0$, we have $y(\gamma(\tau)) = \tau$. Hence we can write $\gamma(y)$ in place of $\gamma(\tau)$. We will consider these curves for as long as they are well-defined, including after they leave the neck region $N$. As $\Sigma_0$ is compact, every curve is defined for $|y|$ small. Let $y_{\min} \in [-\infty, 0)$ and $y_{\max} \in (0, \infty]$ be the minimal and maximal heights such that for every $q \in \Sigma_0$, the curve $\gamma(\cdot, q)$ is defined for $y \in (y_{\min}, y_{\max})$. It is possible for either of $y_{\min}$ or $y_{\max}$ to be infinite because our hypersurface is noncompact. However, we will see that the assumptions on $\omega$ will ensure $y_{\max} < \infty$ and $y_{\min} = -\infty$. For each $y \in (y_{\min}, y_{\max})$, let $\Sigma_y = \{\gamma(y, q) : q \in \Sigma_0\}$. Since our hypersurface is convex, the $\Sigma_y$ are just level sets of the height function. We will say the surfaces $\Sigma_y$ are shrinking if the projection of $\Sigma_{y_2}$ to a fixed hyperplane $y = y'$ is contained in the domain enclosed by the projection of $\Sigma_{y_1}$ to the hyperplane $y = y'$ for any $y_2 \geq y_1$.

Now we give a lemma concerning the behavior integral curves to the height function and the surfaces they define. The lemma is a combination of Proposition 7.18 and Lemma 7.19 in [33]. A slight difference is that our gradient estimate holds at all curvature scales and our hypersurface is noncompact.

**Lemma 2.17.** *Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a noncompact, complete, convex hypersurface satisfying the gradient estimate $|\nabla h| \leq \gamma_1 H$. Let $N \subset M$ be an $(\varepsilon, L)$-neck. Let $\omega$ satisfy the conclusions of Lemma 2.16. Under the hypotheses above, if $0 < \varepsilon < \varepsilon(n)$ is sufficiently small, then there holds:*

*(1) For every $q \in \Sigma_0$, the curve $\gamma(\cdot, q)$ is well-defined for as long as it is contained in the neck $N$.*

*(2) Along a trajectory $\gamma$ (in the direction of $\omega$), we have*

$$\frac{d}{dy}\langle \nu, \omega \rangle \geq \lambda_1 > 0.$$

*(3) Suppose $\langle \nu(q), \omega \rangle > 0$ for all $q \in \Sigma_0$. Then the surfaces $\Sigma_y$ are shrinking for all $y \in [0, y_{\max})$ and $y_{\max} < \infty$. Moreover, there exists a positive constant $\hat{\theta} := \hat{\theta}(n, \gamma_1)$ such that $H(\gamma(y, q)) \geq \hat{\theta}^{-1} H(p)$ for all $q \in \Sigma_0$ and $y \in [0, y_{\max})$.*

*Proof.* Statement $(1)$ is clear from the inequality $0 \le \langle \nu, \omega\rangle \le C\varepsilon$ on $N$ and $(2)$ is a computation as in the proof of Proposition 7.18 in [33]. Statement $(4)$ is essentially Lemma 7.19 in [33], but we verify it here. Because $\Sigma_0$ is compact, we may assume $\langle \nu(q), \omega\rangle \ge \varepsilon' > 0$ for all $q \in \Sigma_0$. By $(2)$, this implies $\langle \nu(\gamma(y,q)), \omega\rangle \ge \varepsilon'$ for all $q \in \Sigma_0$ and all $y \in [0, y_{\max})$. For $\bar{y} \in [0, y_{\max})$, consider the projection of the surface $\Sigma_{\bar{y}}$ to the hyperplane $y = 0$. The outward-pointing normal direction of the projected surface is $\nu - \langle \nu, \omega\rangle\omega$. We compute for any $q \in \Sigma_0$,

$$\Big\langle \dot{\gamma}, \frac{\nu - \langle \nu, \omega\rangle\omega}{|\nu - \langle \nu, \omega\rangle\omega|}\Big\rangle = \Big\langle \frac{\omega^\top}{|\omega^\top|^2}, \frac{\nu - \langle \nu, \omega\rangle\omega}{\sqrt{1 - \langle \nu, \omega\rangle^2}}\Big\rangle = -\frac{\langle \nu, \omega\rangle}{\sqrt{1 - \langle \nu, \omega\rangle^2}} \le -\frac{\varepsilon'}{\sqrt{1 - (\varepsilon')^2}}.$$

This shows the horizontal component of $\dot{\gamma}(\bar{y}, q)$ points towards the interior of $\Sigma_{\bar{y}}$ and has norm at least $\varepsilon'/\sqrt{1 - (\varepsilon')^2}$. This means the surfaces $\Sigma_{\bar{y}}$ are shrinking at definite rate for $y \ge 0$ and hence $y_{\max} < \infty$. The proof of the second statement in $(3)$ follows directly from the proof of Lemma 7.19 in [33]. Because $\omega$ is approximately the axis of our neck, $\Sigma_0$ is very close to a standard $(n-1)$-sphere of radius $\frac{n-1}{H(p)}$. Supposing $\varepsilon$ is sufficiently small, this implies $H(q) \ge \frac{1}{2}H(p)$ for all $q \in \Sigma_0$ and that there exists an $(n-1)$-sphere of radius $R = 2\frac{n-1}{H(p)}$ that encloses $\Sigma_0$ in the hyperplane $y = 0$. Integrating the pointwise gradient estimate, for any points $q, q' \in M$, we have

$$H(q) \ge \frac{1}{H(q')^{-1} + n\gamma_1 d_g(q, q')}.$$

If $\bar{y} \in [0, y_{\max})$ satisfies $\bar{y} < R$, then for any $q \in \Sigma_{\bar{y}}$, it is clear there exists $q' \in \Sigma_0$ such that $d_g(q, q') \le 2R$ (the extrinsic distance between points in $\Sigma_0$ and $q \in \Sigma_{\bar{y}}$ is bounded by $R$ vertically and $R$ horizontally; since our hypersurface is strictly convex, the intrinsic distance is similarly bounded). Since $H(q') \ge \frac{1}{2}H(p)$, we have

$$H(q) \ge \frac{1}{H(q')^{-1} + n\gamma_1 d_g(q, q')} \ge \frac{1}{2H(p)^{-1} + n\gamma_1 2R} = \frac{H(p)}{2 + 4n(n-1)\gamma_1}.$$

On the other hand, if $\bar{y} \ge R$, then we can find $y'$ such that $\bar{y} \in [y', y' + R] \subset [0, y_{\max})$. We can construct a suitable portion of cone with spherical cross-section, axis $\omega$, and bases in the hyperplanes $y = y'$ and $y = y' + R$ of radius $R_1, R_2 \leq R$ respectively. For suitable choices of $R_1, R_2$, we can arrange that the cone touches a point $q' \in \cup_{y \in (y', y'+R)} \Sigma_y$ from the outside. This is possible by convexity of $\cup_{y \in (y', y'+R)} \Sigma_y$ and because the surfaces are shrinking. Now $H(q') \geq \frac{n-1}{R} = \frac{1}{2} H(p)$ by comparison to the cone. If $q \in \Sigma_{\bar{y}}$, noting that the intrinsic diameter of $\Sigma_{\bar{y}}$ is bounded by $\pi R$, then $d_g(q, q') \leq (2 + \pi) R$. Thus the above argument applies. We can take $\hat{\theta} := (2 + 2(2 + \pi) n(n-1)\gamma_1)^{-1}$ to complete the proof. □

In the next step, we prove our ancient solution has a convex cap outside of which every point lies at the center of a neck. For the mean curvature flow of two-convex hypersurfaces, the following key result is often called the Neck Continuation Theorem. See Theorem 8.2 in [33] and also Theorem 3.2 in [14]. The proof of our version of the Neck Continuation Theorem is modeled on the proofs given by Huisken-Sinestrari and Brendle-Huisken. Of course, our argument is also a bit simpler in that we do not need to consider if regions have been previously affected by surgery. Our phrasing of the Neck Continuation Theorem is inspired by similar results established by Perelman in his study of ancient $\kappa$-solutions in the Ricci flow.

Before the theorem, let us point out that if $N$ is an $(\varepsilon, L)$-neck in a noncompact, complete, connected, strictly convex hypersurface $M$, then $M \setminus N$ consists of two connected components, one bounded and the other unbounded. If both components were unbounded, then $M$ would split a line, thereby contradicting strict convexity.

**Theorem 2.18.** *Suppose $\mathcal{M}$ is noncompact and satisfies $(*)$. Given $0 < \varepsilon_0 < \varepsilon(n)$ small and $L \geq 100$, there exist constants $\varepsilon_1 \in (0, \varepsilon_0)$, and $C_0 < \infty$, depending only upon $\varepsilon_0$, $L$, $n$, $\gamma_1$ and $\gamma_2$, so that the following holds. Fix any time $t \in (-\infty, 0]$. Suppose that $p \in M$ is a point which lies at the center of an $(\varepsilon_1, L)$-neck $N$ at time $t$, and suppose further that $p$ does not lie at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck at time $t$. Let $D$ denote the bounded connected component of $M \setminus N$, and let $\tilde{D}$ denote the unbounded connected component of $M \setminus N$. Then:*

*(1) Every point $q \in \tilde{D} \cup N$ lies at the center of an $(\varepsilon_0, L)$-neck at time $t$.*

*(2) $D$ is diffeomorphic to $B^n$.*

*(3)* $\partial D \subset \partial N$ *is a cross-sectional sphere of an* $(\varepsilon_0, L)$*-neck.*

*(4) The (intrinsic) diameter of* $D$ *is bounded by* $C_0 H(p,t)^{-1}$*.*

*(5) Every point* $q \in D$ *satisfies* $C_0^{-1} H(p,t) \leq H(q,t) \leq C_0 H(p,t)$ *and* $\lambda_1(q,t) \geq C_0^{-1} H(q,t)$*.*

*Proof.* Fix a time $t_0$ and for simplicity let us suppress $t_0$ in our notation. For any point $q \in M$, let $r_q := (n-1)H(q)^{-1}$ denote the mean curvature scale. As usual, let $B_g(q,r)$ denote an intrinsic ball of radius $r$ around $q$. Let $\varepsilon_0 > 0$, suitably small, and $L \geq 100$ be given. We can assume that on any $(\varepsilon_0, L)$-neck, the mean curvature satisifes $\frac{9}{10} \leq \frac{H(q_1)}{H(q_2)} \leq \frac{10}{9}$ for any pair of points $q_1, q_2$ on the neck. We will determine the constants $\varepsilon_1 \in (0, \varepsilon_0)$ and $C_0$ via two claims. By Lemma 2.16, we can fix unit vector $\omega \in S^n$ with the property that $\langle \nu, \omega \rangle \geq 0$ everywhere on $M$ and $0 \leq \langle \nu, \omega \rangle \leq C\varepsilon_1$ on any $(\varepsilon_1, L)$-neck $N \subset M$.

**Claim 1:** For any $0 < \varepsilon_1 < \varepsilon(n)$ sufficiently small, there exists $C_0 := C_0(\varepsilon_1, n, L, \gamma_1, \gamma_2) < \infty$ such that if $p$ and $N$ satisfy the assumptions of the theorem, then parts (2), (4), and (5) of the theorem hold.

We will establish this first claim in four steps. We assume $\varepsilon_1 < \varepsilon(n)$ is sufficiently small so that the conclusions of Lemma 2.17 hold.

*Step 1.1:* Given $\varepsilon_1$ small and $p$ lying on $N$ as in the theorem, we can find $\eta_1 > 0$ such that $\lambda_1 > \eta_1 H$ everywhere on the neck $N$. This follows from the assumption that $p$ does not lie at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck. Here is the procedure to choose $\eta_1$:

(a) Choose $\hat{\eta} := \hat{\eta}(\varepsilon_1, n, L, \gamma_1, \gamma_2) \in (0, \frac{\varepsilon_1}{2})$ such that if $\lambda_1(q) \leq \hat{\eta} H(q)$, then $q$ lies at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck.

(b) Choose $\hat{\varepsilon} := \hat{\varepsilon}(\hat{\eta}, n) \in (0, \hat{\eta})$ sufficiently small such if $q$ lies at the center of an $(\hat{\varepsilon}, 3L)$-neck $\hat{N}$, then $\lambda_1 \leq \hat{\eta} H$ everywhere on $\hat{N}$ and $B_g(q, 2Lr_q) \subset \hat{N}$.

(c) Choose $\eta_1 := \eta_1(\hat{\varepsilon}, n, L, \gamma_1, \gamma_2) \in (0, \hat{\varepsilon})$ such that if $\lambda_1(q) \leq \eta_1 H(q)$, then $q$ lies at the center of an $(\hat{\varepsilon}, 3L)$-neck $\hat{N}$.

Choices (a) and (c) are possible via Neck Detection Lemma and (b) is a simple consequence of the definition of a neck. Now let $q \in N$. For $\varepsilon_1$ small (depending only on $n$), $H(q) \leq \frac{10}{9}H(p)$ and $N \subset B_g(p, \frac{3}{2}Lr_p)$. Therefore, $d_g(p,q) < \frac{3}{2}Lr_p < 2Lr_q$, which implies $p \in B_g(q, 2Lr_q)$. Because $p$ does not lie at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck, we must have $\lambda_1(p) > \hat{\eta}H(p)$ given (a). We conclude $\lambda_1(q) > \eta_1 H(q)$ for all $q \in N$, otherwise we obtain a contradiction given (b) and (c).

*Step 1.2:* Strict convexity of the neck $N$ implies implies it must close up. Our primary tool, as in [33], is to analyze the integral curves of the height function. Define $y : M \to \mathbb{R}$ by $y(q) = \langle F(q) - F(p), \omega \rangle$. Our normalization ensures $y(p) = 0$. Let $\Sigma_0 \subset N$ denote the level set $y = 0$. By Lemma 2.16, $\Sigma_0$ is $O(\varepsilon_1)$-close to a round sphere. For every $q \in \Sigma_0$, define $\gamma(y,q)$ to be the integral curve of $\frac{\omega^\top}{|\omega^\top|^2}$ passing through $q$. As in Lemma 2.17, these curves are defined for $y \in (y_{\min}, y_{\max})$, with $y_{\min} < 0 < y_{\max}$, and we know they are well-defined at least as long as they are in $N$. For each $y$, let $\Sigma_y = \{\gamma(y,p) : p \in \Sigma_0\}$ denote the smooth level sets of the height function. For $y_{\min} \leq y_1 < y_2 \leq y_{\max}$, let $\Sigma(y_1, y_2) = \bigcup_{y_1 < y < y_2} \Sigma_y$.

Because $\omega$ is an approximate axis of $N$ and the neck has intrinsic length approximately equal to $2Lr_p$ and $L \geq 100$, we certainly have $\Sigma(-2r_p, 0) \subset N$. Since $\frac{d}{dy}\langle \nu, \omega \rangle \geq \lambda_1$ and $\lambda_1 \geq \eta_1 H \geq \frac{1}{2}\eta_1 H(p)$ on $N$, for any $q \in \Sigma_0$ we have

$$\begin{aligned} \langle \nu(q), \omega \rangle &= \langle \nu(\gamma(-2r_p, q), 0), \omega \rangle + \int_{-2r_p}^{0} \frac{d}{dy}\langle \nu, \omega \rangle \, dy \\ &\geq r_p \eta_1 H(p) = (n-1)\eta_1. \end{aligned}$$

Now by part (3) of Lemma 2.17, we must have $y_{\max} < \infty$. Moreover, $H(\gamma(y,q)) > \hat{\theta}^{-1}H(p)$ for all $q \in \Sigma_0$ and $y \in [0, y_{\max})$. Here $\hat{\theta}$ depends only on $n$ and $\gamma_1$.

*Step 1.3:* We next show the region $\Sigma(0, y_{\max})$ is uniformly convex, depending only upon the given constants. In particular, there exists $\eta_2 := \eta_2(\eta_1, n, L, \gamma_1, \gamma_2) \in (0, \eta_1)$ such that the following four properties hold for all $y \in [0, y_{\max})$:

$$|\langle \nu, \omega \rangle| < 1, \quad \lambda_1 > \eta_2 H, \quad H > \hat{\theta}^{-1} H(p), \quad \langle \nu, \omega \rangle > \eta_1.$$

If $\eta_2 \leq \frac{1}{2}\eta_1$, then these four properties hold for $y$ sufficiently close to zero. If they do not hold until $y_{\max}$, let $\tilde{y} \in (0, y_{\max})$ be the first value of $y$ for which one of these properties fails. The first property must hold for $y < y_{\max}$ by definition. Also the fourth property holds until $y_{\max}$ because $\langle \nu, \omega \rangle > \eta_1$ on $\Sigma_0$ and $\langle \nu, \omega \rangle$ is increasing as $y$ increases. In particular, $\langle \nu, \omega \rangle > \eta_1$ on $\Sigma_{\tilde{y}}$. We have already observed that the third property must hold. So if a property fails at $\tilde{y}$, it must be the second one. If it fails, then there exists $\tilde{p} \in \Sigma_{\tilde{y}}$ such that $\lambda_1(\tilde{p}) \leq \eta_2 H(\tilde{p})$. By the Neck Detection Lemma and Lemma 2.16, if we take $\eta_2$ sufficiently small, then $\tilde{p}$ lies at the center of a very fine neck $\tilde{N}$ and satisfies $\langle \nu(\tilde{p}), \omega \rangle < \eta_1$. This contradicts the fourth inequality above.

*Step 1.4:* With the existence of $\eta_2$ established and the above properties verified, elementary arguments (see the end of the proof of Theorem 8.2 in [33]) imply all integral curves converge to the same critical point of the height function. Thus the region $M \cap \{y > 0\}$ is a uniformly convex cap diffeomorphic to $B^n$. Since $\omega$ is $C\varepsilon_1$-close to the axis of the cylinder over which $N$ is a graph, this implies $D$ is diffeomorphic to $B^n$, which is (2). So far, we have shown that for all $q \in D$, $\lambda_1(q) > \eta_2 H(q)$ and $H(q) > \hat{\theta}^{-1} H(p)$. We now prove a diameter bound and an upper mean curvature bound for points in $D$. Let $q_0$ be an arbitrary point on $D$. $M$ is noncompact, so by Corollary 2.13, we can find constants $\hat{a}$ and $\hat{b}$, depending upon $\eta_2$, $n$, and $\gamma_1$, but independent of $q_0$, such that every point $q$ in the intrinsic ball $B := B_g(q_0, \hat{a} H(q_0)^{-1})$ satisfies $H(q) \geq \hat{b}^{-1} H(q_0)$ and there exists a point $q_1 \in B_g(q_0, \hat{a} H(q_0)^{-1})$ with $\lambda_1(q_1) \leq \eta_2 H(q_1)$. Since $\lambda_1 > \eta_2 H$ on the connected region $D \cup N$, it follows that the point $q_1$ is contained in $\tilde{D}$ and therefore the ball $B$ has nonempty intersection with $N$. The mean curvature of points in $N$ is upper bounded by $2H(p)$ and the mean curvature of points in $B$ is lower bounded by $\hat{b}^{-1} H(q_0)$. Putting these together gives the upper bound $H(q_0) \leq 2\hat{b} H(p)$. The intrinsic distance from $q_0$ to $N$ is bounded by $\hat{a} H(q_0)^{-1} \leq \hat{a}\hat{\theta} H(p)^{-1}$ and the distance between any two points on the neck $N$ is bounded by

$100nLH(p)^{-1}$. Therefore $d_g(q_0, p) \leq (\hat{a}\hat{\theta} + 100nL)H(p)^{-1}$. This implies the diameter bound for $D$. Choosing

$$C_0 := \max\{\eta_2^{-1}, \hat{\theta}, 2\hat{b}, 2(\hat{a}\hat{\theta} + 100nL)\}$$

completes the proofs of parts (4) and (5) of the theorem.

**Claim 2:** There exists $\varepsilon_1 \in (0, \varepsilon_0)$, depending upon $n$, $L$, $\varepsilon_0$, $\gamma_1$, and $\gamma_2$, such that if $p$ and $N$ satisfy the assumptions of the theorem, then parts (1) and (3) of the theorem hold.

Since $\varepsilon_1 \leq \varepsilon_0$, part (3) of the theorem holds. We will establish part of the theorem (1) in two steps.

*Step 2.1:* We first determine $\varepsilon_1$ as follows. Via the Neck Detection Lemma, we first choose $\eta_0 := \eta_0(\varepsilon_0, n, L, \gamma_1, \gamma_2) \in (0, \varepsilon_0)$ so that if $\lambda_1(q) \leq \eta_0 H(q)$, then $q$ lies on an $(\varepsilon_0, L)$-neck. Taking $\hat{\varepsilon}_0 := \hat{\varepsilon}_0(\eta_0, n) \in (0, \eta_0)$ sufficient small and using the Neck Detection Lemma once more, we can find $\hat{\eta}_0 := \hat{\eta}_0(\hat{\varepsilon}_0, n, L, \gamma_1, \gamma_2) \in (0, \eta_0)$ so that if $\lambda_1(q) \leq \hat{\eta}_0 H(q)$, then $q$ lies at the center of an $(\hat{\varepsilon}_0, L)$-neck with the property that $\lambda_1 \leq \frac{1}{2}\eta_0 H$ everywhere on the neck. We can now fix our choice of $\varepsilon_1$. Recall that we have $\langle \nu, \omega \rangle \geq 0$ everywhere on $M$ and $\langle \nu, \omega \rangle \leq C\varepsilon_1$ on the neck $N$. We assume $\varepsilon_1$ is sufficiently small so as to satisfy the following two inequalities:

$$\sup_{q \in N} \langle \nu(q), \omega \rangle \leq \hat{\eta}_0 \quad \text{and} \quad \sup_{q \in N} \frac{\lambda_1(q)}{H(q)} \leq \eta_0.$$

It is clear we have chosen $\varepsilon_1$ in a way that only depends upon the given constants.

*Step 2.2:* As in the proof of the previous claim, we consider the surfaces $\Sigma_y$, but now for $y < 0$. To flip our orientation, define $\tilde{\omega} := -\omega$ and let $z := -y$ so that $z_{\max} = y_{\min}$. Then $\langle \nu, \tilde{\omega} \rangle \leq 0$ everywhere on $M$ and $-\hat{\eta}_0 \leq \langle \nu(q), \tilde{\omega} \rangle$ for all $q \in \Sigma_0$. Consider the curves $\tilde{\gamma}(z, q) := \gamma(-z, q)$ for $z \in [0, z_{\max})$ and all $q \in \Sigma_0$. Let $\tilde{\Sigma}_z := \Sigma_{-z}$ and $\tilde{\Sigma}(z_1, z_2) = \Sigma(-z_2, -z_1)$. First, we observe $z_{\max} = \infty$. This is because $\frac{d}{dz}\langle \nu, \tilde{\omega} \rangle = \frac{d}{dy}\langle \nu, \omega \rangle \geq \lambda_1 > 0$ which shows $\langle \nu, \tilde{\omega} \rangle$ is increasing in $z$. On the other hand $\langle \nu, \tilde{\omega} \rangle \leq 0$ everywhere. In other words, for all $q \in \tilde{\Sigma}(0, z_{\max})$,

$-\hat{\eta}_0 \leq \langle \nu(\tilde{\gamma}(z,q)), \tilde{\omega}\rangle \leq 0$, so we can never encounter a critical point of the height function in the direction of increasing $z$.

By construction, every point in $N$ lies at the center of an $(\varepsilon_0, L)$-neck. For sake of contradiction, suppose that there exist points in $\tilde{D}$ that do not lie at the center of an $(\varepsilon_0, L)$-neck. Let $\tilde{z} \in [0,\infty)$ be maximal among heights such that $\frac{\lambda_1}{H} \leq \eta_0$ holds for every $z \in [0, \tilde{z}]$. We have assumed $\varepsilon_1$ is sufficiently small so that this inequality holds at every point on $N$, which implies $\tilde{z} \geq \frac{1}{2}Lr_p > 0$. By maximality of $\tilde{z}$, we can find $\tilde{q} \in \tilde{\Sigma}_{\tilde{z}}$ such that $\lambda_1(\tilde{q}) = \eta_0 H(\tilde{q})$. In other words, $\tilde{q}$ barely lies at the center of an $(\varepsilon_0, L)$-neck $\tilde{N}$. As we argued in the previous proof of the previous claim, $\tilde{\Sigma}(\tilde{z} - 2r_{\tilde{q}}, \tilde{z}) \subset \tilde{N}$. In view of our choice of $\hat{\eta}_0$, we must have $\lambda_1 > \hat{\eta}_0 H \geq \frac{1}{2}\hat{\eta}_0 H(\tilde{q})$ on $\tilde{\Sigma}(\tilde{z} - 2r_{\tilde{q}}, \tilde{z})$, because otherwise we would contradict $\lambda_1(\tilde{q}) = \eta_0 H(\tilde{q})$. Then for any $q \in \tilde{\Sigma}_0$ we find

$$\begin{aligned}
\langle \nu(\gamma(\tilde{z}, q)), \tilde{\omega}\rangle &= \langle \nu(\gamma(\tilde{z} - 2r_{\tilde{q}}, q)), \tilde{\omega}\rangle + \int_{\tilde{z}-2r_{\tilde{q}}}^{\tilde{z}} \frac{d}{dz}\langle \nu, \tilde{\omega}\rangle \, dz \\
&\geq -\hat{\eta}_0 + (2r_{\tilde{q}})\frac{1}{2}\hat{\eta}_0 H(\tilde{q}) \\
&= (n-2)\hat{\eta}_0 > 0.
\end{aligned}$$

This contradicts our previous observation that $\langle \nu, \tilde{\omega}\rangle \leq 0$. This completes the proof of this claim and the theorem. □

For our final result in this section, we show there exist points satisfying the hypothesis of the theorem above.

**Lemma 2.19.** *Suppose $\mathcal{M}$ is noncompact and satisfies $(*)$. Suppose $L \geq 100$. If $0 < \varepsilon_1 < \varepsilon(n)$ is sufficiently small, then for every $t_0 \in (-\infty, 0]$, we can find a point $p_0 \in M$ that lies at the center of an $(\varepsilon_1, L)$-neck at time $t_0$, but not at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck at time $t_0$.*

*Proof.* Fix a time $t_0$ and for simplicity, let us suppress $t_0$ in our notation. As in Lemma 2.16, let $\omega$ be a unit vector in $\mathbb{R}^{n+1}$ such that $\langle \nu, \omega\rangle \geq 0$ everywhere on $M$. Assume $\varepsilon_1$ is sufficiently small so that $\langle \nu, \omega\rangle \leq \frac{1}{100}$ on any $(2\varepsilon_1, \frac{L}{2})$-neck. First, we claim there exists a point $q_0 \in M$ that does not lie on an $(2\varepsilon_1, \frac{L}{2})$-neck. This is clear for topological reasons. If every point in $M$ lies at the center of an $(2\varepsilon_1, \frac{L}{2})$-neck, then $0 \leq \langle \nu, \omega \rangle \leq \frac{1}{100}$ everywhere on $M$. Let $p \in M$ be an arbitrary point lying at the center of an $(2\varepsilon_1, \frac{L}{2})$-neck $N$. Let $y$ denote the height coordinate defined by $\omega$ normalized so that $y(p) = 0$. The estimate $0 \leq \langle \nu, \omega \rangle \leq \frac{1}{100}$ implies integral curves of $\frac{\omega^\top}{|\omega^\top|^2}$ can be continued indefinitely in either direction of the neck $N$. On the other hand, one connected component of $M \setminus N$ is bounded, and therefore the height $y$ has a one-sided bound. At a critical point of $y$ on the bounded component we find a contradiction.

Now if $q_0$ does not lie at the center of an $(2\varepsilon_1, \frac{L}{2})$-neck, there is an open neighborhood around $q_0$ that does not contain any points at the center of an $(\varepsilon_1, L)$-neck. On the other hand, Corollary 2.13 and the Neck Detection Lemma imply there exist points in $M$ that lie at the center of $(\varepsilon_1, L)$-necks. Among all such points, let $p_0$ be a point of least intrinsic distance to $q_0$. Then $p_0$ lies at the center of an $(\varepsilon_1, L)$-neck, but does not lie at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck. If $p_0$ were at the center of the finer neck, we could express the region around $p_0$ as a graph over an cylinder $\Sigma := S^{n-1} \times [-2L, 2L]$. Following a minimal geodesic connecting $p_0$ to $q_0$ for a little ways and restricting our attention to a suitable subcylinder of $\Sigma$, we will find points that lie at the center of $(\varepsilon_1, L)$-necks closer to $q_0$ than $p_0$, a contradiction. $\square$

## 2.3. Tube and Cap Decomposition: Compact Case

In this section, we will assume $M$ is compact and prove a structure theorem analogous to Theorem 2.18. Our first lemma shows that if $\mathcal{M}$ is not a family of shrinking round spheres, then for sufficiently negative times, the solution must contain a very fine neck.

**Lemma 2.20.** *Suppose $\mathcal{M}$ is compact and satisfies $(*)$. Then either $\mathcal{M}$ is a family of shrinking round spheres or for every $\eta > 0$, we can find a time $T_\eta \in (-\infty, 0]$ such that for every $t \leq T_\eta$, there exists a point $p \in M$ where $\lambda_1(p,t) < \eta H(p,t)$.*

*Proof.* By the classical result of Huisken [31], the inequality $\lambda_1 \geq \eta H$ is preserved by the mean curvature flow for compact initial data in $\mathbb{R}^{n+1}$. Suppose for some $\eta > 0$, there exists a sequence of times $t_j \to -\infty$ such that at each time $t_j$ we have $\lambda_1 \geq \eta H$ on $M$. Since this inequality is preserved forward in time, we conclude $\lambda_1 \geq \eta H$ for all $t \in (-\infty, 0]$. The main result of [34] then implies $\mathcal{M}$ is a family of shrinking round spheres. $\square$

In the compact setting, evidently there is no unit vector $\omega$ that satisfies $\langle \nu, \omega \rangle \geq 0$ everywhere to serve as an approximate axis for every neck. Nevertheless, convexity still implies that the axes of different necks cannot differ by much.

**Lemma 2.21.** *Suppose $F : M \to \mathbb{R}^{n+1}$ is an embedding of a closed, convex hypersurface. Let $L \geq 100$. Suppose $N \subset M$ is an $(\varepsilon, L)$-neck and let $\omega$ denote its axis. There are constants $C := C(n)$ and $\varepsilon(n) > 0$, such that if $0 < \varepsilon < \varepsilon(n)$ and $\tilde{N} \subset M$ is any other $(\varepsilon, L)$-neck, then $|\langle \omega, \nu \rangle| \leq C\varepsilon$ on $\tilde{N}$.*

*Proof.* Let $L \geq 100$ be given and let $N \subset M$ be an $(\varepsilon, L)$-neck. The axis of $N$, denoted $\omega$, is a unit vector parallel to the axis of the cylinder over which $N$ is a graph. Let $p$ be a point on the central sphere of $N$. Let $r_p := \frac{n-1}{H(p)}$. It follows from the definition of a neck that:

(i) $|\langle \nu(q), \omega \rangle| \leq C\varepsilon$ for every $q \in N$.

(ii) For every unit vector $e$ orthogonal to $\omega$, there exists a point $q \in N$ with $d_g(p, q) \leq 2\pi r_p$ where $|\nu(q) - e| \leq C\varepsilon$.

Now suppose $\tilde{N}$ is any other $(\varepsilon, L)$-neck in $M$. Let $\tilde{p}$ be a point on its central sphere and $\tilde{\omega}$ its axis. We may assume $\langle \omega, \tilde{\omega} \rangle \geq 0$. Let $y(q) = \langle F(q) - F(p), \omega \rangle$ denote the height function with respect to $\omega$ and let $\Sigma_0$ denote the level set $y = 0$. It suffices to prove that $|\omega - \tilde{\omega}| \leq C\varepsilon$. Evidently, we may assume $\omega \neq \tilde{\omega}$, otherwise there is nothing to prove.

First, suppose $d_g(\tilde{p}, \Sigma_0) \leq \frac{L}{2} r_p$. This means $\tilde{p} \in N$ and so, by taking $\varepsilon$ small, $H(\tilde{p})$ is as close as we like to $H(p)$. In this case, $B_g(\tilde{p}, 3\pi r_{\tilde{p}}) \subset B_g(\tilde{p}, 4\pi r_p) \subset N$ since $\tilde{p}$ has distance at least $\frac{L}{4} r_p \geq 25 r_p$ from boundary of $N$. Now if $e$ is a unit vector orthogonal to $\tilde{\omega}$, it follows from (ii) that we can find a point $q \in B_g(\tilde{p}, 3\pi r_{\tilde{p}})$ where $|\nu(q) - e| \leq C\varepsilon$. Since $q \in N$, by property (i), we conclude $|\langle e, \omega \rangle| \leq C\varepsilon$. Summing over all directions orthogonal to $\tilde{\omega}$ gives $|\omega - \tilde{\omega}| \leq C\varepsilon$.

Next, suppose $d_g(\tilde{p}, \Sigma_0) > \frac{L}{2} r_p$. We may assume $\omega$ is oriented so that $y(\tilde{p}) > 0$. Let $v := \omega - \langle \omega, \tilde{\omega} \rangle \tilde{\omega}$. By (ii), we can find a point $q \in \tilde{N}$ with $d_g(\tilde{p}, q) \leq 2 r_{\tilde{p}}$ where $|\nu(q) + \frac{v}{|v|}| \leq C\varepsilon$. We cannot have $y(q) < 0$. Otherwise, by convexity of $M$ a minimal geodesic connecting $\tilde{p}$ to $q$ (which must remain in $\tilde{N}$) crosses the neck $N$ and so $d_g(\tilde{p}, q) \geq \frac{L}{2} r_p$. Hence $\frac{L}{2} r_p \leq 2 r_{\tilde{p}}$, or, equivalently, $\frac{H(\tilde{p})}{H(p)} \leq \frac{4\pi}{L} \leq \frac{1}{5}$. However, since $N$ and $\tilde{N}$ intersect (and $\varepsilon$ is small) $\frac{H(\tilde{p})}{H(p)} \geq \frac{1}{2}$, which gives a contradiction. Thus, $y(q) > 0$. In the compact and convex setting, the integral curves of $\frac{\omega^\top}{|\omega^\top|^2}$ emanating from $\Sigma_0$ cover $M \cap \{y \geq 0\}$. Using that $\frac{d}{dy}\langle \nu, \omega \rangle \geq \lambda_1 > 0$, we get $\langle \nu(q), \omega \rangle \geq -C\varepsilon$. Finally, as before

$$-C\varepsilon \leq \langle \nu(q), \omega \rangle = -\Big\langle \frac{v}{|v|}, \omega \Big\rangle + \Big\langle \nu(q) + \frac{v}{|v|}, \omega \Big\rangle \leq -\sqrt{1 - \langle \omega, \tilde{\omega} \rangle^2} + C\varepsilon.$$

Hence $|\omega - \tilde{\omega}| \leq C\varepsilon$. This completes the proof. □

We now prove the structure theorem for compact $\mathcal{M}$ satisfying $(*)$.

**Theorem 2.22.** *Suppose $\mathcal{M}$ is compact, satisfies $(*)$, and is not a family of shrinking round spheres. Given $0 < \varepsilon_0 < \varepsilon(n)$ small and $L \geq 100$, there exist constants $C_0 < \infty$ and $T_0 \leq 0$, depending only upon $\varepsilon_0$, $L$, $n$, $\gamma_1$ and $\gamma_2$, so that the following holds. For any time $t \in (-\infty, T_0]$, we can find disjoint domains $D_1, D_2 \subset M$, and points $p_1, p_2 \in M$ that lie at the center of necks $N_1, N_2$ such that*

*(1) Every point in $M \setminus (D_1 \cup D_2)$ lies at the center of an $(\varepsilon_0, L)$-neck at time $t$.*

*(2) $D_1$ and $D_2$ are diffeomorphic to $B^n$.*

*(3) $\partial D_1 \subset \partial N_1$ and $\partial D_2 \subset \partial N_2$ are cross-sectional spheres of $(\varepsilon_0, L)$-necks.*

*(4) For $i = 1, 2$, the intrinsic diameter of $D_i$ is bounded by $C_0 H(p_i, t)^{-1}$.*

*(5) For $i = 1, 2$, every point $q \in D_i$ satisfies $C_0^{-1} H(p_i, t) \leq H(q, t) \leq C_0 H(p_i, t)$ and $\lambda_1(q, t) \geq C_0^{-1} H(q, t)$.*

*Proof.* Let $0 < \varepsilon_0 < \varepsilon(n)$ small and $L \geq 100$ be given. The argument is very similar to the proof of Theorem 2.18. As in that proof, we begin by fixing certain parameters using the Neck Detection Lemma and the definition of a neck. So that the dependence of the constants is clear, we enumerate our choices.

(a) We assume $\varepsilon_0 < \varepsilon(n)$ is sufficiently small so that if $\omega$ is the axis of an $(\varepsilon_0, L)$-neck, then $|\langle \nu, \omega \rangle| \leq \frac{1}{100}$ everywhere on the neck. Moreover, using Lemma 2.21, we require that if $\tilde{\omega}$ is the axis any other $(\varepsilon_0, L)$-neck, then up to a choice of orientation, we have $|\omega - \tilde{\omega}| \leq \frac{1}{100}$. We also assume parts (i) and (iii) of Lemma 2.17 hold with respect to the axis of any $(\varepsilon_0, L)$-neck (part (ii) holds automatically) and $\frac{9}{10} \leq \frac{H(q_1)}{H(q_2)} \leq \frac{10}{9}$ for any pair of points $q_1, q_2$ on an $(\varepsilon_0, L)$-neck.

(b) Choose $\eta_0 := \eta_0(\varepsilon_0, n, L, \gamma_1, \gamma_2) \in (0, \varepsilon_0)$ so that if $\lambda_1(q) \leq \eta_0 H(q)$, then $q$ lies at the center of an $(\varepsilon_0, L)$-neck.

(c) Choose $\varepsilon_1 := \varepsilon_1(n, \eta_0) \in (0, \eta_0)$ and $\eta_1 := \eta_1(\varepsilon_1, n, L, \gamma_1, \gamma_2) \in (0, \varepsilon_1)$ so that on any $(\varepsilon_1, L)$-neck, we have $\lambda_1 \leq \eta_0 H$ and if $\lambda_1(q) \leq \eta_1 H(q)$, then $q$ lies at the center of an $(\varepsilon_1, L)$-neck.

(d) Choose $\hat{\varepsilon}_1 := \hat{\varepsilon}_1(n, \eta_1) \in (0, \eta_1)$ and $\hat{\eta}_1 := \hat{\eta}_1(\hat{\varepsilon}_1, n, L, \gamma_1, \gamma_2) \in (0, \hat{\varepsilon}_1)$ so that on any $(\hat{\varepsilon}_1, L)$-neck, we have $\lambda_1 \leq \frac{1}{2}\eta_1 H$ and if $\lambda_1(q) \leq \hat{\eta}_1 H(q)$, then $q$ lies at the center of an $(\hat{\varepsilon}_1, L)$-neck.

(e) Choose $\varepsilon_2 := \varepsilon_2(\hat{\eta}_1, \eta_1, n)$ so that the axis $\omega$ of an $(\varepsilon_2, L)$-neck satisfies $|\langle \nu, \omega \rangle| \leq \hat{\eta}_1$ on $N$ and so that $\lambda_1 \leq \frac{1}{2}\eta_1 H$ everywhere on the neck.

(f) Choose $\varepsilon_3 := \varepsilon_3(n, \hat{\eta}_1) \in (0, \hat{\eta}_1)$ and $\eta_3 := \eta_3(\varepsilon_3, n, L, \gamma_1, \gamma_2) \in (0, \varepsilon_3)$ so that if $\lambda_1(q) \leq \eta_3 H(q)$, then $q$ lies at the center of an $(\varepsilon_3, L)$-neck $N$ for which the axis $\omega$ satisfies $|\langle \nu, \omega \rangle| \leq \hat{\eta}_1$.

(g) Finally, using Lemma 2.20 and the Neck Detection Lemma, we choose $T_0 := T_0(\varepsilon_2) \leq 0$, such that if $t \leq T_0$, then we can find a point that lies at the center of an $(\varepsilon_2, L)$-neck.

Fix a time $t \leq T_0$. By (g), we can find a point $p \in M$ that lies at the center of an $(\varepsilon_2, L)$-neck $N$ at time $t$. Having fixed it, let us suppress $t$ from our notation. Let $\omega$ be the axis of $N$ satisfying $|\langle \nu, \omega\rangle| \leq \hat{\eta}_1$ in view of (e). Our goal is to follow the neck $N$ in the direction $\omega$ until we encounter a point which barely lies on an $(\varepsilon_1, L)$-neck. As in the notation of the previous section, let $y$ denote the height function with respect to axis defined by $\omega$; let $\Sigma_y$ for $y \in [0, y_{\max})$ denote the level sets of the height function (with $p \in \Sigma_0$); and let $\gamma(\cdot, y)$ denote the integral curves of $\frac{\omega^\top}{|\omega^\top|^2}$. Note $y_{\max} < \infty$ because $M$ is compact. By definition every integral curve is defined for $y \in [0, y_{\max})$ and there exists some point $q \in \Sigma_0$ such that $\gamma(q, y) \to \tilde{q} \in M$ as $y \to y_{\max}$ satisfying $\langle \nu(\tilde{q}), \omega\rangle = 1$. As we noted in the proof of Theorem 2.18, since our hypersurface is strictly convex, elementary arguments imply $\Sigma_{y_{\max}} := \{\lim_{y\to y_{\max}} \gamma(q, y) : q \in \Sigma_0\} = \{\tilde{q}\}$; so every curve converges to the same point.

Let $\tilde{y} \in [0, y_{\max})$ be the supremum over heights $y$ such that $\lambda_1 \leq \eta_1 H$ holds for every $y \in [0, \tilde{y}]$. Since by (e) we have $\lambda_1 \leq \frac{1}{2}\eta_1 H$ on $N$, this gives $\tilde{y} \geq \frac{1}{2} L r_p > 0$. On the other hand, if $\tilde{q}$ lies on an $(\varepsilon_1, L)$-neck, then the axis of such a neck would be nearly orthogonal to $\omega$. However, this would contradict our application of Lemma 2.21 in (a) above. Hence, we must have $\lambda_1(\tilde{q}) > \eta_0 H(\tilde{q})$ and $\tilde{y} < y_{\max}$.

On $\Sigma_{\tilde{y}}$, we find a point $\tilde{p}$ satisfying $\lambda_1(\tilde{p}) = \eta_1 H(\tilde{p})$, which barely lies at the center of $(\varepsilon_1, L)$-neck $\tilde{N}$. As we have argued before, choices (c) and (d) ensure $\lambda_1 \geq \hat{\eta}_1 H \geq \frac{1}{2}\hat{\eta}_1 H(p)$ on $\Sigma(\tilde{y} - 2r_{\tilde{p}}, \tilde{y})$. Otherwise, there is a point on $\tilde{N}$ near $\tilde{p}$ which lies on an $(\hat{\varepsilon}_1, L)$-neck, which would contradict $\lambda_1(\tilde{p}) = \eta_1 H(\tilde{p})$. Now as in Step 2.2 of Theorem 2.18, integrating this convexity lower bound along the integral curves of $\omega$, we get $\langle \nu, \omega\rangle > (n-2)\hat{\eta}_1$ on $\Sigma_{\tilde{y}}$. We can obtain the uniform convexity, mean curvature, and diameter estimates for the region $\Sigma(\tilde{y}, y_{\max})$ in the same fashion as Step 1.4 in the proof of Theorem 2.18. The estimate $\langle \nu, \omega\rangle > (n-2)\hat{\eta}_1$ holds for $y \in [\tilde{y}, y_{\max}]$ because $\langle \nu, \omega\rangle$ is increasing along integral curves. The mean curvature lower bound, $H \geq \theta^{-1} H(\tilde{p})$, follows from part (iii) of Lemma 2.17. If a point $q$ in the region $\Sigma(\tilde{y}, y_{\max})$ satisfies $\lambda_1(q) \leq \eta_3 H(q)$, then by choice (f) the point $q$ lies at the center of an $(\varepsilon_3, L)$-neck $\hat{N}$. Our assumptions imply axis $\hat{\omega}$ of this neck cannot equal $\omega$ otherwise we contradict the estimate $\langle \nu, \omega\rangle >$

$(n-2)\hat{\eta}_1$. However, if $\omega$ and $\hat{\omega}$ differ, then by arguments we have made before, $(n-2)\hat{\eta}_1 \leq -\sqrt{1-\langle\omega,\hat{\omega}\rangle^2}+\hat{\eta}_1$, which is a contradiction. Therefore $\lambda_1 > \eta_3 H$ for $y \in [\tilde{y}, y_{\max}]$. In Corollary 2.13, choose $\hat{a}$ and $\hat{b}$ with respect to the constants $\frac{1}{2}\eta_3$ and $\gamma_1$. Let $q \in \Sigma(\tilde{y}, y_{\max})$. Then either $M = B_g(q, \hat{a}H(q)^{-1})$ or there exists a point $q_1 \in B_g(q, \hat{a}H(q)^{-1})$ where $\lambda_1(q_1) \leq \frac{1}{2}\eta_3 H(q_1)$. In either case, $B_g(q, \hat{a}H(q)^{-1})$ has nonempty intersection with the neck $\tilde{N}$. Arguing as in Step 1.4 in the proof of Theorem 2.18 yields the desired upper mean curvature bound and the diameter estimate for the region $\Sigma(\tilde{y}, y_{\max})$.

To finish, we let $p_1 := \tilde{p}$ and let $D_1 \subset M \setminus \tilde{N}$ be the connected component containing $\tilde{q}$. Our arguments show claims (2)-(5) in the theorem statement hold for $D_1$ and $p_1$. Repeating the entire argument above for $y \leq 0$, we similarly find $p_2$ and $D_2$. Moreover, since $p_1$ and $p_2$ are the first points that barely lie at the center of $(\varepsilon_1, L)$-necks, our choices (b) and (c) ensures that every point in $M \setminus (D_1 \cup D_2)$ lies at the center of an $(\varepsilon_0, L)$-neck, which is (1). This completes the proof of the theorem. □

Theorem 2.22 addresses the structure of the compact ancient solution for sufficiently negative times. The following corollary addresses the structure of the solution at later times.

**Corollary 2.23.** *Suppose $\mathcal{M}$ is compact and satisfies $(*)$. Given $0 < \varepsilon_0 < \varepsilon(n)$ and $L \geq 100$, there exists a constant $C < \infty$, depending only upon $\varepsilon_0$, $L$, $n$, $\gamma_1$, and $\gamma_2$ so that the following holds. For every time $t \in (-\infty, 0]$, either there exists domains $D_1, D_2$, points $p_1, p_2$, and necks $N_1, N_2$ satisfying the conclusions of Theorem 2.22, or else, for any point $p \in M$, there holds:*

*(1) The intrinsic diameter of $M$ is bounded by $C_0 H(p,t)^{-1}$.*

*(2) Every point $q \in M$ satisfies $C_0^{-1}H(p,t) \leq H(q,t) \leq C_0 H(p,t)$ and $\lambda_1(q,t) \geq C_0^{-1}H(q,t)$.*

*Proof.* Let $0 < \varepsilon_0 < \varepsilon(n)$ and $L \geq 100$ be given (where $\varepsilon(n)$ is determined in the proof of Theorem 2.22). We follow the first part of the proof of Theorem 2.18 to determine a constant $\varepsilon_2$ depending only upon the given constants. Since $t \leq T_0$ is addressed by the previous theorem, suppose $T_0 < t \leq 0$. If there exists a point $p \in M$ that lies at the center of an $(\varepsilon_2, L)$-neck at time $t$, then the proof of the previous theorem goes through unchanged, and we conclude there exists $D_i, p_i$, and $N_i$ for $i = 1, 2$ at time $t$. If, however, there does not exist a point in $M$ that lies at the center of an $(\varepsilon_2, L)$-neck, then by the Neck Detection Lemma, we can find a constant $\eta := \eta(\varepsilon_0, L, n, \gamma_1, \gamma_2) > 0$, such that $\lambda_1 \geq \eta H$ everywhere on $M$. Let $p$ be an arbitrary point in $M$. By Corollary 2.13, we can find constants $\hat{a}$ and $\hat{b}$ depending upon $n, \eta$, and $\gamma_1$ such that $M = B_{g(t)}(p, \hat{a}H(p,t)^{-1})$ and $H(q,t) \geq \hat{b}^{-1}H(p,t)$ for every $q \in M$. We can take $C_0 := \max\{\eta^{-1}, \hat{a}, \hat{b}\}$. This completes the proof. □

## 2.4. Noncollapsing

To complete the proof of our main theorem, we first show how controlled geometry of the cap implies $\alpha$-noncollapsing for an appropriate $\alpha$.

**Theorem 2.24.** *Let $n \geq 2$ and $F : M \to \mathbb{R}^{n+1}$ be an embedding of a closed, convex hypersurface. Given a large positive constant $C_0 \geq n$ such that $C_0^{-1} \leq H \leq C_0$ and $\lambda_1 \geq C_0^{-1}H$ everywhere on $M$, there exists a real number $\alpha := \alpha(C_0) > 0$ such that $M$ is $\alpha$-noncollapsed.*

*Proof.* Let $\Omega$ denote the convex interior of $F(M)$, let $\nu$ denote the outward pointing normal vector, and let $r_{\text{in}}(p)$ denote the inscribe radius at $p \in M$. Take $\alpha := (2C_0)^{-5}$. We claim that $M$ is $\alpha$-noncollapsed. This means we must show $r_{\text{in}}(p) \geq \alpha H(p)^{-1}$ for every $p \in M$. Since by assumption $H(p)^{-1} \leq C_0$, it suffices to show that $r_{\text{in}}(p) \geq \frac{1}{32}C_0^{-4}$.

Define $r_0 := (2C_0)^{-1}$ and note $r_0 < 1$. Fix a point $p_0 \in M$ and let $\omega := -\nu(p_0)$. Without loss of generality, we may assume that $F(p_0) + r_0\omega = 0 \in \mathbb{R}^{n+1}$. Define the height function $f(p) := \langle F(p) - F(p_0), \omega\rangle$.

*Step 1*: We first show that $|F(p)| \geq r_0$ for all $p \in B_g(p_0, r_0)$. Suppose for sake of contradiction that $|F(p_1)| < r_0$ for some $p_1 \in B_g(p_0, r_0)$. Let $s_* := d_g(p_0, p_1) \leq r_0$ and let $\gamma : [0, s_*] \to M$ be a unit-speed geodesic from $p_0$ to $p_1$. Define $\tilde{\gamma}(s) := F(\gamma(s))$ and let $\rho(s) = \frac{1}{2}|\tilde{\gamma}(s)|^2$. By definition

$\rho(0) = \frac{1}{2}|F(p_0)|^2 = \frac{1}{2}r_0^2$ and

$$\rho'(0) = \langle \tilde{\gamma}'(0), F(p_0)\rangle = -r_0\langle \tilde{\gamma}'(0), \omega\rangle = \langle \tilde{\gamma}'(0), \nu(p_0)\rangle = 0.$$

Moreover, since $\gamma$ is a geodesic $\tilde{\gamma}''(s) = -h(\gamma'(s), \gamma'(s))\nu(\gamma(s))$ and hence

$$\rho''(s) = |\tilde{\gamma}'(s)|^2 + \langle \tilde{\gamma}''(s), \tilde{\gamma}(s)\rangle = 1 - h\big(\gamma'(s), \gamma'(s)\big)\langle \nu(\gamma(s)), \tilde{\gamma}(s)\rangle.$$

In particular, by (strict) convexity, $\rho''(s) > 1 - H(\gamma(s))|\tilde{\gamma}(s)| \geq 1 - C_0|\tilde{\gamma}(s)|$. This implies

$$\rho''(0) > 1 - C_0|F(p_0)| = 1 - C_0 r_0 \geq \frac{1}{2}.$$

By assumption $\rho(s_*) = \frac{1}{2}|F(p_1)|^2 < \frac{1}{2}r_0^2 = \rho(0)$. Together with the fact that $\rho'(0) = 0$ and $\rho''(0) > 0$, this means there exists $s_0 \in (0, s_*)$ where $\rho(s)$ attains a local maximum larger than $\frac{1}{2}r_0^2$. At this maximum,

$$0 \geq \rho''(s_0) > 1 - C_0|\tilde{\gamma}(s_0)|,$$

which implies $|\tilde{\gamma}(s_0)| > C_0^{-1} = 2r_0$. On the other hand, since $\tilde{\gamma}(s)$ is a unit-speed curve

$$|\tilde{\gamma}(s_0)| \leq |\tilde{\gamma}(s_0) - \tilde{\gamma}(0)| + |\tilde{\gamma}(0)| \leq s_0 + |F(p_0)| \leq 2r_0,$$

which is a contradiction.

*Step 2:* Next, we show that $\langle \nu(p), \nu(p_0)\rangle \geq \frac{1}{2}$ for all $p \in B_g(p_0, r_0)$. Let $p_1 \in B_g(p_0, r_0)$ and let $\gamma : [0, s_*] \to M$, where $s_* = d_g(p_0, p_1) \leq r_0$ be a minimizing unit-speed geodesic from $p_0$ to $p_1$. As before, let $\tilde{\gamma}(s) = F(\gamma(s))$. Let $e_1, \ldots, e_n$ denote an orthonormal frame of $T_{\gamma(s)}M$. Then

$$\frac{d}{ds}\langle \nu(\gamma(s)), \nu(p_0)\rangle = \langle D_{\tilde{\gamma}'(s)}\nu(\gamma(s)), \nu(p_0)\rangle = \sum_{i=1}^{n} h(\gamma'(s), e_i)\langle e_i, \nu(p_0)\rangle.$$

Consequently,

$$\Big|\frac{d}{ds}\langle \nu(\gamma(s)), \nu(p_0)\rangle\Big| \leq |h|(\gamma(s)) \leq H(\gamma(s)) \leq C_0.$$

Since $s \leq s_* \leq r_0$, this implies that

$$\big|\langle \nu(p_1), \nu(p_0)\rangle - 1\big| \leq C_0 s_* \leq C_0 r_0 \leq \frac{1}{2}.$$

In particular, $\langle \nu(p_1), \nu(p_0)\rangle \geq \frac{1}{2}$.

*Step 3:* Next, we show that $\{p \in M : f(p) \leq r_0^4\} \subset B_g(p_0, r_0)$. In view of the previous step, $\partial B_g(p_0, r_0)$ nonempty and $F(B_g(p_0, r_0))$ is a graph over $dF_{p_0}(T_{p_0}M)$. If the claim is not true, then we can find $p_1 \in \partial B_g(p_0, r_0)$ such that $y_1 := f(p_1) \leq r_0^4$. As before, consider a unit-speed geodesic $\gamma(s)$ between $p_0$ and $p_1$ and let $\tilde{\gamma}(s) = F(\gamma(s))$. Define $k(s) := f(\gamma(s))$. Then by definition $k(0) = 0$ and

$$k'(0) = \langle \tilde{\gamma}'(0), \omega\rangle = 0.$$

Moreover, since $\gamma(s)$ is a geodesic, we have

$$k''(s) = \langle \tilde{\gamma}''(s), \omega\rangle = h(\gamma'(s), \gamma'(s))\langle \nu(\gamma(s)), \nu(p_0)\rangle.$$

From the assumption $\lambda_1 \geq C_0^{-1} H \geq C_0^{-2}$ and the previous step, we deduce that $k''(s) \geq \frac{1}{2} C_0^{-2}$ for all $s \in [0, r_0]$. Integration then implies that $y_1 = k(r_0) \geq \frac{1}{2} C_0^{-2} r_0^2 = 2r_0^4$, a contradiction.

*Step 4:* Finally, we claim the set $U := \{x \in \mathrm{R}^{n+1} : |x| < r_0,\ \langle x - F(p_0), \omega\rangle \leq r_0^4\}$ is contained in $\Omega$. We begin by observing that $\bar{x} = F(p_0) + r_0^4\omega \in \Omega$. Note that $F(p_0) + y\omega \in \Omega$ for $y > 0$ sufficiently small. If $\bar{x} \notin \Omega$, then there is a point $p_1 \in M$, where $\nu(p_1) = -\nu(p_0)$, but $f(p_1) \leq r_0^4$, in contradiction with what have established in the previous two steps. Now suppose there is a point $\bar{y} \in U$ such that $\bar{y} \notin \Omega$. Since $\langle \bar{y}, \omega\rangle \leq r_0^4$, the segment $\{(1-s)\bar{x} + s\bar{y} : s \in [0,1]\}$ must intersect $\partial\Omega$ in a point which lies in $F(B_g(p_0, r_0))$. But the segment $\{(1-s)\bar{x} + s\bar{y} : s \in [0,1]\}$ is contained in the ambient Euclidean $(n+1)$-ball of radius $r_0$ around this origin. So the segment $\{(1-s)\bar{x} + s\bar{y} : s \in [0,1]\}$ cannot intersect $\partial\Omega$, otherwise we would contradict Step 1. Hence

$U \subset \Omega$. Now, it is easy to check that Euclidean $(n+1)$-ball $\{x \in \mathrm{R}^{n+1} : |x - F(p_0) - \frac{1}{2}r_0^4\omega| < \frac{1}{2}r_0^4\} \subset U$. So we conclude that $r_{\text{in}}(p_0) \geq \frac{1}{2}r_0^4 = \frac{1}{32}C_0^{-4}$. As $p_0$ was chosen arbitrarily, this completes the proof of the theorem. □

The proof of the above result is local in the sense that it only depended upon examining the hypersurface $M$ in a small geodesic ball around each point $p$. From the proof, we can deduce the following corollary.

**Corollary 2.25.** *Let $n \geq 2$ and $F : M \to \mathbb{R}^{n+1}$ be an embedding of a (possibly noncompact) complete, convex hypersurface. Let $D \subset M$ be an open region and suppose there exists a constant $C_0$ such that $C_0^{-1} \leq H \leq C_0$ and $\lambda_1 \geq C_0^{-1}H$ everywhere on $D$. Then there exists $\alpha := \alpha(C_0) > 0$ such that if $p \in D$ is any point with $d_g(p, \partial D) > \frac{1}{2}C_0^{-1}$, then the inscribe radius at $p$ is at least $\frac{\alpha}{H(p)}$.*

The corollary shows that points on a cap have a uniform inscribe radius. On the other hand, a point that lies at the center of a fine neck has an inscribe radius that is comparable to the inscribe radius of the cylinder. That is, a point $p$ that lies at the center of a $(\varepsilon, L)$-neck in a convex hypersurface has $r_{\text{in}}(p) \geq \alpha(n)H(p)^{-1}$ for $\varepsilon < \varepsilon(n)$ and $L \geq 100$.

With our structure theorem and these noncollapsing arguments, we can show the time slices of a solution $\mathcal{M}$ are all noncollapsed with a uniform noncollapsing constant. This will complete the proof of Theorem 2.1.

**Theorem 2.26.** *Suppose $\mathcal{M}$ satisfies $(*)$. There exists $\alpha > 0$, depending only upon $n$, $\gamma_1$, and $\gamma_2$, such that $\mathcal{M}$ is $\alpha$-noncollapsed.*

*Proof.* First suppose $\mathcal{M}$ is noncompact. Fix a time $t_0$ and let $\Omega$ denote the convex interior of $F(M, t_0)$. Set $L = 100$ and fix some $0 < \varepsilon_0 < \varepsilon(n)$ sufficiently small for Theorem 2.18 to apply. For these values of $L$ and $\varepsilon_0$, we can find constants $\varepsilon_1 \in (0, \varepsilon_0)$ and $C_0 > n$, depending only upon $n, \gamma_1$, and $\gamma_2$, so that the conclusions of Theorem 2.18 hold. We can assume $\varepsilon_1$ is small enough for Lemma 2.19 to apply. Then there exists a point $p \in M$ that lies at the center of an $(\varepsilon_1, L)$-neck, but not at the center of an $(\frac{\varepsilon_1}{2}, 2L)$-neck (at time $t_0$).

After a rescaling, we may assume $H(p) = 1$. The results of Theorem 2.18 tell us $M$ is the union of a compact connected component $D$, the neck $N$, and an unbounded connected component $\tilde{D}$. In $\tilde{D} \cup N$, every point lies at the center of an $(\varepsilon_0, L)$-neck. Moreover, the estimates $C_0^{-1} \leq H \leq C_0$ and $\lambda_1 \geq C_0^{-1} H$ hold in $D \cup N$.

First, suppose $q \in M$ lies at the center of an $(\varepsilon_0, L)$-neck $N$. An exact cylinder is $\alpha$-noncollapsed for $\alpha = n - 1$. As remarked above, the inscribe radius of $q$ is at least $\frac{\alpha}{H(q)}$ for some $\alpha = \alpha(n)$. Now suppose $q \in D$ is a point on the cap. The intrinsic length of the neck $N$ is approximately $100(n-1)$, where as $\frac{1}{2} C_0^{-1} < \frac{1}{2}$. Thus clearly $d_g(q, \partial \tilde{D}) > \frac{1}{2} C_0^{-1}$. By Corollary 2.25 applied to the region $D \cup N$, the inscribe radius of $q$ is at least $\frac{\alpha(C_0)}{H(q)}$. Since every point in $M$ is either contained in $D$ or at the center of an $(\varepsilon_0, L)$-neck, the hypersurface $M$ is $\alpha$-noncollapsed everywhere for $\alpha$ independent of the time $t_0$. Hence $\mathcal{M}$ is $\alpha$-noncollapsed.

If $\mathcal{M}$ is compact, the argument is similar. We may assume $\mathcal{M}$ is not a family of shrinking round spheres. For $L = 100$, $0 < \varepsilon_0 < \varepsilon(n)$ sufficiently small, and for $t_0$ sufficiently negative, we conclude $M$ is the union of a neck region and two caps $D_1$, $D_2$, which have uniform mean curvature and convexity estimates and whose boundaries are spherical cross-sections of $(\frac{\varepsilon_0}{2}, 2L)$-necks. The arguments above show each region is $\alpha$-noncollapsed and hence $\mathcal{M}$ is $\alpha$-noncollapsed for all sufficiently negative times. Since the noncollapsing constant is preserved forward in time for compact solutions, this completes the proof. □

## 2.5. Strict Maximum Principle for Tensors: An Example

In this section, we explain how to show the sets $E_{p,t}$ introduced in the proof of Lemma 2.11 are invariant under parallel transport (with respect to $g(t)$) using the strict maximum principle. A reference for this argument is [7]. For the convenience of the reader, we make the minor modifications necessary for our setting.

We begin by introducing relevant notation. Suppose we have a local solution to the mean curvature flow $\mathcal{M}$ in $\mathbb{R}^{n+1}$ defined on $\Omega \times (-T, 0)$, for $\Omega$ a smooth domain in $\mathbb{R}^n$ and $T > 0$. Suppose we know our solution is weakly convex and that $\lambda_1 + \lambda_2 \equiv \beta H$ for some $\beta \in (0, \frac{1}{n-1})$.

Define a vector bundle $\mathcal{E}$ over $\Omega \times (-T, 0)$ to be the pullback of the tangent bundle $T\Omega$ under the projection $\Omega \times (-T, 0) \to \Omega$ so that $\mathcal{E}_{(p,t)} = T_p\Omega$. We have a bundle metric $g(t)$ defined on $\mathcal{E}$ and there is a standard compatible connection $D$ on $\mathcal{E}$ that extends the Levi-Civita connection. Namely, for any section $X$ of $\mathcal{E}$, we define $D_{\frac{\partial}{\partial t}} X = \frac{\partial}{\partial t} X - \sum_{k=1}^{n} Hh(X, e_i)e_i$ where $\{e_1, \ldots, e_n\}$ is an orthonormal frame with respect to $g(t)$. It is easy to check that $D_{\frac{\partial}{\partial t}} g(t) = 0$. Finally, let $\mathcal{O}$ denote the orthonormal frame bundle associated to $\mathcal{E}$. For every $(p, t)$, the fiber $\mathcal{O}_{(p,t)}$ consists of orthonormal frames $\underline{e} = \{e_1, \ldots, e_n\}$ of $T_p\Omega$ with respect to $g(t)$. Given a point $(p, t)$ and an orthonormal frame $\underline{e} = \{e_1, \ldots, e_n\}$ with respect to $g(t)$, the tangent space $T_{\underline{e}}\mathcal{O}$ decomposes into the direct sum of vertical and horizontal vector spaces, $\mathcal{V}_{\underline{e}}$ and $\mathcal{H}_{\underline{e}}$. The vertical space is the tangent space to the fiber $\mathcal{O}_{(p,t)}$ and vertical vectors are induced by infinitesimal action of $O(n)$ upon $\underline{e}$. The horizontal space is defined through the connection $D$. To define it, consider any smooth path $\gamma : (-\varepsilon, \varepsilon) \to \Omega \times (-T, 0)$ such that $\gamma(0) = (p, t)$. Extend the frame $\underline{e}$ by parallel transport along $\gamma$ using $D$. This defines a path $\tilde{\gamma} : (-\varepsilon, \varepsilon) \to \mathcal{O}$. The vector $\tilde{\gamma}'(0)$ is defined to be the horizontal lift of $\gamma'(0)$. Define $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $\mathcal{Y}$ in $T_{\underline{e}}\mathcal{O}$ to be the horizontal lifts of $e_1, \ldots, e_n$ and $\frac{\partial}{\partial t}$, respectively. Then one has $T_{\underline{e}}\mathcal{O} = \mathcal{V}_{\underline{e}} \oplus \operatorname{span}\{\mathcal{X}_1, \ldots, \mathcal{X}_n, \mathcal{Y}\}$.

With the notation above, for each orthonormal frame $\underline{e} \subset \mathcal{E}_{(p,t)}$, we define

$$\varphi(\underline{e}) = h_{g(t)}(e_1, e_1) + h_{g(t)}(e_2, e_2) - \beta H(p, t).$$

This defines a smooth, nonnegative function $\varphi : \mathcal{O} \to \mathbb{R}$. Recall the evolution equations $D_{\frac{\partial}{\partial t}} h = \Delta h + |A|^2 h$ and $D_{\frac{\partial}{\partial t}} H = \Delta H + |A|^2 H$. It follows that

$$\begin{aligned}\mathcal{Y}(\varphi) - \mathcal{X}_i(\mathcal{X}_i(\varphi)) &= (D_{\frac{\partial}{\partial t}} h - \Delta h)(e_1, e_1) + (D_{\frac{\partial}{\partial t}} h - \Delta h)(e_2, e_2) - \beta(\frac{\partial}{\partial t} H - \Delta H) \\ &= |A|^2 \varphi \geq 0.\end{aligned}$$

This is a degenerate elliptic equation for $\varphi$. Let $\mathcal{F} = \{\underline{e} \in \mathcal{O} : \varphi(\underline{e}) = 0\}$ denote the zero set of $\varphi$. Fix a time $\tau \in (-T, 0)$. We claim the set of all two-frames $\{e_1, e_2\}$ that are orthonormal with respect to $g(\tau)$ and satisfy $h_{g(\tau)}(e_1, e_1) + h_{g(\tau)}(e_2, e_2) = \beta H_{g(\tau)}$ is invariant under parallel transport.

Let $\gamma : [0,1] \to \Omega$ be a smooth path and let $\underline{v}(s) := \{v_1(s), \ldots, v_n(s)\}$ be a parallel orthonormal frame along $\gamma$ for $s \in [0,1]$ with respect to $g(\tau)$. This defines a smooth path $\underline{v} : [0,1] \to \mathcal{O}$ with the property that $\underline{v}(s)$ lies above $\gamma(s)$ and $\underline{v}'(s)$ is the horizontal lift of $\gamma'(s)$. Evidently, we can find smooth functions $f_1, \ldots, f_n : [0,1] \to \mathbb{R}$ such that $\gamma'(s) = \sum_{j=1}^n f_j(s) v_j(s)$ for $s \in [0,1]$ and this implies that

$$\underline{v}'(s) = \sum_{j=1}^{n} f_j(s) \mathcal{X}_j|_{\underline{v}(s)}.$$

If we assume that $h_{g(\tau)}(v_1(0), v_1(0)) + h_{g(\tau)}(v_2(0), v_2(0)) = \beta H_{g(\tau)}$, then $\underline{v}(0) \in \mathcal{F}$. At this point, all of the assumptions of Bony's strict maximum principle for degenerate elliptic equations (Corollary 9.7 in [7]) have been verified, and thus we conclude $\underline{v}(s) \in \mathcal{F}$ for $s \in [0,1]$. This completes the proof.

# Chapter 3

# Local Compactness for MCFs

For the moment, we now take a short excursion to discuss compactness in the space of solutions of the mean curvature flow. Our eventual goal, in Chapter 4, is to use the classification in Chapter 2 to deduce information about singularity formation on a meaningful scale. The works of Perelman [43, 44] and White [48, 49] are the archetype for this sort of argument. We will do this through a blow-up argument, which requires the appropriate compactness result. The present chapter is meant to serve as a reference for the reader on a local compactness result for the mean curvature flow given uniform local curvature estimates. This type of result is well-known and often used by experts. Our goal is to make precise the local compactness results we use in our proof of the canonical neighborhood theorem in Chapter 4, as well as provide sufficiently many details for the proofs of these results.

In [17], building on work of Langer in [35], Breuning proves a local compactness result for immersions that is close to what we desire here. The key idea behind the result, originally due to Langer, is that estimates for the second fundamental form, and its derivatives, yield estimates for the immersion and its derivatives when the immersion is given as a graph. The result of [17], however, requires estimates for the second fundamental form in extrinsic Euclidean balls rather than in intrinsic geodesic balls, which are more natural here. The result we are after is therefore closer in spirit Hamilton's compactness result [26] for the Ricci flow and its local adaptation by Cao and Zhu in [18]. We found their work to be a good reference for this type of result and our compactness theorem will be a corollary of theirs.

First, a word on notation. Suppose that $F : (M, g) \to (\mathbb{R}^N, g_{\text{flat}})$ is an isometric immersion.

As usual, we let $\langle \cdot\,, \cdot\rangle$ also denote that flat metric on $\mathbb{R}^N$. Let $\bar{g}$ be any other metric on $M$ and let $\nabla_{\bar{g}}$ denote the corresponding covariant derivative. To estimate derivatives of $F$ with respect to $\bar{g}$, we will think of the immersion $F$ as a tuple of $\mathbb{R}$-valued functions $F = (F^1, \ldots, F^N)$ and use the notation

$$|\nabla^k_{\bar{g}} F|^2_{\bar{g}} := \sum_{m=1}^{N} |\nabla^k_{\bar{g}} F^m|^2_{\bar{g}}$$

where $|\nabla^k_{\bar{g}} F^m|_{\bar{g}}$ denotes the usual norm of the $(0, k)$-tensor $\nabla^k_{\bar{g}} F^m$ for each $m$. We will similarly view the second fundamental form $A$ as an $\mathbb{R}^N$-valued $(0, 2)$-tensor and the mean curvature vector, here denoted by $\vec{H}$, as an $\mathbb{R}^N$-valued function. This means we view $A$ as a section of $T^*M^{\otimes 2} \otimes F^*\mathbb{R}^N$ rather than as a section of $T^*M^{\otimes 2} \otimes NM$. Let $\nabla = \nabla_g$ and $\nabla^\perp = \nabla^\perp_g$ denote the induced connections on $T^*M^{\otimes k} \otimes F^*\mathbb{R}^N$ and $T^*M^{\otimes k} \otimes NM$ respectively. Acting on the second fundamental form, these derivatives are related in a local coordinates $x_1, \ldots, x_n$ on $M$ by

$$\nabla_i A_{jk} = \frac{\partial}{\partial x_i} A_{jk} - \Gamma^l_{ij} A_{lk} - \Gamma^l_{ik} A_{jl};$$
$$\nabla^\perp_i A_{jk} = (\nabla_i A_{jk})^\perp = \nabla_i A_{jk} - \sum_{l=1}^{n} \langle A_{jk}, A_{il}\rangle e_l,$$

where $e_1, \ldots, e_n$ is a local orthonormal frame for $TM$ (the equalities above hold as sections of $F^*\mathbb{R}^N = TM \oplus NM$). In particular, this means that $|\nabla A|^2 = |\nabla^\perp A|^2 + |\langle A, A\rangle|^2$ (where $\langle A, A\rangle$ is a $(0, 4)$-tensor). It is perhaps more natural to assume bounds on $|(\nabla^\perp)^k A|$ than it is to assume bounds on $|\nabla^k A|$. However, taking derivatives of the second identity above, we can estimate

$$|\nabla^k A| \leq |(\nabla^\perp)^k A| + C_k \sum_{p+q=k-1} |(\nabla^\perp)^p A||(\nabla^\perp)^q A| + C_k \sum_{p+q+r=k-1} |(\nabla^\perp)^p A||(\nabla^\perp)^q A||(\nabla^\perp)^r A|.$$

for a constant $C_k = C_k(n, N)$. Thus estimates for $|\nabla^k A|$ for $0 \leq k \leq \bar{k}$ are equivalent to estimates for $|(\nabla^\perp)^k A|$ for $0 \leq k \leq \bar{k}$ up to some constants. In the lemmas and propositions below, we will assume estimates on $|\nabla^k A|$ rather than on $|(\nabla^\perp)^k A|$. In part, this is because $\nabla^{k+2} F = \nabla^k A$ as sections of $T^*M^{\otimes k+2} \otimes F^*\mathbb{R}^N$. However, we note that uniform bounds for $|\nabla^k F| = |\nabla^k_{F^*g_{\text{flat}}} F|_{F^*g_{\text{flat}}}$ (a quantity which is invariant under reparametrization) along a sequence of immersions is not enough to deduce compactness via Arzela-Ascoli. We must obtain estimates with respect to a fixed background metric.

We will reserve the notation $|\nabla^k A| = |\nabla^k_g A|_g$ to denote the norm of derivatives of the curvature with respect to the metric $g$ induced by the immersion. For any other fixed background metric $\bar{g}$, we use the notation $|\nabla^k_{\bar{g}} A|_{\bar{g}}$. Estimates with respect to a fixed background metric are needed for extracting limits, but estimates for the second fundamental form are usually obtained with respect to the metric induced by the immersion.

**Definition 3.1** (cf. Definition 4.1.1 in [18])**.** Let $(M_j, g_j, p_j)$ be a sequence of pointed, complete $n$-dimensional Riemannian manifolds and suppose for each $j$ that $F_j : (M_j, g_j, p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ is a pointed isometric immersion. Let $B_j := B_{g_j}(p_j, \rho_j) \subset M_j$ denote the open geodesic ball centered at $p_j \in M_j$ of radius $\rho_j \in (0, \infty]$. Suppose $\rho_j \to \rho^* \in (0, \infty]$. Let $(B_\infty, g_\infty, p_\infty)$ be a (possibly incomplete) Riemannian manifold such that $B_\infty = B_{g_\infty}(p_\infty, \rho^*)$, the open geodesic ball centered at $p_\infty$ of radius $\rho^*$ with respect to $g_\infty$. We say the sequence of pointed immersions $F_j|_{B_j}$ converges in $C^\infty_{\text{loc}}$ to a pointed isometric immersion $F_\infty : (B_\infty, g_\infty, p_\infty) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ if:

- We can find a sequence of smooth, relatively compact open sets $U_j$ in $B_\infty$ satisfying $p_\infty \in U_j$, $U_j \subset U_{j+1}$, and $\cup_j U_j = B_\infty$.
- We can find a sequence of diffeomorphisms $\phi_j : U_j \to \phi_j(U_j) \subset B_j$ satisfying $\phi_j(p_\infty) = p_j$. Moreover, for all $\rho \in (0, \rho^*)$, if $j$ is sufficiently large, then $B_{g_j}(p_j, \rho) \subset \phi_j(U_j)$.
- The sequence of immersions $F_j \circ \phi_j$ converges to $F_\infty$ smoothly with respect to $g_\infty$ on every compact subset of $B_\infty$.

To be specific, in the definition above, the sequence $F_j \circ \phi_j : U_j \to \mathbb{R}^N$ converges to $F_\infty : B_\infty \to \mathbb{R}^N$ smoothly with respect to $g_\infty$ on compact subsets as $\mathbb{R}^N$-valued functions. This means that for every nonnegative integer $\bar{k}$, every compact subset $K \subset B_\infty$, and every positive real number $\varepsilon > 0$, there exists $j_0 := j_0(\bar{k}, K, \varepsilon)$ such that $K \subset U_{j_0}$ and, for $j \geq j_0$,

$$\sup_K \sum_{k=0}^{\bar{k}} \left| \nabla^k_{g_\infty} \left( (F_j \circ \phi_j) - F_\infty \right) \right|^2_{g_\infty} < \varepsilon.$$

In particular, the pullback metrics $\phi_j^* g_j = (F_j \circ \phi_j)^* g_{\text{flat}}$ converge smoothly on compact subsets of $B_\infty$ to $g_\infty = F_\infty^* g_{\text{flat}}$. If $\rho^* = \infty$, then this means the sequence $(M_j, g_j, p_j)$ converges in the traditional pointed Cheeger-Gromov sense to a complete Riemannian manifold $(M_\infty, g_\infty, p_\infty)$.

The following proposition is the analogue of Theorem 4.1.2 in [18].

**Proposition 3.2** (Local compactness of pointed immersions)**.** *Let $(M_j, g_j, p_j)$ be a sequence of pointed, complete $n$-dimensional Riemannian manifolds and suppose for each $j$ that $F_j : (M_j, g_j, p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ is a pointed isometric immersion. Consider a sequence of radii $\rho_j \in (0, \infty]$ such that $\rho_j \to \rho^* \in (0, \infty]$ and let $B_j := B_{g_j}(p_j, \rho_j)$. Suppose that for every radius $0 < \rho < \rho^*$ and every integer $k \geq 0$, there exists a constant $\Lambda_k(\rho)$, independent of $j$, and a positive integer $j_0(k, \rho)$ such that for every $j \geq j_0(k, \rho)$ the $k^{th}$ covariant derivative of the second fundamental form $A_j$ of the immersion $F_j$ satisfies the pointwise estimate*

$$\sup_{B_{g_j}(p_j, \rho)} |\nabla^k A_j|_{g_j} \leq \Lambda_k(\rho).$$

*Then there exists a subsequence of the immersions $F_j|_{B_j}$ which converges in $C^\infty_{\text{loc}}$ to a pointed isometric immersion $F_\infty : (B_\infty, g_\infty, p_\infty) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ of an open geodesic ball $B_\infty = B_{g_\infty}(p_\infty, \rho^*)$. If $\rho^* = \infty$, then the limiting Riemannian manifold is complete.*

Before we discuss the proof of the proposition, let us show how to deduce a local convergence result for the mean curvature flow as a corollary. We will take the following definition for local convergence of flows.

**Definition 3.3** (cf. Definition 4.1.3 in [18])**.** Fix $\tau < 0$. Let $(M_j, g_j(t), p_j)$ be a sequence of evolving, pointed, complete $n$-dimensional Riemannian manifolds for $t \in (\tau, 0]$. Suppose $F_j(\cdot, t) : (M_j, g_j(t), p_j) \to \mathbb{R}^N$, for $t \in (\tau, 0]$, is a sequence of smoothly evolving immersions satisfying $F_j(p_j, 0) = 0$. Consider a sequence of radii $\rho_j \in (0, \infty]$ such that $\rho_j \to \rho^* \in (0, \infty]$ and let $P_j := B_{g_j(0)}(p_j, \rho_j) \times (-\tau, 0]$. We say the sequence of pointed evolving immersions $F_j|_{P_j}$ converges in $C^\infty_{\text{loc}}$ to a pointed evolving immersion $F_\infty(\cdot, t) : (B_\infty, g_\infty(t), p_\infty) \to \mathbb{R}^N$ for $t \in (\tau, 0]$ of an evolving, pointed, $n$-dimensional Riemannian manifold $B_\infty = B_{g_\infty(0)}(p_\infty, \rho^*)$, if:

- We can find a sequence of (time-independent) smooth, relatively compact open sets $U_j$ in $B_\infty$ satisfying $p_\infty \in U_j$, $U_j \subset U_{j+1}$, and $\cup_j U_j = B_\infty$.

- We can find a sequence of (time-independent) diffeomorphisms $\phi_j : U_j \to \phi_j(U_j) \subset B_{g_j}(p_j, \rho_j)$ satisfying $\phi_j(p_\infty) = p_j$. Moreover, for all $\rho \in (0, \rho^*)$, if $j$ is sufficiently large, then $B_{g_j}(p_j, \rho) \subset \phi_j(U_j)$.

- The sequence of evolving immersions $F_j(\cdot, t) \circ \phi_j$ converges to $F_\infty(\cdot, t)$ smoothly with respect to $g_\infty(0)$ on every compact subset of $B_\infty \times (\tau, 0]$.

To be specific, in the definition above, the sequence $F_j(\cdot, t) \circ \phi_j : U_j \to \mathbb{R}^N$ converges to $F_\infty : B_\infty \times (\tau, 0] \to \mathbb{R}^N$ smoothly with respect to $g_\infty(0)$ on compact subsets if for every nonnegative integer $k$, compact subset $K \subset B_\infty$, compact subset $[\tilde{\tau}, 0] \subset (\tau, 0]$, and positive real number $\varepsilon > 0$, there exists $j_0 := j_0(k, K, \tilde{\tau}, \varepsilon)$ such that $K \subset U_{j_0}$ and, for $j \geq j_0$,

$$\sup_{K \times [\tilde{\tau}, 0]} \sum_{m=0}^{k} \left| \nabla^m_{g_\infty(0)} \big( (F_j(\cdot, t) \circ \phi_j) - F_\infty(\cdot, t) \big) \right|^2_{g_\infty(0)} < \varepsilon.$$

Here is the compactness result for local solutions of the mean curvature flow that we are after. This is the analogue of Theorem 4.1.5 in [18].

**Corollary 3.4** (Local compactness of mean curvature flow)**.** *Fix $\tau < 0$. Let $(M_j, g_j(t), p_j)$ be a sequence of evolving, pointed, complete $n$-dimensional Riemannian manifolds for $t \in (\tau, 0]$. Suppose $F_j(\cdot, t) : (M_j, g_j(t), p_j) \to \mathbb{R}^N$, for $t \in (\tau, 0]$, is a sequence of smoothly evolving immersions satisfying $F_j(p_j, 0) = 0$. Consider a sequence of radii $\rho_j \in (0, \infty]$ such that $\rho_j \to \rho^* \in (0, \infty]$. Suppose $F_j$ is a solution to the mean curvature flow on the parabolic neighborhood $P_j := B_{g_j(0)}(p_j, \rho_j) \times (\tau, 0]$. Finally, suppose that for every radius $0 < \rho < \rho^*$ there exists a constant $\Lambda(\rho)$, independent of $j$, and a positive integer $j_0(\rho)$ such that for every $j \geq j_0(\rho)$, the second fundamental form $A_j$ of the evolving immersion $F_j(\cdot, t)$ satisfies the pointwise estimate*

$$\sup_{B_{g_j(0)}(p_j, \rho) \times (\tau, 0]} |A_j|_{g_j} \leq \Lambda(\rho).$$

*Then there exists a subsequence of solutions $F_j$ such that $F_j|_{P_j}$ converge in $C^\infty_{\rm loc}$ to a solution of the mean curvature flow $F_\infty : B_\infty \times (\tau, 0] \to \mathbb{R}^N$ with $g_\infty(t) = F_\infty(\cdot, t)^* g_{\rm flat}$ and $B_\infty = B_{g_\infty(0)}(p_\infty, \rho^*)$. If $\rho^* = \infty$, the limiting solution is complete at time $t = 0$.*

Note that if $\rho^* = \infty$ and if the bounds for the second fundamental form in the proposition above can be taken to be independent of $\rho$, then the solution will be complete on every time-slice.

It is straightforward to see that the assumptions of the corollary together with Proposition 3.2 allow us to extract a limiting immersion at the time $t = 0$. In order to extend the convergence backwards in time, we will use the follow lemma, which is an adaptation of Lemma 4.1.4 in [18] to our setting.

**Lemma 3.5** (cf. Lemma 4.1.4 in [18])**.** *Let $(B, g, p)$ be a pointed Riemannian manifold, $K \subset B$ be a compact subset, and $\tilde{F}_j(\cdot, t)$ be a sequence of pointed (i.e, $\tilde{F}_j(p, 0) = 0$) solutions of the mean curvature flow defined on $K \times [\tilde{\tau}, 0]$. Let $\tilde{g}_j(t) = \tilde{F}_j(\cdot, t)^* g_{\rm flat}$. Let $\nabla$ denote covariant derivative of $g$ and $\tilde{\nabla} = \tilde{\nabla}_{\tilde{g}_j}$ denote the covariant derivative of $\tilde{g}_j(t)$ for each $j$. Suppose there exist constants $C_k$ (independent of $j$) for each integer $k \geq 0$ such that*

*(a) $C_0^{-1} g \leq \tilde{g}_j(0) \leq C_0 g$ on $K$ for all $j$;*

*(b) for each $k \geq 0$, $|\nabla^k \tilde{F}_j(\cdot, 0)|_g \leq C_k$ on $K$ for all $j$;*

*(c) $|\nabla^k_{\tilde{g}_j} \tilde{A}_j|_{\tilde{g}_j} \leq C_k$ on $K \times [\tilde{\tau}, 0]$ for all $j$.*

*Then there exists constants $\tilde{C}_k$ (independent of $j$) for each integer $k \geq 0$ such that*

$$\tilde{C}_0^{-1} g \leq \tilde{g}_j(t) \leq \tilde{C}_0 g; \qquad |\nabla^k \tilde{F}_j|_g \leq \tilde{C}_k \ \ (k \geq 0)$$

*on $K \times [\tilde{\tau}, 0]$ for all $j$.*

*Proof.* The proof of this lemma differs very little from the proof of Lemma 4.1.4 in [18]. This lemma also follows from the results in Appendix A of Brendle's book [7]. So we will just highlight the key points. Let $p \in K$, $v \in T_pB$, and define $\mu(t) = \tilde{g}_j(t)(v,v)$. Since

$$\Big|\frac{\partial}{\partial t}\tilde{g}_j(v,v)\Big| = 2\Big|\langle \tilde{A}_j(v,v), \vec{\tilde{H}}_j\rangle\Big| \leq C(n)C_0^2\tilde{g}_j(v,v),$$

(recall $\langle \cdot\,,\cdot\rangle = g_{\text{flat}}$) we have $|\mu'(t)| \leq C\mu(t)$. From this and (a), the uniform equivalence of the metrics $\tilde{C}_0^{-1}g \leq \tilde{g}_j(t) \leq \tilde{C}_0 g$ readily follows. Let $\tilde{\Gamma}_j$ and $\Gamma$ denote the connection coefficients of $\tilde{g}_j$ and $g$ respectively. The expression $\frac{\partial}{\partial t}(\tilde{\Gamma}_j - \Gamma) = \frac{\partial}{\partial t}\tilde{\Gamma}_j$ is tensorial and, since $\frac{\partial}{\partial t}\tilde{\Gamma}_j = \tilde{g}_j^{-1} * \tilde{\nabla}_{\tilde{g}_j}(\frac{\partial}{\partial t}\tilde{g}_j)$, assumption (c) implies an estimate $\big|\frac{\partial}{\partial t}(\tilde{\Gamma}_j - \Gamma)\big|_g \leq C$ on $K \times [\tilde{\tau}, 0]$. On the other hand, assumption (b) implies (for each $k$) that $|\nabla^k \tilde{g}_j(0)| \leq C$ on $K$, which implies $|\tilde{\Gamma}_j(0) - \Gamma|_g \leq C$. By integration one concludes $|\tilde{\Gamma}_j - \Gamma|_g \leq C$ on $K \times [\tilde{\tau}, 0]$. Now that we have estimates for the difference of the connection coefficients, we estimate

$$\Big|\frac{\partial}{\partial t}\nabla^k \tilde{g}_j\Big|_g = \Big|\nabla^k \frac{\partial}{\partial t}\tilde{g}_j\Big|_g \leq \Big|\tilde{\nabla}^k \frac{\partial}{\partial t}\tilde{g}_j\Big|_g + \Big|(\nabla^k - \tilde{\nabla}^k)\frac{\partial}{\partial t}\tilde{g}_j\Big|_g.$$

In light of the evolution equation for $\tilde{g}_j$ and assumption (c), we can bound $|\tilde{\nabla}^k \frac{\partial}{\partial t}\tilde{g}_j|_g \leq C$. By induction, we can bound the second term $|(\nabla^k - \tilde{\nabla}^k)\frac{\partial}{\partial t}\tilde{g}_j|_g$ by $C + C|\nabla^k \tilde{g}_j|_g$. Hence, we have $\frac{\partial}{\partial t}|\nabla^k \tilde{g}_j|_g \leq C + C|\nabla^k \tilde{g}_j|_g$ By integration and assumption (b), we obtain the estimate $|\nabla^k \tilde{g}_j|_g \leq C$ on $K \times [\tilde{\tau}, 0]$ for each $k$. See Lemma A.3 in [7] and the proof in [18] for further details. In a similar fashion, we have

$$\Big|\frac{\partial}{\partial t}(\nabla^k \tilde{F}_j)\Big|_g = \Big|\nabla^k \frac{\partial}{\partial t}\tilde{F}_j\Big|_g = \Big|\nabla^k \vec{\tilde{H}}_j\Big|_g \leq \Big|\tilde{\nabla}^k \vec{\tilde{H}}_j\Big|_g + \Big|(\nabla^k - \tilde{\nabla}^k)\vec{\tilde{H}}_j\Big|_g.$$

By assumption (c) the first term is bounded by a constant. Our estimates for the metric and its derivatives together with assumption (c) give control of the second term. For example, because

$\tilde{\nabla}\tilde{\vec{H}}_j = \nabla\tilde{\vec{H}}_j$, we have

$$
\begin{aligned}
(\nabla^2 - \tilde{\nabla}^2)\tilde{\vec{H}}_j &= (\nabla - \tilde{\nabla})(\nabla\tilde{\vec{H}}_j) + \tilde{\nabla}(\nabla - \tilde{\nabla})\tilde{\vec{H}}_j \\
&= (\nabla - \tilde{\nabla})(\tilde{\nabla}\tilde{\vec{H}}_j) \\
&= (\Gamma - \tilde{\Gamma}_j) * \tilde{\nabla}\tilde{\vec{H}}_j
\end{aligned}
$$

Note that $\Gamma - \tilde{\Gamma}_j = \tilde{g}_j^{-1} * \nabla\tilde{g}_j$. For the general case, see Lemma A.4 in [7]. By integration and assumption (b), we obtain the desired estimates. □

Now we can give a proof of Corollary 3.4.

*Proof of Corollary 3.4.* We have uniform estimates for $A_j$ on the parabolic neighborhood $B_{g_j(0)}(p_j, \rho) \times (\tau, 0]$ for each $0 < \rho < \rho^*$ (assuming $j \geq j_0(\rho)$). Therefore, by standard interior estimates for the mean curvature flow, for each integer $k \geq 0$ there exists a constant $C_k(\rho)$ such that

$$\sup_{B_{g_j(0)}(p_j,\rho)} |\nabla^k_{g_j(0)} A_j(0)|_{g_j(0)} \leq C_k(\rho)$$

for each $0 < \rho < \rho^*$ (assuming $j \geq j_0(\rho)$). In particular, the sequence $F_j(\cdot, 0) : (M_j, g_j(0), p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ satisfies the hypotheses of Proposition 3.2. Thus after passing to a subsequence, which we still denote by $F_j$, we can find a pointed Riemannian manifold $(B_\infty, g_\infty, p_\infty)$, a sequence of domains $U_j$ exhausting $B_\infty$, and injective smooth maps $\phi_j : U_j \to B_{g_j(0)}(p_j, \rho_j)$ such that $\tilde{F}_j := F_j \circ \phi_j$ converges smoothly on compact subsets of $B_\infty$ with respect to $g_\infty$ to a pointed isometric immersion $F_\infty : (B_\infty, g_\infty, p_\infty) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$. Here $B_\infty$ is the open geodesic ball $B_{g_\infty}(p_\infty, \rho^*)$.

Now we can apply Lemma 3.5. Let $K \subset B_\infty$ be any compact subset and $[\tilde{\tau}, 0] \subset (\tau, 0]$. Let $\tilde{g}_j(t) := \tilde{F}_j(\cdot, t)^* g_{\text{flat}}$. To simplify notation, let $F := F_\infty$ and $g := g_\infty$. Let $\tilde{\nabla}_{\tilde{g}_j}$ and $\nabla$ denote the covariant derivatives of $\tilde{g}_j$ and $g$ respectively. By the definition of convergence, after passing to a suitable diagonal subsequence, for every integer $k \geq 0$, there exists a constant $C_k$ such that

$$|\nabla^k \tilde{F}_j(\cdot,0)|_g \leq \big|\nabla^k\big(\tilde{F}_j(\cdot,0) - F\big)\big|_g + |\nabla^k F\big|_g \leq C_k$$

on $K$ for all $j$. Clearly, we also have $C_0^{-1}g \leq \tilde{g}_j(0) \leq C_0 g$ on $K$ for all $j$. Finally, by diffeomorphism invariance and interior estimates we obtain $|\tilde{\nabla}^k_{\tilde{g}_j}\tilde{A}_j|_{\tilde{g}_j} = |\nabla^k_{g_j}A_j|_{g_j} \leq C_k$ on $K \times [\tilde{\tau}, 0]$ for all $j$. Thus, assumptions (a) - (c) of Lemma 3.5 are satisfied and consequently, we have uniform estimates $\tilde{C}_0^{-1}g \leq \tilde{g}_j(t) \leq \tilde{C}_0 g$ and $|\nabla^k \tilde{F}_j|_g \leq \tilde{C}_k$ for $k \geq 0$ with respect to the fixed background metric $g$. We can now use Arzela-Ascoli with a standard diagonalization argument to extract a subsequence which converges uniformly on compact subsets of $B_\infty \times (\tau, 0]$. In the limit, we obtain a family of smooth maps $F_\infty(\cdot,t) : B_\infty \to \mathbb{R}^N$ for $t \in (\tau, 0]$. Evidently, $F_\infty(\cdot, 0) = F = F_\infty$ since we already have convergence at time $t = 0$. It remains to verify the family $F_\infty(\cdot,t)$ is a solution of the mean curvature flow. The (0,2)-tensor $g_\infty(t) = F_\infty(\cdot,t)^* g_{\text{flat}}$ is the limit of the nondegenerate metrics $\tilde{g}_j(t)$ each of which are uniformly equivalent to the Riemannian metric $g = g_\infty(0)$. In particular, $g_\infty(t)$ is itself a Riemannian metric and $F_\infty(\cdot,t)$ is a family of immersions. Finally, as a limit of solutions of the mean curvature flow, it is clear that $F_\infty(\cdot,t)$ satisfies the same equation. □

In the remainder of this appendix, we will give a proof of Proposition 3.2. Our approach will be first to extract an intrinsic limit along a suitable subsequence, using the work of Hamilton and Cao-Zhu's localization of it. To do so, we will need an injectivity radius estimate, which we obtain from the following lemma.

**Lemma 3.6.** *Let $F : (M, g, p) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ be a pointed isometric immersion of a smooth connected complete Riemannian manifold. Suppose that for some $\rho \in (0,\infty)$,*

$$\sup_{B_g(p,\rho)} |A| \leq \Lambda.$$

*There exists a positive constant $\delta := \delta(n,\Lambda,\rho) > 0$ such that $\text{inj}(M,p) \geq \delta$.*

*Proof.* After composing $F$ with an isometry of $\mathbb{R}^N$, we may assume that $F(0) = 0$ and $dF_p(T_pM) =$

$\mathbb{R}^n \times \{0\} \subset \mathbb{R}^N$. Let $\pi : \mathbb{R}^N \to \mathbb{R}^n$ denote the projection onto the first $n$-coordinates of $\mathbb{R}^N$. Let

$$r_0 := \frac{1}{500}\min\{\Lambda^{-1}, \rho\}.$$

We will first show that the immersion can be expressed as a graph over a ball in $dF_p(T_pM)$ of radius proportional to $r_0$. Then, in the graphical parametrization we can estimate the intrinsic volume, and from this the local injectivity radius estimate will follow.

*Step 1:* First, we show that $\pi \circ F$ is injective on $B_g(p, 15r_0)$. If not, then there exists distinct points $q_0, q_1 \in B_g(p, 15r_0)$ such that $\pi \circ F(q_0) = \pi \circ F(q_1)$. Let $\gamma : [0,1] \to M$ be a minimizing geodesic with $\gamma(0) = q_0$ and $\gamma(1) = q_1$. Now $d_g(q_0, q_1) < 30r_0$, which implies

$$d_g(p, \gamma(t)) \le d_g(p, q_0) + d_g(q_0, \gamma(t)) < 45r_0 < \rho.$$

So $\gamma([0,1]) \subset B_g(p, \rho)$ and therefore $|A|(\gamma(t)) \le \Lambda$ for each $t \in [0,1]$. Let $\tilde{\gamma}(t) = F \circ \gamma(t)$ and $\hat{\gamma}(t) = \pi \circ \tilde{\gamma}(t)$. Because $\gamma(t)$ is a geodesic, we have $\tilde{\gamma}'' = A(\gamma', \gamma')$ and $|\tilde{\gamma}'| = |\gamma'| = d_g(q_0, q_1)$. We claim that $|\hat{\gamma}'(t)|^2 \ge \frac{1}{10} d_g(q_0, q_1)^2$ for all $t \in [0,1]$. To see this fix some $t \in [0,1]$ and let $X(s)$ be the parallel transport of $\gamma'(t)$ along a minimal geodesic $\sigma(s)$ connecting $\sigma(0) = p$ to $\sigma(1) = \gamma(t)$. Note $|\sigma'| \le 45r_0$. Let $\tilde{X}(s) = dF_{\sigma(s)}(X(s))$ and note that $|\pi(\tilde{X}(0))| = |\tilde{X}(0)| = |X(0)| = d_g(q_0, q_1)$. Now

$$\left|\frac{d}{ds}\frac{1}{2}|\pi(\tilde{X})|^2\right| = \left|\langle \pi(A(\sigma', X)), \pi(\tilde{X})\rangle\right| \le |A||\sigma'||X|^2 \le 45\Lambda r_0 d_g(q_0, q_1)^2 \le \frac{9}{20} d_g(q_0, q_1)^2.$$

Therefore,

$$\frac{1}{2}|\hat{\gamma}'(t)|^2 = \frac{1}{2}|\pi(\tilde{X}(1))|^2 \ge \frac{1}{2} d_g(q_0, q_1)^2 - \frac{9}{20} d_g(q_0, q_1)^2 \ge \frac{1}{20} d_g(q_0, q_1)^2,$$

which implies $|\hat{\gamma}'(t)|^2 \ge \frac{1}{10} d_g(q_0, q_1)^2$. Now we consider the function $f(t) = \frac{1}{2}|\hat{\gamma}(t) - \hat{\gamma}(0)|^2$.

Then $f' = \langle \hat{\gamma}', \hat{\gamma} - \hat{\gamma}(0) \rangle$ and $f'' = \langle \hat{\gamma}'', \hat{\gamma} - \hat{\gamma}(0) \rangle + |\hat{\gamma}'|^2$. Also, $f(0) = f(1) = 0$, $f'(0) = 0$ and $f''(0) = |\hat{\gamma}'(0)|^2 \geq \frac{1}{10} d_g(q_0, q_1) > 0$. Consequently, $f$ attains its maximum at some point $t_0 \in (0, 1)$. Since $|\hat{\gamma}''| = |\pi(A(\gamma', \gamma'))| \leq \Lambda d_g(q_0, q_1)^2$ and $|\hat{\gamma} - \hat{\gamma}(0)| \leq 30 r_0$, at the point $t_0$, we obtain the inequality

$$0 \geq f''(t_0) \geq -30 r_0 \Lambda d_g(q_0, q_1)^2 + \frac{1}{10} d_g(q_0, q_1)^2.$$

But this implies that $r_0 \geq \frac{1}{300} \Lambda^{-1}$, in contradiction with its definition.

*Step 2:* Next, we show that $c(n) r_0^n \leq \mathrm{Vol}(B_g(p, 15 r_0)) \leq C(n) r_0^n$. Let $\Omega := \pi(F(B_g(p, 15 r_0))) \subset \mathbb{R}^n$. Let $|| \cdot ||$ denote the Euclidean norm on $\mathbb{R}^n$ and $D$ its standard derivative. Now since $\pi$ is injective on $F(B_g(p, 15 r_0))$, we can find a smooth function $f : \Omega \to \mathbb{R}^{N-n}$ such that $f(0) = 0$ , $Df(0) = 0$, and $\mathrm{graph}(f) = F(B_g(p, 15 r_0))$. For graphical parametrizations, it is straightforward to show (see Lemma 2.2 in [17]) the inequality

$$||D^2 f|| \leq (1 + ||Df||^2)^{\frac{3}{2}} (|A|_g \circ F^{-1})$$

holds on $\Omega$. Let $\tilde{r} > 0$ be the maximal radius such that $B_{\tilde{r}}^n \subset \Omega$, where $B_{\tilde{r}}^n$ denote the Euclidean $n$-ball of radius $r$ centered at the origin. Clearly, $\tilde{r} \leq 15 r_0$. For $x \in B_{\tilde{r}}^n \setminus \{0\}$, write $x = r\omega$ where $\omega$ is a unit vector and $r \in (0, \tilde{r})$. Consider the function $\mu(t) = ||Df||^2(t\omega)$ for $t \in [0, r]$. Noting that $\mathrm{graph}(f) \subset F(B_g(p, \rho))$, the inequality above gives

$$\mu'(t) \leq 2||D^2 f||(t\omega)\, \mu(t)^{\frac{1}{2}} \leq 2(1 + \mu(t))^{\frac{3}{2}} \mu(t)^{\frac{1}{2}} \Lambda.$$

We can rewrite this as

$$\frac{d}{dt} \Big( \frac{\mu(t)}{1 + \mu(t)} \Big)^{\frac{1}{2}} \leq \Lambda$$

Since $\mu(0) = 0$, integrating from $0$ to $r$ gives

$$\Big(\frac{||Df||^2(x)}{1 + ||Df||^2(x)}\Big)^{\frac{1}{2}} \leq r\Lambda \leq 15\, r_0 \Lambda < \frac{1}{20}.$$

Therefore, we conclude that $||Df|| \leq \frac{1}{100}$ on $B^n_{\tilde{r}}$. Finally, using this slope bound, we can derive the volume estimate. Consider a direction $\omega$ such that $\tilde{r}\omega \in \partial\Omega$. Using our estimate for the slope of $f$, the path $\tilde{\sigma}(t) = (t\omega, f(t\omega))$ clearly has length bounded by $\frac{3}{2}\tilde{r}$. Since the path $\sigma(t) := F^{-1} \circ \tilde{\sigma}(t)$ is a path in $M$ from $p$ to the boundary of $B_g(p, 15r_0)$ and $F$ preserves lengths, we conclude $\tilde{r} \geq 10r_0$. Since $B^n_{\tilde{r}} \subset \Omega$, using the graphical parametrization, we obtain $\mathrm{Vol}(B_g(p, 15r_0)) \geq c(n)r_0^n$. On the other hand, the slope bound for $f$ gives the reverse inequality $\mathrm{Vol}(B_g(p, 15r_0)) \leq C(n)r_0^n$.

*Conclusion:* Using the Gauss equation, we can bounded the absolute value of the sectional curvature in $B_g(p, \rho)$ by $2\Lambda$. Now it follows from Theorem 4.2.2. in [18] (which is a local injectivity radius estimate due to Cheeger-Gromov-Taylor in [19]) together with the volume estimates from Step 2, that $\mathrm{inj}(M, p) \geq c(n, \Lambda, \rho)r_0$. □

We can now prove Proposition 3.2.

*Proof of Proposition 3.2.* We will complete the proof in two steps.

*Step 1:* We will first take an intrinsic limit in the sense of Definition 4.1.1 in [18] by applying Theorem 4.1.2. in [18]. To that end, we consider the sequence of geodesic balls $B_j = B_{g_j}(p_j, \rho_j) \subset M_j$ and verify two conditions.

(a) Consider a radius $\rho < \rho^*$ and an integer $k \geq 0$. Via the Gauss equation, estimates for covariant derivatives of the second fundamental form yield corresponding estimates for the Riemannian curvature tensor $\mathrm{Rm}(g_j)$ of the metric $g_j$. In particular, $|\nabla^k \mathrm{Rm}(g_j)|$ can be bounded pointwise by an expression in $|\nabla^l A_j|$ for $0 \leq l \leq k$. Thus we can find a constant

$\tilde{\Lambda}_k(\rho)$, independent of $j$ and a positive integer $j_1(k,\rho)$ such that $j \geq j_1(k,\rho)$

$$\sup_{B_{g_j}(p_j,\rho)} |\nabla^k \mathrm{Rm}(g_j)| \leq \tilde{\Lambda}_k(\rho).$$

(b) Let $\tilde{\rho} = \min\{\frac{1}{2}\rho^*, 1\}$. After passing to a subsequence, we have $|A_j| \leq \Lambda_0(\tilde{\rho})$ on the geodesic ball $B_{g_j}(p_j, \tilde{\rho})$ for all $j$. Thus, by Lemma 3.6, there exists a positive constant $\delta := \delta(n, \Lambda_0(\tilde{\rho}), \tilde{\rho})$, independent of $j$, such that the injectivity radius of $M_j$ at $p_j$ in the metric $g_j$ satisfies

$$\mathrm{inj}(M_j, p_j) \geq \delta.$$

Having verified conditions (a) and (b), we may now apply Theorem 4.1.2. in [18] to obtain the following conclusion: there exists a subsequence of pointed geodesic balls $(B_j, g_j, p_j)$ which converge to a pointed geodesic ball $(B_\infty, g_\infty, p_\infty)$ centered at a point $p_\infty$ of radius $\rho^*$ (that is, $B_\infty = B_{g_\infty}(p_\infty, \rho^*)$) in the intrinsic $C^\infty_{\mathrm{loc}}$ topology. This means, we can find a sequence of exhausting open sets $U_j$ in $B_\infty$, each containing $p_\infty$, and a sequence of diffeomorphisms $\phi_j : U_j \to \phi_j(U_j) \subset B_j \subset M_j$ such that $\phi_j(p_\infty) = p_j$ and the metrics $\tilde{g}_j := \phi_j^* g_j$ converge to $g_\infty$ in the smooth topology on every compact subset of $B_\infty$. Moreover, the proof of Theorem 4.1.2. in [18] implies $U_j \subset U_{j+1}$ and for every $\rho \in (0, \rho^*)$, if $j$ is sufficiently large then $B_{g_j}(p_j, \rho) \subset \phi_j(U_j)$.

*Step 2:* Now that we have an intrinsic limit, it is straightforward to show that a subsequence of the immersions $\tilde{F}_j := F_j \circ \phi_j : U_j \to \mathbb{R}^N$ converges to a limit $F_\infty : B_\infty \to \mathbb{R}^N$, smoothly with respect to $g_\infty$ on compact subsets of $B_\infty$. Consider any compact subset $K \subset B_\infty$. If $j$ is sufficiently large, then $K \subset U_j$ and $\tilde{F}_j$ is defined on $K$. Moreover, there exists $\rho := \rho(K) \in (0, \rho^*)$ such that $\phi_j(K) \subset B_{g_j}(p_j, \rho)$ if $j$ is large enough. By diffeomorphism invariance, $|\nabla^k_{\tilde{g}_j} \tilde{F}_j|_{\tilde{g}_j} = |\nabla^k_{g_j} F_j|_{g_j}$. Recall from the discussion at the beginning of the appendix that $\nabla^k_{g_j} F_j = \nabla^{k-2}_{g_j} A_j$. Therefore, given $K$ and $k \geq 2$, our uniform bounds for the second fundamental form and its covariant derivatives imply uniform bounds for $|\nabla^k_{\tilde{g}_j} \tilde{F}_j|_{\tilde{g}_j}$ on $K$ once $j$ is sufficiently large. On the other hand, for $k = 1$, we have $|\nabla_{\tilde{g}_j} \tilde{F}_j|_{\tilde{g}_j} = n$ since $\tilde{g}_j = \phi^* g_j = \tilde{F}_j^* g_{\mathrm{flat}}$. Thus, since the metrics $\tilde{g}_j$

converge to $g_\infty$ on $K$, given a compact $K \subset B_\infty$ and an integer $k \geq 1$, we obtain uniform estimates for $|\nabla^k_{g_\infty} \tilde{F}_j|_{g_\infty}$ if $j$ is sufficiently large. By assumption $\tilde{F}_j(p_\infty) = F_j(p_j) = 0$ and so with the first derivative estimate, we conclude $|\tilde{F}_j| \leq C$ on $K$ as well for $j$ large enough. By the classical Arzela-Ascoli and diagonalization, we can find a subsequence of the $\tilde{F}_j$ which converge smoothly with respect to $g_\infty$ on every compact subset of $B_\infty$ to smooth limit $F_\infty : B_\infty \to \mathbb{R}^N$. Since $g_\infty = \lim_{j\to\infty} \tilde{g}_j = \lim_{j\to\infty} \tilde{F}_j^* g_{\text{flat}} = F_\infty^* g_{\text{flat}}$, the limiting $\mathbb{R}^N$-valued function is an immersion, completing the proof. $\square$

# Chapter 4

# Singularity Formation and Canonical Neighborhoods

Having discussed the necessary compactness theory in the previous chapter, we can now return out attention to singularity formation of the mean curvature flow in higher codimension for initial data that satisfies a curvature pinching condition. The work in this chapter is taken from [40]. Recall our running assumption from Chapter 2: we assume that $n \geq 5$, $N > n$, and define $\tilde{c}_2 = \min\{\frac{1}{n-2}, \frac{3(n+1)}{2n(n+2)}\}$. We consider compact solutions of the mean curvature flow, $F : M \times [0, T) \to \mathbb{R}^N$, for which the initial immersion satisfies $|A|^2 < \tilde{c}_2|H|^2$. In the previous chapters, we have shown that the only blow-up models at the first singular time of a $\tilde{c}_2$-pinched solution of the mean curvature flow are the codimension one shrinking round spheres, shrinking round cylinders, and translating bowl solitons. This suitably extends the classification result of Brendle and Choi [13].

The purpose of this chapter to upgrade our description of the infinitesimal scale at spacetime points of infinite curvature to a description of small scales near spacetime points of high curvature. Since Perelman's work [43] on the Ricci flow in three dimensions, such results are nowadays known as canonical neighborhood theorems. In the mean curvature flow, Huisken and Sinestrari essentially proved a canonical neighborhood theorem through their Neck Detection Lemma and Neck Continuation Theorem (Lemma 7.4 and Theorem 8.2 in [33]). The first of these two results has been extended to higher codimension by Nguyen in [41]. It is an interesting problem to prove a version of the Neck Continuation Theorem in higher codimension. Huisken and Sinestrari's proof of this theorem makes significant use of convexity, which is absent in higher codimension, so some new ideas will be needed. At the same time as this work was completed, Nguyen has published a preprint where he has addressed this problem [42] (as well as developed a surgery procedure). We have a different approach here and prove the following theorem, which is much closer in spirit to Perelman's result.

**Theorem 4.1.** *Suppose $F_0 : M \to \mathbb{R}^N$ is an immersion of a closed manifold of dimension $n \geq 5$ satisfying $|A|^2 < \tilde{c}_2|H|^2$, where $\tilde{c}_2 := \min\{\frac{1}{n-2}, \frac{3(n+1)}{2n(n+2)}\}$. There exist constants $\tilde{\varepsilon}$ and $\tilde{K}$, depending only upon the dimension and $F_0$, with the following property. Let $F : M \times [0,T) \to \mathbb{R}^N$ denote the solution of mean curvature flow with initial immersion given by $F_0$. Given $\varepsilon_0 \in (0, \tilde{\varepsilon})$ and $K_0 \in (\tilde{K}, \infty)$, there exists a positive number $\hat{r} := \hat{r}(n, F_0, \varepsilon_0, K_0) > 0$ with the following property. If $(p_0, t_0)$ is a spacetime point such that $Q_0 := |H|(p_0, t_0) \geq \hat{r}^{-1}$, then the solution is $\varepsilon_0$-close in the intrinsic parabolic neighborhood $B_{g(t_0)}(p_0, Q_0^{-1}K_0) \times [t_0 - K_0^2 Q_0^{-2}, t_0]$ to an ancient model solution in the sense of Definition 4.3.*

In fact, one can show that the constants $\tilde{\varepsilon}$ and $\tilde{K}$ depend only upon the dimension. Here, an ancient model solution is an $n$-dimensional ancient, nonflat, complete, codimension one (i.e. lying in an $(n + 1)$-dimensional plane) solution of mean curvature flow in $\mathbb{R}^N$ that is uniformly two-convex and noncollapsed. The notion "$\varepsilon_0$-close" roughly means that our flow is close to a model flow after a suitable reparametrization. A detailed definition is given in the next section.

To prove this theorem, we will follow the original strategy of Perelman [43]. In particular, we will adapt a variation of Perelman's proof given by Brendle in [10] to our setting. In Section 4.1, we establish definitions and results we will use in the proof of the main theorem. In Section 4.2, we give the proof of the canonical neighborhood theorem.

## 4.1. Preliminaries

Henceforth, we will let $H$ denote the scalar mean curvature and $\vec{H}$ denote mean curvature vector. In higher codimension, this means $H = |\vec{H}|$. Since we assume $H > 0$, we may define a (0,2)-tensor $h = \langle A, H^{-1}\vec{H}\rangle$, where $A$ is the full vector-valued version of second fundamental form. In codimension one, $h$ is just the usual scalar-valued version of the second fundamental form. We will adopt the notations $P(p,t,r,\theta) := B_{g(t)}(p,r) \times [t-\theta, t]$ and

$$\hat{P}(p,t,r,\theta) := P(p,t,H(p,t)^{-1}r, H(p,t)^{-2}\theta),$$

as in [33], to denote intrinsic parabolic neighborhoods.

We will use the following definitions for $\varepsilon$-necks and $\varepsilon$-caps.

**Definition 4.2.** Let $\varepsilon > 0$ be a small positive constant and $F : M \to \mathbb{R}^N$ an isometric immersion of a complete Riemannian manifold. Let $\bar{g}$ denote the standard metric on the round cylinder $S^{n-1} \times \mathbb{R}$ of radius 1 (or, equivalently, of constant scalar curvature $(n-1)(n-2)$).

- An $\varepsilon$-neck is a compact region $N \subset M$ for which there exists a diffeomorphism $\phi : S^{n-1} \times [-\varepsilon^{-1}, \varepsilon^{-1}] \to N$, a positive constant $r > 0$, and an isometric embedding $\bar{F} : S^{n-1} \times [-\varepsilon^{-1}, \varepsilon^{-1}] \to \mathbb{R}^N$ (with respect to $\bar{g}$) such that the immersion $r^{-1}(F \circ \phi)$ is $\varepsilon$-close in $C^{[1/\varepsilon]}$ on $S^{n-1} \times [-\varepsilon^{-1}, \varepsilon^{-1}]$ to the embedding $\bar{F}$ with respect to the metric $\bar{g}$. The constant $r$ is called the radius of the neck $N$. For any $z \in [-\varepsilon^{-1}, \varepsilon^{-1}]$, we call $\phi(S^{n-1} \times \{z\}) \subset N$ a cross-sectional sphere of the neck.

- We say a point $p_0 \in M$ lies at the center of an $\varepsilon$-neck if $p_0$ lies on the central cross-sectional sphere, $\phi(S^{n-1} \times \{0\})$, of a neck of radius $\frac{n-1}{H(p_0)}$.

- An $\varepsilon$-cap is a compact region $D \subset M$ diffeomorphic to a closed $n$-dimensional ball with the property that $\partial D$ is the central cross-sectional sphere of an $\varepsilon$-neck.

Recently, there has been significant progress in the classification of ancient solutions of mean curvature flow that are two-convex and noncollapsed. There has also been great progress in classifying ancient solutions of the Ricci flow in three dimensions. See [12, 13, 11, 16]. In the mean curvature flow, by the aforementioned results of Brendle and Choi and the works of Angenent, Daskalopoulos, and Sesum [3, 4], an $n$-dimensional, ancient, uniformly two-convex, noncollapsed and nonflat solution of mean curvature flow in $\mathbb{R}^{n+1}$ is either a family of shrinking round spheres, a family of shrinking round cylinders, a translating bowl soliton, or an ancient oval. Only the first three can arise as blow-up limits before the first singular time. It is still an open problem to determine whether an ancient oval can occur as a singularity model at subsequent singular times. In any case, for the canonical neighborhood theorem, it is important to include the ancient ovals in our class of model solutions since it is possible for regions of high curvature to be modeled on domains within ancient ovals. A similar phenomenon occurs in the Ricci flow in three dimensions. To the author's knowledge, Perelman [44] gave the first construction of ancient ovals for the Ricci flow in three dimensions and White [49] gave the analogous construction for the mean curvature flow.

In the Ricci flow, Perelman did not use an explicit classification of model solutions to prove his canonical neighborhood theorem. We do not need an explicit classification either. So that our results do not depend upon a classification, we opt to use the following definition for ancient model solutions.

**Definition 4.3.** An ancient model solution is an $n$-dimensional ancient, nonflat, complete, connected, codimension one solution of mean curvature flow in $\mathbb{R}^N$ that is uniformly two-convex and noncollapsed.

Note that since the model flow is contained in an $(n+1)$-dimensional plane, here noncollapsing is meant with respect to this plane.

By Theorem 1.11 in Haslhofer-Kleiner [29], any ancient, noncollapsed, mean-convex solution of mean curvature flow is automatically weakly convex. Moreover, by Theorem 1.8 in [29], given a noncollapsing constant $\alpha > 0$, there exists constants $\gamma_1 := \gamma_1(n, \alpha)$ and $\gamma_2 := \gamma_2(n, \alpha)$ with the property that any ancient, $\alpha$-noncollapsed, mean-convex solution of mean curvature flow satisfies the pointwise derivative estimates $|\nabla A| \leq \gamma_1 |H|^2$ and $|\nabla^2 A| \leq \gamma_2 |H|^3$. In Chapter 2, we proved two structure theorems for weakly convex, uniformly two-convex, ancient solutions satisfying the derivative estimates $|\nabla A| \leq \gamma_1 |H|^2$ and $|\nabla^2 A| \leq \gamma_2 |H|^3$. Our work also shows that these two derivative estimates together with convexity and uniform two-convexity imply noncollapsing (a sort of converse to the Haslhofer-Kleiner result). The following proposition is a straightforward corollary of the Theorem 2.18 and Theorem 2.22 in Chapter 2. It is also a consequence of the works [4, 13].

**Proposition 4.4.** *Given $\varepsilon > 0$ and $\alpha > 0$, there exist positive constants $C_1 := C_1(n, \alpha, \varepsilon)$ and $C_2 := C_2(n, \alpha, \varepsilon)$ with the following property. Assume $\bar{F} : \bar{M} \times (-\infty, T) \to \mathbb{R}^N$ is an ancient model solution which is $\alpha$-noncollapsed. Then for each space-time point $(p_0, t_0)$, there exists a closed neighborhood $B \subset \bar{M}$ containing $p_0$ such that $B_{g(t_0)}(p_0, C_1^{-1} H(p_0, t_0)^{-1}) \subset B \subset B_{g(t_0)}(p_0, C_1 H(p_0, t_0)^{-1})$ and $C_2^{-1} H(p_0, t_0) \leq H(p, t_0) \leq C_2 H(p_0, t_0)$ for every $p \in B$. Moreover, the neighborhood $B$ is either an $\varepsilon$-neck, an $\varepsilon$-cap, or a closed manifold diffeomorphic to $S^n$.*

The constants $\tilde{K}$ and $\tilde{\varepsilon}$ from Theorem 4.1 will depend upon $n, C_1$, and $C_2$. The initial immersion $F_0$ determines the scale-invariant derivative estimate bounds $\gamma_1, \gamma_2$ satisfied by the singularity models (by [41]). These constant $\gamma_1$ and $\gamma_2$ in turn determine the noncollapsing constant $\alpha$ satisfied by blow-up limits (by our work in Chapter 2). In this way, $\tilde{K}$ and $\tilde{\varepsilon}$ will depend upon $n$ and $F_0$. Note, however, the classification of ancient model solutions reveals they are universally noncollapsed for some $\alpha = \alpha(n)$. By relying on the classification, we could remove the dependence of the constants $C_1$ and $C_2$ upon $\alpha$ and consequently $\tilde{K}$ and $\tilde{\varepsilon}$ on $F_0$, if we desired.

There are a few reasonable topologies one could use to say a given solution is $\varepsilon_0$-close to a model solution on a parabolic neighborhood. Based on the compactness result of Chapter 3, we will use the following definition. Fix a small constant $\varepsilon_0 > 0$ and a large constant $K_0 < \infty$. Suppose $F : M \times [0, T) \to \mathbb{R}^N$ is a solution of the mean curvature and $(p_0, t_0)$ is a spacetime point. Set $Q_0 := H(p_0, t_0)$. Suppose that $F$ is defined in the intrinsic parabolic neighborhood $\hat{P}(p_0, t_0, K_0, K_0)$. Consider the rescaled solution

$$\tilde{F}(p, t) := Q_0 F(p, t_0 + Q_0^{-2}(t - t_0)).$$

After rescaling, we have $\tilde{g}(t_0) = Q_0^2 g(t_0)$, $\tilde{H}(p_0, t_0) = 1$, and the solution $\tilde{F}$ is defined in the intrinsic parabolic neighborhood $P(p_0, t_0, K_0, K_0)$.

**Definition 4.5.** Let $F : M \times [0,T) \to \mathbb{R}^N$ be an $n$-dimension solution to mean curvature flow and $(p_0, t_0)$ a spacetime point satisfying $H(p_0, t_0) = 1$. Suppose $F$ is defined in the parabolic neighborhood $P(p_0, t_0, K_0, K_0)$ (i.e. $\partial M$, if it exists, satisfies $d_{g(t_0)}(p_0, \partial M) > K_0$ and $[t_0 - K_0, t_0] \subset [0,T)$). We will say the solution $F$ is $\varepsilon_0$-close in $P(p_0, t_0, K_0, K_0)$ to an ancient model solution if the following holds. We can find

- an ancient model solution $\bar{F} : \bar{M} \times (-\infty, t_0] \to \mathbb{R}^N$;
- a point $\bar{p}_0 \in \bar{M}$ with $H(\bar{p}_0, t_0) = 1$;
- a smooth relatively compact domain $V$ such that $B_{\bar{g}(t_0)}(\bar{p}_0, K_0) \subset V \subset \bar{M}$;
- and a diffeomorphism $\Phi : V \to \Phi(V) \subset M$ such that $\Phi(\bar{p}_0) = p_0$ and $B_{g(t_0)}(p_0, K_0) \subset \Phi(V) \subset M$.

Moreover, for each $t \in [t_0 - K_0, t_0]$, the immersions $F(\cdot, t) \circ \Phi$ and $\bar{F}(\cdot, t)$ are $\varepsilon_0$-close in $C^{[1/\varepsilon_0]}$ on $V$ with respect to the metric $\bar{g} := \bar{g}(t_0)$. Specifically, we think of $F(\cdot, t) \circ \Phi$ and $\bar{F}(\cdot, t)$ as $\mathbb{R}^N$-valued functions on $V$ and we require

$$\sup_{V \times [t_0 - K_0, t_0]} \sum_{m=0}^{[1/\varepsilon_0]} \left| \bar{\nabla}^m \big( F(\cdot, t) \circ \Phi - \bar{F}(\cdot, t) \big) \right|_{\bar{g}}^2 < \varepsilon_0,$$

where $\bar{\nabla}$ denotes the Levi-Civita connection of $\bar{g}$ on $\bar{M}$.

Of course, the definition applies to the rescaled solution if $H(p_0, t_0) \neq 1$. Note that it follows from the evolution equation of $F$ that if $2\ell + m \leq [1/\varepsilon_0]$, then the derivatives of the form $\frac{\partial}{\partial t}^{\ell} \bar{\nabla}^m (F(\cdot, t) \circ \Phi)$ are $O(\varepsilon_0)$-close to the corresponding derivatives of the model solution.

In the final step of the proof of the canonical neighborhood theorem, we will need to make use of the follow distance distortion estimate. The lemma is analogous to Lemma 8.3 in Perelman's work [43]. To the author's knowledge, this type of argument is originally due to Hamilton. By our pinching estimate, the product $Hh$ is comparable to $\mathrm{Ric}$ and this allows us follow Perelman's proof of the estimate for the Ricci flow.

**Lemma 4.6.** *Suppose $F : M \times (t_1, t_2) \to \mathbb{R}^{n+1}$ is a complete, convex, $n$-dimensional solution of mean curvature flow satisfying $|h|^2 \leq \beta^2 H^2$ for some $0 < \beta < 1$. Suppose we have a bound $H(\cdot, t) \leq \Lambda(t)$ for each $t \in (t_1, t_2)$. Then for any $p, q \in M$ and $t \in (t_1, t_2)$, we have*

$$0 \leq -\frac{d}{dt} d_{g(t)}(p, q) \leq C(n, \beta)\Lambda(t).$$

*Proof.* Recall that the evolution of the metric is given by $\frac{\partial}{\partial t} g_{ij} = -2Hh_{ij}$. Let us fix a time $t_0$ and restrict our attention to $(M, g(t_0))$. Fix some points $p, q \in M$. Set $\ell := d_{g(t_0)}(p, q)$ and let $\gamma : [0, \ell] \to M$ be a minimizing, unit-speed geodesic between $p$ and $q$. Set $X(s) = \gamma'(s)$. Then the derivative of the distance is given by

$$\frac{d}{dt} d_{g(t)}(p, q)\Big|_{t=t_0} = -\int_0^\ell Hh(X, X)\, ds.$$

We have assumed $h \geq 0$ and the inequality $|h|^2 \leq \beta^2 H^2$ implies $\lambda_n \leq \beta H$, where $\lambda_n$ denotes the maximum eigenvalue of $h$. From the identity above, together with $0 \leq Hh(X, X) \leq \beta H^2 < \Lambda(t_0)^2$, we first derive the crude estimate

$$0 \leq -\frac{d}{dt} d_{g(t)}(p, q)\Big|_{t=t_0} \leq \Lambda(t_0)^2 \ell.$$

By the Gauss equation that $Hh(X, X) = \mathrm{Ric}(X, X) + h^2(X, X)$. On the one hand, this gives $\mathrm{Ric}(X, X) \leq Hh(X, X) < \Lambda(t_0)^2$. On the other hand, since $h^2(X, X) \leq \beta Hh(X, X)$, it follows that $Hh(X, X) \leq \frac{1}{1-\beta}\mathrm{Ric}(X, X)$, and we have

$$0 \leq -\frac{d}{dt} d_{g(t)}(p, q)\Big|_{t=t_0} \leq \frac{1}{1-\beta}\int_0^\ell \mathrm{Ric}(X, X)\, ds.$$

Having established the inequality above and the upper bound $\mathrm{Ric}(X, X) \leq \Lambda(t_0)^2$, we now proceed as Perelman does in [43]. Since $\gamma$ is a minimizing geodesic, for any vector field $V(s)$ defined along $\gamma$ and orthogonal to $X$, the second variation formula for the energy of $\gamma$ gives

$$0 \leq \int_0^\ell |\nabla_X V|^2 - R(X,V,X,V)\, ds.$$

Consider a distance $r_0 \in (0, \frac{\ell}{2}]$. Let $e_1(s), \ldots, e_n(s)$ be a parallel orthonormal frame along $\gamma$ with $e_n(s) = X(s)$. Define a function

$$f(s) = \begin{cases} \frac{s}{r_0} & s \in [0, r_0] \\ 1 & s \in [r_0, \ell - r_0] \\ \frac{\ell - s}{r_0} & s \in [\ell - r_0, \ell] \end{cases}.$$

Plugging $V_i(s) = f(s)e_i(s)$ for $i = 1, \ldots, n-1$ into the second variation formula and summing implies

$$0 \leq \int_0^{r_0} \frac{n-1}{r_0^2} - \frac{s^2}{r_0^2}\mathrm{Ric}(X,X)\, ds - \int_{r_0}^{\ell - r_0} \mathrm{Ric}(X,X)\, ds + \int_{\ell - r_0}^{\ell} \frac{n-1}{r_0^2} - \frac{(\ell - s)^2}{r_0^2}\mathrm{Ric}(X,X)\, ds.$$

Rearranging and using our upper bound for $\mathrm{Ric}$, we get

$$\begin{aligned} \int_0^\ell \mathrm{Ric}(X,X)\, ds &\leq 2(n-1)r_0^{-1} + \int_0^{r_0} (1 - \frac{s^2}{r_0^2})\mathrm{Ric}(X,X)\, ds + \int_{\ell - r_0}^{\ell} (1 - \frac{(\ell - s)^2}{r_0^2})\mathrm{Ric}(X,X)\, ds \\ &\leq 2(n-1)r_0^{-1} + 2\Lambda(t_0)^2 \int_0^{r_0} (1 - \frac{s^2}{r_0^2})\, ds \\ &= 2(n-1)r_0^{-1} + \frac{4}{3}\Lambda(t_0)^2 r_0. \end{aligned}$$

Therefore, for any $0 < r_0 \leq \frac{\ell}{2}$, we have

$$0 \leq -\frac{d}{dt} d_{g(t)}(p,q)\Big|_{t=t_0} \leq \frac{1}{1-\beta}\big(2(n-1)r_0^{-1} + \frac{4}{3}n\Lambda(t_0)^2 r_0\big).$$

If $\Lambda(t_0)^{-1} \leq \frac{\ell}{2}$, then taking $r_0 = \Lambda(t_0)^{-1}$ gives the desired conclusion. If $\Lambda(t_0)^{-1} > \frac{\ell}{2}$, then $\ell < 2\Lambda(t_0)^{-1}$ and our crude estimate established earlier suffices. This completes the proof. $\square$

We will use the lemma above in conjunction with Hamilton's Harnack inequality for mean curvature flow.

**Theorem 4.7** (R. Hamilton [25]). *Assume $F : M \times (0,T) \to \mathbb{R}^{n+1}$ is a complete, weakly convex solution of mean curvature flow such that*

$$\sup_{(p,t)\in M\times(\tau,T)} H(p,t) < \infty$$

*for all $\tau \in (0,T)$. Then, for all $(p,t) \in M \times (0,T)$ and any tangent vector $v \in T_pM$,*

$$\frac{\partial}{\partial t}H + \frac{1}{2t}H + 2\langle \nabla H, v\rangle + h(v,v) \geq 0.$$

*In particular, the quantity $\sqrt{t}H(p,t)$ is nondecreasing for each point $p \in M$ along the flow.*

## 4.2. Canonical Neighborhood Theorem

In this section, we give a proof of the main theorem:

**Theorem 4.8.** *Suppose $F_0 : M \to \mathbb{R}^N$ is an immersion of a closed manifold of dimension $n \geq 5$ satisfying $|A|^2 < \tilde{c}_2|H|^2$, where $\tilde{c}_2 = \min\{\frac{1}{n-2}, \frac{3(n+1)}{2n(n+2)}\}$. There exist constants $\tilde{\varepsilon}$ and $\tilde{K}$, depending only upon the dimension and $F_0$, with the following property. Let $F : M \times [0,T) \to \mathbb{R}^N$ denote the solution of mean curvature flow with initial immersion given by $F_0$. Given $\varepsilon_0 \in (0,\tilde{\varepsilon})$ and $K_0 \in (\tilde{K},\infty)$, there exists a positive number $\hat{r} > 0$, depending upon, $n$, $F_0$, $\varepsilon_0$, and $K_0$, with the following property. If $(p_0,t_0)$ is a spacetime point such that $|H|(p_0,t_0) \geq \hat{r}^{-1}$, then the solution is $\varepsilon_0$-close in $\hat{P}(p_0,t_0,K_0,K_0)$ to an ancient model solution.*

*Proof.* Our proof closely follows the proof of Theorem 7.2 in [10], which is an adaptation of Perelman's work by Brendle. We will argue by contradiction and induction on scales, as introduced by Perelman. The theorem is stated only for $\varepsilon_0$ sufficiently small and $K_0$ sufficiently large depending upon the dimension and $F_0$. Recall that the initial immersion $F_0$ determines a noncollapsing constant $\alpha > 0$ for singularity models. We begin by fixing some $\varepsilon > 0$ small and letting

$C_1 = C_1(n, F_0, \varepsilon)$ and $C_2 = C_2(n, F_0, \varepsilon)$ be the constants appearing in Proposition 4.4. We will assume through the proof that $\varepsilon$ is sufficiently small depending upon certain universal constants that arise in the proof. We let $C(n)$ denote an arbitrary constant depending upon the dimension, which may change from line to line. We will prove the theorem assuming $\varepsilon_0$ is much smaller than $\varepsilon$ and $K_0 \geq 16C_1$.

The trace of the Gauss equation gives $\mathrm{scal} = H^2 - |A|^2$. Together with the pinching assumption $\frac{1}{n}H^2 \leq |A|^2 \leq \tilde{c}_2 H^2$, this implies that $(1-\tilde{c}_2)H^2 \leq \mathrm{scal} \leq (1-\frac{1}{n})H^2$. Hence bounds for extrinsic curvature give bounds for the intrinsic curvature and vice versa.

Now, if the assertion of the theorem is false, then there exists a sequence of spacetime points $(p_j, t_j)$ with the following properties:

(i) $Q_j := H(p_j, t_j) \geq j$.

(ii) The solution is not $\varepsilon_0$-close in the parabolic neighborhood $\hat{P}(p_j, t_j, K_0, K_0)$ to any ancient model solution.

After point-picking process, we can further assume that the spacetime points $(p_j, t_j)$ have the additional property:

(iii) If $t \leq t_j$ and $(p, t)$ satisfies $H(p, t) \geq 4Q_j$, then the solution is $\varepsilon_0$-close in the parabolic neighborhood $\hat{P}(p, t, K_0, K_0)$ to an ancient model solution.

If $j \geq j_0(F_0)$ is sufficiently large and $H(p, t) \geq 2j$, then we must have $t \geq \frac{T}{2}$ since the flow is smooth and the curvature is bounded on $M \times [0, \frac{T}{2}]$. Consequently, $t - K_0 H(p, t)^{-2} \geq \frac{T}{2} - K_0 \frac{1}{4j^2} \geq \frac{T}{4}$ if $j \geq \max\{j_0, \sqrt{K_0 T}\}$. Restricting our attention to such $j$ ensures $\hat{P}(p, t, K_0, K_0) \subset M \times [0, T)$. Since the curvature is bounded on $M \times [0, t_j]$ and any point $(p, t)$ with $t \leq t_j$ and $H(p, t) \geq 4Q_j$ must have $t \geq \frac{T}{4}$, property (iii) must hold (for each $j$) after replacing $(p_j, t_j)$ finitely many times if necessary.

Under assumptions (i), (ii) and (iii), we will show that, after dilating the solution $F$ around the point $(p_j, t_j)$ by the factor $Q_j$, a subsequence of the rescaled solutions converges to an ancient model solution.

*Step 1:* We begin by recalling the derivative and cylindrical estimates established by Nguyen in [41]. We can find $a := a(F_0) > 0$ so that the estimate $|A|^2 + a \leq \tilde{c}_2 H^2$ holds initially. This inequality is preserved along the flow and hence $H^2 \geq \frac{a}{\tilde{c}_2}$ holds for all $t \in [0, T)$. Together with the derivative estimates established in Theorems 3.1 and 3.2 in [41], this implies we can find a large positive constant $\eta \geq 100$, depending only upon $n$ and the initial immersion $F_0$, such that the pointwise derivative estimates $|\nabla A| \leq \eta H^2$ and $|\nabla^2 A| \leq \eta H^3$ hold for all $t \in [0, T)$. Using Kato's inequality and the evolution equation for $\vec{H}$, we can choose $\eta$ so that $|\nabla H| \leq \eta H^2$ and $|\frac{\partial}{\partial t} H| \leq \eta H^3$ hold for $t \in [0, T)$.

In addition, Theorem 4.1 in [41] gives a cylindrical pinching estimate for our solution: for every $\delta > 0$, there exists a constant $C_\delta$ such that $|A|^2 \leq \left(\frac{1}{n-1} + \delta\right) H^2 + C_\delta$.

*Step 2:* In this step, we use the derivative estimates of the previous step to establish short range curvature estimates. We then obtain higher order pointwise derivative estimates for the second fundamental form. Suppose $\tilde{t} \leq t_j$ and $\tilde{p} \in M$. Assume $H(\tilde{p}, \tilde{t}\,) = r_0^{-1}$. The estimates of the previous step imply $|\nabla H^{-1}| \leq \eta$ and $|\frac{\partial}{\partial t} H^{-2}| \leq \eta$. Suppose $(p, t)$ is a spacetime point satisfying $d_{g(\tilde{t})}(p, \tilde{p}) \leq \frac{1}{4\eta} r_0$ and $0 \leq \tilde{t} - t \leq \frac{1}{4\eta} r_0^2$. Integrating the spacial derivative estimate over small distances gives

$$\Big|\frac{1}{H(\tilde{p}, \tilde{t})} - \frac{1}{H(p, \tilde{t})}\Big| \leq \eta \, d_{g(\tilde{t})}(p, \tilde{p}) \leq \frac{1}{4} r_0.$$

Integrating the time derivative estimate over small time intervals gives

$$\Big|\frac{1}{H(p, \tilde{t})^2} - \frac{1}{H(p, t)^2}\Big| \leq \eta \, |\tilde{t} - t| \leq \frac{1}{4} r_0^2.$$

Combining these estimates with the assumption $H(\tilde{p}, \tilde{t}) = r_0^{-1}$ shows that $\frac{1}{4} r_0^{-1} \leq H(p, t) \leq 4 r_0^{-1}$ for all $(p, t)$ in the parabolic neighborhood $P(\tilde{p}, \tilde{t}, \frac{r_0}{4\eta}, \frac{r_0^2}{4\eta}) = \hat{P}(\tilde{p}, \tilde{t}, \frac{1}{4\eta}, \frac{1}{4\eta})$. Now standard interior estimates for mean curvature flow imply that

$$|\nabla^k A|(\tilde{p},\tilde{t}) \leq C(k,n,\eta)H(\tilde{p},\tilde{t})^{k+1}.$$

*Step 3:* Next, we establish uniform bounds for $Q_j^{-1}H$ at bounded distance from $p_j$ at time $t_j$. This long-range curvature estimate is by far the most involved step in the proof. The goal, as usual, is to show that if the curvature blows up in finite distance, then the singularity must be modeled on a cone. Then we derive a contradiction using the strong maximum principle for the flow.

For all $\rho > 0$, let

$$\mathbb{M}(\rho) := \limsup_{j\to\infty} \sup_{p\in B_{g(t_j)}(p_j, Q_j^{-1}\rho)} Q_j^{-1}H(p,t_j).$$

Note $\mathbb{M}(\rho)$ is monotone increasing. Our goal is to show $\mathbb{M}(\rho) < \infty$ for all $\rho > 0$. By the previous step (setting $(\tilde{p},\tilde{t}\,) = (p_j,t_j)$ and $r_0 = Q_j^{-1}$), we have $\mathbb{M}(\rho) \leq 4$ for $0 < \rho < \frac{1}{4\eta}$. Define

$$\rho_* := \sup\{\rho \geq 0 : \mathbb{M}(\rho) < \infty\}.$$

Evidently, $\rho_* \geq \frac{1}{4\eta}$. Suppose, for sake of contradiction, that $\rho_* < \infty$. This step of the proof is involved, so we argue in five sub-steps.

*Step 3.1.* We begin by extracting a local geometric limit. By definition of $\rho_*$, for every $0 < \rho < \rho_*$, we have uniform bounds

$$\sup_{p\in B_{g(t_j)}(p_j, Q_j^{-1}\rho)} Q_j^{-1}H(p,t_j) \leq C(\rho).$$

The higher order derivative estimates in Step 2 imply uniform bounds

$$\sup_{p\in B_{g(t_j)}(p_j, Q_j^{-1}\rho)} Q_j^{-k-1}|\nabla^k A|(p,t_j) \leq C(k,n,\eta,\rho)$$

for each $0 < \rho < \rho_*$. For each $j$, define an immersion $\tilde{F}_j : B_{g(t_j)}(p_j, Q_j^{-1}\rho_*) \to \mathbb{R}^N$ by

$$\tilde{F}_j(p) = Q_j\big(F(p,t_j) - F(p_j,t_j)\big).$$

Let $\tilde{g}_j := Q_j^2 g(t_j)$ denote the rescaled metric. Then $B_{g(t_j)}(p_j, Q_j^{-1}\rho_*) = B_{\tilde{g}_j}(p_j, \rho_*)$. So $\tilde{F}_j : (M, \tilde{g}_j, p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ is a sequence of pointed isometric immersions and, for all $0 < \rho < \rho_*$ and for each $j$, we have uniform estimates for the second fundamental form of $\tilde{F}_j$ and its derivatives on $B_{\tilde{g}_j}(p_j, \rho)$. By Proposition 3.2, we can pass to a local limit. After passing to a subsequence, the sequence of immersions $\tilde{F}_j : (B_{\tilde{g}_j}(p_j, \rho_*), \tilde{g}_j, p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ converges on compact subsets to a locally defined pointed isometric immersion $F_\infty : (B_\infty, g_\infty, p_\infty) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$, where $B_\infty = B_{g_\infty}(p_\infty, \rho_*)$ is a geodesic ball in the metric in $g_\infty$. In particular, $d_{g_\infty}(p_\infty, \partial B_\infty) \geq \rho_*$.

The limit must satisfy $H_\infty > 0$. Indeed, since the derivative estimate $|\nabla H_\infty| \leq \eta H_\infty^2$ holds in the limit and $H_\infty(p_\infty) = 1$, this follows from integration. Now because the limit satisfies $H_\infty > 0$, the planarity estimate in of Chapter 1 (in particular, see Proposition 1.7) implies $F_\infty(B_\infty)$ is contained in an affine $(n+1)$-dimensional subspace. So we may write $F_\infty : B_\infty \to \mathbb{R}^{n+1}$. Since the limit is codimension one, the cylindrical estimate, $|A_\infty|^2 \leq \frac{1}{n-1} H_\infty^2$, implies the limit is weakly convex.

By assumption $\mathbb{M}(\rho_*) = \infty$. Thus, we can find a sequence of points $q_j \in B_{g(t_j)}(p_j, Q_j^{-1}\rho_*)$ such that

$$Q_j^{-1} H(q_j, t_j) \to \infty \quad \text{and} \quad \rho_j := Q_j d_{g(t_j)}(p_j, q_j) \to \rho_*.$$

For each $j$, let $\gamma_j : [0, Q_j^{-1}\rho_j] \to B_{g(t_j)}(p_j, Q_j^{-1}\rho_*)$ be a unit-speed length-minimizing $g(t_j)$-geodesic between the points $p_j$ and $q_j$. The geodesics $s \mapsto \gamma_j(Q_j^{-1}s)$ for $s \in [0, \rho_j]$ converge locally to a unit-speed geodesic $\gamma_\infty : [0, \rho_*) \to B_\infty$ missing its terminal point.

*Step 3.2* The pointwise derivative estimate, $|\nabla H_\infty| \leq \eta H_\infty^2$, implies

$$H_\infty(\gamma_\infty(s)) = \lim_{j\to\infty} Q_j^{-1} H(\gamma_j(Q_j^{-1}s), t_j) \geq (2\eta(\rho_* - s))^{-1} \geq 16$$

for $s \in [\rho_* - \frac{1}{32\eta}, \rho_*)$. Indeed, suppose for some $\tilde{s} \in [\rho_* - \frac{1}{32\eta}, \rho_*)$, we have $H_\infty(\gamma_\infty(\tilde{s})) < (2\eta(\rho_* - \tilde{s}))^{-1}$. Arguing as in Step 2, we find that

$$H_\infty(\gamma_\infty(s)) \leq \frac{H_\infty(\gamma_\infty(\tilde{s}))}{1 - \eta(s - \tilde{s})H_\infty(\gamma_\infty(\tilde{s}))} \leq (\eta(\rho_* - \tilde{s}))^{-1} < \infty$$

for $s \in (\tilde{s}, \rho_*)$. In particular, $\lim_{s \to \rho_*} H_\infty(\gamma_\infty(s)) < \infty$, which contradicts $\lim_{j\to\infty} Q_j^{-1} H(q_j, t_j) \to \infty$.

*Step 3.3.* Next, we show that for all $s$ sufficiently close to $\rho_*$, the point $\gamma_\infty(s)$ lies at the center of a neck. Let

$$s_* := \max\Big\{\frac{32C_1\eta}{32C_1\eta + 1}\rho_*, \rho_* - \frac{1}{32\eta}\Big\}.$$

Consider $s \in [s_*, \rho_*)$ so that $32C_1\eta(\rho_* - s) \leq s$. Consider the point $x_j := \gamma_j(Q_j^{-1}s)$. By Step 3.2, $H(x_j, t_j) \geq 4Q_j$ if $j$ is sufficiently large. By property (iii), we can find a neighborhood of $x_j$ containing $B_{g(t_j)}(x_j, \frac{1}{2}K_0 H(x_j, t_j)^{-1})$ on which the immersion $F(\cdot, t_j)$ is $\varepsilon_0$-close to a time slice of an embedding of a model solution. Under the assumptions $K_0 \geq 16C_1$ and $\varepsilon_0 \ll \varepsilon$ (by a factor depending on the dimension) Definition 4.2, Definition 4.5, and Proposition 4.4 imply the point $x_j$ has a canonical neighborhood $U_j$ with the properties:

- $U_j$ is either a $2\varepsilon$-neck, a $2\varepsilon$-cap, or a closed manifold diffeomorphic to $S^n$.
- $B_{g(t_j)}(x_j, (2C_1)^{-1}H(x_j, t_j)^{-1}) \subset U_j \subset B_{g(t_j)}(x_j, 2C_1 H(x_j, t_j)^{-1})$.
- In $U_j$, the mean curvature satisfies $(2C_2)^{-1}H(x_j, t_j) \leq H \leq 2C_2 H(x_j, t_j)$.

Let us show $U_j$ must be a $2\varepsilon$-neck. Since $\mathbb{M}(s) < \infty$, the ratio $H(q_j, t_j)/H(x_j, t_j) \to \infty$ as $j \to \infty$. Therefore, $H(q_j, t_j) \geq 4C_2 H(x_j, t_j)$ if $j$ is sufficiently large and it follows from the third item above that $q_j \notin U_j$. The estimate in Step 3.2 gives $8C_1 H(\gamma_\infty(s))^{-1} \leq 16C_1\eta(\rho_* - s) < s$, by definition of $s_*$. Hence $4C_1 H(x_j, t_j)^{-1} \leq Q_j^{-1}s = d_{g(t_j)}(p_j, x_j)$ if $j$ is sufficiently large and it follows from the second item above that $p_j \notin U_j$. These considerations evidently imply $U_j$ is not a closed manifold diffeomorphic to $S^n$. If $U_j$ is a $2\varepsilon$-cap, then the geodesic $\gamma_j$ enters and exits the cap. However, by definition, the boundary of a $2\varepsilon$-cap is a central cross-sectional sphere of a $2\varepsilon$-neck. If $\varepsilon$ is sufficiently small depending upon the dimension, this contradicts the fact that $\gamma_j$ is minimizing.

In summary, for each $s \in [s_*, \rho_*)$, the point $\gamma_j(Q_j^{-1}s)$ has a canonical neighborhood which is a $2\varepsilon$-neck for $j$ sufficiently large (depending upon $s$). Moreover, on each $2\varepsilon$-neck, we have the improved gradient estimate $|\nabla H| \leq C(n)\varepsilon H^2$. Passing to the limit for each such $s \in (s_*, \rho_*)$, we conclude the point $\gamma_\infty(s)$ lies at the center of a $C(n)\varepsilon$-neck in $(B_\infty, g_\infty)$. As in Step 3.2, the improved gradient estimate gives

$$H_\infty(\gamma_\infty(s)) \geq (C(n)\varepsilon(\rho_* - s))^{-1}$$

for $s \in [s_*, \rho_*)$.

*Step 3.4.* For the last two parts of Step 3, we follow the argument of Brendle in [9]. By Step 3.3, every point $\gamma_\infty(s)$, for $s \in [s_*, \rho_*)$, lies at the center of an $C(n)\varepsilon$-neck. In the present context, a $C(n)\varepsilon$-neck is an extrinsic notion given by Definition 4.2, but the definition implies each point lies at the center of an intrinsic $C(n)\varepsilon$-neck in the sense used by Perelman in [44]. Let $U \subset B_\infty$ denote the connected region obtained by taking the union over all of the $C(n)\varepsilon$-necks centered at the points $\gamma_\infty(s)$ for $s \in [s_*, \rho_*)$.

The work of Hamilton [27] shows that each of these $C(n)\varepsilon$-necks admits a canonical foliation by constant mean curvature spheres (at least away from the boundary of the neck). Moreover, the foliations of overlapping necks must agree and can be joined together. Consequently, the domain $U$ admits a foliation by a one-parameter family of CMC spheres, which we denote by $\Sigma_u$. We can arrange the parameter $u$ so that the CMC spheres are defined for $u \in (0, u_*]$ and so that as $u \to 0$, the spheres $\Sigma_u$ move away from the point $p_\infty$ and towards the end of the horn. We let $v : \Sigma_u \to \mathbb{R}$ denote the lapse function of the foliation. We can parametrize the foliation $P : S^{n-1} \times (0, u_*] \to U$ so that $v = |\frac{\partial P}{\partial u}|_{g_\infty}$. We can also express the leaves of the foliation as the level sets of the projection

$\pi : U \to (0, u_*]$. In this case, $v = |\nabla \pi|^{-1}_{g_\infty}$. After reparametrization, we may assume the average of $v$ over $\Sigma_u$ is 1. Note that $\sup_{\Sigma_u} |v-1| \leq C(n)\varepsilon$. Finally, let $\nu := -v^{-1}\frac{\partial P}{\partial u} = -v\nabla\pi$ denote the unit normal to the foliation.

For each $s \in [s_*, \rho_*)$, the point $\gamma_\infty(s)$ contained in a CMC sphere $\Sigma_{u(s)}$ where $u(s) := \pi(\gamma_\infty(s))$. We may assume $u(s_*) = u_*$. Since $\gamma_\infty$ is the limit of a sequence of length minimizing geodesics $\gamma_j$ on $2\varepsilon$-necks, we must have $1 - C(n)\varepsilon \leq g_\infty(\nu, \gamma'_\infty) \leq 1$.

Since

$$\frac{du}{ds} = g_\infty(\nabla\pi, \gamma'_\infty) = -v^{-1} g_\infty(\nu, \gamma'_\infty),$$

this gives $|u'(s) + 1| \leq C(n)\varepsilon$ and $1 - C(n)\varepsilon \leq u(s)(\rho_* - s)^{-1} \leq 1 + C(n)\varepsilon$ for $s \in [s_*, \rho_*)$. Because each CMC sphere $\Sigma_u$ must be close to a round sphere on a $C(n)\varepsilon$-neck, we have the estimates

$$\sup_{q\in\Sigma_u} \mathrm{scal}_{g_\infty}(q) \leq (1 + C(n)\varepsilon) \inf_{q\in\Sigma_u} \mathrm{scal}_{g_\infty}(q)$$

and

$$\frac{1}{C(n)}\Big(\sup_{q\in\Sigma_u} \mathrm{scal}_{g_\infty}(q)\Big)^{-\frac{n-1}{2}} \leq \mathrm{area}_{g_\infty}(\Sigma_u) \leq C(n)\Big(\inf_{q\in\Sigma_u} \mathrm{scal}_{g_\infty}(q)\Big)^{-\frac{n-1}{2}}$$

for $u \in (0, u_*]$.

Recall that the cylindrical estimate $|A_\infty|^2 \leq \frac{1}{n-1} H^2_\infty$ implies $H^2_\infty$ and $\mathrm{scal}_{g_\infty}$ are comparable. In particular, from Step 3.3 we obtain

$$(\rho_* - s)^2 \mathrm{scal}_{g_\infty}(\gamma_\infty(s)) \geq (C(n)\varepsilon)^{-2} > 0$$

for $s \in [s_*, \rho_*)$. Since $u(s)$ and $\rho_* - s$ are comparable, we get

$$u^2 \inf_{q\in\Sigma_u} \mathrm{scal}_{g_\infty}(q) \geq (C(n)\varepsilon)^{-2} > 0$$

for $u \in (0, u_*]$. This implies

$$\frac{\mathrm{area}_{g_\infty}(\Sigma_u)}{u^{n-1}} \leq C(n)\varepsilon^{n-1} < \infty$$

for $u \in (0, u_*]$. In particular, this implies that $\text{area}_{g_\infty}(\Sigma_u) \to 0$ as $u \to 0$.

We will now consider the extrinsic geometry of each of the leaves $\Sigma_u$ as hypersurfaces within $(B_\infty, g_\infty)$. Let $\mathcal{H}$ and $\mathcal{A}$ denote the scalar mean curvature and second fundamental form of $\Sigma_u$ with respect to $g_\infty$ (not to be mistaken for $H_\infty$ and $A_\infty$). For each $u$, the mean curvature $\mathcal{H} = \mathcal{H}(u)$ of $\Sigma_u$ is constant. As we are on a $C(n)\varepsilon$-neck, the ratio $\big(\sup_{q \in \Sigma_u} \text{scal}_{g_\infty}(q)\big)^{-\frac{1}{2}} \mathcal{H}(u)$ is close to zero. Using the first variation formula for the mean curvature and the fact that $(B_\infty, g_\infty)$ has nonnegative Ricci curvature, we obtain

$$-\mathcal{H}'(u) = \Delta_{\Sigma_u} v + (|\mathcal{A}|^2 + \text{Ric}_{g_\infty}(\nu, \nu))v \geq \Delta_{\Sigma_u} v + \frac{1}{n-1}\mathcal{H}(u)^2 v.$$

Taking the average over $\Sigma_u$ on both sides gives

$$-\mathcal{H}'(u) \geq \frac{1}{n-1}\mathcal{H}(u)^2.$$

This inequality implies that either $\mathcal{H}(u) \leq 0$ for all $u$, or $\mathcal{H}(u)$ is positive and satisfies

$$\mathcal{H}(u) \leq \frac{n-1}{u}.$$

The former cannot occur since

$$\frac{d}{du}\text{area}_{g_\infty}(\Sigma_u) = \mathcal{H}(u) \int_{\Sigma_u} v = \mathcal{H}(u)\, \text{area}_{g_\infty}(\Sigma_u),$$

and we know $\text{area}_{g_\infty}(\Sigma_u) \to 0$ as $u \to 0$. Consequently,

$$\frac{d}{du}\big(u^{1-n}\text{area}_{g_\infty}(\Sigma_u)\big) = u^{1-n}\text{area}_{g_\infty}(\Sigma_u)\big(\mathcal{H}(u) - \frac{n-1}{u}\big) \leq 0.$$

Now because the function $u \mapsto u^{1-n}\text{area}_{g_\infty}(\Sigma_u)$ is monotone decreasing, clearly

$$\liminf_{u \to 0} \frac{\text{area}_{g_\infty}(\Sigma_u)}{u^{n-1}} > 0.$$

This implies

$$\limsup_{u\to 0}\left(u^2\sup_{q\in\Sigma_u}\operatorname{scal}_{g_\infty}(q)\right)<\infty.$$

To summarize, $u(s)^2\operatorname{scal}_{g_\infty}(\gamma_\infty(s))$, and hence $u(s)H_\infty(\gamma_\infty(s))$, are bounded above and below as $s\to\rho_*$. Moreover, the ratio $u^{1-n}\operatorname{area}_{g_\infty}(\Sigma_u)$ is bounded and monotone, and therefore

$$\lim_{u\to 0}\frac{\operatorname{area}_{g_\infty}(\Sigma_u)}{u^{n-1}}=\kappa$$

for some positive constant $\kappa\in(0,C(n)\varepsilon^{n-1})$. In other words, the geometry of our local limit is asymptotically conical as $u\to 0$. For later use, let us choose a positive constant $L$ such that

$$u\sup_{q\in\Sigma_u}H_\infty(q)\le L$$

for $u\in(0,u_*]$.

*Step 3.5:* Now we can find a suitable sequence of rescalings of our original flow which converge to a local flow for which the final time slice is a metric cone. This will contradict the strong maximum principle. Here are the details. Choose a sequence of distances $s_\ell\in[s_*,\rho_*)$ with $s_\ell\to\rho_*$. Let

$$u_\ell:=u(s_\ell),\qquad \hat{x}_\ell:=\gamma_\infty(s_\ell)\in\Sigma_{u_\ell}.$$

Then $u_\ell\to 0$. After passing to a subsequence, we can assume

$$u_\ell^{-1}(\rho_*-s_\ell)\to\hat{\rho}$$

where $\hat{\rho}$ is a constant which is close to $1$. For $j$ sufficiently large, let

$$x_{\ell,j}:=\gamma_j(Q_j^{-1}s_\ell)\in B_{g(t_j)}(p_j,Q_j^{-1}\rho_*)$$

so that $x_{\ell,j}\to\hat{x}_\ell$ as $j\to\infty$. Define

$$R_{\ell,j}:=u_\ell Q_j^{-1}.$$

Recall $\tilde{g}_j = Q_j^2 g(t_j)$ and the metrics $\tilde{g}_j$ converge to $g_\infty$. Now consider the rescaled metrics $\hat{g}_{\ell,j} := R_{\ell,j}^{-2} g(t_j) = u_\ell^{-2}\tilde{g}_j$ on the metric balls

$$\hat{B}_{\ell,j} := B_{\hat{g}_{\ell,j}}\Big(x_{\ell,j}, \frac{\rho_* - s_\ell}{2u_\ell}\Big) = B_{\tilde{g}_j}\Big(x_{\ell,j}, \frac{\rho_* - s_\ell}{2}\Big) \subset B_{\tilde{g}_j}(p_j, \rho_*).$$

By our work in Step 3.1, $(\hat{B}_{\ell,j}, \hat{g}_{\ell,j})$ converges locally smoothly to $\big(B_{g_\infty}\big(\hat{x}_\ell, \frac{\rho_*-s_\ell}{2}\big), u_\ell^{-2}g_\infty\big)$ for each fixed $\ell$ as we let $j \to \infty$. By our work in Step 3.3, if $j$ is large depending upon $\ell$, then every point $x \in \hat{B}_{\ell,j}$ lies at the center of a $C(n)\varepsilon$-neck and therefore on a canonical CMC sphere. So $\hat{B}_{\ell,j}$ is contained a tube which is canonically foliated by CMC spheres. Because the foliation by CMC spheres is uniquely determined by the metrics (via the inverse function theorem) and because the metrics $\tilde{g}_j = Q_j^2 g(t_j)$ converge locally smoothly to $g_\infty$, each CMC sphere contained in $\hat{B}_{\ell,j}$ converges to a unique CMC sphere in $B_{g_\infty}\big(\hat{x}_\ell, \frac{\rho_*-s_\ell}{2}\big)$ as we let $j \to \infty$. It follows from comparability of $u(s)$ and $\rho_* - s$ that $B_{g_\infty}\big(\hat{x}_\ell, \frac{\rho_*-s_\ell}{2}\big) \subset \bigcup_{u\in(\frac{1}{3}u_\ell, \frac{5}{3}u_\ell)} \Sigma_u$ and conversely, if $\frac{2}{3}u_\ell < u < \frac{4}{3}u_\ell$, then $\Sigma_u \subset B_{g_\infty}\big(\hat{x}_\ell, \frac{\rho_*-s_\ell}{2}\big)$.

Let $\hat{\Sigma}_{\hat{u}}^{(\ell,j)}$ denote the one-parameter family of CMC spheres that foliate $\hat{B}_{\ell,j}$. After a translation, choice of sign, and suitable reparametrization (so that the lapse function $\hat{v}^{(\ell,j)}$ has average 1 on each CMC sphere) the foliation is defined at least for $\hat{u} \in (\frac{1}{3}, \frac{5}{3})$ with $\hat{\Sigma}_{\hat{u}}^{(\ell,j)} \subset \hat{B}_{\ell,j}$ if $\hat{u} \in (\frac{2}{3}, \frac{4}{3})$ and $\hat{B}_{\ell,j} \subset \bigcup_{\hat{u}\in(\frac{1}{3},\frac{5}{3})} \Sigma_{\hat{u}}^{(\ell,j)}$. In particular, these choices determine the parameter $\hat{u}$ and consequently $\hat{\Sigma}_{\hat{u}}^{(\ell,j)}$ converges to $\Sigma_{u_\ell \hat{u}}$ as $j \to \infty$ for $\hat{u} \in (\frac{1}{3}, \frac{5}{3})$. Therefore

$$\begin{aligned}
&\lim_{j\to\infty} \frac{\mathrm{area}_{\hat{g}_{\ell,j}}(\hat{\Sigma}_u^{(\ell,j)})}{\hat{u}^{n-1}} = \frac{\mathrm{area}_{u_\ell^{-2}g_\infty}(\Sigma_{u_\ell\hat{u}})}{\hat{u}^{n-1}} = \frac{\mathrm{area}_{g_\infty}(\Sigma_{u_\ell\hat{u}})}{(u_\ell\hat{u})^{n-1}},\\
&\lim_{\ell\to\infty} \frac{\mathrm{area}_{g_\infty}(\Sigma_{u_\ell\hat{u}})}{(u_\ell\hat{u})^{n-1}} = \kappa.
\end{aligned}$$

Set $\tau = -\frac{1}{288\eta L^2}$. Finally, consider the sequence of flows $\hat{F}_{\ell,j} : \hat{B}_{\ell,j} \times (\tau, 0] \to \mathbb{R}^N$ defined by

$$\hat{F}_{\ell,j}(p,t) = R_{\ell,j}^{-1}\big(F(p, t_j + R_{\ell,j}^2 t) - F(x_\ell, t_j)\big).$$

For each $p \in \hat{B}_{\ell,j}$, we can find $\hat{u} \in (\frac{1}{3}, \frac{5}{3})$ with $p \in \Sigma_{\hat{u}}^{(\ell,j)}$. Assuming that $j$ is large enough depending upon $\ell$, it follows from our work in Step 3.4 that

$$Q_j^{-1} H(p, t_j) \le 2L u_\ell^{-1} \hat{u}^{-1} \le 6L u_\ell^{-1}.$$

For $t \in (\tau, 0]$, this implies that

$$R_{\ell,j}^2(-t) < R_{\ell,j}^2(-\tau) = \frac{1}{288\eta} u_\ell^2 Q_j^{-2} L^{-2} \le \frac{1}{8\eta H(p, t_j)^2}.$$

So we may use the short-range curvature estimates in Step 2 to obtain

$$R_{\ell,j} H(p, t_j + R_{\ell,j}^2 t) \le 4 R_{\ell,j} H(p, t_j) \le 24L$$

for $(p, t) \in \hat{B}_{\ell,j} \times (\tau, 0]$. The estimate above implies we have uniform bounds for the second fundamental forms of the flows $\hat{F}_{\ell,j}$ on the domains $\hat{B}_{\ell,j} \times (\tau, 0]$ assuming $j$ is sufficiently large depending upon $\ell$. For a suitable diagonal subsequence, we can now apply Corollary 3.4 to obtain a locally defined solution of mean curvature flow $\hat{F}_\infty$ in $\mathbb{R}^{n+1}$ which is defined on a parabolic neighborhood $\hat{B}_\infty \times (\tau, 0]$, where $\hat{B}_\infty = B_{\hat{g}_\infty}(\hat{x}_\infty, \frac{1}{2}\hat{\rho})$. If along the subsequence $j$ is taken sufficiently large for each $\ell$, then the CMC foliations of $\hat{B}_{\ell,j}$ converge to a CMC foliation $\hat{\Sigma}_{\hat{u}}$ of suitable domain $\hat{U}$ in $\hat{B}_\infty$ with the property that

$$\frac{\mathrm{area}_{\hat{g}_\infty}(\hat{\Sigma}_{\hat{u}})}{\hat{u}^{n-1}} = \kappa$$

independent of $\hat{u}$.

Let $\hat{v}$ and $\hat{\nu}$ denote the lapse function and inward-pointing unit normal of the foliation $\hat{\Sigma}_{\hat{u}}$. If $\hat{\mathcal{H}}(\hat{u})$ and $\hat{\mathcal{A}}$ denote the the mean curvature and second fundamental form of $\hat{\Sigma}_{\hat{u}}$ with respect to $\hat{g}_\infty$, then the identities in Step 3.4 imply $\hat{\mathcal{H}}(\hat{u}) = \frac{n-1}{\hat{u}}$, $-\hat{\mathcal{H}}'(\hat{u}) = \frac{1}{n-1}\hat{\mathcal{H}}(\hat{u})^2$, $|\hat{\mathcal{A}}|^2 = \frac{1}{n-1}\hat{\mathcal{H}}(\hat{u})^2$, and $\mathrm{Ric}_{\hat{g}_\infty}(\hat{\nu}, \hat{\nu}) = 0$. In particular, $\Delta_{\hat{\Sigma}_{\hat{u}}}(\hat{v} - 1) + \frac{n-1}{\hat{u}^2}(\hat{v} - 1) = 0$. The sphere $\hat{\Sigma}_{\hat{u}}$ is $C(n)\varepsilon$-close to a round sphere of radius $\lesssim \varepsilon \hat{u}$. This implies the operator $\Delta_{\hat{\Sigma}_{\hat{u}}} + \frac{n-1}{\hat{u}^2}$ is a small perturbation of

$\Delta_{S^{n-1}}$, after a suitable rescaling. Because the average of $\hat{v} - 1$ is zero, this in turn implies $\hat{v} \equiv 1$, and hence $\hat{U}$ must be a piece of metric cone. Note, the estimate $\hat{u}^{1-n}\text{area}_{\hat{g}_\infty}(\hat{\Sigma}_{\hat{u}}) \leq C(n)\varepsilon^{n-1}$ implies opening angle of the cone must be very small.

Since the solution $\hat{F}_\infty$ is codimension 1, weakly convex, and we have $\text{Ric}_{\hat{g}_\infty}(\hat{\nu}, \hat{\nu}) = 0$, the Gauss equation implies that the first eigenvalue of the second fundamental form of $F_\infty$ must vanish at time $t = 0$. By the strong maximum principle, this implies the solution locally splits a line. In particular, the metric induced by the immersion is locally a product, which is not compatible with the conclusion that geometry at time zero is conical. This gives us the desired contradiction (see, for example, Appendix A in [29]). We conclude

$$\mathbb{M}(\rho) := \limsup_{j\to\infty} \sup_{p\in B_{g(t_j)}(p_j, Q_j^{-1}\rho)} Q_j^{-1} H(p, t_j) < \infty$$

for each $\rho > 0$. This completes the proof of the long range curvature estimate.

*Step 4:* Using the previous steps, we extract a complete limit and show it must have bounded curvature. By the previous step, for every $\rho \in (0, \infty)$ and every integer $k = 0, 1, \ldots$, there exists a positive constant $C(n, k, \eta, \rho)$ such that $Q_j^{-k-1}|\nabla^k A|(p, t_j) \leq C(n, k, \eta, \rho)$ if $d_{g(t_j)}(p, p_j) < \rho Q_j^{-1}$. Hence for each of the rescaled pointed isometric immersions $\tilde{F}_j : (M, \tilde{g}_j, p_j) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$ in Step 3.1, we have uniform estimates for the second fundamental form and each of its derivatives at bounded distance. After passing to a subsequence, by Proposition 3.2 we have

- the sequence $(M, \tilde{g}_j, p_j)$ converges smoothly in the pointed Cheeger-Gromov sense to a complete Riemannian manifold $(M_\infty, g_\infty, p_\infty)$;
- the sequence of immersions $\tilde{F}_j$ converges smoothly to a pointed isometric immersion $F_\infty : (M_\infty, g_\infty, p_\infty) \to (\mathbb{R}^N, g_{\text{flat}}, 0)$.

In the limit, we have the estimate $|A_\infty|^2 \leq \frac{1}{n-1}H_\infty^2$. The codimension estimate implies that $F_\infty(M_\infty)$ is contained in an $(n+1)$-dimensional affine subspace of $\mathbb{R}^N$. In codimension one, the estimate $|A_\infty|^2 \leq \frac{1}{n-1}H_\infty^2$ implies the limit has nonnegative sectional curvature.

By condition (iii) and Proposition 4.4, either $M_\infty$ is a closed manifold diffeomorphic to $S^n$ or every point $p \in M_\infty$ where $H_\infty(p) \geq 8$ has a canonical neighborhood which is either $2\varepsilon$-neck or a $2\varepsilon$-cap. Note that if a point $p \in M_\infty$ lies on a $2\varepsilon$-cap, then the estimates in Proposition 4.4 imply the cap is incident to a $2\varepsilon$-neck of radius bounded by $C(n)C_2 H_\infty(p)^{-1}$. Thus, if the curvature of $(M_\infty, g_\infty)$ is unbounded, then the limit must contain a sequence of $2\varepsilon$-necks of radii tending to zero. However, by an argument of Perelman, this is impossible in a complete Riemannian manifold with nonnegative sectional curvature. See Proposition 2.2 in [20] for a detailed version of the argument. Consequently, $(M_\infty, g_\infty)$ has bounded curvature.

*Step 5:* By the previous step, the mean curvature of the immersion $F_\infty$ is bounded from above by a constant $\Lambda > 8$. Since the sequence of immersions $\tilde{F}_j$ converges locally smoothly to $F_\infty$, for every $\rho > 1$, we have

$$\limsup_{j\to\infty} \sup_{p \in B_{g(t_j)}(p_j, \rho Q_j^{-1})} Q_j^{-1} H(p, t_j) \leq 2\Lambda.$$

The short range curvature estimates in Step 2 then imply that for every $\rho > 1$,

$$\limsup_{j\to\infty} \sup_{(p,t) \in \hat{P}(p_j, t_j, \rho, \frac{1}{32\eta\Lambda^2})} Q_j^{-1} H(p, t) \leq 8\Lambda,$$

where recall $\hat{P}(p_j, t_j, \rho, \frac{1}{32\eta\Lambda^2}) = B_{g(t_j)}(p_j, \rho Q_j^{-1}) \times [t_j - \frac{1}{32\eta\Lambda^2} Q_j^{-2}, t_j]$. By a minor abuse of notation, let $\tilde{F}_j : M \times [-Q_j^2 t_j, 0] \to \mathbb{R}^N$ denote the flow

$$\tilde{F}_j(p, t) = Q_j\big(F(p, t_j + Q_j^{-2} t) - F(p_j, t_j)\big).$$

If we take $\tau_1 := -\frac{1}{64\eta\Lambda^2}$ and $j$ sufficiently large, then for every $\rho > 1$, we have uniform estimates for the second fundamental form of $\tilde{F}_j$ on the parabolic neighborhood $P(p_j, 0, \rho, \tau_1)$. Note these uniform estimates are independent of $\rho$. By Corollary 3.4, a subsequence of these flows converge to a complete solution of mean curvature flow $F_\infty(\cdot, t)$ defined for $t \in [\tau_1, 0]$ with $F_\infty(\cdot, 0) = F_\infty$.

Moreover, the limiting solution satisfies

$$\Lambda_1 := \sup_{(p,t)\in M_\infty\times[\tau_1,0]} H(p,t) \leq 8\Lambda.$$

Now set $\tau_2 := \tau_1 - \frac{1}{64\eta\Lambda_1^2}$. Using the short range curvature estimates once more and passing to a further subsequence, we can extend the solution $F_\infty(\cdot,t)$ to the interval $t \in [\tau_2, 0]$ and the solution will satisfy $\Lambda_2 := \sup_{(p,t)\in M_\infty\times[\tau_2,0]} H(p,t) \leq 8\Lambda_1$. Analogously, for each $m \geq 1$, with $\tau_{m+1} = \tau_m - \frac{1}{64\eta\Lambda_m^2}$ and $\Lambda_{m+1} := \sup_{(p,t)\in M_\infty\times[\tau_{m+1},0]} H(p,t) \leq 2\Lambda_m$, we get a complete solution of mean curvature flow $F_\infty : M_\infty \times [\tau_m, 0] \to \mathbb{R}^N$.

Let $\tau^* := \lim_{m\to\infty} \tau_m$. Taking the limit along a suitable diagonal sequence of the flows $\tilde{F}_j$, we obtain a complete solution of mean curvature flow $F_\infty : M_\infty \times (\tau^*, 0] \to \mathbb{R}^N$. By construction, the solution has bounded mean curvature for each $t \in (\tau^*, 0]$. The solution satisfies the estimates $|\nabla A_\infty| \leq \eta H_\infty^2$, $|\nabla^2 A| \leq \eta H_\infty^3$, and $|A_\infty|^2 \leq \frac{1}{n-1} H_\infty^2$. By the planarity estimate, the solution is contained in an affine $(n+1)$-dimensional subspace of $\mathbb{R}^N$, so without loss of generality we may write $F_\infty : M_\infty \times (\tau^*, 0] \to \mathbb{R}^{n+1}$. Finally, by property (iii) together with Proposition 4.4, any spacetime point $(p,t) \in M_\infty \times (\tau^*, 0]$ where $H_\infty(p,t) \geq 8$ has a canonical neighborhood which is either a $2\varepsilon$-neck, a $2\varepsilon$-cap, or a closed manifold diffeomorphic to $S^n$.

*Step 6:* In this final step, we show that $\tau^* = -\infty$ using Hamilton's Harnack inequality for mean curvature flow. Suppose $\tau^* > -\infty$. This implies $\lim_{m\to\infty}(\tau_{m+1} - \tau_m) \to 0$ and consequently $\lim_{m\to\infty} \Lambda_m = \infty$. Thus, as we go backward in time to $\tau^*$, the mean curvature of $F_\infty(\cdot,t)$ blows up.

By the Harnack inequality, $(t-\tau^*)^{\frac{1}{2}} H_\infty(p,t)$ is nondecreasing for each $p \in M_\infty$. Since $H_\infty(p,0) \leq \Lambda$, this gives

$$H_\infty(p,t) \leq \sqrt{\frac{-\tau^*}{t-\tau^*}}\Lambda$$

for all $t \in (\tau^*, 0]$ and $p \in M_\infty$. Applying Lemma 4.6, we get

$$0 \le -\frac{d}{dt} d_{g_\infty(t)}(p,q) \le C(n) \sqrt{\frac{-\tau^*}{t-\tau^*}} \Lambda$$

for all $t \in (\tau^*, 0]$ and $p, q \in M_\infty$. The key point is that the right hand side of the inequality above is integrable in $t$. Integrating this inequality gives

$$d_{g_\infty(0)}(p,q) \le d_{g_\infty(t)}(p,q) \le d_{g_\infty(0)}(p,q) + C(n)(-\tau^*)\Lambda$$

for all times $t \in (\tau^*, 0]$ and all points $p, q \in M_\infty$.

By our rescaling procedure, clearly $H_\infty(p_\infty, 0) = 1$. The maximal principle implies the infimum of the mean curvature is nondecreasing and hence

$$\inf_{p \in M_\infty} H_\infty(p,t) \le \inf_{p \in M_\infty} H_\infty(p,0) \le 1$$

for all $t \in (\tau^*, 0]$. It follows that we can find a point $q_\infty \in M_\infty$ where $H_\infty(q_\infty, t) \le 2$ for $t = \tau^* + \frac{1}{64\eta}$. By the short range curvature estimates of Step 2, $H_\infty(q_\infty, t) \le 8$ for $t \in (\tau^*, \tau^* + \frac{1}{64\eta}]$. In particular, if $m$ is sufficiently large, this implies $H_\infty(q_\infty, \tau_m) \le 8$. We claim

$$\limsup_{m \to \infty} \sup_{p \in B_{g_\infty(\tau_m)}(q_\infty, \rho)} H_\infty(p, \tau_m) < \infty$$

for every $\rho > 1$. This follows from the long range curvature estimate proven in Step 3. For the argument in Step 3 to work, we need only the pointwise derivative estimates and condition (iii). Both of these properties are satisfied by the limit. Since the mean curvature is bounded at bounded distance, the sequence of immersions $\hat{F}_m : (M_\infty, g_\infty(\tau_m), q_\infty) \to (\mathbb{R}^{n+1}, g_{\text{flat}}, 0)$, where $\hat{F}_m(p) = F_\infty(p, \tau_m) - F_\infty(q_\infty, \tau_m)$, subsequentially converges to a smooth limit. By the argument in Step 4, this limit has bounded curvature. It follows that there exists a constant $\Lambda_* > \Lambda$, independent of $\rho$, such that

$$\liminf_{m\to\infty} \sup_{p\in B_{g_\infty(\tau_m)}(q_\infty,\rho)} H_\infty(p,\tau_m) \leq \Lambda_*,$$

for every $\rho > 1$. The geodesic balls $B_{g_\infty(\tau_m)}(q_\infty,\rho)$ may change in size as $\tau_m \to \tau^*$. By our distance estimate, however, if $\rho$ is sufficiently large, then

$$B_{g_\infty(0)}(q_\infty, \rho - C(n)(-\tau^*)\Lambda) \subset B_{g_\infty(\tau_m)}(q_\infty,\rho).$$

Consequently,

$$\liminf_{m\to\infty} \sup_{p\in B_{g_\infty(0)}(q_\infty,\rho)} H_\infty(p,\tau_m) \leq \Lambda_*,$$

for every $\rho > 1$.

In summary, for every $\rho > 1$, we can find a large integer $m$ such that

$$\sup_{p\in B_{g_\infty(0)}(q_\infty,\rho)} H_\infty(p,\tau_m) \leq 2\Lambda_*.$$

By the short range curvature estimates (both forwards and backwards in time), taking $m$ sufficiently large, this implies

$$\sup_{t\in(\tau^*,\tau^*+\frac{1}{16\eta\Lambda_*^2}]} \sup_{p\in B_{g_\infty(0)}(q_\infty,\rho)} H_\infty(p,t) \leq 8\Lambda_*.$$

Since $\Lambda^*$ is independent of $\rho$, this gives

$$\sup_{t\in(\tau^*,\tau^*+\frac{1}{16\eta\Lambda_*^2}]} \sup_{p\in M_\infty} H_\infty(p,t) \leq 8\Lambda_*.$$

This contradicts our observation that $\lim_{m\to\infty}\Lambda_m = \infty$. Therefore, we must have $\tau^* = -\infty$. From the derivative estimates, the codimension estimate, the cylindrical estimate, and the main result of Chapter 2, the solution $F_\infty : M_\infty \times (-\infty, 0] \to \mathbb{R}^N$ is an ancient model solution.

In conclusion, a subsequence of flows obtained by rescaling the solution $F$ around the points

$(p_j, t_j)$ by $H(p_j, t_j)$ must converge to an ancient model solution. This, of course, contradicts property (ii), and thereby completes the proof of the theorem. □

# Bibliography


[1] *S. J. Altschuler*, Singularities of the curve shrinking flow for space curves, J. Differential Geom. **34** (1991), 491–514.

[2] *B. Andrews* and *C. Baker*, Mean curvature flow of pinched submanifolds to spheres, J. Differential Geom. **85** (2010), 357–395.

[3] *S. Angenent*, *P. Daskalopoulos*, and *N. Šešum*, Unique asymptotics of ancient convex mean curvature flow solutions, J. Differential Geom. **111** (2019), 381–455.

[4] *S. Angenent*, *P. Daskalopoulos*, and *N. Šešum*, Uniqueness of two-convex closed ancient solutions to the mean curvature flow Ann. of Math. (2) **192** (2020), 353–436.

[5] *T. Bourni* and *M. Langford*, Type-II singularities of two-convex immersed mean curvature flow, Geom. Flows **2** (2017), 1–17.

[6] *S. Brendle*, A general convergence result for the Ricci flow in higher dimensions, Duke Math. J. **145** (2008), 585–601.

[7] *S. Brendle*, Ricci flow and the sphere theorem, Grad. Stud. Math. **111** American Mathematical Society, Providence, RI 2010.

[8] *S. Brendle*, Rotational symmetry of self-similar solutions to the Ricci flow, Invent. Math. **194** (2013), 731–764.

[9] *S. Brendle*, Ricci flow with surgery in higher dimensions, Ann. of Math. (2) **187** (2018), 263–299.

[10] *S. Brendle*, Ricci flow with surgery on manifolds with positive isotropic curvature, Ann. of Math. (2) **190** (2019), 465–559.

[11] *S. Brendle*, Ancient solutions to the Ricci flow in dimension 3, Acta Math. **225** (2020), 1–102.

[12] *S. Brendle* and *K. Choi*, Uniqueness of convex ancient solutions to the mean curvature flow in $\mathbb{R}^3$, Invent. Math. **217** (2019), 35–76.

[13] *S. Brendle* and *K. Choi*, Uniqueness of convex ancient solutions to mean curvature flow in higher dimensions, Geom. Topol. To appear. (2021)

[14] *S. Brendle* and *G. Huisken*, Mean curvature flow with surgery of mean convex surfaces in $\mathbb{R}^3$, Invent. Math. **203** (2016), 615–654.

[15] *S. Brendle* and *P.-K. Hung*, A sharp inscribed radius estimate for fully nonlinear flows, Amer. J. Math. **141** (2019), 41–53.

[16] *S. Brendle, S*, *P. Daskalopoulos*, and *N. Šešum*, Uniqueness of compact ancient solutions to three-dimensional Ricci flow, Preprint, prepring 2020 `http://arxiv.org/abs/2002.12240`.

[17] *P. Breuning* Immersions with bounded second fundamental form, J. Geom. Anal. **25** (2015), 1344–1386.

[18] *H.-D. Cao* and *X.-P. Zhu*, A complete proof of the Poincaré and geometrization conjectures - application of the Hamilton-Perelman theory of the Ricci flow, Asian J. Math **10** (2006), 165–492.

[19] *J. Cheeger*, *M. Gromov*, and *M. Taylor*, Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds, J. Differenital Geom. **17** (1982), 15–53.

[20] *B.-L. Chen* and *X.-P. Zhu*, Ricci flow with surgery on four-manifolds with positive isotropic curvature, J. Differential Geom. **74** (2006),177–264.

[21] *T. H. Colding* and *W. P. Minicozzi II*, Complexity of parabolic systems, Publ. Math. IHES. **132** (2020), 83–135.

[22] *K. Ecker* and *G. Huisken*. Interior estimates for hypersurfaces moving by mean curvature, Invent. Math. **105** (1991), 547–569.

[23] *R. S. Hamilton*, Three-manifolds with positive Ricci curvature, J. Differential Geom. **17** (1982), 255–306.

[24] *R. S. Hamilton*, Convex hypersurfaces with pinched second fundamental form, Comm. Anal. Geom. **2** (1994), 167–172.

[25] *R. S. Hamilton*, A compactness property for solutions of the Ricci flow, Amer. J. Math. **117** (1995), 454–572.

[26] *R. S. Hamilton*, Harnack estimate for the mean curvature flow, J. Differential Geom. **41** (1995), 215–226.

[27] *R. S. Hamilton*, Four-manifolds with positive isotropic curvature, Comm. Anal. Geom. **5** (1997), 1–92.

[28] *R. Haslhofer*, Uniqueness of the bowl soliton, Geom. Topol. **19** (2015), 2393–2406.

[29] *R. Haslhofer* and *B. Kleiner*, Mean curvature flow of mean convex hypersurfaces, Comm. Pure Appl. Math. **70** (2017), 511–546.

[30] *R. Haslhofer* and *B. Kleiner*, Mean curvature flow with surgery, Duke Math. J. **166** (2017), 1591–1626.

[31] *G. Huisken*, Flow by mean curvature of convex surfaces into spheres, J. Differential Geom. **20** (1984), 237–266.

[32] *G. Huisken* and *C. Sinestrari*, Convexity estimates for mean curvature flow and singularities of mean convex surfaces, Acta Math. **183** (1999), 45–70.

[33] *G. Huisken* and *C. Sinestrari*, Mean curvature flow with surgeries of two-convex hypersurfaces, Invent. Math. **175** (2009), 137–221.

[34] *G. Huisken* and *C. Sinestrari*, Convex ancient solutions of the mean curvature flow, J. Differential Geom. **101** (2015), 267–287.

[35] *J. Langer*, A compactness theorem for surfaces with $l_p$-bounded second fundamental form, Math. Ann. **270** (1985), 223–234.

[36] *A-M. Li* and *J. Li*, An intrinsic rigidity theorem for minimal submanifolds in a sphere, Arch. Math. (Basel) **58** (1992), 582–594.

[37] *M. J. Micallef* and *J. D. Moore*, Minimal two-spheres and the topology of manifolds with positive curvature on totally isotropic two-planes, Ann. of Math. (2) **127** (1988), 199–227.

[38] *K. Naff*, A planarity estimate for pinched solutions of mean curvature flow, Duke Math. J. To appear. (2021)

[39] *K. Naff*, Singularity models of pinched solutions of the mean curvature flow in higher codimension, Preprint, preprint 2019, `http://arxiv.org/abs/1910.03968`.

[40] *K. Naff*, A canonical neighborhood theorem for the mean curvature in higher codimension, Preprint, preprint 2020, `http://arxiv.org/abs/2004.08060`.

[41] *H. T. Nguyen*, Cylindrical estimates for high codimension mean curvature flow, Preprint, preprint 2018, `http://arxiv.org/abs/1805.11808`.

[42] *H. T. Nguyen*, High codimension mean curvature flow with surgery, Preprint, preprint 2020, `http://arxiv.org/abs/2004.07163`.

[43] *G. Perelman*, The entropy formula for the Ricci flow and its geometric applications, Preprint, preprint 2002, `http://arxiv.org/abs/math/0211159`.

[44] *G. Perelman*, Ricci flow with surgery in 3 dimensions, Preprint, preprint 2002, `http://arxiv.org/abs/math/0303109`.

[45] *R. Sacksteder*, On hypersurfaces with no negative sectional curvatures, Am. J. Math **92** (1960), 609–630.

[46] *K. Smoczyk*, Mean curvature flow in higher codimension: introduction and survey Global differential geometry, Spring Proc. Math. **17** Springer, Heidelberg (2012), 231–274.

[47] *W. Sheng* and *X.-J. Wang*, Singularity profile in the mean curvature flow, Methods Appl. Anal. **16** (2009), 139–155.

[48] *B. White*, The size of the singular set in the mean curvature flow of mean convex sets, J. Amer. Math. Soc. **13** (2000), 665–695.

[49] *B. White*, The nature of singularities in mean curvature flow of mean convex sets, J. Amer. Math. Soc. **16** (2003), 123–138.