JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 27, Number 3, July 2014, Pages 613–638
S 0894-0347(2014)00792-X
Article electronically published on March 19, 2014

UNIQUENESS OF SELF-SIMILAR SHRINKERS WITH ASYMPTOTICALLY CONICAL ENDS

LU WANG

1. Introduction

Self-shrinkers are a special class of solutions to the mean curvature flow in $\mathbb{R}^{n+1}$, in which a later time slice is a scaled down copy of an earlier slice. More precisely, a hypersurface Σ in $\mathbb{R}^{n+1}$ is said to be a self-shrinker if it satisfies

$$H = \frac{1}{2}\langle x, \mathbf{n}\rangle. \tag{1.1}$$

Here $H = \operatorname{div}(\mathbf{n})$ is the mean curvature, $\mathbf{n}$ is the outward unit normal, x is the position vector and $\langle , \rangle$ denotes the Euclidean inner product. One reason that self-shrinking solutions to the mean curvature flow are particularly interesting is that they provide singularity models of the flow; see [20, 21], [24] and [48].

It is a fundamental problem to classify smooth complete embedded self-shrinkers in $\mathbb{R}^{n+1}$. And the new families of examples constructed by Kapouleas, Kleene and Møller [28], Nguyen [38–40], and Møller [36], as well as the numerical examples produced by Chopp [4], indicate that it is very difficult to establish a full classification. However, under certain conditions, several results in this direction have been obtained since the 1980s. For $n = 1$, Abresch and Langer [1] had already shown that the circle is the only simple closed self-shrinking curve. For higher dimensions, Colding and Minicozzi [5], proved that the only smooth complete embedded self-shrinkers in $\mathbb{R}^{n+1}$ that are mean convex and have polynomial volume growth are $S^k(\sqrt{2k}) \times \mathbb{R}^{n-k}$, $0 \le k \le n$, which generalized an earlier result of Huisken [20, 21]. Here $S^k(\sqrt{2k})$ denotes the k-dimensional round sphere centered at the origin with radius $\sqrt{2k}$. Moreover, in the same paper, they showed that those self-shrinkers are the only entropy stable ones under the mean curvature flow. Also, in [47], we established a Bernstein type theorem for smooth entirely graphical self-shrinkers in $\mathbb{R}^{n+1}$, which generalized a result of Ecker and Huisken [11]. Besides, by imposing symmetries, Kleene and Møller [29] classified all smooth complete embedded self-shrinking hypersurfaces of revolution in $\mathbb{R}^{n+1}$.

Our specific interest is in smooth noncompact embedded self-shrinkers. Excluding products, to the author's knowledge, the only smooth complete noncompact properly embedded self-shrinkers in the literature belong to the family of examples in [28] and [38–40], each of which has a conical structure at infinity. In fact, it is a

Received by the editors October 3, 2011 and, in revised form, June 27, 2013, and October 19, 2013.

2010 *Mathematics Subject Classification.* Primary 53C44, 53C24, 35J15; Secondary 35B60.

Key words and phrases. Self-shrinkers, mean curvature flow, backwards uniqueness.

©2014 American Mathematical Society

well-known conjecture (see page 39 of [23]) that

Conjecture 1.1. *Let $\Sigma \subset \mathbb{R}^3$ be a smooth complete properly embedded self-shrinker of finite genus.*[1] *Then Σ outside a sufficiently large ball decomposes into a finite number of ends U_j and for each j, either*

(a) *As $\lambda \longrightarrow 0+$, λU_j converges locally smoothly to a cone C_j such that C_j is smooth except at its vertex.*
(b) *There is a vector v_j such that, as $\tau \longrightarrow +\infty$, $U_j - \tau v_j$ converges to the cylinder $\left\{x : \operatorname{dist}(x, \operatorname{span}(v_j)) = \sqrt{2}\right\}$.*

Moreover, there is another conjecture on the rigidity of self-shrinking cylinders (see page 39 of [23]), which asserts that

Conjecture 1.2. *Let $\Sigma \subset \mathbb{R}^3$ be a smooth complete properly embedded self-shrinker of finite genus. If one end of Σ is asymptotic to a cylinder (see* (b) *of Conjecture* 1.1*), then Σ is isometric to the self-shrinking cylinder.*

It is the previously mentioned noncompact examples and Conjectures 1.1 and 1.2 that motivate our investigation of the rigidity of asymptotically conical structures. We approach this by proving the uniqueness for smooth properly embedded self-shrinking ends asymptotic to a given cone.

We first fix some notations and conventions to present the main results of the paper. Throughout, O is the origin of $\mathbb{R}^{n+1}$; B_R denotes the open ball in $\mathbb{R}^{n+1}$ centered at O with radius R and $S_R = \partial B_R$; D_R denotes the open disk in $\mathbb{R}^n \times \{0\}$ centered at the origin with radius R. We say that $C \subset \mathbb{R}^{n+1}$ is a regular cone with vertex at O, if

$$C = \{l\Gamma,\ 0 \le l < \infty\}, \tag{1.2}$$

where Γ is a smooth connected closed (compact without boundary) embedded submanifold of S_1 with codimension one. Note that the normal component of the position vector on C vanishes and $C \backslash \{O\}$ is smooth.

The main theorem of this paper states that

Theorem 1.3. *Let $C \subset \mathbb{R}^{n+1}$ be a regular cone with vertex at O and R_0 a positive constant. Suppose that Σ and $\tilde{\Sigma}$ are smooth, connected, properly embedded self-shrinkers in $\mathbb{R}^{n+1} \backslash B_{R_0}$ with their boundaries in S_{R_0}. If Σ and $\tilde{\Sigma}$ are asymptotic to the same cone C, i.e., $\lambda\Sigma$ and $\lambda\tilde{\Sigma}$ converge to C locally smoothly as $\lambda \longrightarrow 0+$*[2]*, then Σ coincides with $\tilde{\Sigma}$.*

There are several interesting applications of our uniqueness theorem. For instance, by Theorem 1.3, it is not hard to show that not every regular cone with vertex at O has a smooth complete properly embedded self-shrinker asymptotic to it. Consider a rotationally symmetric regular cone $C \subset \mathbb{R}^{n+1}$ with vertex at O. In [29], Kleene and Møller constructed a smooth embedded self-shrinking end of revolution in $\mathbb{R}^{n+1}$ that is asymptotic to C. Thus, Theorem 1.3 implies that any smooth, connected, properly embedded self-shrinker that is asymptotic to C must

[1] These assumptions on self-shrinkers in Conjectures 1.1 and 1.2 are implied from the content of Lecture 3 in [23] and the equivalence of Euclidean volume growth and properness of self-shrinkers demonstrated in [9] and [3].

[2] More precisely, we mean that, for $\forall R > 0$ and $k \in \mathbb{N}$, $\lambda\Sigma \cap \left(\bar{B}_R \backslash B_{1/R}\right)$ and $\lambda\tilde{\Sigma} \cap \left(\bar{B}_R \backslash B_{1/R}\right)$ converge to $C \cap \left(\bar{B}_R \backslash B_{1/R}\right)$ in the C^k topology, as $\lambda \longrightarrow 0+$.

have a rotational symmetry. Hence, using the classification result for smooth complete embedded self-shrinkers with a rotational symmetry (see [29]), we conclude that

Corollary 1.4. *Let $C \subset \mathbb{R}^{n+1}$ be a rotationally symmetric regular cone with vertex at O. Assume that C is not a hyperplane. Then there do not exist smooth complete properly embedded self-shrinkers in $\mathbb{R}^{n+1}$ with an end asymptotic to C.*

On the other hand, by Huisken's monotonicity formula (see [20,21]), self-shrinkers are hypersurfaces in the Euclidean space that are minimal with respect to the Gaussian conformally changed metric; see [2] and [5, 6]. It is well known that one can construct many smooth complete properly embedded minimal surfaces with asymptotically planar ends, e.g., Costa-Hoffman-Meeks minimal surfaces (see [8], [27] and [19]). However, in contrast to the minimal surfaces theory, it follows from Theorem 1.3 that the only smooth complete properly embedded self-shrinkers with ends asymptotic to hyperplanes are hyperplanes. This gives an explanation of the dramatic change of asymptotics of the noncompact ends occurring in the desingularization construction of self-shrinkers in [28] and [38–40].[3]

Also, assuming that Conjecture 1.1 is true, a gap theorem for self-shrinkers can be established using Theorem 1.3.

Corollary 1.5. *Let $\Sigma \subset \mathbb{R}^3$ be a smooth complete properly embedded self-shrinker. Assume that one end of Σ is asymptotic to a regular cone and Conjecture* 1.1 *is true. Then any smooth complete properly embedded self-shrinker in $\mathbb{R}^3$ that has sublinear growth of Hausdorff distance from Σ must coincide with Σ.*

We give an overview of the proof of Theorem 1.3 in the following. Broadly speaking, our strategy is a reduction of the problem in Theorem 1.3 to a backwards uniqueness problem for parabolic differential inequalities. First, outside some compact set $K \subset \mathbb{R}^{n+1}$, we write $\tilde{\Sigma}\backslash K$ as the graph of a function v, which is defined over Σ and vanishes at infinity with a certain rate; see Section 2. Next, we derive the differential equation for the function v (see (2.24)) which involves the Ornstein-Uhlenbeck type operator. One main difficulty comes from the fact that, when changing infinity to the origin, the resulting equation is elliptic but highly singular and degenerate at the origin; see the discussion in the beginning of Section 3. Besides, under our assumptions, v need not vanish to infinite order asymptotically. Thus, we cannot apply the strong unique continuation theorems in [17, 18], [25], [26] or [41] to conclude that v is identically zero. Instead, we introduce a new function w (see (3.4)), which is defined by a suitable scaling of v on a domain of the space-time. Then the decay of v in space is converted to that of w as time approaches zero. Thus it suffices to show the backwards uniqueness for the parabolic differential inequality for w (see (3.5)) on a noncompact incomplete domain.

The subtlety of our backwards uniqueness problem is that the values of w at the parabolic boundary of the domain are not controlled by the assumptions of Theorem 1.3. We overcome this issue by extending the Carleman inequalities of Escauriaza, Seregin and Šverák [15] to the special cases of mean curvature flow in question, in which the backwards uniqueness was proved for a class of linear perturbations

[3]In those papers, one starts with surfaces by desingularizing the intersection of a sphere and a plane, but the resulting self-similar surfaces are asymptotic to cones (not planes). This phenomenon does not happen in the desingularization construction of minimal surfaces.

of the standard heat equation in the Euclidean space with a ball removed. As applications, they settled a long-standing question concerning sufficient conditions for regularity of solutions for the Navier-Stokes equations in $\mathbb{R}^3$; see [15]. For more bibliographic information on the backwards uniqueness for parabolic equations, we recommend the readers refer to [13–16], [31–35], [37], and [42–45].

Our approach above can be embedded into an abstract scheme for general geometric flows. Namely, let us consider two geometric objects moving by the same flow of second order parabolic type. Assume that they have the same limit at some terminal time. One might expect to address that the two must coincide with each other for all times, using similar Carleman type estimates as in our cases of mean curvature flow. However, besides the usual gauge degeneracy of geometric evolution equations, which requires different techniques to deal with, there are possible serious obstacles to constructing weight functions with desired convexity in the Carleman inequalities. Such constructions for the mean curvature flow discussed in the current paper depend heavily on the asymptotically conical structure of the flow and the fact that the equation for the evolving distance function is similar to the Euclidean distance. Indeed, in [46], we proved a weaker but optimal uniqueness theorem for asymptotically cylindrical self-shrinkers, on which we had to impose the infinite order rate assumption (i.e., (1.4) in Theorem 1.1 of [46]).

2. Notations and auxiliary lemmas

In this section, we set up the notations for the rest of the paper and prove several auxiliary lemmas. In particular, we deduce from the assumption that Σ and $\tilde{\Sigma}$ are asymptotic to the same cone that, outside some compact set, $\tilde{\Sigma}$ can be written as the graph of a function v over Σ decaying with a certain rate, and we derive the differential equation for v.

Throughout, for any multi-index $\alpha = \left(\alpha_1, \ldots, \alpha_{|\alpha|}\right)$, $\partial_\alpha^{|\alpha|}$ denotes the $|\alpha|$-th order partial derivative with respect to the α_i-th, $1 \leq i \leq |\alpha|$, coordinates of the Euclidean space; the meaning of ∂_t may vary in lemmas and propositions, and we will clarify it in the content; ∇, ∇^2 and Δ denote the gradient, Hessian and Laplacian on the hypersurfaces appearing in the subscripts, respectively; A and $\nabla^i A$ are the second fundamental form and its ith covariant derivative, respectively; constants in lemmas and propositions depend only on n, Σ, $\tilde{\Sigma}$ and C, unless it is specified; constants in proofs are not preserved when crossing lemmas and propositions.

Denote $\lambda\Sigma \cap \left(\bar{B}_R \backslash B_{1/R}\right)$ by $\Sigma_{R,\lambda}$ and $C \cap \left(\bar{B}_R \backslash B_{1/R}\right)$ by C_R. We recall that $\Sigma_{R,\lambda}$ converges to C_R in the C^k topology as $\lambda \longrightarrow 0+$, if $\Sigma_{R,\lambda}$ converges to C_R in the Hausdorff topology, and for any $x \in C_R$ and $\lambda > 0$ small, $\Sigma_{R,\lambda}$ locally (near x) is a graph over the tangent hyperplane $T_x C_R$ and the graph of $\lambda_{R,\lambda}$ converges to the graph of C_R in the usual C^k topology.

Since Σ is a self-shrinker under the mean curvature flow, $\left\{\Sigma_t = \sqrt{t}\ \Sigma\right\}_{t\in(0,1]}$ is a solution to the backwards mean curvature flow, which means that, parametrizing Σ_t in a suitable way, for $x \in \Sigma_t$ and $t \in (0, 1]$,

$$\partial_t x = H\mathbf{n}. \tag{2.1}$$

Here, $\partial_t x$ stands for the normal velocity of the hypersurface. We recommend the readers refer to Chapter 2 of [10] for more details and other equivalent definitions. We begin with the following elementary lemma on the geometry of Σ_t.

Lemma 2.1. *There exist $C_1 > 0$ and $R_1 \geq R_0$ such that for $x \in \Sigma_t \backslash B_{R_1}$, $t \in (0,1]$ and $0 \leq i \leq 2$,*

$$\left|\nabla^i A(x)\right| \leq C_1 \left|x\right|^{-i-1}. \tag{2.2}$$

Proof. By (1.2) and the fact that Γ is a smooth closed embedded submanifold of S_1, the second fundamental form and all its covariant derivatives of C are bounded inside the annulus $\bar{B}_2 \backslash B_{1/2}$. Furthermore, since $\lambda\Sigma \longrightarrow C$ locally smoothly as $\lambda \longrightarrow 0+$, there exist $\delta_1, \lambda_1 > 0$ such that if $0 < \lambda < \delta_1$, then for $0 \leq i \leq 2$, $\left|\nabla^i A\right| \leq \lambda_1$ on $\lambda\Sigma \cap \left(\bar{B}_2 \backslash B_{1/2}\right)$. Thus, for $x \in \Sigma$ with $|x| \geq R = 2\max\{R_0, 1/\delta_1\}$, we can choose $\lambda = 1/\left|x\right|$ such that for $0 \leq i \leq 2$,

$$\left|\nabla^i A(x)\right| = \lambda^{i+1} \left|\nabla^i A(\lambda x)\right| \leq \lambda_1 \left|x\right|^{-i-1}, \tag{2.3}$$

where we note that $\lambda x \in \lambda\Sigma$ is inside $\bar{B}_2 \backslash B_{1/2}$.

Since $\Sigma_t = \sqrt{t}\,\Sigma$, we conclude that for $x \in \Sigma_t \backslash B_R$ and $0 \leq i \leq 2$,

$$\left|\nabla^i A(x)\right| = t^{\frac{-i-1}{2}} \left|\nabla^i A\left(\frac{x}{\sqrt{t}}\right)\right| \leq \lambda_1 \left|x\right|^{-i-1}. \tag{2.4}$$

□

Next, it follows from the assumption of Theorem 1.3 that, outside a compact set, Σ_t is given by the graph of a smooth function over C. On the other hand, using the proof of Lemma 2.2 on page 30 of [7], we can write Σ_t locally as the graph of a smooth function over a fixed hyperplane.

Lemma 2.2. *There exist $R_2 > R_1$ and compact sets $K_t \subset \Sigma_t$, $0 < t \leq 1$, such that $K_t \subset \Sigma_t \cap B_{2R_2}$, and $\Sigma_t \backslash K_t$ is given by the graph of a smooth function $U(\cdot, t) : C \backslash \bar{B}_{R_2} \longrightarrow \mathbb{R}$. Moreover, there exist $0 < \varepsilon_0 < 1$ and $C_2 > 0$ such that for $z_0 \in C \backslash B_{2R_2}$ and $t \in (0,1]$, the component of $\Sigma_t \cap B_{\varepsilon_0 |z_0|}(z_0)$ containing $z_0 + U(z_0,t)\mathbf{n}(z_0)$ can be written as the graph of a smooth function $u(\cdot, t)$ over the tangent hyperplane $T_{z_0}C$ of C at z_0 satisfying that, for $i = 0, 1, 2$,*

$$\left|D^{i+1} u\right| \leq C_2 \left|z_0\right|^{-i} \quad \textit{and} \quad \left|D^i \partial_t u\right| \leq C_2 \left|z_0\right|^{-1-i}. \tag{2.5}$$

Here D and D^2 are the Euclidean gradient and Hessian on $\mathbb{R}^n$, respectively, and ∂_t denotes the partial derivative with respect to t fixing points in $T_{z_0}C$.

Proof. Given $\delta > 0$, by the assumption of Theorem 1.3, there exists $t_0 = t_0(\delta) \in (0, R_1^{-2}]$ such that if $0 < t \leq t_0$, then $\Sigma_t \cap B_3 \backslash B_{1/3}$ can be written as the graph of a function $V(\cdot, t) : \Omega_t \longrightarrow \mathbb{R}$ such that $\|V(\cdot,t)\|_{C^1} \leq \delta$. Here, the domain Ω_t of C satisfies that $B_2 \backslash B_{1/2} \subset \Omega_t \subset B_4 \backslash B_{1/4}$ and $\Omega_t \longrightarrow C \cap \left(B_3 \backslash B_{1/3}\right)$ as $t \longrightarrow 0$. Thus, we can choose δ sufficiently small, depending on C inside the annulus $\bar{B}_4 \backslash B_{1/4}$, such that for $z \in C \cap \left(B_2 \backslash B_{1/2}\right)$, $x_t = z + V(z,t)\mathbf{n}(z) \in \Sigma_t$ and $0 < t \leq t_0$, the distance $\mathrm{dist}(x_t, C)$ from x_t to C, which is achieved uniquely at z and is equal to $|V(z,t)|$, is less than $1/100$, and $\langle \mathbf{n}(x_t), \mathbf{n}(z) \rangle > 99/100$. Hence, if $x \in \Sigma_t$ with $|x|^2 > t_0^{-1}$ and $0 < t \leq 1$, then, by the homogeneity of C, $\mathrm{dist}(x, C) < |x|/100$, which is attained at a unique point z on C, and $\langle \mathbf{n}(x), \mathbf{n}(z) \rangle > 99/100$. This implies that the nearest point projection $\Pi_t : \Sigma_t \backslash B_{2/\sqrt{t_0}} \longrightarrow C$ is well defined for each $t \in (0,1]$ and $|x|/2 < |\Pi_t(x)| < 2|x|$ for $x \in \Sigma_t \backslash B_{2/\sqrt{t_0}}$. Moreover, the map Π_t is injective and the image of Π_t contains $C \backslash \bar{B}_{4/\sqrt{t_0}}$. This proves the first part of Lemma 2.2 with $R_2 = 4/\sqrt{t_0}$ and $K_t = \Sigma_t \backslash \Pi_t^{-1}(C \backslash \bar{B}_{4/\sqrt{t_0}})$.

It follows from the discussion in the previous paragraph that, given $\delta > 0$, there exists $r = r(\delta) > 0$ such that for $z_0 \in C \backslash B_r$ and $0 < t \leq 1$, $\left|\Pi_t^{-1}(z_0) - z_0\right| < \delta \left|z_0\right|$. Thus, by the proof of Lemma 2.2 in [7] and Lemma 2.1, there exist $\delta_1 \in (0,1)$ and $r_1, \lambda_1 > 0$, depending only on C_1, such that for $z_0 \in C \backslash B_{r_1}$ and $t \in (0,1]$, the component of $\Sigma_t \cap B_{\delta_1 |z_0|}(z_0)$ containing $\Pi_t^{-1}(z_0)$ can be written as the graph of a smooth function $u(\cdot, t)$ over the tangent hyperplane of C at z_0 and $u(\cdot, t)$ satisfies that $|Du(\cdot, t)| \leq \lambda_1$ and $\left|D^2 u(\cdot, t)\right| \leq \lambda_1 \left|z_0\right|^{-1}$. In fact, Lemma 2.1 implies that $\left|D^3 u(\cdot, t)\right| \leq \lambda_2 \left|z_0\right|^{-2}$ and $\left|D^4 u(\cdot, t)\right| \leq \lambda_2 \left|z_0\right|^{-3}$, where $\lambda_2 > 0$ depends only on C_1 and λ_1. Moreover, since Σ_t varies smoothly with respect to t, the function U in this lemma depends on t smoothly and so does u. In the following, we parametrize $T_{z_0} C$ by $F : \mathbb{R}^n \longrightarrow T_{z_0} C$, $F(p) = z_0 + \sum_i p_i e_i$, where $p = (p_1, \ldots, p_n)$ and $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $T_{z_0} C - z_0$. And we identify $u(p, t)$ with $u(F(p), t)$. Note that ∂_t means the partial derivative with respect to t fixing p. Since $\{\Sigma_t\}_{t \in (0,1]}$ moves by mean curvature backwards, we derive the differential equation for u:

$$-\partial_t u = \sqrt{1 + |Du|^2} \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right). \tag{2.6}$$

Thus, $|\partial_t u| \leq \lambda_3 \left|z_0\right|^{-1}$, where $\lambda_3 > 0$ depends only on λ_1. Next, differentiating once with respect to the ith coordinate on both sides of the equation (2.6) gives

$$-\partial_i \partial_t u = \frac{\sum_k \partial_k u \partial_{ik}^2 u}{\sqrt{1 + |Du|^2}} \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) + \sqrt{1 + |Du|^2}\, \partial_i \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right).$$

Note that the divergence term on the right-hand side of the equation (2.6) is the mean curvature of Σ_t at Graph(u). Thus, Lemma 2.1 implies that

$$\left|\partial_i \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right)\right| \leq \lambda_4 \left|z_0\right|^{-2}, \tag{2.7}$$

where $\lambda_4 > 0$ depends only on λ_1 and C_1. Hence, there exists $\lambda_5 > 0$, depending only on λ_1, λ_2 and λ_4, such that $|\partial_i \partial_t u| \leq \lambda_5 \left|z_0\right|^{-2}$. Similarly, differentiating twice with respect to the ith and jth coordinates on the both sides of the equation (2.6), we get

$$\begin{aligned}
-\partial_{ij}^2 \partial_t u = {} & \left(\frac{\sum_k \partial_{ik}^2 u \partial_{jk}^2 u + \partial_k u \partial_{ijk}^3 u}{\sqrt{1 + |Du|^2}} - \frac{\sum_{k,l} \partial_k u \partial_l u \partial_{ik}^2 u \partial_{jl}^2 u}{(1 + |Du|^2)^{3/2}}\right) \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) \\
& + \frac{\sum_k \partial_k u \partial_{ik}^2 u}{\sqrt{1 + |Du|^2}}\, \partial_j \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) \\
& + \frac{\sum_k \partial_k u \partial_{jk}^2 u}{\sqrt{1 + |Du|^2}}\, \partial_i \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) \\
& + \sqrt{1 + |Du|^2}\, \partial_{ij}^2 \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right).
\end{aligned}$$

Note that, by Lemma 2.1, we have that

$$\left|\partial_{ij}^2 \mathrm{div}\left(\frac{Du}{\sqrt{1+|Du|^2}}\right)\right| \leq \lambda_6 |z_0|^{-3}, \tag{2.8}$$

where $\lambda_6 > 0$ depends on λ_1 and λ_2. Hence, there exists $\lambda_7 > 0$, depending only on λ_1, λ_2, λ_4 and λ_6, such that $\left|\partial_{ij}^2 \partial_t u\right| \leq \lambda_7 |z_0|^{-3}$. □

Thus, by Lemmas 2.1 and 2.2, we can write $\tilde{\Sigma}$ outside a compact set as the graph of a function over Σ. Namely,

Lemma 2.3. *There exist $R_3 > R_2$ and $C_3 > 0$ such that, outside a compact set, $\tilde{\Sigma}$ is given by the graph of a smooth function $v : \Sigma \backslash \bar{B}_{R_3} \longrightarrow \mathbb{R}$ satisfying that for $x \in \Sigma \backslash \bar{B}_{R_3}$ and $0 \leq i \leq 2$,*

$$\left|\nabla_\Sigma^i v(x)\right| \leq C_3 |x|^{-1-i}. \tag{2.9}$$

Proof. First, using the same arguments in Lemma 2.2, there exist $\tilde{R}_2 > R_1$ and compact sets $\tilde{K}_t \subset \tilde{\Sigma}_t$, $0 < t \leq 1$, such that $\tilde{K}_t \subset B_{2\tilde{R}_2}$, and $\tilde{\Sigma}_t \backslash \tilde{K}_t$ is given by the graph of a smooth function $\tilde{U}(\cdot, t) : C \backslash \bar{B}_{\tilde{R}_2} \longrightarrow \mathbb{R}$. Moreover, there exist $0 < \tilde{\varepsilon}_0 < 1$ and $\tilde{C}_2 > 0$ such that for $z_0 \in C \backslash B_{2\tilde{R}_2}$ and $t \in (0, 1]$, the component of $\tilde{\Sigma}_t \cap B_{\tilde{\varepsilon}_0 |z_0|}(z_0)$ containing $z_0 + \tilde{U}(z_0, t)\mathbf{n}(z_0)$ can be written as the graph of a smooth function $\tilde{u}(\cdot, t)$ over the tangent hyperplane of C at z_0 satisfying that, for $i = 0, 1, 2$,

$$\left|D^{i+1}\tilde{u}\right| \leq \tilde{C}_2 |z_0|^{-i} \quad \text{and} \quad \left|D^i \partial_t \tilde{u}\right| \leq \tilde{C}_2 |z_0|^{-1-i}. \tag{2.10}$$

For $r_1 > 2\max\{R_2, \tilde{R}_2, \sqrt{C_2}, \sqrt{\tilde{C}_2}\}$, let $\Pi : \Sigma \backslash B_{r_1} \longrightarrow C$ and $\tilde{\Pi} : \tilde{\Sigma} \backslash B_{r_1} \longrightarrow C$ be the nearest point projections. Thus, if $y \in \tilde{\Sigma} \backslash B_{4r_1}$, then $z = \tilde{\Pi}(y) \in C \backslash B_{2r_1}$ and $\mathrm{dist}(y, \Sigma) \leq |y - z| + \left|z - \Pi^{-1}(z)\right| \leq 2(C_2 + \tilde{C}_2)|y|^{-1}$. Hence, by Lemma 2.1, we can choose $r_1 \gg 1$, depending only on C_1, C_2 and $\tilde{C}_2$, such that for $y \in \tilde{\Sigma} \backslash B_{4r_1}$, there exists a unique $x \in \Sigma$ satisfying that $|y - x| = \mathrm{dist}(y, \Sigma)$.

Next, for $x \in \Sigma \backslash B_{2r_1}$, Lemma 2.2 implies that $z' = \Pi(x) \in C \backslash B_{r_1}$, $|x - z'| \leq C_2 |z'|^{-1}$ and $\langle \mathbf{n}(x), \mathbf{n}(z') \rangle > 1 - \mu_1 |z'|^{-4}$, where $\mu_1 > 0$ depends on C_2. Thus, if r_1 is sufficiently large, depending only on C_2 and $\tilde{C}_2$, then the component of $\tilde{\Sigma} \cap B_{\tilde{\varepsilon}_0 |z'|/2}(z')$ containing $\tilde{\Pi}^{-1}(z')$ can be written as the graph of the function f over the tangent hyperplane of Σ at x. And $|f(x)| \leq \mu_2 |z'|^{-1}$ for some μ_2, depending on C_2 and $\tilde{C}_2$. Hence, there exists $r_2 > 2r_1$ large, depending only on μ_2 and C_1, such that the image of the nearest point projection from $\tilde{\Sigma} \backslash B_{4r_1} \longrightarrow \Sigma$ contains $\Sigma \backslash B_{r_2}$.

Finally, it is easy to show that there exists $\delta_1 > 0$, depending only on C, such that if $z_1, z_2 \in C \backslash B_{r_1}$ and $|z_1 - z_2| < \delta_1$, then the geodesic distance between z_1 and z_2 on C, $\mathrm{dist}_C(z_1, z_2)$, is less than $2|z_1 - z_2|$. Note that for $y \in \tilde{\Sigma} \backslash B_{2r_1}$, $\tilde{\Pi}(y) \in C \backslash B_{r_1}$, $|\tilde{U}(\tilde{\Pi}(y), 1)| \leq \tilde{C}_2 / |\tilde{\Pi}(y)|$ and $|\nabla_C \tilde{U}(\cdot, 1)|$ is uniformly small on $C \backslash B_{r_1}$. Thus, there exist $\delta_2 > 0$, $r_3 \gg 1$ and $\mu_3 > 1$, depending only on $\tilde{C}_2$, δ_1 and r_1, such that if y_1, $y_2 \in \tilde{\Sigma} \backslash B_{r_3}$ and $|y_1 - y_2| \leq \delta_2$, then the geodesic distance

$\mathrm{dist}_{\tilde{\Sigma}}(y_1, y_2) \leq \mu_3 |y_1 - y_2|$. Hence, by the discussion in the previous paragraph, the nearest point projection from $\tilde{\Sigma} \backslash B_{r_4}$ to Σ is injective, for some large $r_4 > 4\max\{r_1, r_3\}$, depending on C_2 and $\tilde{C}_2$. Therefore, choosing $r_5 = 2\max\{r_2, r_4\}$, we conclude that, outside a compact set, $\tilde{\Sigma}$ is given by the graph of a smooth function v over $\Sigma \backslash \bar{B}_{r_5}$.

In the following, we derive the bounds of v and its derivatives up to the second order. Fix x in $\Sigma \backslash \bar{B}_{r_5}$. It is clear that $|v(x)| \leq \lambda_1 |x|^{-1}$, where $\lambda_1 > 0$ depends only on C_2 and $\tilde{C}_2$. Let $z' = \Pi(x)$ and $y = x + v(x)\mathbf{n}(x)$. Thus, by Lemma 2.2, $|x - z'|$ and $|y - \tilde{\Pi}^{-1}(z')|$ are bounded by multiples of $1/|x|$. Hence, the geodesic distance $\mathrm{dist}_{\tilde{\Sigma}}(y, \tilde{\Pi}^{-1}(z'))$ is also a multiple of $1/|x|$. Therefore, we can write the component of $\Sigma \cap B_{\varepsilon_0|z|}(z')$ containing x and the component of $\tilde{\Sigma} \cap B_{\tilde{\varepsilon}_0|z|}(z')$ containing y as the graphs of smooth functions $u_1 = u(\cdot, 1)$ and $\tilde{u}_1 = \tilde{u}(\cdot, 1)$ over the tangent hyperplane $T_{z'}C$ of C at z', respectively. We parametrize $T_{z'}C$ by $F : \mathbb{R}^n \longrightarrow T_{z'}C$, $F(p) = z' + \sum_i p_i e_i$, where $p = (p_1, \ldots, p_n)$ and $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $T_{z'}C - z'$. For $p, q \in \mathbb{R}^n$, we identify $u_1(p)$, $\tilde{u}_1(q)$ and $v(p)$ with $u_1(F(p))$, $\tilde{u}_1(F(q))$ and $v(F(p))$, respectively. Let e_{n+1} be the unit normal of the hyperplane $T_{z'}C$. Note that the unit normal $\mathbf{n}$ of Σ at a point $F(p) + u_1(p)e_{n+1}$ is given by

$$\mathbf{n}(p) = \sum_{h=1}^{n} \frac{-\partial_h u_1}{\sqrt{1 + |Du_1|^2}}\, e_h + \frac{1}{\sqrt{1 + |Du_1|^2}}\, e_{n+1}. \tag{2.11}$$

Thus, for $h = 1, 2 \ldots, n$,

$$q_h = p_h - \frac{\partial_h u_1}{\sqrt{1 + |Du_1|^2}}\, v, \tag{2.12}$$

$$\tilde{u}_1(q) = u_1(p) + \frac{1}{\sqrt{1 + |Du_1|^2}}\, v. \tag{2.13}$$

Differentiating the equations (2.12) and (2.13) with respect to p_i gives

$$\begin{aligned} \partial_i q_h = \delta_{hi} &- \frac{\partial^2_{hi} u_1}{\sqrt{1 + |Du_1|^2}}\, v + \frac{\partial_h u_1}{(1 + |Du_1|^2)^{3/2}} \left(\sum_l \partial_l u_1 \partial^2_{li} u_1 \right) v \\ &- \frac{\partial_h u_1}{\sqrt{1 + |Du_1|^2}}\, \partial_i v, \end{aligned} \tag{2.14}$$

$$\sum_k \partial_k \tilde{u}_1 \partial_i q_k = \partial_i u_1 + \frac{\partial_i v}{\sqrt{1 + |Du_1|^2}} - \frac{v}{(1 + |Du_1|^2)^{3/2}} \sum_l \partial_l u_1 \partial^2_{li} u_1. \tag{2.15}$$

Furthermore, plugging the equation (2.14) into the equation (2.15) gives

$$\begin{aligned} \frac{1 + \sum_k \partial_k \tilde{u}_1 \partial_k u_1}{\sqrt{1 + |Du_1|^2}}\, \partial_i v = (\partial_i \tilde{u}_1 - \partial_i u_1) &- \frac{\sum_k \partial_k \tilde{u}_1 \partial^2_{ki} u_1}{\sqrt{1 + |Du_1|^2}}\, v \\ &+ \frac{1 + \sum_k \partial_k \tilde{u}_1 \partial_k u_1}{(1 + |Du_1|^2)^{3/2}} \left(\sum_l \partial_l u_1 \partial^2_{li} u_1 \right) v. \end{aligned} \tag{2.16}$$

Note that $|q(0)| \leq \lambda_1 |x|^{-1}$ and $|x|$, $|y|$ and $|z'|$ are comparable. Thus, by the assumption of Theorem 1.3, (2.5) and (2.10), there exists $\lambda_2 > 0$, depending only on λ_1, C_2 and $\tilde{C}_2$ such that at $p = 0$,

$$|\partial_i \tilde{u}_1 \circ q - \partial_i u_1| \leq |\partial_i \tilde{u}_1 \circ q - \partial_i \tilde{u}_1| + |\partial_i \tilde{u}_1 - \partial_i u_1| \leq \lambda_2 |x|^{-2}. \tag{2.17}$$

Hence, it follows from the upper bound of v, (2.5), (2.10) and (2.17) that $|\partial_i v(0)| \leq \lambda_3 |x|^{-2}$ for some $\lambda_3 > 0$ depending only on λ_2, C_2 and $\tilde{C}_2$, and thus the same holds true for $|\nabla_\Sigma v|$ at x.

Next, differentiating the equations (2.14) and (2.15) with respect to p_j gives that

$$\begin{aligned}
\partial^2_{ij} q_h = &-\frac{\partial^3_{hij} u_1}{\sqrt{1+|Du_1|^2}}\, v + \frac{\partial^2_{hi} u_1}{(1+|Du_1|^2)^{3/2}} \left(\sum_l \partial_l u_1 \partial^2_{lj} u_1\right) v \\
&- \frac{3\partial_h u_1}{(1+|Du_1|^2)^{5/2}} \left(\sum_l \partial_l u_1 \partial^2_{li} u_1\right)\left(\sum_m \partial_m u_1 \partial^2_{mj} u_1\right) v \\
&+ \frac{\partial_h u_1}{(1+|Du_1|^2)^{3/2}} \left(\sum_l \partial^2_{li} u_1 \partial^2_{lj} u_1 + \partial_l u_1 \partial^3_{lij} u_1\right) v \\
&+ \frac{\partial^2_{hj} u_1}{(1+|Du_1|^2)^{3/2}} \left(\sum_l \partial_l u_1 \partial^2_{li} u_1\right) v - \frac{\partial^2_{hi} u_1 \partial_j v + \partial^2_{hj} u_1 \partial_i v}{\sqrt{1+|Du_1|^2}} \\
&+ \frac{\partial_h u_1}{(1+|Du_1|^2)^{3/2}} \sum_l \partial_l u_1 (\partial^2_{li} u_1 \partial_j v + \partial^2_{li} u_1 \partial_i v) \\
&- \frac{\partial_h u_1}{\sqrt{1+|Du_1|^2}}\, \partial^2_{ij} v,
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
&\sum_{k,h} \left(\partial^2_{kh} \tilde{u}_1 \partial_i q_k \partial_j q_h + \partial_k \tilde{u}_1 \partial^2_{ij} q_k\right) \\
&\quad = \partial^2_{ij} u_1 + \frac{3}{(1+|Du_1|^2)^{5/2}} \left(\sum_l \partial_l u_1 \partial^2_{li} u_1\right)\left(\sum_m \partial_m u_1 \partial^2_{mj} u_1\right) v \\
&\qquad - \frac{1}{(1+|Du_1|^2)^{3/2}} \left(\sum_l \partial^2_{li} u_1 \partial^2_{lj} u_1 + \partial_l u_1 \partial^3_{lij} u_1\right) v \\
&\qquad - \frac{\sum_l \partial_l u_1 \partial^2_{lj} u_1}{(1+|Du_1|^2)^{3/2}}\, \partial_i v - \frac{\sum_l \partial_l u_1 \partial^2_{li} u_1}{(1+|Du_1|^2)^{3/2}}\, \partial_j v + \frac{\partial^2_{ij} v}{\sqrt{1+|Du_1|^2}}.
\end{aligned} \tag{2.19}$$

It follows from Lemma 2.2 and the bounds for $|v(0)|$ and $|Dv(0)|$ that, for $h = 1, 2, \ldots, n$,

$$\left|\partial^2_{ij} q_h(0) + \frac{\partial_h u_1(0)}{\sqrt{1+|Du_1(0)|^2}}\, \partial^2_{ij} v(0)\right| \leq \lambda_4 |x|^{-3}, \tag{2.20}$$

where $\lambda_4 > 0$ depends only on C_2, λ_1 and λ_3. Thus, by the assumption of Theorem 1.3, (2.5), (2.10) and $|q(0)| = \lambda_1 |x|^{-1}$, we estimate, at $p = 0$,

$$\left|\partial^2_{ij} \tilde{u}_1 \circ q - \partial^2_{ij} u_1\right| \leq \left|\partial^2_{ij} \tilde{u}_1 \circ q - \partial^2_{ij} \tilde{u}_1\right| + \left|\partial^2_{ij} \tilde{u}_1 - \partial^2_{ij} u_1\right| \leq \lambda_5 |x|^{-3}. \tag{2.21}$$

Here $\lambda_5 > 0$ depends only on C_2, $\tilde{C}_2$ and λ_1. Since $|D\tilde{u}_1(q(0))| \leq \tilde{C}_2$, $\left|D^2\tilde{u}_1(q(0))\right| \leq \tilde{C}_2 |x|^{-1}$ and there exists $\lambda_6 > 0$, depending only on C_2, λ_1 and λ_3, such that

$$|\partial_i q_h(0) - \delta_{ih}| \leq \lambda_6 |x|^{-2}, \tag{2.22}$$

and simplifying the equation (2.19) gives that

$$\left| \frac{1 + \sum_k \partial_k u_1(0) \partial_k \tilde{u}_1(0)}{\sqrt{1 + |Du_1(0)|^2}} \, \partial^2_{ij} v(0) - \left(\partial^2_{ij}\tilde{u}_1(0) - \partial^2_{ij} u_1(0)\right) \right| \leq \lambda_7 |x|^{-3}. \tag{2.23}$$

Here $\lambda_7 > 0$ depends only on C_2, $\tilde{C}_2$, λ_1 and λ_3. Therefore, it follows from (2.17), (2.21), (2.23) and Lemma 2.2 that $\left|\nabla^2_\Sigma v(x)\right| \leq \lambda_8 |x|^{-3}$, where $\lambda_8 > 0$ depends on C_2, λ_2, λ_5 and λ_7. □

Finally, we conclude this section by deriving the differential equation for v from the definition of self-shrinkers. Note that the operator L_Σ in Lemma 2.4 is a small perturbation of the stability operator introduced in [5].

Lemma 2.4. *There exists $C_4 > 0$ such that at $x \in \Sigma \backslash B_{R_3}$,*

$$L_\Sigma v = \Delta_\Sigma v - \frac{1}{2}\langle x, \nabla_\Sigma v\rangle + \left(|A|^2 + \frac{1}{2}\right) v + Q\left(x, v, \nabla_\Sigma v, \nabla^2_\Sigma v\right) = 0, \tag{2.24}$$

where the function Q satisfies that

$$\left|Q\left(x, v, \nabla_\Sigma v, \nabla^2_\Sigma v\right)\right| \leq C_4 |x|^{-2} \left(|v| + |\nabla_\Sigma v|\right). \tag{2.25}$$

Proof. Fix $x_0 \in \Sigma \backslash B_{R_3}$. We choose a local parametrization of Σ in a neighborhood of x_0, $F : \Omega \longrightarrow \Sigma$, satisfying that $F(0) = x_0$, $\langle \partial_i F(0), \partial_j F(0)\rangle = \delta_{ij}$ and $\partial^2_{ij} F(0) = a_{ij}(0)\mathbf{n}(x_0)$ with $a_{ij} = A(\partial_i F, \partial_j F)$ and $a_{ij}(0) = 0$ if $i \neq j$. Here Ω is a domain in $\mathbb{R}^n$ containing 0. Thus, in a neighborhood of $y_0 = x_0 + v(x_0)\mathbf{n}(x_0)$, there exists a local parametrization of $\tilde{\Sigma}$, $\tilde{F} : \Omega \longrightarrow \tilde{\Sigma}$, such that for $p \in \Omega$,

$$\tilde{F}(p) = F(p) + v(p)\mathbf{n}(p). \tag{2.26}$$

Here, we identify $v(p)$ and $\mathbf{n}(p)$ with $v(F(p))$ and $\mathbf{n}(F(p))$, respectively.

First, we calculate the tangent vectors $\partial_i \tilde{F}$ for $1 \leq i \leq n$. Namely,

$$\partial_i \tilde{F} = \partial_i F + (\partial_i v)\mathbf{n} + v\partial_i \mathbf{n}. \tag{2.27}$$

Thus, the unit normal vector $\mathbf{n}(\tilde{F}(0))$ to $\tilde{\Sigma}$ at $\tilde{F}(0)$ parallels to the following vector, that is,

$$\mathbf{N} = -\sum_k \left[\prod_{l \neq k}(1 - a_{ll} v)\right] (\partial_k v)\, \partial_k F + \left[\prod_k (1 - a_{kk} v)\right] \mathbf{n}. \tag{2.28}$$

Next, we calculate the second derivatives of $\tilde{F}$ at $p=0$. It follows from (2.27) that

$$\partial^2_{ij}\tilde{F} = \partial^2_{ij}F + \left(\partial^2_{ij}v\right)\mathbf{n} + \left(\partial_i v\right)\partial_j\mathbf{n} + \left(\partial_j v\right)\partial_i\mathbf{n} + v\partial^2_{ij}\mathbf{n}. \tag{2.29}$$

Note that, at $p=0$,

$$\begin{aligned}\partial^2_{ij}\mathbf{n} &= \sum_k \left\langle \partial^2_{ij}\mathbf{n}, \partial_k F\right\rangle \partial_k F + \left\langle \partial^2_{ij}\mathbf{n}, \mathbf{n}\right\rangle \mathbf{n}\\ &= -\sum_k \left(\partial_j a_{ik}\right)\partial_k F - a_{ii}a_{jj}\delta_{ij}\mathbf{n}.\end{aligned} \tag{2.30}$$

Plugging (2.30) into (2.29) gives that, at $p=0$,

$$\begin{aligned}\partial^2_{ij}\tilde{F} = &-a_{ii}\left(\partial_j v\right)\partial_i F - a_{jj}\left(\partial_i v\right)\partial_j F - \sum_k \left(\partial_j a_{ik}\right) v\partial_k F\\ &+ \left(a_{ij} - a_{ii}a_{jj}\delta_{ij}v + \partial^2_{ij}v\right)\mathbf{n}.\end{aligned} \tag{2.31}$$

Thus, making an inner product with the vector $\mathbf{N}$, we obtain that, at $p=0$,

$$\begin{aligned}\left\langle \partial^2_{ij}\tilde{F}, \mathbf{N}\right\rangle = &\, a_{ii}\left(\partial_i v\right)\left(\partial_j v\right)\prod_{k\neq i}(1-a_{kk}v) + a_{jj}\left(\partial_i v\right)\left(\partial_j v\right)\prod_{k\neq j}(1-a_{kk}v)\\ &+ \left(a_{ij} - a_{ii}a_{jj}\delta_{ij}v + \partial^2_{ij}v\right)\prod_k (1-a_{kk}v)\\ &+ v\sum_k \left[\prod_{l\neq k}(1-a_{ll}v)\right]\left(\partial_j a_{ik}\right)\partial_k v.\end{aligned} \tag{2.32}$$

Next, it is easy to calculate the pullback metric $g=(g_{ij})$ from $\tilde{\Sigma}$, that is, at $p=0$,

$$g_{ij} = \langle \partial_i F, \partial_j F\rangle = (1-a_{ii}v)(1-a_{jj}v)\delta_{ij} + (\partial_i v)(\partial_j v). \tag{2.33}$$

Thus, at $p=0$, the determinant $\det(g)$ and the inverse $g^{-1} = \left(g^{ij}\right)$ of g are

$$\det(g) = 1 - 2\sum_k a_{kk}v + Q_1(p,v,Dv), \tag{2.34}$$

$$g^{ij}\det(g) = \begin{cases} Q_{2ij}(p,v,Dv)\partial_i v & \text{if } i\neq j\\ 1 - 2\sum_{k\neq i} a_{kk}v + Q_{2ij}(p,v,Dv) & \text{if } i=j.\end{cases} \tag{2.35}$$

Here, by Lemmas 2.1 and 2.3, there exists $\lambda_1 > 0$, depending only on C_1 and C_3, such that $|Q_1(0,v,Dv)| \le \lambda_1 |x_0|^{-2}\left(|v(0)| + |Dv(0)|\right)$ with similar bounds holding for $|Q_{2ij}(0,v,Dv)|$ for $1\le i,j\le n$.

Finally, we compute, at $p=0$,

$$\left\langle \tilde{F}, \mathbf{N}\right\rangle = -\sum_k \left[\prod_{l\neq k}(1-a_{ll}v)\right]\langle F, \partial_k F\rangle\,\partial_k v + \left(\langle F,\mathbf{n}\rangle + v\right)\prod_k (1-a_{kk}v). \tag{2.36}$$

Since $\tilde{\Sigma}$ is a self-shrinker under the mean curvature flow, we have that

$$\sum_{i,j} g^{ij}\left\langle \partial^2_{ij}\tilde{F}, \mathbf{N}\right\rangle = -\frac{1}{2}\left\langle \tilde{F}, \mathbf{N}\right\rangle. \tag{2.37}$$

Substituting all the previous computation into the equation (2.37) gives that, at $p = 0$,

$$\sum_k \partial^2_{kk} v - \frac{1}{2}\sum_k \langle F, \partial_k F\rangle \partial_k v + \left(|A|^2 + \frac{1}{2}\right) v = \tilde{Q}\left(p, v, Dv, D^2 v\right), \tag{2.38}$$

where the function $\tilde{Q}$ satisfies that

$$\left|\tilde{Q}\left(0, v, Dv, D^2 v\right)\right| \leq \lambda_2 |x_0|^{-2} \left(|v(0)| + |Dv(0)|\right). \tag{2.39}$$

Here $\lambda_2 > 0$ depends only on C_1 and C_3 in Lemma 2.3. Therefore, by the choice of the parametrization in a neighborhood of x_0, Lemma 2.4 follows immediately from (2.38). □

3. Proof of Theorem 1.3

By the discussion in Section 2, Theorem 1.3 can be restated as the unique continuation property for the elliptic equation (2.24). Let us consider the simplest case that $\Sigma = \mathbb{R}^n \times \{0\}$. We define $\bar{v}(r,\theta) = v(1/r,\theta)$, where $(r,\theta) \in (0,+\infty) \times S^{n-1}$ are the spherical coordinates of $\mathbb{R}^n \times \{0\}$. Then, the equation (2.24) gives that

$$\Delta_{\mathbb{R}^n} \bar{v} + \frac{1}{2}\left[r^{-3} + 4(2-n)r^{-1}\right] \partial_r \bar{v} + \frac{1}{2} r^{-4} \bar{v} + \bar{Q}\left(r, \theta, \bar{v}, D\bar{v}, D^2\bar{v}\right) = 0, \tag{3.1}$$

or equivalently,

$$\operatorname{div}\left(r^{4-2n} \mathrm{e}^{-\frac{1}{4r^2}} D\bar{v}\right) + \frac{1}{2} r^{-2n} \mathrm{e}^{-\frac{1}{4r^2}} \bar{v} + r^{4-2n} \mathrm{e}^{-\frac{1}{4r^2}} \bar{Q}\left(r, \theta, \bar{v}, D\bar{v}, D^2\bar{v}\right) = 0. \tag{3.2}$$

Here the function $\bar{Q}$ satisfies that

$$\left|\bar{Q}\left(r, \theta, \bar{v}, D\bar{v}, D^2\bar{v}\right)\right| \leq C_4 r^{-2}\left(|\bar{v}| + r^2 |D\bar{v}|\right). \tag{3.3}$$

Hence, we cannot apply some well-known strong unique continuation theorems in [25], [26] and [17, 18] to the equations (3.1) or (3.2). In [41], Pan and Wolff proved the unique continuation results for operators in the general form, $\Delta_{\mathbb{R}^n} + \langle W, \nabla_{\mathbb{R}^n}\rangle + V$, where $\left|r^2 V(r,\theta)\right|$ is bounded and $|rW(r,\theta)|$ is small. They also constructed counter examples when the conditions are violated. However, it seems impossible to transform the equation (3.2) to that of their form by changing coordinates. Besides, $\bar{v}$ need not vanish of infinite order at the origin under our assumption. For general Σ, global coordinates of Σ near infinity may not even exist.

Instead, we associate the elliptic differential equation (2.24) to the parabolic equation in the following lemma. Define a function $w : \cup_{t\in(0,1]} \Sigma_t \times \{t\} \longrightarrow \mathbb{R}$ by

$$w(x,t) = \sqrt{t}\, v\left(\frac{x}{\sqrt{t}}\right). \tag{3.4}$$

Let d/dt denote the total derivative with respect to time t. Then, by Lemma 2.4, w satisfies the following equation:

Lemma 3.1. *Given $\varepsilon_1 > 0$, there exists $R > R_3$ such that on $\cup_{t\in(0,1]}\Sigma_t\backslash B_R \times \{t\}$,*

$$\left|\frac{dw}{dt} + \Delta_{\Sigma_t} w\right| \leq \varepsilon_1 \left(|w| + |\nabla_{\Sigma_t} w|\right). \tag{3.5}$$

Proof. By a straightforward computation, we get that

$$\begin{aligned}\frac{dw}{dt} + \Delta_{\Sigma_t} w = {} & \frac{1}{2\sqrt{t}}\, v\left(\frac{x}{\sqrt{t}}\right) + \sqrt{t}\left\langle \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right), \frac{1}{\sqrt{t}}\frac{\partial x}{\partial t} - \frac{x}{2t\sqrt{t}}\right\rangle \\ & + \frac{1}{\sqrt{t}}\, \Delta_\Sigma v\left(\frac{x}{\sqrt{t}}\right).\end{aligned} \tag{3.6}$$

Note that for $x \in \Sigma_t$ and $t \in (0, 1]$,

$$\partial_t x = H\mathbf{n} \quad \text{and} \quad H = \frac{\langle x, \mathbf{n}\rangle}{2t}. \tag{3.7}$$

Hence, the equation (3.6) and (3.7) give that

$$\frac{dw}{dt} + \Delta_{\Sigma_t} w = \frac{1}{\sqrt{t}}\left(\Delta_\Sigma v\left(\frac{x}{\sqrt{t}}\right) - \frac{1}{2}\left\langle \frac{x}{\sqrt{t}}, \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right)\right\rangle + \frac{1}{2} v\left(\frac{x}{\sqrt{t}}\right)\right). \tag{3.8}$$

Therefore, Lemma 3.1 follows immediately from Lemmas 2.1 and 2.4. □

It is implied from (3.4) and (2.9) that w is asymptotic to zero as $t \longrightarrow 0+$. Thus, appealing to the strong unique continuation theorems in [17,18], it is enough to show that there exist $T \in (0, 1)$ and $R > 0$ such that $w \equiv 0$ in $\cup_{t\in(0,T)}\Sigma_t\backslash B_R\times\{t\}$ so as to prove Theorem 1.3. One of the standard methods to prove backwards uniqueness is the so-called Carleman type estimates, i.e., a family of uniform weighted L^2 estimates. The key to derive such estimates is the construction of suitable weight functions satisfying a convexity property, which is closely related to the distance function and heat kernel on the underlying manifold.

As no control of the parabolic boundary value is made on w, our backwards uniqueness problem is local near infinity and thus more delicate than the usual global ones. Fortunately, some precedent in the case of heat type inequalities in the Euclidean space with a ball removed has been obtained by Escauriaza, Seregin and Šverák in [15]. Since the radial distance function on any cone satisfies the same Hessian equation as that on the Euclidean space, we are able to modify their arguments to apply to asymptotically conical manifolds.

First, we prove the following key identity (see (3.9) below) which is a generalization of that in Lemma 1 of [15] to our geometric setting.[4] Namely,

Lemma 3.2. *Assume that ϕ and $\mathcal{G}$ are smooth functions on $\cup_{t\in(0,1]}\Sigma_t \times \{t\}$ and that $\mathcal{G}$ is positive. And let $\mathcal{F} = (d\mathcal{G}/dt - \Delta_{\Sigma_t}\mathcal{G})/\mathcal{G}$. Then, the following identity*

[4]Under slightly different notations, a similar identity might have already been implicitly obtained in [13].

holds on Σ_t*:*

$$\begin{aligned}
&\operatorname{div}_{\Sigma_t}\Big\{2\mathcal{G}\frac{d\phi}{dt}\nabla_{\Sigma_t}\phi+\left|\nabla_{\Sigma_t}\phi\right|^2\nabla_{\Sigma_t}\mathcal{G}-2\left\langle\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\mathcal{G}\right\rangle\nabla_{\Sigma_t}\phi+\mathcal{F}\mathcal{G}\phi\nabla_{\Sigma_t}\phi\\
&\quad+\frac{1}{2}\phi^2\mathcal{F}\nabla_{\Sigma_t}\mathcal{G}-\frac{1}{2}\phi^2\mathcal{G}\nabla_{\Sigma_t}\mathcal{F}\Big\}-\frac{d}{dt}\left(\left|\nabla_{\Sigma_t}\phi\right|^2\mathcal{G}+\frac{1}{2}\phi^2\mathcal{F}\mathcal{G}\right)\\
&=-2\nabla^2_{\Sigma_t}\log\mathcal{G}\left(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi\right)\mathcal{G}-2\left(\frac{d\phi}{dt}-\left\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle+\frac{1}{2}\phi\mathcal{F}\right)^2\mathcal{G}\\
&\quad+2\left(\frac{d\phi}{dt}-\left\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle+\frac{1}{2}\phi\mathcal{F}\right)\left(\frac{d\phi}{dt}+\Delta_{\Sigma_t}\phi\right)\mathcal{G}\\
&\quad-2HA\left(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi\right)\mathcal{G}-\frac{1}{2}\phi^2\left(\frac{d\mathcal{F}}{dt}+\Delta_{\Sigma_t}\mathcal{F}\right)\mathcal{G}.
\end{aligned}\tag{3.9}$$

Proof. First, a straightforward calculation gives that

$$\begin{aligned}
&2\left(\frac{d\phi}{dt}-\left\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle+\frac{1}{2}\phi\mathcal{F}\right)\left(\frac{d\phi}{dt}+\Delta_{\Sigma_t}\phi\right)\mathcal{G}\\
&\quad-2\left(\frac{d\phi}{dt}-\left\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle+\frac{1}{2}\phi\mathcal{F}\right)^2\mathcal{G}\\
&=\operatorname{div}_{\Sigma_t}\left(2\mathcal{G}\frac{d\phi}{dt}\nabla_{\Sigma_t}\phi+\phi\mathcal{F}\mathcal{G}\nabla_{\Sigma_t}\phi+\frac{1}{2}\phi^2\mathcal{F}\nabla_{\Sigma_t}\mathcal{G}-\frac{1}{2}\phi^2\mathcal{G}\nabla_{\Sigma_t}\mathcal{F}\right)\\
&\quad-\frac{d}{dt}\left(\left|\nabla_{\Sigma_t}\phi\right|^2\mathcal{G}+\frac{1}{2}\phi^2\mathcal{F}\mathcal{G}\right)+\frac{1}{2}\phi^2\mathcal{G}\left(\frac{d\mathcal{F}}{dt}+\Delta_{\Sigma_t}\mathcal{F}\right)+\left|\nabla_{\Sigma_t}\phi\right|^2\Delta_{\Sigma_t}\mathcal{G}\\
&\quad-2\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\Delta_{\Sigma_t}\phi-2\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\rangle^2\mathcal{G}\\
&\quad+\left(\frac{d}{dt}\left|\nabla_{\Sigma_t}\phi\right|^2-2\left\langle\nabla_{\Sigma_t}\left(\frac{d\phi}{dt}\right),\nabla_{\Sigma_t}\phi\right\rangle\right)\mathcal{G}.
\end{aligned}\tag{3.10}$$

Here we use the fact that, given functions f and g on a hypersurface N,

$$f\Delta_N g=g\Delta_N f+\operatorname{div}_N\left(f\nabla_N g-g\nabla_N f\right).\tag{3.11}$$

On the other hand, we note that

$$\begin{aligned}
2\left\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle\Delta_{\Sigma_t}\phi&=\operatorname{div}_{\Sigma_t}\left(2\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\nabla_{\Sigma_t}\phi\right)-2\nabla_{\Sigma_t}\phi\left(\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\right)\\
&=\operatorname{div}_{\Sigma_t}\left(2\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\nabla_{\Sigma_t}\phi\right)-2\left\langle\nabla_{\nabla_{\Sigma_t}\phi}\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\right\rangle\\
&\quad-2\left\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\nabla_{\Sigma_t}\phi}\nabla_{\Sigma_t}\phi\right\rangle\\
&=\operatorname{div}_{\Sigma_t}\left(2\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\nabla_{\Sigma_t}\phi\right)-2\nabla^2_{\Sigma_t}\mathcal{G}(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi)\\
&\quad-\left\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\left|\nabla_{\Sigma_t}\phi\right|^2\right\rangle\\
&=\operatorname{div}_{\Sigma_t}\left(2\langle\nabla_{\Sigma_t}\mathcal{G},\nabla_{\Sigma_t}\phi\rangle\nabla_{\Sigma_t}\phi-\left|\nabla_{\Sigma_t}\phi\right|^2\nabla_{\Sigma_t}\mathcal{G}\right)\\
&\quad-2\nabla^2_{\Sigma_t}\mathcal{G}(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi)+\left|\nabla_{\Sigma_t}\phi\right|^2\Delta_{\Sigma_t}\mathcal{G}.
\end{aligned}\tag{3.12}$$

Also, it is easy to check that

$$\nabla^2_{\Sigma_t}\log\mathcal{G}(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi)\mathcal{G}=\nabla^2_{\Sigma_t}\mathcal{G}(\nabla_{\Sigma_t}\phi,\nabla_{\Sigma_t}\phi)-\langle\nabla_{\Sigma_t}\log\mathcal{G},\nabla_{\Sigma_t}\phi\rangle^2\mathcal{G}.\tag{3.13}$$

Finally, let $\Omega\subset\mathbb{R}^n$ and $F:\Omega\times(0,1]\longrightarrow\mathbb{R}^{n+1}$ be the local parametrization of $\{\Sigma_t\}_{t\in(0,1]}$ such that $F(\Omega,t)\subset\Sigma_t$ and $dF/dt=H\mathbf{n}$. Let g_t be the pullback metric

on Ω from Σ_t. Thus, we have that

(3.14)
$$\frac{d}{dt}\left|\nabla_{\Sigma_t}\phi\right|^2 - 2\left\langle \nabla_{\Sigma_t}\left(\frac{d\phi}{dt}\right), \nabla_{\Sigma_t}\phi\right\rangle = \sum_{i,j}\partial_t g_t^{ij}\partial_i\phi\partial_j\phi = 2HA(\nabla_{\Sigma_t}\phi, \nabla_{\Sigma_t}\phi),$$

where $\partial_t g_t^{ij}$ denotes the time derivative fixing a point in Ω, and we use the following equation in the second equality (see Appendix B in [10]):

$$\partial_t g_t^{ij} = 2H\sum_{k,l} g_t^{ik} g_t^{jl} A(\partial_k F, \partial_l F). \tag{3.15}$$

Therefore, Lemma 3.2 follows from (3.10) and (3.12)–(3.15). □

Let g_t be the metric on Σ_t induced from $\mathbb{R}^{n+1}$ and $d\mu_t$ the Hausdorff measure on Σ_t. Thus we obtain an integral inequality from the divergence identity (3.9).

Lemma 3.3. *Given $R > 2R_2$ and $0 < T \leq 1$, let $\mathcal{G}$ be a smooth positive function in*

$$Q_{R,T} = \left\{(x,t) \,\middle|\, x \in \Sigma_t \backslash \bar{B}_R, t \in (0,T]\right\}. \tag{3.16}$$

As in Lemma 3.2, set $\mathcal{F} = (d\mathcal{G}/dt - \Delta_{\Sigma_t}\mathcal{G})/\mathcal{G}$. Then the following inequality holds for all $\phi \in C_c^\infty(Q_{R,T})$:

$$\begin{aligned}
&\int_0^T\int_{\Sigma_t}\left(2\nabla^2_{\Sigma_t}\log\mathcal{G} + 2HA + H^2 g_t\right)\left(\nabla_{\Sigma_t}\phi, \nabla_{\Sigma_t}\phi\right)\mathcal{G}d\mu_t dt \\
&\quad + \int_0^T\int_{\Sigma_t}\frac{1}{2}\left(\frac{d\mathcal{F}}{dt} + \Delta_{\Sigma_t}\mathcal{F} + H^2\mathcal{F}\right)\phi^2\mathcal{G}d\mu_t dt \\
&\leq \int_0^T\int_{\Sigma_t}\left(\frac{d\phi}{dt} + \Delta_{\Sigma_t}\phi\right)^2\mathcal{G}d\mu_t dt + \int_{\Sigma_T}\left(|\phi|^2\mathcal{F} + |\nabla_{\Sigma_T}\phi|^2\right)\mathcal{G}d\mu_T.
\end{aligned} \tag{3.17}$$

Proof. Since $\phi \in C_c^\infty(Q_{R,T})$, $\phi(\cdot,t)$ is compactly supported in $\Sigma_t\backslash\bar{B}_R$ for each $t \in (0,T]$ and $\phi(\cdot,t) \equiv 0$ when $t < t_0$ for some small $t_0 > 0$. Thus, integrating (3.9) against the measure $d\mu_t dt$, the inequality (3.17) follows from the divergence theorem and integration by parts together with applications of the Cauchy-Schwarz inequality and the basic property for backwards mean curvature flow (see Appendix B of [10])

$$\frac{d}{dt}d\mu_t = H^2 d\mu_t. \tag{3.18}$$

Namely, it follows from the Cauchy-Schwarz inequality, i.e., $2ab \leq \varepsilon^{-2}b^2 + \varepsilon^2 a^2$, that

$$\begin{aligned}
&2\left(\frac{d\phi}{dt} - \langle\nabla_{\Sigma_t}\log\mathcal{G}, \nabla_{\Sigma_t}\phi\rangle + \frac{1}{2}\phi\mathcal{F}\right)\left(\frac{d\phi}{dt} + \Delta_{\Sigma_t}\phi\right) \\
&\leq 2\left(\frac{d\phi}{dt} - \langle\nabla_{\Sigma_t}\log\mathcal{G}, \nabla_{\Sigma_t}\phi\rangle + \frac{1}{2}\phi\mathcal{F}\right)^2 + \frac{1}{2}\left(\frac{d\phi}{dt} + \Delta_{\Sigma_t}\phi\right)^2.
\end{aligned} \tag{3.19}$$

Recall that, in our notations, d/dt is the total of the time derivatives, and thus, by the evolution of the volume element of mean curvature flow,

$$\begin{aligned}
&\int_0^T \int_{\Sigma_t} \frac{d}{dt}\left(|\nabla_{\Sigma_t}\phi|^2 \mathcal{G} + \frac{1}{2}\phi^2 \mathcal{F}\mathcal{G}\right) d\mu_t dt \\
&= \int_0^T \frac{d}{dt}\left\{\int_{\Sigma_t}\left(|\nabla_{\Sigma_t}\phi|^2 + \frac{1}{2}\phi^2\mathcal{F}\right)\mathcal{G} d\mu_t\right\} dt \\
&\quad - \int_0^T \int_{\Sigma_t}\left(|\nabla_{\Sigma_t}\phi|^2 + \frac{1}{2}\phi^2\mathcal{F}\right) H^2 \mathcal{G} d\mu_t dt \\
&= \int_{\Sigma_T} (|\nabla_{\Sigma_T}\phi|^2 + |\phi|^2 \mathcal{F})\mathcal{G} d\mu_T - \int_0^T \int_{\Sigma_t}\left(|\nabla_{\Sigma_t}\phi|^2 + \frac{1}{2}\phi^2\mathcal{F}\right) H^2 \mathcal{G} d\mu_t dt.
\end{aligned}$$

Therefore Lemma 3.3 follows from collecting terms involving ϕ and $\nabla_{\Sigma_t}\phi$ in the integrand. □

Next, we choose the function $\mathcal{G}$ in the previous lemma, which is a suitable variation of the choice in [15]. From now on, we fix a $\delta \in (0,1)$. Given $\alpha > 0$, $0 < T \le 1$ and $R > 2R_2$, we set $\mathcal{G}$ in Lemma 3.3 to be

$$\mathcal{G}_{\alpha,T,R}(x,t) = \exp\left[2\alpha(T-t)(|x|^{1+\delta} - R^{1+\delta}) + 2|x|^2\right]. \tag{3.20}$$

Similarly, set $\mathcal{F}_{\alpha,T,R} = (d\mathcal{G}_{\alpha,T,R}/dt - \Delta_{\Sigma_t}\mathcal{G}_{\alpha,T,R})/\mathcal{G}_{\alpha,T,R}$. Then, by straightforward calculations, we have that

Lemma 3.4. *There exist $\alpha_0 > 0$, $R_4 > 2R_2$ and $C_5 > 0$ such that for $\alpha \ge \alpha_0$, $0 < T \le 1$, $R > R_4$ and $(x,t) \in \cup_{t\in(0,T]}\Sigma_t \backslash B_R \times \{t\}$,*

$$\mathcal{F}_{\alpha,T,R}(x,T) < 0, \tag{3.21}$$

$$\frac{d}{dt}\mathcal{F}_{\alpha,T,R} + \Delta_{\Sigma_t}\mathcal{F}_{\alpha,T,R} + H^2\mathcal{F}_{\alpha,T,R} \ge 2, \tag{3.22}$$

$$2\nabla^2_{\Sigma_t}\log\mathcal{G}_{\alpha,T,R} + 2HA + H^2 g_t \ge C_5 \mathrm{Id}. \tag{3.23}$$

Proof. To prove the first property (3.21), we compute $\mathcal{F}_{\alpha,T,R}$ explicitly as follows. First, note that

$$\mathcal{F}_{\alpha,T,R} = \frac{d}{dt}\log\mathcal{G}_{\alpha,T,R} - \Delta_{\Sigma_t}\log\mathcal{G}_{\alpha,T,R} - |\nabla_{\Sigma_t}\log\mathcal{G}_{\alpha,T,R}|^2. \tag{3.24}$$

Then, we calculate each term on the right-hand side of the equation (3.24) accordingly. Since Σ is a self-shrinker of the mean curvature flow, we have that, for $x \in \Sigma_t$ and $0 < t \le 1$,

$$\partial_t x = H\mathbf{n} = \frac{\langle x, \mathbf{n}\rangle}{2t}\,\mathbf{n}. \tag{3.25}$$

Thus,

$$\begin{aligned}
\frac{d}{dt}\log\mathcal{G}_{\alpha,T,R} &= -2\alpha\left(|x|^{1+\delta} - R^{1+\delta}\right) + 2\alpha(1+\delta)(T-t)|x|^{\delta-1}\langle x, \partial_t x\rangle + 4\langle x, \partial_t x\rangle \\
&= -2\alpha\left(|x|^{1+\delta} - R^{1+\delta}\right) + 2\alpha(1+\delta)(T-t)|x|^{\delta-1} H\langle x, \mathbf{n}\rangle + 4H\langle x, \mathbf{n}\rangle \\
&= -2\alpha\left(|x|^{1+\delta} - R^{1+\delta}\right) + 4\alpha(1+\delta)t(T-t)|x|^{\delta-1} H^2 + 8tH^2. \qquad (3.26)
\end{aligned}$$

Note that

$$\nabla_{\Sigma_t}|x|^\beta = \beta|x|^{\beta-2}x^T, \tag{3.27}$$

where x^T is the tangential part of $x \in \Sigma_t$. Thus, by the equation (3.25),

$$\begin{aligned} |\nabla_{\Sigma_t} \log \mathcal{G}_{\alpha,T,R}|^2 &= \left[2\alpha(1+\delta)(T-t)\,|x|^{\delta-1}+4\right]^2 \left|x^T\right|^2 \\ &= \left[2\alpha(1+\delta)(T-t)\,|x|^{\delta-1}+4\right]^2 \left(|x|^2-4t^2H^2\right). \end{aligned} \tag{3.28}$$

It follows from (3.25) that

$$\mathrm{div}_{\Sigma_t} x^T = n - \langle x, \mathbf{n}\rangle \mathrm{div}_{\Sigma_t}\mathbf{n} = n - 2tH^2. \tag{3.29}$$

Thus, by (3.25), (3.27) and (3.29),

$$\begin{aligned} \Delta_{\Sigma_t} |x|^\beta &= \mathrm{div}_{\Sigma_t}\left(\beta\,|x|^{\beta-2}\,x^T\right) \\ &= \beta(\beta-2)\,|x|^{\beta-4}\left|x^T\right|^2 + \beta\,|x|^{\beta-2}\,\mathrm{div}_{\Sigma_t} x^T \\ &= \beta(\beta-2+n)\,|x|^{\beta-2} - 2\beta t\,|x|^{\beta-2}\,H^2 - 4\beta(\beta-2)t^2\,|x|^{\beta-4}\,H^2. \end{aligned} \tag{3.30}$$

Hence, (3.30) with $\beta = 1+\delta$ or 2 gives that

$$\begin{aligned} \Delta_{\Sigma_t} \log \mathcal{G}_{\alpha,T,R} &= 2\alpha(T-t)\Delta_{\Sigma_t}|x|^{1+\delta} + 2\Delta_{\Sigma_t}|x|^2 \\ &= 2\alpha(1+\delta)(\delta-1+n)(T-t)\,|x|^{\delta-1} - 4\alpha(1+\delta)t(T-t)\,|x|^{\delta-1}\,H^2 \\ &\quad + 8\alpha(1-\delta^2)t^2(T-t)\,|x|^{\delta-3}\,H^2 + 4n - 8tH^2. \end{aligned} \tag{3.31}$$

Therefore, combining the equations (3.26), (3.28) and (3.31), we conclude that

$$\begin{aligned} \mathcal{F}_{\alpha,T,R}(x,t) &= -2\alpha\left(|x|^{1+\delta} - R^{1+\delta}\right) - 2\alpha(1+\delta)(\delta-1+n)(T-t)\,|x|^{\delta-1} - 4n \\ &\quad - \left[2\alpha(1+\delta)(T-t)\,|x|^\delta + 4\,|x|\right]^2 + 16tH^2 \\ &\quad + 8\alpha(1+\delta)t(T-t)\,|x|^{\delta-1}\,H^2 - 8\alpha(1-\delta^2)t^2(T-t)\,|x|^{\delta-3}\,H^2 \\ &\quad + 4t^2\left[2\alpha(1+\delta)(T-t)\,|x|^{\delta-1}+4\right]^2 H^2. \end{aligned} \tag{3.32}$$

In particular, by (3.32) and Lemma 2.1, there exists $\lambda_1 > 0$, depending only on C_1 and n, such that when $t = T$, $R > 2R_2$ and $x \in \Sigma_T \backslash B_R$,

$$\mathcal{F}_{\alpha,T,R}(x,T) < -4n - 16\,|x|^2 + 16TH^2 + 64T^2H^2 \le -\lambda_1 R^2. \tag{3.33}$$

Next, in order to estimate derivatives of $\mathcal{F}_{\alpha,T,R}$, we group the terms in $\mathcal{F}_{\alpha,T,R}$ according to their powers of α. Thus we write

$$\mathcal{F}_{\alpha,T,R} = B_0 + \alpha B_1 + \alpha^2 B_2, \tag{3.34}$$

and we will analyze each B_i in the following. Namely,

$$B_0 = -4n - 16\,|x|^2 + 16tH^2 + 64t^2H^2, \tag{3.35}$$

and we observe that

$$\frac{d\,|x|}{dt} = 2t\,|x|^{-1}\,H^2 \quad \text{and} \quad \frac{dH}{dt} + \Delta_{\Sigma_t} H = -\,|A|^2\,H. \tag{3.36}$$

Here the evolution equation for H is obtained by the computation in Appendix B of [10]. Thus, it follows from Lemma 2.1 that there exists $\lambda_2 > 0$, depending only on R_2 and C_1, such that for $x \in \Sigma_t \backslash B_{2R_2}$,

$$\left|\frac{d\,|x|}{dt}\right| + \left|\frac{dH}{dt}\right| \le \lambda_2\,|x|^{-3}, \tag{3.37}$$

together with (3.35), implies that

$$\left|\frac{dB_0}{dt}\right| \leq O(|x|^{-2}). \tag{3.38}$$

For B_1, we extract out the constant term and the term with the highest power of $|x|$ and denote the rest terms by P_1. That is,

$$B_1 = -2\left[1 + 8(1+\delta)(T-t)\right]|x|^{1+\delta} + 2R^{1+\delta} + P_1, \tag{3.39}$$

where the highest power of $|x|$ in P_1 is equal to $\delta - 1$. Note that P_1 can be viewed as a polynomial of t, $|x|$ and H. Also, by Lemma 2.1 and (3.36), the time derivatives of $|x|$ and H are uniformly bounded in some neighborhood of infinity. Thus, $|dP_1/dt|$ decays at least as fast as $|x|^{\delta-1}$. On the other hand,

$$\begin{aligned} &\frac{d}{dt}\left\{-2\left[1+8(1+\delta)(T-t)\right]|x|^{1+\delta}\right\} \\ &= 16(1+\delta)|x|^{1+\delta} - 4(1+\delta)\left[1+8(1+\delta)(T-t)\right]t\,|x|^{\delta-1}H^2. \end{aligned} \tag{3.40}$$

Hence there exists $r_1 > 2R_2$, depending only on R_2 and C_1, such that for $R > r_1$ and $x \in \Sigma_t \backslash B_R$,

$$\frac{dB_1}{dt} \geq 8(1+\delta)|x|^{1+\delta}. \tag{3.41}$$

By similar reasoning, one can enlarge r_1 (if necessary) to conclude from (3.36) that for $R > r_1$ and $x \in \Sigma_t \backslash B_R$,

$$\frac{dB_2}{dt} \geq 4(1+\delta)^2(T-t)|x|^{2\delta}, \tag{3.42}$$

where B_2 is given by

$$B_2 = -4(1+\delta)^2(T-t)^2|x|^{2\delta} + 16(1+\delta)^2(T-t)^2t^2|x|^{2\delta-2}H^2. \tag{3.43}$$

Therefore, combining (3.38), (3.41) and (3.42), there exist $\alpha_1 > 1$ and $r_1 > 2R_2$, depending only on C_1 and R_2, such that for $\alpha > \alpha_1$, $R > r_1$ and $x \in \Sigma_t \backslash B_R$,

$$\frac{d}{dt}\mathcal{F}_{\alpha,T,R}(x,t) \geq 4\alpha(1+\delta)|x|^{1+\delta}\left[\alpha(1+\delta)(T-t)|x|^{\delta-1} + 1\right]. \tag{3.44}$$

To estimate $\Delta_{\Sigma_t}\mathcal{F}_{\alpha,T,R}$, we first note that

$$\Delta_{\Sigma_t}H^2 = 2H\Delta_{\Sigma_t}H + 2\left|\nabla_{\Sigma_t}H\right|^2, \tag{3.45}$$

and for $\beta \neq 0$,

$$\Delta_{\Sigma_t}(|x|^\beta H^2) = H^2\Delta_{\Sigma_t}|x|^\beta + 2\left\langle \nabla_{\Sigma_t}H^2, \nabla_{\Sigma_t}|x|^\beta\right\rangle + |x|^\beta\Delta_{\Sigma_t}H^2. \tag{3.46}$$

Assuming that $R > 2R_2$ and $x \in \Sigma_t \backslash B_R$, then Lemma 2.1, (3.27) and (3.30) give that, at x,

$$\left|\Delta_{\Sigma_t}H^2\right| \leq \lambda_3|x|^{-4} \quad \text{and} \quad \left|\Delta_{\Sigma_t}\left(|x|^\beta H^2\right)\right| \leq \lambda_4|x|^{\beta-4}, \tag{3.47}$$

where $\lambda_3 > 0$ depends only on C_1 and $\lambda_4 > 0$ depends on C_1, n and β. Since each nonconstant term in $\mathcal{F}_{\alpha,T,R}$ is either $|x|^\beta$, H^2, or $|x|^\beta H^2$, a straightforward computation and (3.47) give that, for $x \in \Sigma_t \backslash B_{2R_2}$,

$$|\Delta_{\Sigma_t}\mathcal{F}_{\alpha,T,R}| \leq \lambda_5\alpha^2(T-t)|x|^{2\delta-2} + \lambda_5\alpha, \tag{3.48}$$

where $\lambda_5 > 0$ depends only on C_1, R_2, λ_3 and λ_4. Also, for $x \in \Sigma_t \backslash B_{2R_2}$, it follows from Lemma 2.1 that

$$\left|H^2 \mathcal{F}_{\alpha,T,R}\right| \leq \lambda_6 \alpha^2 (T-t) \left|x\right|^{2\delta-2} + \lambda_6 \alpha, \tag{3.49}$$

where $\lambda_6 > 0$ depends on R_2 and C_1. Hence, the second property (3.22) is verified by (3.44), (3.48) and (3.49), when $t \in (0,1]$.

Finally, we estimate $\nabla^2_{\Sigma_t} \left|x\right|^\beta$ for $\beta \in (1,2]$. Fix $x \in \Sigma_t$, $0 < t \leq 1$. We choose a local geodesic orthonormal frame $\{e_1, \ldots, e_n\}$ of Σ_t at x. Thus, (3.25) gives that, at x,

$$\begin{aligned}
&\nabla^2_{\Sigma_t} \left|x\right|^\beta (e_i, e_j) \\
&= \left\langle \nabla_{e_i} \nabla_{\Sigma_t} \left|x\right|^\beta, e_j \right\rangle = \beta \left\langle \nabla_{e_i} \left(\left|x\right|^{\beta-2} x^T\right), e_j \right\rangle \\
&= \beta \left\langle \nabla_{e_i} \left(\left|x\right|^{\beta-2} x\right), e_j \right\rangle - \beta \left|x\right|^{\beta-2} \left\langle x, \mathbf{n}\right\rangle \left\langle \nabla_{e_i} \mathbf{n}, e_j \right\rangle \\
&= \beta \left\langle D \left|x\right|^{\beta-2}, e_i \right\rangle \left\langle x, e_j \right\rangle + \beta \left|x\right|^{\beta-2} \left\langle \nabla_{e_i} x, e_j \right\rangle + 2\beta t \left|x\right|^{\beta-2} HA(e_i, e_j) \\
&= \beta(\beta-2) \left|x\right|^{\beta-4} \left\langle x, e_i \right\rangle \left\langle x, e_j \right\rangle + \beta \left|x\right|^{\beta-2} \delta_{ij} + 2\beta t \left|x\right|^{\beta-2} HA(e_i, e_j).
\end{aligned} \tag{3.50}$$

Now, given $\eta \in T_x \Sigma_t$, $\eta = \sum_k \eta_k e_k$ and at x, it follows from (3.50) that

$$\begin{aligned}
&\nabla^2_{\Sigma_t} \left|x\right|^\beta (\eta, \eta) \\
&\geq \beta \left|x\right|^{\beta-4} \sum_{i,j} \left[(\beta-2)\langle x, e_i\rangle\langle x, e_j\rangle + \left|x^T\right|^2 \delta_{ij}\right] \eta_i \eta_j + 2\beta t \left|x\right|^{\beta-2} HA(\eta, \eta) \\
&= \beta(\beta-2) \left|x\right|^{\beta-4} \left(\sum_i \langle x, e_i\rangle \eta_i\right)^2 + \beta \left|x\right|^{\beta-4} \left|x^T\right|^2 \left|\eta\right|^2 + 2\beta t \left|x\right|^{\beta-2} HA(\eta, \eta) \\
&\geq \beta(\beta-1) \left|x\right|^{\beta-4} \left|x^T\right|^2 \left|\eta\right|^2 + 2\beta t \left|x\right|^{\beta-2} HA(\eta, \eta) \\
&= \beta(\beta-1) \left|x\right|^{\beta-2} \left|\eta\right|^2 - 4\beta(\beta-1) t^2 \left|x\right|^{\beta-4} H^2 \left|\eta\right|^2 + 2\beta t \left|x\right|^{\beta-2} HA(\eta, \eta).
\end{aligned} \tag{3.51}$$

Hence, by Lemma 2.1 and the assumption that $\beta \in (1,2]$, there exists $r_2 > 2R_2$, depending only on β, C_1 and R_2, such that for $x \in \Sigma_t \backslash B_{r_2}$,

$$\nabla^2_{\Sigma_t} \left|x\right|^\beta (\eta, \eta) \geq \frac{1}{2} \beta(\beta-1) \left|x\right|^{\beta-2} \left|\eta\right|^2. \tag{3.52}$$

Hence, the third property (3.23) follows immediately from (3.52) with $\beta = 1 + \delta$ or 2, and Lemma 2.1, when $t \in (0,1]$. □

Finally, combining Lemma 3.3 and Lemma 3.4, we establish the following Carleman inequality.

Proposition 3.5. *Let $\alpha > \alpha_0$, $0 < T \leq 1$, $R > R_4$ and $Q_{R,T}$ as in Lemma 3.3. Then for all $\phi \in C_c^\infty(Q_{R,T})$,*

$$\begin{aligned}
\int_0^T \int_{\Sigma_t} \left(\left|\phi\right|^2 + C_5 \left|\nabla_{\Sigma_t} \phi\right|^2\right) \mathcal{G}_{\alpha,T,R} d\mu_t dt \leq &\int_0^T \int_{\Sigma_t} \left(\frac{d\phi}{dt} + \Delta_{\Sigma_t} \phi\right)^2 \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&+ \int_{\Sigma_T} \left|\nabla_{\Sigma_T} \phi\right|^2 \mathcal{G}_{\alpha,T,R} d\mu_T.
\end{aligned} \tag{3.53}$$

Since the weight function $\mathcal{G}_{\alpha,T,R}$ in Proposition 3.5 grows very fast in space, in order to apply Proposition 3.5 to conclude backwards uniqueness, we need to show that $|w|$ and its gradient $|\nabla_{\Sigma_t} w|$ decay at least as rapidly as $\mathcal{G}_{\alpha,T,R}$.

Lemma 3.6. *There exist $R_5 > R_3$ and $M > 0$ such that for $x \in \Sigma_t \backslash B_{R_5}$ and $t \in (0,1]$,*

$$|w(x,t)| + |\nabla_{\Sigma_t} w(x,t)| \leq \exp\left(-\frac{M\,|x|^2}{t}\right). \tag{3.54}$$

Proof. Fix $z_0 \in C \backslash B_{2R_3}$. By Lemma 2.2, for each $t \in (0,1]$, the component of $\Sigma_t \cap B_{\varepsilon_0|z_0|}(z_0)$ containing $z_0 + U(z_0,t)\mathbf{n}(z_0)$ can be written as the graph of a smooth function $u(\cdot,t)$ over the tangent plane $T_{z_0}C$ of C at z_0 satisfying the property (2.5). Moreover, there exists $\delta_1 \in (0,\varepsilon_0]$, depending only on C_2, such that the image of the orthogonal projection of $\Sigma_t \cap B_{\varepsilon_0|z_0|}(z_0)$ to $T_{z_0}C$, $0 < t \leq 1$ contains the disk in $T_{z_0}C$ centered at z_0 with radius $\delta_1\,|z_0|$. We parametrize $T_{z_0}C$ by $F : \mathbb{R}^n \longrightarrow T_{z_0}C$, $F(p) = z_0 + \sum_i p_i e_i$, where $p = (p_1, \ldots, p_n)$ and $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $T_{z_0}C - z_0$. And we identify $u(p,t)$ with $u(F(p),t)$ and define $\bar{w}(p,t) = w(F(p) + u(p,t)\mathbf{n}(z_0), t)$. Let g_t be the pullback metric on $D_{\delta_1|z_0|}$ from Σ_t via the map $p \mapsto F(p) + u(p,t)\mathbf{n}(z_0)$. In the following, ∂_t denotes the partial derivative with respect to time t fixing p. Then, it follows from Lemma 2.4 that, on $D_{\delta_1|z_0|} \times (0,1]$,

$$\begin{aligned}
&\partial_t \bar{w} + \Delta_{g_t} \bar{w} \\
&\quad = \frac{1}{2\sqrt{t}}\, v\left(\frac{x}{\sqrt{t}}\right) + \sqrt{t}\left\langle \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right), \frac{1}{\sqrt{t}}\frac{\partial x}{\partial t} - \frac{x}{2t\sqrt{t}}\right\rangle + \frac{1}{\sqrt{t}}\,\Delta_\Sigma v\left(\frac{x}{\sqrt{t}}\right) \\
&\quad = \frac{1}{\sqrt{t}}\left[\frac{1}{2} v\left(\frac{x}{\sqrt{t}}\right) - \frac{1}{2}\left\langle \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right), \frac{x}{\sqrt{t}}\right\rangle + \Delta_\Sigma v\left(\frac{x}{\sqrt{t}}\right)\right] \\
&\qquad + \left\langle \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right), \frac{\partial x}{\partial t}\right\rangle \\
&\quad = \frac{1}{\sqrt{t}}\left[-|A|^2\, v\left(\frac{x}{\sqrt{t}}\right) + Q\left(\frac{x}{\sqrt{t}}, v, \nabla_\Sigma v, \nabla^2_\Sigma v\right)\right] + \left\langle \nabla_\Sigma v\left(\frac{x}{\sqrt{t}}\right), \frac{\partial x}{\partial t}\right\rangle.
\end{aligned} \tag{3.55}$$

Here, we use the backwards mean curvature flow equation for graphs, that is, for $x = F(p) + u(p,t)\mathbf{n}(z_0)$ and $p \in D_{\delta_1|z_0|}$,

$$\langle \partial_t x, \mathbf{n}\rangle = H, \tag{3.56}$$

and thus equivalently,

$$\partial_t u = -\sqrt{1 + |Du|^2}\ \mathrm{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right). \tag{3.57}$$

Hence, by the calculation (3.55), Lemmas 2.1, 2.2 and 2.4, we have that, on $D_{\delta_1|z_0|} \times (0,1]$,

$$|\partial_t \bar{w} + \Delta_{g_t} \bar{w}| \leq \lambda_1\,|z_0|^{-2}\left(|\bar{w}| + |z_0|\,|\nabla_{g_t} \bar{w}|\right), \tag{3.58}$$

where $\lambda_1 > 0$ depends only on C_1, C_2, C_4 and δ_1. It is easy to verify that g_t has the following properties: $g_0(0) = \mathrm{Id}$,

$$\lambda_2^{-1} |\xi|^2 \le \sum_{i,j} g_t^{ij} \xi_i \xi_j \le \lambda_2 |\xi|^2 \quad \text{for all} \quad \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n, \tag{3.59}$$

$$\left| D g_t^{ij} \right| \le \lambda_2 |z_0|^{-1} \quad \text{and} \quad \left| \partial_t g_t^{ij} \right| \le \lambda_2 |z_0|^{-2} \quad \text{for} \quad 1 \le i, j \le n, \tag{3.60}$$

where $\lambda_2 > 0$ depends only on n and C_2. Thus, (3.58) gives that

$$\left| \partial_t \bar{w} + \operatorname{div} \left(g_t^{-1} D \bar{w} \right) \right| \le \lambda_3 |z_0|^{-2} \left(|\bar{w}| + |z_0| |D\bar{w}| \right), \tag{3.61}$$

where $\lambda_3 > 0$ depends only on λ_1 and λ_2.

For the convenience of the readers, we present the arguments in pages 2877–2879 of [37] here to conclude the proof of Lemma 3.6. Let $R = \delta_1 |z_0| / 2$ and $\bar{w}_R$ be a rescaling of $\bar{w}$, $\bar{w}_R(q, s) = \bar{w}(Rq, R^2 s)$. Then $\bar{w}_R$ satisfies

$$\left| \partial_s \bar{w}_R + \operatorname{div} \left(g_R^{-1} D \bar{w}_R \right) \right| \le \lambda_3 \left(|\bar{w}_R| + |D\bar{w}_R| \right), \tag{3.62}$$

where $g_R^{-1}(q, s) = (g_R^{ij}) = g_{R^2 s}^{-1}(Rq)$. It follows from (3.60) that, in $D_2 \times [0, 1/R^2]$,

$$\left| D g_R^{ij} \right| \le \lambda_2 \quad \text{and} \quad \left| \partial_s g_R^{ij} \right| \le \lambda_2 \quad \text{for} \quad 1 \le i, j \le n. \tag{3.63}$$

Furthermore, by Lemma 2.3 and (3.59),

$$|\bar{w}_R| + |D\bar{w}_R| \le \lambda_4 |z_0|^{-1} \tag{3.64}$$

in $D_2 \times [0, 1/R^2]$, where $\lambda_4 > 0$ depends only on C_3 and λ_2. And $\bar{w}_R, D\bar{w}_R \in C^0(D_2 \times [0, 1/R^2])$ with $\bar{w}_R(\cdot, 0) = 0$, and $\bar{w}_R$ is smooth in $D_2 \times (0, 1/R^2]$. Define $\phi = \psi(q)\eta(s)\bar{w}_R$, where $\chi_{D_1} \le \psi \le \chi_{D_2}$ and $\chi_{[0,1/\alpha]} \le \eta \le \chi_{[0,2/\alpha)}$ are bump functions, and $\alpha \ge 2R^2$ is a positive constant to be chosen. Then

$$\left| \partial_s \phi + \operatorname{div} \left(g_R^{-1} D \phi \right) \right| \le \lambda_3 \left(|\phi| + |D\phi| \right) + \lambda_5 \alpha \left(|\bar{w}_R| + |D\bar{w}_R| \right) \chi_E. \tag{3.65}$$

Here $E = (D_2 \times [0, 2/\alpha)) \setminus (D_1 \times [0, 1/\alpha])$, and $\lambda_5 > 0$ depends only on R_3, δ_1, λ_2 and λ_3. Hence, using the Carleman inequality in [13] (see also Lemma 2.1 in [37]), there exists $M_1 > 1$, depending only on n and λ_2, such that for $0 < a < 1/\alpha$,

$$\begin{aligned} &\int_{\mathbb{R}^{n+1}} \left(\alpha^2 \phi^2 + \alpha \sigma_a |D\phi|^2 \right) \sigma_a^{-\alpha} \mathcal{G}_a dq ds \\ &\quad \le \alpha^\alpha M_1^\alpha \sup_{s \ge 0} \int_{\mathbb{R}^n \times \{s\}} \left(\phi^2 + |D\phi|^2 \right) dq + M_1 \lambda_3^2 \int_{\mathbb{R}^{n+1}} \left(|\phi| + |D\phi| \right)^2 \sigma_a^{1-\alpha} \mathcal{G}_a dq ds \\ &\qquad + M_1 \lambda_5^2 \alpha^2 \int_E \left(|\bar{w}_R| + |D\bar{w}_R| \right)^2 \sigma_a^{1-\alpha} \mathcal{G}_a dq ds. \end{aligned}$$

Here $\mathcal{G}_a(q, s) = (s + a)^{-n/2} \exp\left[-|q|^2 / 4(s + a) \right]$, $\sigma_a(s) = \sigma(s + a)$ and $\sigma : (0, 4/\alpha) \longrightarrow (0, +\infty)$ satisfying that $M_1^{-1} s \le \sigma(s) \le s$. If $\alpha \ge 2M_1 \lambda_3$, then the second term on the right-hand side can be absorbed by the left-hand side. In E, $\sigma_a^{-\alpha} \mathcal{G}_a \le \alpha^{\alpha + \frac{n}{2}} M_1^\alpha$. Hence, by (3.64), we get

$$\int_{\mathbb{R}^{n+1}} \left(\alpha^2 \phi^2 + \alpha \sigma_a |D\phi|^2 \right) \sigma_a^{-\alpha} \mathcal{G}_a dq ds \le \alpha^{\alpha + \frac{n}{2}} M_2^\alpha, \tag{3.66}$$

where $M_2 > M_1$ depends only on R_3, λ_4, λ_5 and M_1. Let $\rho = 1/(M_2 e)$ and $a = \rho^2/(2\alpha)$. Then, in $D_{2\rho} \times [0, \rho^2/(2\alpha)]$,

$$\sigma_a^{1-\alpha} \mathcal{G}_a \ge \alpha^{\alpha + \frac{n}{2} - 1} M_2^{2\alpha + n - 2}. \tag{3.67}$$

Therefore, we deduce from (3.66) that

$$\int_{D_{2\rho}\times[0,\rho^2/(2\alpha)]} \left(\phi^2 + |D\phi|^2\right) dqds \le M_2^{2-\alpha-n}. \tag{3.68}$$

We now choose $\alpha = M_3R^2$, where $M_3 > 1$ depends only on R_3, δ_1, λ_3 and M_2, so that

$$\int_{D_{2\rho}\times[0,\rho^2/(2M_3R^2)]} \left(\phi^2 + |D\phi|^2\right) dqds \le M_2^{2-n}\mathrm{e}^{-2R^2}. \tag{3.69}$$

By Lemma 4.1 in [37], this implies that, in $D_\rho \times [0, \rho^2/(4M_3R^2)]$,

$$|\phi| + |D\phi| \le \lambda_6 R^c \mathrm{e}^{-R^2}, \tag{3.70}$$

where $\lambda_6 > 0$ depends on n, λ_2 and λ_3, and $c > 0$ depends on n. Undoing the change of variables, we get

$$|w(z_0,t)| + |\nabla_{\Sigma_t} w(z_0,t)| \le \lambda_6 \left|z_0\right|^c \mathrm{e}^{-\delta_1^2|z_0|^2}, \tag{3.71}$$

if $0 \le t \le \rho^2/(4M_3)$.

Therefore, it follows from the definition of w that there exist $M_4 > 0$ and $r_1 > 0$, depending only on R_3, δ_1, M_2, M_3, λ_6 and c, such that for $x \in \Sigma\backslash B_{r_1}$,

$$|v(x)| + |\nabla_\Sigma v(x)| \le \exp\left(-M_4 \left|x\right|^2\right). \tag{3.72}$$

It is clear that Lemma 3.6 follows immediately from the inequality (3.72). □

Now, we are ready to conclude the proof of Theorem 1.3.

Proof of Theorem 1.3. We basically follow the arguments in [15]. Choose ε_1 in Lemma 3.1 to be $(1 + C_5)/4$. Then, choose R large such that each lemma and proposition can be applied, and $T = M/16$.

Given $a \in (0,1)$ and $r \gg 1$, we consider $\psi_{a,r} \in C_c^\infty(\mathbb{R}^{n+1})$ satisfying that $\psi_{a,r} \equiv 1$ when $(1+2a)R \le |x| \le r$, $\psi_{a,r} \equiv 0$ when $|x| \le (1+a)R$ or $|x| \ge 2r$, $0 \le \psi_{a,r} \le 1$ and $|D\psi_{a,r}|$, $\left|D^2\psi_{a,r}\right|$ are bounded from above by a function of a and R. We also need to introduce a cutoff function in the time direction: given $0 < \varepsilon < T/8$, $\eta_\varepsilon \in C_c^\infty(\mathbb{R})$ satisfies that $\eta_\varepsilon \equiv 1$ when $t \ge 2\varepsilon$, $\eta_\varepsilon \equiv 0$ when $t \le \varepsilon$, $0 \le \eta_\varepsilon \le 1$ and $|\eta_\varepsilon'| \le 2\varepsilon^{-1}$. Then we choose the test function ϕ in Proposition 3.5 to be $\phi_{\varepsilon,a,r} = \eta_\varepsilon \psi_{a,r} w$. Thus, we get that

$$\begin{aligned}\frac{d}{dt}\phi_{\varepsilon,a,r} + \Delta_{\Sigma_t}\phi_{\varepsilon,a,r} &= \eta_\varepsilon\psi_{a,r}\left(\frac{dw}{dt} + \Delta_{\Sigma_t} w\right) + 2\langle\nabla_{\Sigma_t}\psi_{a,r}, \nabla_{\Sigma_t} w\rangle\eta_\varepsilon \\ &\quad + \left(\langle D\psi_{a,r}, \partial_t x\rangle + \Delta_{\Sigma_t}\psi_{a,r}\right)\eta_\varepsilon w + \eta_\varepsilon'\psi_{a,r} w.\end{aligned} \tag{3.73}$$

Hence, by Lemma 3.1,

$$\begin{aligned}\left|\frac{d}{dt}\phi_{\varepsilon,a,r} + \Delta_{\Sigma_t}\phi_{\varepsilon,a,r}\right|^2 &\le \frac{1}{2}\left(\left|\phi_{\varepsilon,a,r}\right|^2 + C_5\left|\nabla_{\Sigma_t}\phi_{\varepsilon,a,r}\right|^2\right) \\ &\quad + \lambda\left(|w|^2 + |\nabla_{\Sigma_t} w|^2\right)\chi_{a,R,r} + 16\varepsilon^{-2}|w|^2\tilde{\chi}_{\varepsilon,R,r},\end{aligned} \tag{3.74}$$

where $\lambda > 0$ depends only on n, C_1, a and R; $\chi_{a,R,r}$ is the characteristic function supported in $(B_{2r}\backslash B_r) \cup \left(B_{(1+2a)R}\backslash B_{(1+a)R}\right)$; and $\tilde{\chi}_{\varepsilon,R,r}$ is the characteristic function supported in $(B_{2r}\backslash B_R) \times (\varepsilon, 2\varepsilon)$.

Hence, by Proposition 3.5 and the inequality (3.74), for $\alpha > \alpha_0$,

$$\begin{aligned}
&\int_0^T \int_{\Sigma_t} |\phi_{\varepsilon,a,r}|^2 \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\lesssim \int_0^T \int_{\Sigma_t \cap B_{(1+2a)R} \backslash B_{(1+a)R}} \left(|w|^2 + |\nabla_{\Sigma_t} w|^2\right) \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\quad + \int_0^T \int_{\Sigma_t \cap B_{2r} \backslash B_r} \left(|w|^2 + |\nabla_{\Sigma_t} w|^2\right) \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\quad + \varepsilon^{-2} \int_\varepsilon^{2\varepsilon} \int_{\Sigma_t \cap B_{2r} \backslash B_R} |w|^2 \mathcal{G}_{\alpha,T,R} d\mu_t dt + \int_{\Sigma_T} |\nabla_{\Sigma_T} \phi_{\varepsilon,a,r}|^2 \mathcal{G}_{\alpha,T,R} d\mu_T,
\end{aligned} \tag{3.75}$$

where $\lesssim$ stands for being less than some positive multiple of the quantities that follow, which depends only on n, λ, a, R and T. By the assumption of Theorem 1.3, $\Sigma_t \backslash B_R$ flows backwards by mean curvature smoothly to the cone $C \backslash B_R$ as $t \longrightarrow 0$. This implies that there exists $V_0 > 0$, depending only on the geometry of C, such that if t is sufficiently small, then $\mathrm{Vol}(\Sigma_t \cap B_{2r} \backslash B_R) \leq V_0 r^n$. Also, it follows from (2.9) that $|w|(x,t) \leq C_3 t$ for $x \in \Sigma_t \cap B_{2r} \backslash B_R$. Thus, letting $\varepsilon \longrightarrow 0$, the penultimate integral on the right-hand side of (3.75) converges to zero and the monotone convergence theorem implies that

$$\begin{aligned}
&\int_0^T \int_{\Sigma_t} |\psi_{a,r} w|^2 \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\lesssim \int_0^T \int_{\Sigma_t \cap B_{(1+2a)R} \backslash B_{(1+a)R}} \left(|w|^2 + |\nabla_{\Sigma_t} w|^2\right) \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\quad + \int_0^T \int_{\Sigma_t \cap B_{2r} \backslash B_r} \left(|w|^2 + |\nabla_{\Sigma_t} w|^2\right) \mathcal{G}_{\alpha,T,R} d\mu_t dt \\
&\quad + \int_{\Sigma_T} |\nabla_{\Sigma_T} (\psi_{a,r} w)|^2 \mathcal{G}_{\alpha,T,R} d\mu_T.
\end{aligned} \tag{3.76}$$

Furthermore, it follows from the fact that $M/T > 8$, Lemma 3.6 and the inequality (3.76) that

$$\begin{aligned}
&\exp\left\{\left[(1+17a)^{1+\delta} - 1\right] \alpha T R^{1+\delta}\right\} \int_0^{T/2} \int_{\Sigma_t \cap B_r \backslash B_{(1+17a)R}} w^2 d\mu_t dt \\
&\lesssim r^2 \exp\left(2^{2+\delta} \alpha T r^{1+\delta} - \frac{M r^2}{2T}\right) + \exp\left\{2\left[(1+2a)^{1+\delta} - 1\right] \alpha T R^{1+\delta}\right\} + 1.
\end{aligned} \tag{3.77}$$

Note that $(1+17a)^{1+\delta} - 1 \geq 17a$ and $(1+2a)^{1+\delta} - 1 \leq 8a$. Thus, letting $r \longrightarrow \infty$,

$$\int_0^{T/2} \int_{\Sigma_t \backslash B_{(1+17a)R}} w^2 d\mu_t dt \lesssim \exp\left(-\alpha a T R^{1+\delta}\right). \tag{3.78}$$

Then, let $\alpha \longrightarrow \infty$ and thus,

$$\int_0^{T/2} \int_{\Sigma_t \backslash B_{(1+17a)R}} w^2 d\mu_t dt \leq 0. \tag{3.79}$$

By the arbitrariness of a, we conclude that $w \equiv 0$ on $Q_{R,T/2}$ and thus $v \equiv 0$ on $\Sigma \backslash B_{2R/\sqrt{T}}$.

Therefore, Theorem 1.3 follows immediately by applying the strong continuation theorem (cf. [17, 18]) to L_Σ inside the compact set, and an open and closed argument. □

Acknowledgments

The author is indebted to Brett Kotschwar for bringing the paper [15] to her attention and many inspiring discussions which stimulated the present paper. And the author is very grateful to Tom Ilmanen for many useful suggestions which led to the proof of Lemma 2.2. The author would also like to thank Jacob Bernstein, Tobias Colding and the anonymous referees for their interest in this work and many constructive comments to improve the expositions. Finally, special thanks to the author's parents for their love and support. The author was supported in part by the postdoctoral fellowship at MSRI in fall 2011 and the AMS-Simons Travel Grant 2012–2014.

References

[1] U. Abresch and J. Langer, *The normalized curve shortening flow and homothetic solutions*, J. Differential Geom. **23** (1986), no. 2, 175–196. MR845704 (88d:53001)

[2] Sigurd B. Angenent, *Shrinking doughnuts*, (Gregynog, 1989), Progr. Nonlinear Differential Equations Appl. vol. 7, Birkhäuser Boston, Boston, MA, 1992, pp. 21–38. MR1167827 (93d:58032)

[3] Xu Cheng and Detang Zhou, *Volume estimate about shrinkers*, Proc. Amer. Math. Soc. **141** (2013), no. 2, 687–696, DOI 10.1090/S0002-9939-2012-11922-7. MR2996973

[4] David L. Chopp, *Computation of self-similar solutions for mean curvature flow*, Experiment. Math. **3** (1994), no. 1, 1–15. MR1302814 (95j:53006)

[5] Tobias H. Colding and William P. Minicozzi II, *Generic mean curvature flow I: generic singularities*, Ann. of Math. (2) **175** (2012), no. 2, 755–833, DOI 10.4007/annals.2012.175.2.7. MR2993752

[6] Tobias H. Colding and William P. Minicozzi II, *Smooth compactness of self-shrinkers*, Comment. Math. Helv. **87** (2012), no. 2, 463–475, DOI 10.4171/CMH/260. MR2914856

[7] Tobias H. Colding and William P. Minicozzi II, *Minimal surfaces*, Courant Lecture Notes in Mathematics, vol. 4, New York University Courant Institute of Mathematical Sciences, New York, 1999. MR1683966 (2002b:49072)

[8] Celso J. Costa, *Example of a complete minimal immersion in* $\mathbf{R}^3$ *of genus one and three embedded ends*, Bol. Soc. Brasil. Mat. **15** (1984), no. 1-2, 47–54, DOI 10.1007/BF02584707. MR794728 (87c:53111)

[9] Q. Ding and Y.L. Xin, *Volume growth, eigenvalue and compactness for self-shrinkers*, Asian J. Math. **17** (2013), no. 3, 443–456.

[10] Klaus Ecker, *Regularity theory for mean curvature flow*, Progress in Nonlinear Differential Equations and their Applications, 57, Birkhäuser Boston Inc. Boston, MA, 2004. MR2024995 (2005b:53108)

[11] Klaus Ecker and Gerhard Huisken, *Mean curvature evolution of entire graphs*, Ann. of Math. (2) **130** (1989), no. 3, 453–471, DOI 10.2307/1971452. MR1025164 (91c:53006)

[12] Klaus Ecker and Gerhard Huisken, *Interior estimates for hypersurfaces moving by mean curvature*, Invent. Math. **105** (1991), no. 3, 547–569, DOI 10.1007/BF01232278. MR1117150 (92i:53010)

[13] Luis Escauriaza and Francisco Javier Fernández, *Unique continuation for parabolic operators*, Ark. Mat. **41** (2003), no. 1, 35–60, DOI 10.1007/BF02384566. MR1971939 (2004b:35136)

[14] L. Escauriaza, C. E. Kenig, G. Ponce, and L. Vega, *Decay at infinity of caloric functions within characteristic hyperplanes*, Math. Res. Lett. **13** (2006), no. 2-3, 441–453, DOI 10.4310/MRL.2006.v13.n3.a8. MR2231129 (2007e:35125)

[15] L. Escauriaza, G. Seregin, and V. Šverák, *Backward uniqueness for parabolic equations*, Arch. Ration. Mech. Anal. **169** (2003), no. 2, 147–157, DOI 10.1007/s00205-003-0263-8. MR2005639 (2005j:35097)

[16] L. Escauriaza, G. Seregin, and V. Šverák, *On backward uniqueness for parabolic equations* (English, with English and Russian summaries), Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **288** (2002), no. Kraev. Zadachi Mat. Fiz. i Smezh. Vopr. Teor. Funkts. 32, 100–103, 272, DOI 10.1023/B:JOTH.0000041475.11233.d8; English transl., J. Math. Sci. (N.Y.) **123** (2004), no. 6, 4577–4579. MR1923546 (2003i:35225)

[17] Nicola Garofalo and Fang-Hua Lin, *Unique continuation for elliptic operators: a geometric-variational approach*, Comm. Pure Appl. Math. **40** (1987), no. 3, 347–366, DOI 10.1002/cpa.3160400305. MR882069 (88j:35046)

[18] Nicola Garofalo and Fang-Hua Lin, *Monotonicity properties of variational integrals, A_p weights and unique continuation*, Indiana Univ. Math. J. **35** (1986), no. 2, 245–268, DOI 10.1512/iumj.1986.35.35015. MR833393 (88b:35059)

[19] David Hoffman and William H. Meeks III, *Embedded minimal surfaces of finite topology*, Ann. of Math. (2) **131** (1990), no. 1, 1–34, DOI 10.2307/1971506. MR1038356 (91i:53010)

[20] Gerhard Huisken, *Local and global behaviour of hypersurfaces moving by mean curvature*, Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), Proc. Sympos. Pure Math. vol. 54, Amer. Math. Soc. Providence, RI, 1993, pp. 175–191. MR1216584 (94c:58037)

[21] Gerhard Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geom. **31** (1990), no. 1, 285–299. MR1030675 (90m:53016)

[22] T. Ilmanen, Personal communication.

[23] T. Ilmanen, *Lectures on mean curvature flow and related equations* (1995). unpublished notes.

[24] T. Ilmanen, *Singularities of mean curvature flow of surfaces* (1995). preprint.

[25] David Jerison, *Carleman inequalities for the Dirac and Laplace operators and unique continuation*, Adv. in Math. **62** (1986), no. 2, 118–134, DOI 10.1016/0001-8708(86)90096-4. MR865834 (88b:35218)

[26] David Jerison and Carlos E. Kenig, *Unique continuation and absence of positive eigenvalues for Schrödinger operators*, Ann. of Math. (2) **121** (1985), no. 3, 463–494, DOI 10.2307/1971205. With an appendix by E. M. Stein. MR794370 (87a:35058)

[27] Nikolaos Kapouleas, *Complete embedded minimal surfaces of finite total curvature*, J. Differential Geom. **47** (1997), no. 1, 95–169. MR1601434 (99a:53008)

[28] Nikolaos Kapouleas, S.J. Kleene, and N.M. Møller, *Mean curvature self-shrinkers of high genus: non-compact examples*, J. Reine Angew. Math., to appear, arXiv: 1106.5454.

[29] S.J. Kleene and N.M. Møller, *Self-shrinkers with a rotational symmetry*, Trans. Amer. Math. Soc., to appear, arXiv: 1008.1609.

[30] B. Kotschwar, Personal communication.

[31] B. Kotschwar, *Ricci flow and the holonomy group.* J. Reine Angew. Math., to appear, arXiv: 1105.3722.

[32] Brett L. Kotschwar, *Backwards uniqueness for the Ricci flow*, Int. Math. Res. Not. IMRN **21** (2010), 4064–4097, DOI 10.1093/imrn/rnq022. MR2738351 (2012c:53100)

[33] Lu Li and Vladimír Šverák, *Backward uniqueness for the heat equation in cones*, Comm. Partial Differential Equations **37** (2012), no. 8, 1414–1429, DOI 10.1080/03605302.2011.635323. MR2957545

[34] Fang-Hua Lin, *A uniqueness theorem for parabolic equations*, Comm. Pure Appl. Math. **43** (1990), no. 1, 127–136, DOI 10.1002/cpa.3160430105. MR1024191 (90j:35106)

[35] S. Micu and E. Zuazua, *On the lack of null-controllability of the heat equation on the half space*, Port. Math. (N.S.) **58** (2011), no. 1, 1–24. arXiv: 1111.7318. MR1820835 (2002a:93011)

[36] N.M. Møller, *Closed self-shrinking surfaces in $\mathbb{R}^3$ via the torus* (2001). preprint.

[37] Tu A. Nguyen, *On a question of Landis and Oleinik*, Trans. Amer. Math. Soc. **362** (2010), no. 6, 2875–2899, DOI 10.1090/S0002-9947-10-04733-1. MR2592940 (2011b:35186)

[38] Xuan Hien Nguyen, *Construction of complete embedded self-similar surfaces under mean curvature flow. III*, Duke Math. J., to appear, arXiv: 1106.5272.

[39] Xuan Hien Nguyen, *Construction of complete embedded self-similar surfaces under mean curvature flow. II*, Adv. Differential Equations **15** (2010), no. 5-6, 503–530. MR2643233 (2011f:53153)

[40] Xuan Hien Nguyen, *Construction of complete embedded self-similar surfaces under mean curvature flow. I*, Trans. Amer. Math. Soc. **361** (2009), no. 4, 1683–1701, DOI 10.1090/S0002-9947-08-04748-X. MR2465812 (2009k:53172)

[41] Yifei Pan and Thomas Wolff, *A remark on unique continuation*, J. Geom. Anal. **8** (1998), no. 4, 599–604, DOI 10.1007/BF02921714. MR1724207 (2000m:35032)

[42] Chi-Cheung Poon, *Unique continuation for parabolic equations*, Comm. Partial Differential Equations **21** (1996), no. 3-4, 521–539, DOI 10.1080/03605309608821195. MR1387458 (97f:35081)

[43] David L. Russell, *A unified boundary controllability theory for hyperbolic and parabolic partial differential equations*, Studies in Appl. Math. **52** (1973), 189–211. MR0341256 (49 #6006)

[44] Gregory Seregin and Vladimir Šverák, *The Navier-Stokes equations and backward uniqueness*, Nonlinear problems in mathematical physics and related topics, II, Int. Math. Ser. (N. Y.) vol. 2, Kluwer/Plenum, New York, 2002, pp. 353–366, DOI 10.1007/978-1-4615-0701-7_19. MR1972005 (2005j:35173)

[45] C. D. Sogge, *A unique continuation theorem for second order parabolic differential operators*, Ark. Mat. **28** (1990), no. 1, 159–182, DOI 10.1007/BF02387373. MR1049649 (91m:35051)

[46] L. Wang, *Uniqueness of self-similar shrinkers with asymptotically cylindrical ends*, J. Reine Angew. Math. (2013). in press.

[47] Lu Wang, *A Bernstein type theorem for self-similar shrinkers*, Geom. Dedicata **151** (2011), 297–303, DOI 10.1007/s10711-010-9535-2. MR2780753 (2012e:53130)

[48] Brian White, *Partial regularity of mean-convex hypersurfaces flowing by mean curvature*, Internat. Math. Res. Notices **4** (1994), 186 ff., approx. 8 pp. (electronic), DOI 10.1155/S1073792894000206. MR1266114 (95b:58042)

Department of Mathematics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, Maryland 21218

E-mail address: coral0426@gmail.com