Convexity and gradient estimates for fully nonlinear curvature flows

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

Vorgelegt von

Stephen Lynch

aus Southport/Australien

Tübingen
2020

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen

Tag der mündlichen Qualifikation:	13.11.2020
Stellvertretender Dekan:	Prof. Dr. József Fortágh
1. Berichterstatter:	Prof. Dr. Gerhard Huisken
2. Berichterstatter:	Prof. Dr. Simon Brendle
2. Berichterstatter:	Prof. Dr. Robert Haslhofer

Abstract

We study deformations of hypersurfaces with normal velocity given by a smooth symmetric increasing function of the principal curvatures. Specifically we study flows where the speed is a nonlinear concave function, so that at the coordinate level the evolution is governed by a fully nonlinear parabolic PDE. For each $k \geq 3$ we construct the first flows of this kind which smoothly deform any compact k-convex hypersurface of Euclidean space through a family of hypersurfaces which are also k-convex, before forming finite-time singularities which are necessarily convex (by k-convexity we mean that the sum of the smallest k principal curvatures is everywhere positive). That is, we show that k-convexity is preserved and establish an analogue of the Huisken-Sinestrari convexity estimate, which implies convexity of singularities for mean-convex mean curvature flow.

In contrast to the mean curvature flow, the fully nonlinear flows constructed here also preserve k-convexity in a Riemannian background, and we show that the convexity estimate carries over to this setting as long as the ambient curvature is suitably pinched.

We then employ our convexity estimate to prove Harnack and derivative estimates for the second fundamental form of solutions which are embedded. These results imply for example that sequences of rescalings about a singularity satisfy universal bounds for the second fundamental form and all of its higher derivatives on compact subsets of spacetime. The estimates are obtained by generalising an induction on scales technique introduced by Brendle-Huisken for two-convex flows to the k-convex setting. Our arguments apply to a general class of flows including mean-convex mean curvature flow, and in this case we recover the influential global Harnack inequality of Haslhofer-Kleiner, but without using Huisken's monotonicity formula.

Zusammenfassung

Wir untersuchen Deformationen von Hyperflächen, wobei die Geschwindigkeit in Richtung der Normalen durch eine glatte, symmetrische und monoton wachsende Funktion der Hauptkrümmungen gegeben ist. Insbesondere werden nichtlineare, konkave Geschwindigkeiten betrachtet, wodurch nach der Wahl von Koordinaten die Evolutionsgleichung der Hyperfläche zu einer voll nichtlinearen partiellen Differentialgleichung wird. Zu jedem $k \geq 3$ konstuieren wir erstmals Flüsse dieser Art, die eine k-konvexe Hyperfläche im Euklidischen Raum durch eine Familie k-konvexer Hyperflächen glatt deformiert, bis zu der Entstehung von Singularitäten, die wiederum zwingend konvex sind (mit k-Konvexität ist gemeint, dass die summe der kleinsten k Hauptkrümmungen überall positiv ist). Das heißt, wir zeigen dass die k-Konvexität der Anfangsfläche erhalten bleibt, und leiten eine Konvexitäts-Abschätzung her. Letztere verallgemeinert die Konvexitäts-Abschätzung, die für Lösungen des mittleren Krümmungsflusses mit positiver mittleren Krümmung von Huisken-Sinestrari bewiesen wurde.

Im Gegensatz zum mittleren Krümmungsfluss erhalten die hier konstruierten voll nichtlinearen Flüsse k-Konvexität auch dann, wenn als Anfangswert eine Hyperfläche einer Riemannschen Mannigfaltigkeit genommen wird. Auch in diesem Fall gilt unsere Konvexitäts-Abschätzung, solange der Krümmungstensor des umgebenden Raumes bestimmte Bedingungen erfüllt.

Die Konvexitäts-Abschätzung wird verwendet, um Harnack-Ungleichungen sowie Gradienten-Abschätzungen für eingebettete Lösungen zu zeigen. Aus diesen Resultaten folgt, zum Beispiel, dass die zweiten Fundamentalformen von Aufblasungen einer Singularität auf kompakten Untermengen gleichmäßig beschränkter Ableitungen aller Ordnungen besitzen. Für die Beweise wird eine Methode, mit welcher Brendle-Huisken schon Gradienten-Abschätzungen für zwei-konvexe Lösungen zeigen konnten, so weiterentwickelt, dass sie auch unter der schwächeren Annahme von k-Konvexität anwendbar ist. Wegen der Allgemeinheit der Argumente erhalten wir Abschätzungen für eine große Klasse von Flüssen, die insbesondere den mittleren Krümmungsfluss enthält. In diesem Fall finden wir einen neuen Beweis für die einflussreiche Harnack-Ungleichung von Haslhofer-Kleiner, und zwar ohne Huiskens Monotonie-Formel zu verwenden.

Acknowledgements

I would like to express my deepest thanks to my supervisor Gerhard Huisken for the generosity with which he has taught me over the course of this project. The images and ideas I have received through our many conversations will continue to accompany me.

I am grateful to Simon Brendle for excellent suggestions which shortened and clarified the arguments in Chapter 4, and for reading over a draft of the thesis. Thanks also to Klaus Ecker for a great deal of encouragement and help early on, and to Mat Langford, for the time he spent patiently teaching me about hypersurface flows whilst we were in Berlin. It was also a great help to be able to discuss with Ben Andrews his work on the inverse-concavity condition, which in the end played an important role in this work.

For lots of encouragement, and for hosting me on a number of trips to London, I would like to express gratitude to Huy The Nguyen. I am also very grateful to Joe Grotowski for hosting me in Brisbane and for a well-timed pep talk.

My thanks to Joe Ashley for his mateship, which was particularly visible on many visits to Tübingen. Finally, I am deeply grateful to Orla Gill, whose company enriches everything else.

Contents

CHAPTER 1

Introduction

A geometric evolution equation is a rule for deforming geometric objects where the infinitesimal change at any given time is determined by the geometry at that instant. When the evolving object is a Riemannian manifold or submanifold, the infinitesimal change or velocity at each point might be determined by the intrinsic or extrinsic curvature, and there are in fact many interesting equations of this form. Since curvature can be expressed in terms of the second derivatives of a parameterisation, at the coordinate level, deformations by curvature are governed by partial differential equations.

We will be concerned with geometric flows, that is evolution problems for which the underlying PDE is parabolic. The canonical parabolic equation is the heat equation of classical physics, which models the diffusion of heat in physical objects. Intuitively, we understand that in an otherwise isolated system, unevenly distributed heat will rapidly average itself out, and this is reflected in the strong regularity properties of the heat equation and of parabolic equations more generally. At the geometric level, the effects of diffusion manifest in many interesting ways; parabolic flows tend to exhibit strong regularising properties, yet preserve important aspects of the underlying geometric structure. For these reasons, apart from arising naturally in areas of physics and applied mathematics, parabolic geometric flows have led to striking solutions of difficult problems in pure mathematics.

Eells and Sampson introduced the harmonic map heat flow [**ES64**], which evolves mappings between Riemannian manifolds with pointwise velocity given by the Laplace-Beltrami operator applied to the map. Using this flow, they were able to prove the existence of a harmonic representative in every homotopy class of smooth maps between compact manifolds where the target has nonnegative sectional curvature. Inspired by this work, Hamilton was led to define the Ricci flow [**Ham82**], which is a rule for deforming Riemannian manifolds. At each point on a solution, the infinitesimal change in the metric tensor is a negative multiple of the Ricci curvature. In local normal coordinates, the highest-order part of the Ricci tensor is the coordinate Laplacian of the metric, so the Ricci flow is a direct analogue of the heat equation for Riemannian manifolds. In a certain sense, just as the heat equation attempts to average temperature, the Ricci flow attempts to average out the curvature of the manifold. Unlike the heat equation, however, the equations governing the Ricci flow are nonlinear, and alongside diffusion, the curvature is subject to complicated dynamical effects. This can result in the curvature of a solution becoming unbounded as it evolves, in which case we say the solution encounters a singularity.

Understanding how singularities form, and how a solution can be continued past singularities, are fundamental problems in the study of geometric flows in general. For three-dimensional Ricci flow, Perelman was able to complete Hamilton's 'surgery' procedure for extending the flow past singularities, and in doing so proved Thurston's geometrisation conjecture [**Per02**], [**Per03**]. Perelman's results have been generalised to a natural class of solutions in higher dimensions by Brendle [**Bre18**]. A further example of a deep geometric result obtained using Ricci flow is the differentiable sphere theorem due to Brendle and Schoen [**BS09**], [**BS08**]. This rules out the existence of exotic smooth structures on Riemannian spheres whose sectional curvatures are pointwise 1/4-pinched.

The mean curvature flow is the canonical example of a parabolic flow of submanifolds in a Riemannian background space. The extrinsic curvature of a submanifold is encoded in the second fundamental form (this is a vector-valued bilinear form defined on the tangent space at each point), and the mean curvature vector is the trace of this form. A one-parameter family of submanifolds is said to solve mean curvature flow if the normal velocity at every point is equal to the mean curvature vector. The second fundamental form can be found by computing the Hessian of the position vector of the submanifold and projecting onto the normal bundle, so in normal coordinates, the mean curvature vector is equal to the Laplacian of the position vector. The behaviour of submanifolds moving by their mean curvature vector can vary greatly depending on the structure of the ambient space and the codimension, but even in the simplest case of a hypersurface evolving in Euclidean space, many interesting phenomena occur.

For a smooth orientable hypersurface the normal bundle is trivial, and given a choice of global unit normal (we take the outward normal when this makes sense), the mean curvature vector is captured by a scalar quantity, which we call the mean curvature. If we write the hypersurface locally as the graph of a function, then the mean curvature is given by a second-order quasilinear elliptic operator applied to this function, and in this sense, mean curvature flow is governed by a quasilinear parabolic equation. As with the Ricci flow, the geometric nature of mean curvature flow can lead to solutions forming singularities. It is not difficult to see an example of this: one can compute explicitly that a round sphere in Euclidean space shrinks to a point in finite time under the flow, and if one compact solution sits inside another at some time, a maximum principle argument shows that this remains true while both solutions are smooth. Any compact solution can be surrounded by a shrinking round sphere, so we see that every compact solution must form a singularity no later than when the surrounding sphere vanishes.

Mean curvature flow also arises naturally as the flow of steepest descent for the area functional, so it is clear that understanding the formation of singularities, and continuing the flow past them, are fundamental problems. Generally speaking, it is common in the study of PDE to handle singularities by introducing notions of weak solution, which allow for objects of low regularity to be interpreted as solutions. This was the original approach to mean curvature flow taken by Brakke [**Bra78**], who worked with varifold solutions and tools from geometric measure theory. Another approach, known as level-set flow, was introduced independently by Evans and Spruck in [**ES91**], and by Chen, Giga, and Goto in [**CGG91**]. These authors view a family of embedded hypersurfaces evolving by mean curvature flow as the level sets of a scalar function solving (in the sense of viscosities) a nonlinear degenerate elliptic equation. The level sets of a differentiable function need not be differentiable, or of fixed topology, so the level-set formulation encodes the onset of singularities very efficiently.

Both the level-set flow and Brakke flow are important tools in the study of mean curvature flow, but in this work we adopt the viewpoint first taken by Huisken in [**Hui84**]. Here, smooth solutions are studied in a parameterisation, allowing for the use of powerful PDE methods to extract precise quantitative estimates controlling the geometry. By the combined work of Huisken and others, for certain classes of solutions, it is now possible to give an essentially complete description of the formation of singularities and consequently to continue the flow past them. For example, for embedded solutions in $\mathbb{R}^4$ with two-positive second fundamental form, Huisken and Sinestrari [**HS09**] were able to implement a surgery procedure like the one developed by Hamilton and Perelman for Ricci flow. In this way they confirmed that the Schoenflies conjecture, which asks whether every embedded three-sphere in $\mathbb{R}^4$ bounds a smooth four-ball, holds true if the embedding has positive scalar curvature.

Apart from the mean curvature, there are many other scalar geometric quantities which can be computed from the second fundamental form of a hypersurface. In fact, inserting the principal curvatures (eigenvalues of the second fundamental form) into any symmetric function yields a scalar quantity which captures information about the curvature and is invariant under isometries of the ambient space. For example, the Gauss curvature and scalar curvature of a hypersurface can be obtained in this way. Such curvature quantities then give rise to geometric flows, where the inward normal velocity of an evolving hypersurface is given by a symmetric 'speed' function of the principal curvatures. If the speed is increasing in each of the principal curvatures, then the flow is parabolic. The mean curvature flow is the only quasilinear equation of this kind - all others are fully nonlinear.

These kinds of fully nonlinear hypersurface flows may exhibit behaviour similar to the mean curvature flow, or have different properties, depending on the chosen speed function and ambient manifold. In particular, different speed functions give rise to flows which preserve different classes of hypersurfaces. This fact has been applied in interesting ways, for example, by Andrews in [**And03**] and [**And94b**], and Brendle-Huisken [**BH17**]. In each of these works, a fully nonlinear flow is used to prove statements about hypersurfaces in a curved background space which could not have been obtained using mean curvature flow.

In the present work, we make new contributions to the theory of fully nonlinear parabolic hypersurface flows, with emphasis placed on speed functions which are one-homogeneous and concave in the principal curvatures. Our new results include a general pinching theorem (**Theorem 2.6**), a convexity estimate for certain flows of k-convex hypersurfaces (**Theorem 3.1**), and a Harnack inequality for the curvature of embedded solutions (**Theorem 4.17**).

Let us describe these results in more detail before placing them in context by surveying the earlier works on which we build. Readers to whom the terminology used in the following outline is unfamiliar may prefer to first read Section 2 of the introduction and then return here.

1. Main results and outline

In this work we establish new curvature pinching and regularity results for fully nonlinear hypersurface flows. Specifically, we study flows where the normal velocity at each point is given by a smooth symmetric function of the principal curvatures which is also positive, one-homogeneous and concave.

In Chapter 2 we describe the proof of a cylindrical estimate which is essentially due to Brendle-Huisken [**BH17**][Theorem 3.1]. As it is written, their result only applies to uniformly two-convex solutions, but the same approach works in the k-convex case as well, and we state and prove the result at this level of generality. In summary, the estimate says that as the curvature of a compact uniformly k-convex solution becomes unbounded, the second fundamental form either becomes strictly $(k-1)$-convex, or approaches the second fundamental form of a cylinder of the form $\mathbb{R}^{k-1} \times S^{n-k+1}$ (for the precise statement see Theorem 2.4). There are different quantities which one can estimate to reach this conclusion, but here we work with the ratio of the mean curvature to the speed. For this quantity to contain useful information, we need to assume that the speed function is strictly concave (in off-radial directions - its Hessian vanishes in radial directions by the one-homogeneity). A more general cylindrical estimate, which also works for speeds that are only weakly concave, can be found in work of Langford and the author [**LL**].

The proof of the cylindrical estimate is by Stampacchia iteration and follows the general scheme pioneered by Huisken in [**Hui84**]. Since Huisken's work, this technique has been expanded upon and used to prove convexity and cylindrical estimates for the mean curvature flow, and other flows where the speed is one-homogeneous. What is perhaps interesting about the discussion in Chapter 2 is that (using ideas from [**BH17**]) we carry out the Stampacchia iteration for a general function satisfying some structural conditions, and recover in one go all of the convexity and cylindrical estimates previously established using this method (this is Theorem 2.6).

In Chapter 3 we turn to the problem of establishing convexity estimates for families of k-convex hypersurfaces moving by a concave curvature function. In the $k = 2$ case considered by Brendle-Huisken, such an estimate already follows from the cylindrical estimate, but this is no longer the case if $k \geq 3$. The methods used to prove convexity estimates for the mean curvature flow and for flows by other convex speeds do not seem to work either. The reason is that when the speed is concave, the evolution of the smallest principal curvature contains a nonpositive gradient term which is difficult to overcome. It remains an interesting open problem to determine exactly which algebraic property of the speed ensures that a convexity estimate holds, but we provide here the first examples of flows where the speed is a nonlinear concave function supported on the k-positive cone, and for which compact solutions satisfy a convexity estimate.

To be more specific, for each $n \geq 4$ and $3 \leq k \leq n-1$ we look at the family of speeds

$$\lambda \mapsto \left(\sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\rho}{\lambda_{i_1} + \cdots + \lambda_{i_k}} + \frac{1-\rho}{\lambda_1 + \cdots + \lambda_n} \right)^{-1}$$

for $\rho \in (0, 1]$, and show that if ρ is sufficiently small depending on n and k, compact solutions satisfy a convexity estimate (see Theorem 3.1 for the precise statement).

This family of functions constitutes a nonlinear interpolation between the k-harmonic mean and the sum of the entries. We choose ρ with the a priori cylindrical estimate from Chapter 2 in mind, so that at points where the curvature is very large, the speed is approximately equal to the mean curvature. We present detailed arguments for solutions in a Euclidean background, but the same techniques also work in certain curved background spaces, as discussed in Section 5.

In Chapter 4 we adapt a technique introduced in [**BH17**] to prove a global Harnack inequality for the curvature of embedded k-convex solutions (Theorem 4.17). Combining this estimate with the regularity theory for fully nonlinear parabolic PDE, we establish pointwise scaling-invariant estimates for all of the derivatives of the second fundamental form (Corollary 4.18). These results apply to any flow where the speed is inverse-concave on the positive cone (we define what this means just below), and for which compact solutions satisfy a convexity estimate. This class includes the speeds constructed in Chapter 3, the two-harmonic mean, and the mean curvature. In this last case, we essentially recover the global convergence theorem of Haslhofer-Kleiner [**HK17a**][Theorem 1.12], but our proof has the advantage of not needing Huisken's monotonicity formula or exterior noncollapsing (neither of which is available for concave nonlinear speeds). We note that the results of this chapter use in a crucial way the interior noncollapsing estimate due to Andrews-Langford-McCoy [**ALM13**]. Another important ingredient is a new strong maximum principle for the smallest principal curvature established in Corollary 3.8.

Further discussion of these main results can be found at the beginning of the chapters in which they are contained.

2. Background

Our results build on a host of earlier works for both the mean curvature flow and fully nonlinear flows more generally, which we now survey. We cannot hope to describe all of the important contributions made in these areas, so the focus will be on results that directly relate to the contents of the thesis. We begin with a discussion of preserved curvature conditions and the phenomenon of curvature blow-up. We then discuss the theory of convex and mean-convex solutions of mean curvature flow and other flows, with particular attention given to two-convex solutions and noncollapsing estimates.

2.1. Preserved curvature cones. For scalar parabolic equations, the maximum principle provides a mechanism for identifying properties of solutions that are preserved over the course of their evolution. For a family of hypersurfaces moving by a parabolic curvature flow, the second fundamental form satisfies a parabolic equation, so maximum principle arguments can be used to identify curvature conditions which are preserved by the flow.

A family of compact hypersurfaces which is evolving by its mean curvature can be parameterised by a family of immersions

$$F : M \times [0,T) \to \mathbb{R}^{n+1}$$

which satisfy

$$\partial_t F(x,t) = -H(x,t)\nu(x,t)$$

for each $(x,t) \in M \times [0,T)$. Here M is a compact smooth n-manifold. Writing g_{ij} for the induced metric and g^{ij} for its inverse, the second fundamental form A_{ij} and mean curvature are given by

$$A_{ij} = \langle D_i F, D_j \nu \rangle, \qquad H = g^{ij} A_{ij},$$

where D is the Euclidean connection on $\mathbb{R}^{n+1}$. The mean curvature can also be expressed as the sum of the principal curvatures,

$$H(x,t) = \lambda_1(x,t) + \cdots + \lambda_n(x,t).$$

In [**Hui84**] Huisken demonstrated that under the above parameterisation, the Weingarten map (obtained from A by raising an index) satisfies the following reaction diffusion equation:

$$(\partial_t - \Delta) A^i_j = |A|^2 A^i_j,$$

where Δ is the Laplace-Beltrami operator of the induced metric on M. Since the cubic reaction term is just a nonnegative multiple of A, a version of the maximum principle for tensors due to Hamilton [**Ham82**] implies that every convex curvature cone is preserved by the flow. To be precise, by this we mean that if the principal curvatures initially sit inside some convex symmetric cone in $\mathbb{R}^n$, then they cannot exit this cone for as long as the solution is evolving smoothly. This result provides us with a wealth of curvature conditions which are preserved by the flow, some first examples being convexity ($A > 0$), and mean-convexity ($H > 0$). More generally, if we label principal curvatures so that $\lambda_1 \leq \cdots \leq \lambda_n$ then each of the conditions

$$\lambda_1 + \cdots + \lambda_k > 0$$

is preserved. A hypersurface satisfying this condition at every point is said to be k-convex.

For speeds other than the mean curvature, the evolution of the second fundamental form is more complicated, and identifying preserved curvature conditions becomes more subtle. If γ is a smooth symmetric function of the principal curvatures (or equivalently of the second fundamental form), and the family of immersions F now satisfies

$$\partial_t F(x,t) = -\gamma(A(x,t))\nu(x,t),$$

then there holds

$$\left(\partial_t - \frac{\partial \gamma}{\partial A_{pq}} \nabla_p \nabla_q\right) A^i_j = \frac{\partial \gamma}{\partial A_{pq}} A^r_p A_{rq} A^i_j + \frac{\partial^2 \gamma}{\partial A_{pq} \partial A_{rs}} \nabla^i A_{pq} \nabla_j A_{rs}.$$

The operator on the left-hand side is weakly parabolic if γ is increasing in A, and the cubic curvature term on the right points in the direction of A. If γ is convex in A then the gradient term is favourable, and it is easy to find preserved cones - for example, the flow will preserve k-convexity for each $1 \leq k \leq n$. If γ is instead concave in A, then the Hessian of γ seems to have the wrong sign, and roughly speaking, a curvature cone will only be preserved if its boundary is convex enough to counteract the effect of the nonpositive gradient term. It is always true however that positivity of the speed is preserved: if we set $G(x,t) := \gamma(A(x,t))$, then in an orthonormal frame there holds

$$\left(\partial_t - \frac{\partial \gamma}{\partial A_{pq}} \nabla_p \nabla_q\right) G = \frac{\partial \gamma}{\partial A_{pq}} A_{pr} A_{rq} G,$$

and we see that the right-hand side is positive as long as $G > 0$.

This last observation makes it possible to design speeds which preserve a given curvature condition, even in a curved background space. If the solution sits inside a Riemannian manifold $(N, \bar{g})$, then the speed satisfies the same equation as before, with an extra term involving the curvature tensor of N, which we denote by $\bar{R}$. In a local orthonormal frame,

$$\left(\partial_t - \frac{\partial \gamma}{\partial A_{pq}} \nabla_p \nabla_q\right) G = \frac{\partial \gamma}{\partial A_{pq}} A_{pr} A_{rq} G + \frac{\partial \gamma}{\partial A_{pq}} \bar{R}(e_p, \nu, e_q, \nu) G.$$

If the first derivatives of γ and the curvature of N are bounded, then the maximum principle implies that G can decay at most exponentially in time. In particular, if G is positive initially, then it cannot become zero in finite-time, so to construct a flow which preserves a given curvature cone, we only need to choose a speed γ which vanishes at its boundary. For example, taking the speed to be the k-harmonic mean,

$$\gamma(\lambda) = \left(\sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{1}{\lambda_{i_1} + \cdots + \lambda_{i_k}} \right)^{-1},$$

one obtains a flow which preserves k-convexity in any curved ambient space with bounded curvature. The mean curvature flow does not have this property. The $k = 1$ case was studied by Andrews in [**And94b**], and the $k = 2$ case by Brendle-Huisken [**BH17**]. We elaborate on their results further below.

We note that for the k-harmonic means (and many other interesting speeds) the matrix of first derivatives of γ becomes degenerate as the second fundamental form approaches the boundary of the cone where γ is positive. Hence the parabolic operator appearing in the evolution of A can also become degenerate, unless we can establish that A remains at a controlled distance from the boundary. It is clear then that the identification of preserved curvature cones is often also essential for establishing that a flow has good analytic properties.

2.2. Curvature blow-up. Like solutions of scalar parabolic PDE, solutions of parabolic hypersurface flows tend to have very strong regularity properties. For example, if $F : M \times [0, T) \to \mathbb{R}^{n+1}$ is a compact solution of mean curvature flow, and $|A|$ is uniformly bounded over $M \times [0, T)$, then all of the higher covariant derivatives of A are bounded as well (see for example [**EH91**]). From the curvature derivative bounds it is possible to prove uniform bounds for all of the derivatives of F over $M \times [0, T)$, so no singularity can be forming as $t \to T$. Conversely, singularity formation is characterised in terms of blow-up of the quantity $|A|$. If the velocity is instead given by a nonlinear function γ of A, then this result only holds under conditions on γ. Typically, unless M is two-dimensional (see [**And04**]), one assumes γ is concave or convex in the second fundamental form. The reason is that for fully nonlinear parabolic operators which are concave or convex in the Hessian, there is a well developed regularity theory due to Krylov [**Kry82**] and Evans [**Eva82**]. In particular, for such equations, it is possible to pass from a C^2-estimate to a $C^{2,\alpha}$-estimate (at which point the Schauder estimates can be used to bootstrap and get higher regularity). Identifying further classes of operators with this property remains a major open problem.

When it is possible to characterise singularity formation in terms of curvature blow-up, understanding how singularities form largely reduces to determining the structure of those regions of the solution where $|A|$ is extremely large (relative to the initial hypersurface). This in turn is facilitated by the scaling properties of the flow - particularly interesting is the case where γ is one-homogeneous in A. Then the first derivatives of γ are scaling-invariant, and at the level of the curvature, the diffusive effects of the flow act equally at all curvature scales. Also, the space of solutions is closed under parabolic rescaling: if a family of hypersurfaces $\{M_t : t \in [-T, 0]\}$ is moving with inward normal velocity $\gamma(A(x,t))$, then so is the family of hypersurfaces

$$\{rM_{r^{-2}t} : t \in [-r^2T, 0]\},$$

where r can be any positive number.

The scaling-invariance gives us a way of magnifying solutions to better understand their properties at different curvature scales. In particular, if $\{M_t : t \in [0, T)\}$ is a solution and (x_j, t_j) any sequence of spacetime points, then we can shift in space and time to send (x_j, t_j) to the spacetime origin, and rescale by $r_j = |A|(x_j, t_j)$ to normalise the curvature at this point:

$$M_t^j := r_j(M_{r_j^{-2}t+t_j} - x_j), \qquad t \in [-r_j^2 t_j, 0].$$

If $r_j \to \infty$ we call the sequence of solutions M_t^j a blow-up sequence. If r_j is comparable to the maximum of $|A|$ up to time t_j, then the curvature of the rescaled solution is uniformly bounded over all of spacetime, and by parabolic regularity and the Arzela-Ascoli theorem, it is possible to extract a smooth limiting solution. Since $r_j^2 t_j \to \infty$, this limiting flow is defined for all negative times. Such solutions are called ancient solutions, and are much more rigid than solutions on finite time intervals - this should be likened to the situation for elliptic equations, where entire solutions are often also very rigid.

This is an indication that the regions of a solution where the curvature is extremely large might have special geometric and regularity properties. On the other hand, a priori, the curvature might be blowing up at drastically different rates even at nearby points on the solution, so a general blow-up sequence may not even converge locally about the spacetime origin. Understanding when blow-up sequences converge locally and globally, and the possible geometries of limits thus obtained, is an important and difficult problem. The geometry is controlled by proving curvature pinching estimates, which control the position of the second fundamental in curvature space at high curvature scales, and compactness is established by proving scaling-invariant estimates for the derivatives of A. These two tasks are intimately related - often pinching implies regularity, and vice versa. For the mean curvature flow, the first results of this kind were established for convex solutions by Huisken in [**Hui84**].

2.3. Convex solutions. In [**Hui84**], Huisken considered the evolution of compact convex hypersurfaces of dimension at least two under the mean curvature flow. This work introduced fundamental tools which have continued to be applied to the study of hypersurface flows, also for solutions which are not convex. As mentioned above, the maximum principle implies that convexity is preserved, and this implies that

$$|A|^2 \leq H^2,$$

so as the maximal time is approached, H blows up at the same rate as $|A|$. However, as Huisken observes, not all scalar functions of the second fundamental form have this property, and indeed, the traceless second fundamental form satisfies the following a priori estimate:

$$|A|^2 - \frac{1}{n}|H|^2 \leq CH^{2-\sigma},$$

where C and $\sigma \in (0, 2)$ depend only on the solution at the initial time. The right-hand side can be written as

$$\sum_{i<j}(\lambda_i - \lambda_j)^2,$$

so this estimate says that at points where H is very large, the difference between the principal curvatures is extremely small relative to H. In particular, if a blow-up sequence converges smoothly on some spacetime neighbourhood, the limit must be totally umbilic ($\lambda_1 = \cdots = \lambda_n$), and hence a piece of a shrinking sphere solution.

The pinching estimate is proven using a Stampacchia iteration scheme, which takes as inputs a novel Poincaré inequality derived from Simons' identity, and the Michael-Simon Sobolev inequality. This procedure has since been used to prove various curvature pinching estimates for the mean curvature flow and fully-nonlinear flows where the speed is one-homogeneous, also for solutions which are not convex. It lies at the heart of this work, and is discussed in detail in Chapter 1 (see in particular Theorem 2.6). We note that the ideas underlying Stampacchia iteration were introduced by De Giorgi to solve Hilbert's 19th problem in [**DG60**] (Nash solved the same problem independently in [**Nas58**]).

Using the pinching estimate and the maximum principle, Huisken then derives a pointwise gradient estimate for the mean curvature and combines this with Myers's theorem to prove a global Harnack inequality for the curvature, which says that

$$\max_{M_t} H \leq C \min_{M_t} H$$

on every timeslice where the maximum of H is sufficiently large. The constant C depends only on the initial data, so this estimate ensures that the curvature blows up at the same rate over the whole solution. Combined with the pinching estimate, this shows that the whole solution is simultaneously becoming spherical as the maximal time is approached. Another way to say this is that every blow-up sequence converges to a shrinking sphere solution. Colloquially, the solution is said to shrink to a round point.

There are also fully nonlinear flows for which convex solutions behave in this way. Chow [**Cho85**] considered the flow by the n-th root of the Gauss curvature, and adapted Huisken's arguments to show that the flow contracts strictly convex initial hypersurfaces to a round point. Andrews later showed that for any smooth speed γ which is one-homogeneous and convex, convex solutions shrink to a round point [**And94a**]. The same holds if γ is instead concave and the flow preserves uniform convexity ($A \geq \varepsilon H g$ with $\varepsilon > 0$), which is for example the case when γ vanishes at the boundary of the positive cone. In [**And94b**], Andrews then showed that in a Riemannian ambient space with curvature bounded from above, and sectional curvature bounded below by $-\kappa^2$, the concave speed

$$\gamma(\lambda) = \left(\sum_i \frac{1}{\lambda_i - \kappa} \right)^{-1}$$

gives rise to a flow which contracts any compact initial hypersurface satisfying $\lambda_i > \kappa$ to a round point.

We note that Andrews uses convexity in a more crucial way than Huisken, appealing in particular to the fact that a compact convex n-dimensional hypersurface admits special parameterisations over the unit n-sphere. This gives the space of solutions good compactness properties, which is important since in the fully nonlinear case, it does not seem to be possible to establish gradient estimates for the curvature via the maximum principle alone.

So far we have only discussed flows of convex hypersurfaces where the singular behaviour is very well understood. Flows of non-convex hypersurfaces may exhibit much more complicated behaviour, but it is still possible to make general statements about mean-convex solutions, some of which we will now survey.

2.4. Mean-convex solutions. A remarkable result (of the combined efforts of Gage [**Gag84**], Gage-Hamilton [**GH86**] and Grayson [**Gra87**]) says that when $n = 1$, every compact embedded solution of mean curvature flow (in this case commonly called curve-shortening flow) becomes convex in finite time, and shrinks to a round point. In contrast, when $n \geq 2$ a compact solution of mean curvature flow which is non-convex can form local singularities, where the curvature blows up in some regions, and remains bounded elsewhere. An instructive example is the dumbbell, which is a surface of rotation constructed by attaching two large spheres in $\mathbb{R}^3$ by a long thin cylinder. If this is done correctly, then the mean curvature of the cylinder or 'neck' will be much larger than that of the spherical regions, forcing it to shrink much more quickly and become singular whilst the curvature of the spherical regions remains bounded. If the initial dumbbell is mean-convex, then this remains true by the maximum principle, and blowing up the singularity, one sees a homothetically shrinking cylinder of the form $\mathbb{R} \times S^1$.

In higher dimensions, it is possible to construct compact mean-convex solutions which form singularities modeled on any cylinder of the form $\mathbb{R}^{n-k} \times S^k$ with $k \geq 1$. Other singularity models also occur, such as the bowl soliton, which is a noncompact convex solution with rotational symmetry that moves by translation. Each spatial slice of the bowl soliton is asymptotic to a paraboloid, and so contains arbitrarily large regions modeled on a cylinder of the form $\mathbb{R} \times S^{n-1}$. In general, it is believed that for mean-convex solutions, many more complicated models can also occur (the exception to this is the two-convex case, discussed further below). We note that the class of mean-convex hypersurfaces in $\mathbb{R}^{n+1}$ can be quite complicated. Indeed, if Σ is any compact embedded k-dimensional submanifold in $\mathbb{R}^{n+1}$ with $k \leq n-1$, then the set

$$\Sigma_\varepsilon := \{x \in \mathbb{R}^{n+1} : \mathrm{dist}(x, \Sigma) = \varepsilon\}$$

is a mean-convex hypersurface, provided that ε is small enough.

On the other hand, under the assumption of mean-convexity, powerful curvature pinching and regularity results have been established, and these place marked restrictions on the geometric properties of singularity models. We now survey some these results, and discuss also generalisations to fully nonlinear flows.

2.4.1. *Convexity estimates.* In [**HS99b**] and [**HS99a**], Huisken and Sinestrari established what is referred to as the convexity estimate for mean-convex mean curvature flow. This says that the negative part of the second fundamental form blows up at a strictly slower rate than the full curvature: for each $\varepsilon > 0$,

$$\lambda_1 \geq -\varepsilon H - C,$$

where $C > 0$ depends only on ε and the solution at the initial time. Like Huisken's pinching estimate for convex solutions, the proof is by Stampacchia iteration, but contains many new ideas needed to go beyond the convex case. A consequence of this estimate is that any solution obtained by blowing up a singularity must have nonnegative second fundamental form. We note that convexity of blow-up limits was established independently by White in [**Whi03**], but his methods do not seem to generalise to the fully nonlinear case.

Convexity estimates have also been established for flows where the speed is one-homogeneous and convex in dimension $n \geq 3$ [**ALM14**], or simply one-homogeneous in case $n = 2$ [**ALM15**]. For uniformly two-convex solutions moving by concave nonlinear speed functions, a convexity estimate follows from the cylindrical estimate in [**BH17**], which is further discussed just below. All of these results are established using Stampacchia iteration, as in [**HS99b**] - the hard work always lies in identifying an appropriate function to which this can be applied.

2.4.2. *Cylindrical estimates.* The cylindrical estimate for mean curvature flow is a generalisation of Huisken's pinching estimate for convex solutions to the k-convex case. The statement is that, on a compact k-convex solution, for each $\varepsilon > 0$ there holds

$$|A|^2 - \frac{1}{n-k+1}H^2 \leq \varepsilon H + C,$$

where the C depends only on ε and the solution at the initial time. To say this in words, as the curvature blows up, the quantity $|A|^2/H^2$ improves towards the value it takes on a cylinder $\mathbb{R}^{k-1} \times S^{n-k+1}$. In general, one cannot expect any further improvement, since there are k-convex solutions which form singularities modeled on $\mathbb{R}^{k-1} \times S^{n-k+1}$.

From the cylindrical estimate we conclude that the second fundamental form of any solution obtain by blowing up a singularity satisfies

$$|A|^2 - \frac{1}{n-k+1}H^2 \leq 0.$$

This inequality implies that the sum of the smallest $k-1$ principle curvatures is nonnegative, and vanishes if and only if

$$\lambda_1 = \cdots = \lambda_{k-1} = 0, \qquad \lambda_k = \cdots = \lambda_n.$$

Note that if $n = 2$ and $k = 2$, then the cylindrical estimate is simply the convexity estimate. If $n \geq 3$ and $k = 2$, then the cylindrical estimate is significantly stronger than the convexity estimate.

The cylindrical estimate was first established in Huisken-Sinestrari [**HS09**] (see [Theorem 5.3]). To be precise, they only consider the case of two-convex solutions, but it is clear that their method also applies in the k-convex case. Their work was then generalised to the case of convex one-homogeneous speeds by Andrews-Langford-McCoy [**ALM14**]. Rather than comparing $|A|^2$ with H^2, Andrews and Langford consider a different quantity which is adapted to the speed, but their estimate contains roughly the same information - the second fundamental form either approaches that of an $\mathbb{R}^{k-1} \times S^{n-k+1}$, or becomes strictly $(k-1)$-positive, as $|A|$ blows up.

The proofs of Huisken-Sinestrari and Andrews-Langford both make use of a convexity estimate, but with some more careful analysis, Brendle-Huisken show that this is not necessary in [**BH17**]. With this observation they are able to prove a cylindrical estimate for a large class of flows by concave speed functions, without having to first prove a convexity estimate. They only write down a proof in the two-convex case, but again, only minor modifications are required to get the analogous estimate for k-convex flows by concave speed functions - the details can be found in the literature in [**LL**][Theorem 1.1], or in Chapter 2 of this thesis.

In the case of two-convex mean curvature flow, the cylindrical estimate has lead to a fairly complete picture of the kinds of singularities which can form.

2.4.3. *Two-convex solutions.* Using their cylindrical estimate, Huisken and Sinestrari were able to derive an extremely detailed picture of the singular behaviour of two-convex solutions of mean curvature flow, at least for solutions of dimension at least three. A further key ingredient in their work is a pointwise gradient estimate for the curvature:

$$|\nabla A|^2 \leq CH^4 + C,$$

where C depends only on the initial data. This is proven by a delicate maximum principle argument, applied to a quantity built from $|\nabla A|^2$ and the quantity appearing in the cylindrical estimate. The same kind of argument is also used to prove higher derivative estimates of the form

$$|\nabla^k A|^2 \leq CH^{2k+2} + C,$$

which can be integrated over a small spacetime neighbourhood to prove a local Harnack inequality for the curvature, valid at points where the curvature is sufficiently large.

With these estimates in hand, Huisken and Sinestrari are able to show that if the curvature is sufficiently large at (x_0, t_0), then there are only a few possibilites for the geometry of the solution near this point. Either: the whole solution is convex at time t_0; the point x_0 sits in a convex cap attached to an extremely long neck region modeled on $\mathbb{R} \times S^{n-1}$; or the point (x_0, t_0) sits itself inside an extremely long neck. With further analysis, they determine exactly how these local pieces can fit together, and so obtain an extremely precise picture of the entire high-curvature region of a solution. They then show that the high-curvature region can actually be excised and replaced by finitely many smooth disks of controlled curvature. This procedure is known as surgery, and allows for the flow to be continued until, after a bounded number of surgeries, the hypersurface has been decomposed into finitely many recognisable pieces. The flow with surgeries produces a classification of two-convex immersions up to diffeomorphism, and as mentioned above, this proves the famous Schoenflies conjecture in the special case that the embedding is two-convex (when $n = 3$, every hypersurface which has positive scalar curvature is two-convex).

A particular consequence of the analysis in [**HS09**] is that the local Harnack inequality becomes valid over larger and larger spacetime regions about (x_0, t_0) as the curvature at (x_0, t_0) increases. Another way of saying this is that for a blow-up sequence, rescaled to make the curvature bounded at the spacetime origin, the curvature is universally bounded on compact subsets of spacetime. Hence an Arzela-Ascoli argument shows that any such sequence subconverges smoothly to a convex ancient solution. The picture of the high-curvature regions drawn by Huisken-Sinestrari has since been reinforced by Brendle-Choi [**BC19**], [**BC18**], who show that the only possible limits that can be obtain in this way are shrinking spheres or cylinders of the form $\mathbb{R} \times S^{n-1}$, and the bowl soliton (see also [**ADS18**]).

It is natural to ask whether the surgery procedure can also be carried out for mean convex solutions in $\mathbb{R}^3$. Here the cylindrical estimate contains less information, and for immersed solutions, counterexamples show that the pointwise gradient estimate cannot hold in general. It turns out that this behaviour is ruled out if the solution is embedded, and in this case, it is also possible to define a flow with surgeries.

2.4.4. *Non-collapsing.* Compact solutions of mean curvature flow which are embedded at the initial time remain embedded up to their maximal time of existence. It is not difficult to see why this is true - if two nonequal points on the solution touch for the first time, then the mean curvature vectors at these two points are pointing in opposite directions, so the two points are moving away from each other. This of course means that any rescaled version of the solution is also embedded, but along a blow-up sequence, where the scaling factor tends to infinity, the region inside the solution at time zero may be 'collapsing' onto a set of lower-dimension.

Huisken gave an interesting proof that this kind of behaviour cannot occur for embedded solutions of curve-shortening flow [**Hui98**]. He used a maximum principle argument to show that the intrinsic and extrinsic distance between any two points on the solution remain uniformly comparable along the flow. The first noncollapsing result in higher dimensions was established by White [**Whi00**], [**Whi03**], who uses in a crucial way Huisken's monotonicity formula [**Hui90**]. White's result says that on an embedded solution, if the mean curvature is normalised at a point, then there is a nearby ball of controlled radius which is contained in the region bounded by the solution. Sheng and Wang [**SW09**] later gave a different proof of the same result. The results of White and Sheng-Wang both make use of blow-up arguments and the convexity estimate.

Andrews later established a noncollapsing estimate for embedded mean-convex solutions using a direct maximum principle argument [**And12**]. The argument draws inspiration from Huisken's work on curve-shortening flow, in that the maximum principle is applied to a 'two-point' function, which is defined on the product of the solution with itself. The conclusion is the following: for any point (x, t) on an embedded solution, the radius $\bar{r}(x, t)$ of the largest ball which makes interior contact with the solution at (x, t) is bounded from below in terms of the initial hypersurface and $H(x, t)$. To be precise,

$$\bar{r}(x, t)H(x, t) \geq \min_M \bar{r}(\cdot, 0)H(\cdot, 0).$$

We call this the interior noncollapsing estimate. There is a similar exterior noncollapsing estimate, which says that

$$\underline{r}(x,t)H(x,t) \geq \min_M \underline{r}(\cdot,0)H(\cdot,0),$$

where $\underline{r}(x,t)$ denotes the largest ball making exterior contact with (x,t). Andrews' arguments have been carried over to the fully nonlinear case by Andrews-Langford-McCoy [**ALM13**]: they show using the maximum principle that one-homogeneous concave speeds admit interior noncollapsing, whilst convex speeds exhibit exterior noncollapsing. If the solution is convex and the speed is both concave and has an additional property known as inverse-concavity, then both the interior and exterior noncollapsing estimates are true [**AL16**].

The noncollapsing estimates have significantly advanced the study of embedded solutions of hypersurface flows. For the mean curvature flow, Haslhofer-Kleiner [**HK17a**] combined the noncollapsing with White's ε-regularity theorem [**Whi05**] to derive a new gradient estimate for embedded mean-convex solutions. Taking inspiration from Perelman's work on three-dimensional Ricci flow, they then combine their gradient estimate with the convexity estimate to prove a powerful global Harnack inequality. In particular, this says that for a compact embedded solution, any blow-up sequence satisfies universal curvature bounds on compact subsets of spacetime, and hence subconverges to a convex ancient solution. In another important advancement, Brendle [**Bre15**] used Stampacchia iteration to show that the noncollapsing estimates become optimal, in an appropriate sense, at a singularity. In particular, this improving noncollapsing estimate can be used as a subsitute for the cylindrical estimate in two dimensions, allowing Brendle-Huisken [**BH16**] to carry out surgery for embedded mean convex solutions in $\mathbb{R}^3$, and later in three-manifolds more generally [**BH18**]. The result in Euclidean space was obtained independently by Haslhofer-Kleiner in [**HK17b**].

The interior non-collapsing also plays a role in [**BH17**], where Brendle and Huisken use a fully nonlinear flow to study two-convex embeddings in a Riemannian background space. If the ambient space satisfies the curvature condition

$$\bar{R}_{jiji} + \bar{R}_{kiki} \geq -2\kappa^2,$$

then they define a flow with surgeries for the concave speed

$$\lambda \mapsto \bigg(\sum_{1\leq i<j\leq n} \frac{1}{\lambda_i + \lambda_j - 2\kappa}\bigg)^{-1}.$$

In an important step the authors carry out an induction on scales argument, combining their cylindrical estimate with the interior non-collapsing property, to prove a pointwise gradient estimate for the curvature. The surgery construction then proceeds in much the same way as for two-convex mean curvature flow.

This brings us to the present work, where we generalise some of the results in [**BH17**] to flows of k-convex hypersurfaces. We hope that the new estimates obtained here will play a role in future work on the geometric and topological structure of interesting classes of hypersurfaces in Riemannian background spaces.

3. Notation and preliminary results

In the rest of the introduction we lay down some basic definitions, notation and results which will be used frequently in the rest of the thesis.

3.1. Notation for solutions. We will be concerned with smooth one-parameter families of immersions $F : M \times I \to \mathbb{R}^{n+1}$, which evolve according to

$$\partial_t F(x,t) = -G(x,t)\nu(x,t), \tag{CF}$$

where M is a smooth n-manifold, $I \subset \mathbb{R}$ is an interval, and ν is a globally defined unit normal vector. The function $G(x,t)$ is given by applying a smooth symmetric speed function γ to the n-tuple of principal curvatures $\lambda(x,t)$. The principal curvatures will always be labeled so that

$$\lambda_1(x,t) \leq \cdots \leq \lambda_n(x,t).$$

We define the class of speed functions γ under consideration just below. The principal curvatures are the eigenvalues of the second fundamental form A with respect to the induced metric g. To be precise, $g(x,t)$ is the pullback of the Euclidean metric by $F(\cdot,t)$ at x, and A is defined as follows:

$$\begin{aligned} A(x,t) : T_xM \times T_xM &\to \mathbb{R} \\ (v,w) &\mapsto \langle D_v\nu, w\rangle, \end{aligned}$$

where D denotes the Euclidean connection. With this convention, a round sphere equipped with the outward pointing unit normal has positive principal curvatures.

At times, it will be convenient to work with a slightly more general definition of solution:

$$(\partial_t F(x,t))^\perp = -G(x,t). \tag{1}$$

If F solves (CF), then composing F with any time-dependent diffeomorphism of M yields a solution of (1). If M is compact and I is bounded from below, as will typically be the case, then every solution of (1) can be composed with a time-dependent diffeomorphism of M to get a solution of (CF) (see for example Chapter 1 in [**Eck12**]).

When considering embedded solutions, we write M_t for the smooth hypersurface $F(M,t)$, and also use the notation

$$\mathbf{M} := \{(x,t) \in \mathbb{R}^{n+1} \times I : x \in M_t\}.$$

When M_t bounds an open region in $\mathbb{R}^{n+1}$, we denote this region by Ω_t, and write also

$$\boldsymbol{\Omega} := \{(x,t) \in \mathbb{R}^{n+1} \times I : x \in \Omega_t\}.$$

3.2. Admissible speed functions. Throughout this work, $\gamma : \Gamma \to \mathbb{R}$ will always denote a smooth symmetric function, and Γ an open, convex cone in $\mathbb{R}^n$ which is also symmetric (closed under permutations on n-elements). We say that γ is an admissible speed if in addition to these properties, the following hold:

1. Positivity - for each $\lambda \in \Gamma$, $\gamma(\lambda) > 0$;
2. Ellipticity - the first derivatives $\frac{\partial\gamma}{\partial\lambda_i}(\lambda)$ are all positive for each $\lambda \in \Gamma$;
3. Symmetry - the value of $\gamma(\lambda)$ does not change if the entries of λ are permuted;
4. One-homogeneity - for each positive r and each $\lambda \in \Gamma$, there holds

$$\gamma(r\lambda) = r\gamma(\lambda).$$

One consequence of these properties is that the ray parallel to $(1,\ldots,1)$ is in Γ. The simplest example of an admissible speed is the mean, which we denote by

$$\begin{aligned}\mathrm{tr} : \mathbb{R}^n &\to \mathbb{R}\\ \lambda &\mapsto \lambda_1 + \cdots + \lambda_n,\end{aligned}$$

restricted to its support. Of course, the function tr is smooth and satisfies conditions (2)-(4) on all of $\mathbb{R}^n$, but other key examples which we study, such as the harmonic mean

$$\gamma(\lambda) := \left(\frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_n}\right)^{-1},$$

do not admit smooth extensions beyond the cone on which they are positive. The ellipticity condition is so named because, in local coordinates, this is precisely the property which makes (CF) weakly parabolic. Symmetry is necessary to make sure that $\gamma(\lambda)$ is geometric, and the one-homogeneity ensures that the space of solutions is closed under parabolic rescaling. We note however that there are interesting geometric flows by speeds with other scaling properties, such as the one used in [**And03**].

We discussed in the introduction the necessity of imposing some kind of concavity on the speed, and the differences between flows by concave and convex speed functions. Our discussion will focus on concave speeds.

In addition to these properties, we will often assume one or more of the following additional conditions on the speed:

(1) Zero at the boundary - we say γ vanishes at the boundary if it admits a continuous extension to $\bar{\Gamma}$ which vanishes identically on $\partial\Gamma$;
(2) Strict concavity in off-radial directions - For each $\xi \in \mathbb{R}^n$ and $\lambda \in \Gamma$, there holds
$$\frac{\partial^2 \gamma}{\partial \lambda_i \partial \lambda_j}(\lambda)\xi_i\xi_j \le 0,$$
with equality if and only if ξ is a multiple of λ.

Finally, we say that an admissible speed defined on the positive cone,

$$\Gamma_+ := \{\lambda \in \mathbb{R}^n : \min_{1\le i\le n} \lambda_i > 0\},$$

is inverse-concave if the function

$$\gamma_*(\lambda) := \gamma(\lambda_1^{-1}, \ldots, \lambda_n^{-1})^{-1}$$

is concave in λ on Γ_+.

3.3. Smooth symmetric functions. For each $1 \le k \le n$, let $\sigma_k : \mathbb{R}^n \to \mathbb{R}$ denote the elementary symmetric polynomial

$$\sigma_k(\lambda) := \sum_{1<i_1<\cdots<i_k\le n} \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_k},$$

and let $S : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$S : \lambda \mapsto (\sigma_1(\lambda), \ldots, \sigma_n(\lambda)).$$

Newton proved that every symmetric polynomial in the entries of λ can be expressed as a symmetric polynomial in $\sigma_1(\lambda), \ldots, \sigma_n(\lambda)$. In [**Gla63**], Glaeser proved a far-reaching generalisation of this result: if $U \subset \mathbb{R}^n$ is a symmetric open set and $\varphi : U \to$

$\mathbb{R}$ is a smooth symmetric function, then there is a smooth function $\hat{\varphi}$ defined on $S(U)$ such that

$$\varphi(\lambda) = \hat{\varphi}(S(\lambda))$$

holds for each $\lambda \in U$.

This gives us a way of viewing φ as a smooth function on the set of symmetric matrices with eigenvalues in U. For an open symmetric set U in $\mathbb{R}^n$ we write

$$\mathrm{O}\,U := \{O\,\mathrm{diag}(\lambda)O^T : \lambda \in U,\ O \in O(n)\},$$

where $\mathrm{diag}(\lambda)$ is the diagonal matrix with entries λ on the diagonal, and $O(n)$ is the orthogonal group. This is equivalent to saying that

$$\mathrm{O}\,U = \{A \in \mathrm{Sym}(n) : \lambda(A) \in U\},$$

where λ is the map taking A to its eigenvalues (labeled so that $\lambda_1 \leq \cdots \leq \lambda_n$). The characteristic polynomial of a diagonalisable matrix admits the expansion

$$\det(A - tI) = \prod_{1\leq k\leq n} (t - \lambda_i) = \sum_{1\leq k\leq n} (-1)^k \sigma_k(\lambda) t^{n-k},$$

and from this we see that $\sigma_k(\lambda)$ can be expressed as a smooth function $\hat{\sigma}_k$ of A by defining

$$\hat{\sigma}_k(A) := \frac{1}{(-1)^k (n-k)!} \frac{d^{n-k}}{dt^{n-k}} \det(A - tI)\bigg|_{t=0}.$$

Therefore, the function φ defined on $\mathrm{O}\,U$ by

$$\Phi(A) := \hat{\varphi}(\hat{\sigma}_1(A), \ldots \hat{\sigma}_n(A))$$

is smooth in A, and agrees with φ applied to $\lambda(A)$.

Going the other way, Schwarz [**Sch75**] showed that given a smooth function Φ defined on $\mathrm{O}\,U$ which satisfies

$$\Phi(OAO^{-1}) = \Phi(A) \qquad \forall\ A \in \mathrm{O}\,U,\ O \in O(n),$$

there is a smooth function $\varphi : U \to \mathbb{R}$ such that

$$\varphi(\lambda(A)) = \Phi(A).$$

In practice, if φ is a smooth symmetric function, we view it as either a function of eigenvalues or of matrix entries as is convenient. If we write $\varphi(\lambda)$, where $\lambda \in \mathbb{R}^n$, then it is understood that we are viewing φ as a function of eigenvalues, whereas if we write $\varphi(A)$ and A is a matrix, then we are viewing φ as a function of the entries of the matrix. Also, to ease notation, we often write

$$\dot{\varphi}^i(\lambda) := \frac{\partial \varphi}{\partial \lambda_i}(\lambda), \qquad \ddot{\varphi}^{ij}(\lambda) := \frac{\partial^2 \varphi}{\partial \lambda_i \partial \lambda_j}(\lambda),$$

and

$$\ddot{\varphi}^{ij}(A) := \frac{\partial \varphi}{\partial A_{ij}}(A), \qquad \ddot{\varphi}^{ij,kl}(A) := \frac{\partial^2 \varphi}{\partial A_{ij} \partial A_{kl}}(A).$$

If A is diagonal then

$$\dot{\varphi}^{ij}(A) = \dot{\varphi}^i(\lambda)\delta^{ij},$$

and if in addition the eigenvalues of A satisfy $\lambda_1 < \cdots < \lambda_n$, then

$$\ddot{\varphi}^{ij,kl}(A)T_{ij}T_{kl} = \ddot{\varphi}^{ij}(\lambda)T_{ii}T_{jj} + 2\sum_{i<j}\frac{\dot{\varphi}^j(\lambda)-\dot{\varphi}^i(\lambda)}{\lambda_j-\lambda_i}|T_{ij}|^2$$

holds for every symmetric T. For proofs of these identities we refer to [**And07**][Theorem 5.1]. With these formulae we can show that concavity with respect to eigenvalue and matrix variables are equivalent when U is convex:

LEMMA 1.1. *Let $U \subset \mathbb{R}^n$ be open, convex and symmetric, and let $\varphi : U \to \mathbb{R}$ be a smooth symmetric function. Then the conditions*

$$\ddot{\varphi}^{ij}(\lambda)\xi_i\xi_j \le 0, \qquad \forall\, \lambda \in U,\ \xi \in \mathbb{R}^n$$

and

$$\ddot{\varphi}^{ij,kl}(A)T_{ij}T_{kl} \le 0 \qquad \forall\, A \in \mathrm{O}\,U,\ T \in \mathrm{Sym}(n)$$

are equivalent.

PROOF. That the second inequality implies the first is clear - one only has to apply the second derivative identity above with $A = \mathrm{diag}(\lambda)$ and $B = \mathrm{diag}(\xi)$. To get the second implication, note that since φ is smooth, it suffices to show that the second inequality is true for all A with distinct eigenvalues, since the general case then follows by approximation. In this case, choosing $O \in O(n)$ so that $OAO^T = \tilde{A} = \mathrm{diag}(\lambda)$, and writing also $\tilde{T} = OBO^T$, we compute

$$\ddot{\varphi}^{ij,kl}(A)T_{ij}T_{kl} = \frac{d^2}{dt^2}\bigg|_{t=0}\varphi(OAO^T + tOTO^T) = \ddot{\varphi}^{ij,kl}(\tilde{A})\tilde{T}_{ij}\tilde{T}_{kl}.$$

Now we can invoke the identity above to get

$$\ddot{\varphi}^{ij,kl}(A)T_{ij}T_{kl} = \ddot{\varphi}^{ij}(\lambda)T_{ii}T_{jj} + 2\sum_{i<j}\frac{\dot{\varphi}^j(\lambda)-\dot{\varphi}^i(\lambda)}{\lambda_j-\lambda_i}|T_{ij}|^2.$$

We are assuming that the first term is nonpositive, and nonpositivity of the second term follows if we can show that

$$\frac{\dot{\varphi}^j(\lambda)-\dot{\varphi}^i(\lambda)}{\lambda_j-\lambda_i} \le 0$$

whenever $\lambda_j > \lambda_i$. To this end, we let s denote the permutation which swaps the i-th and j-th entries and note that by symmetry and convexity, $(1-t)\lambda + ts(\lambda)$ is in U for each $t \in [0,1]$. Thus the function

$$t \mapsto \varphi((1-t)\lambda + ts(\lambda))$$

is even and concave, and consequently there holds

$$0 \le \frac{d}{dt}\bigg|_{t=0}\varphi((1-t)\lambda + ts(\lambda)) = \dot{\varphi}^i(\lambda)(\lambda_i - \lambda_j) + \dot{\varphi}^j(\lambda)(\lambda_j - \lambda_i),$$

which we can rearrange to get the desired inequality. □

3.4. Evolution equations. Fix an admissible speed $\gamma : \Gamma \to (0, \infty)$ and let $F : M \to \mathbb{R}^n$ be a smooth immersed hypersurface such that $\lambda(x) \in \Gamma$ for each $x \in M$. Given a a smooth symmetric function φ defined on Γ, we define a smooth function Φ on M by setting $\Phi(x) := \varphi(\lambda(x))$. In a smooth local orthonormal frame, we can view A as a smooth map from M into $\mathrm{O}\,\Gamma$, and write $\Phi(x) = \varphi(A(x))$. Since the coefficients of A are smooth, this shows that $G(x) := \gamma(A(x))$ is smooth as a function of x, and we have

$$\nabla_p G = \frac{\partial \gamma}{\partial A_{ij}} \nabla_p A_{ij}$$

and

$$\nabla_p \nabla_q G = \frac{\partial \gamma}{\partial A_{ij}} \nabla_p \nabla_q A_{ij} + \frac{\partial^2 \gamma}{\partial A_{kl} \partial A_{ij}} \nabla_p A_{ij} \nabla_q A_{kl}.$$

In a slight abuse of notation, we will write $\varphi(A)$ to mean the quantity returned when an orthonormal basis is chosen and φ is evaluated on the coefficient matrix of A. In the same vein, we will write $\frac{\partial \gamma}{\partial A_{ij}}(A)$ or just $\dot{\gamma}^{ij}(A)$ for the $(2,0)$ tensor whose coefficient matrix has these entries with respect to any orthonormal basis, and the analogous convention applies to the second derivatives.

We state here the evolution equations for various geometric quantities along a solution of (CF). Huisken derived these identities for the mean curvature flow in [**Hui84**][Section 3], and only minor modifications are necessary to cover more general speed functions. With respect to a local frame for the tangent bundle of M, the induced metric g and outward pointing unit normal ν satisfy

$$\begin{aligned} \partial_t g_{ij} &= -2GA_{ij}; \\ \partial_t \nu &= \nabla G. \end{aligned}$$

The speed G satisfies

$$\left(\partial_t - \frac{\partial \gamma}{\partial A_{ij}} \nabla_i \nabla_j \right) G = \frac{\partial \gamma}{\partial A_{ij}} g^{kl} A_{ik} A_{lj} G,$$

and for the Weingarten map (this is just the name given to $g^{pr} A_{rq}$) we have

$$\left(\partial_t - \frac{\partial \gamma}{\partial A_{ij}} \nabla_i \nabla_j \right) A^p_q = \frac{\partial \gamma}{\partial A_{ij}} g^{kl} A_{ik} A_{lj} A^p_q + \frac{\partial^2 \gamma}{\partial A_{kl} \partial A_{ij}} \nabla^p A_{ij} \nabla_q A_{kl}.$$

To ease notation, we will write

$$\Delta_\gamma := \frac{\partial \gamma}{\partial A_{ij}} \nabla_i \nabla_j, \qquad |A|^2_\gamma := \frac{\partial \gamma}{\partial A_{ij}} g^{kl} A_{ik} A_{lj}.$$

If φ is a smooth symmetric function on Γ, and is σ-homogeneous (meaning that $\varphi(r\lambda) = r^\sigma \varphi(\lambda)$ for every $r > 0$), then for $\Phi(x,t) := \varphi(A(x,t))$ there holds

$$(\partial_t - \Delta_\gamma)\Phi = \sigma |A|^2_\gamma \Phi + \left(\frac{\partial \varphi}{\partial A_{pq}} \frac{\partial^2 \gamma}{\partial A_{kl} \partial A_{ij}} - \frac{\partial \gamma}{\partial A_{pq}} \frac{\partial^2 \varphi}{\partial A_{kl} \partial A_{ij}} \right) \nabla_p A_{ij} \nabla_q A_{kl}.$$

Notice that if γ is the mean, then the term involving its Hessian vanishes, and the gradient term has a sign as long as φ is convex or concave. This fact is extremely useful in the study of k-convex solutions of mean curvature flow, since there are many interesting concave functions defined on the k-positive cone. In contrast, when γ is nonlinear, it is much more difficult to find curvature functions for which the two gradient terms combine to give something with a sign or otherwise favourable structure. Of course, as well as the speed itself, the mean curvature clearly always satisfies a nice equation.

3.5. Pinching and uniform parabolicity. For many of the speeds we consider, the first derivatives with respect to the principal curvatures are not uniformly positive on Γ. This means that the parabolic operator appearing in the evolution of the second fundamental form and other curvature quantities can become degenerate if λ approaches $\partial\Gamma$ somewhere on a solution. To rule out this behaviour, we usually impose a uniform parabolicity assumption on solutions, by requiring that

$$\lambda(x,t) \in \Gamma', \qquad \forall\ (x,t) \in M \times [0,T),$$

where Γ' is a symmetric cone which is compactly contained in Γ. By this we mean that $\bar{\Gamma}' \cap \partial B(0,1)$ is a compact subset of Γ, and as a shorthand, if Γ' has this property then we write $\Gamma' \Subset \Gamma$.

In fact, if $\lambda \in \Gamma' \Subset \Gamma$, then we get bounds for all of the derivatives of γ at λ, as follows.

Lemma 1.2. *Let $\gamma : \Gamma \to (0,\infty)$ be an admissible speed, and suppose $\Gamma' \Subset \Gamma$. Then there is a constant $C = C(n,\gamma,\Gamma')$ such that if $A \in \mathrm{O}\,\Gamma'$, there holds*

$$C^{-1}\delta_{ij} \le \dot{\gamma}^{ij}(A) \le C\delta_{ij},$$

and

$$|A||\ddot{\gamma}^{ij,kl}(A)T_{ij}T_{kl}| \le C|T|^2$$

for each symmetric T. More generally, if k is an $(n\times n)$-matrix with entries in $\mathbb{N}$ and $|k| := \sum_{i,j} k_{ij}$ then we have

$$\mathrm{tr}(A)^{k-1}\left|\frac{\partial^{|k|}\gamma}{\partial^{k_{11}}A_{11}\dots\partial^{k_{nn}}A_{nn}}\right| \le C(n,k,\gamma,\Gamma').$$

Proof. Differentiation of the one-homogeneity condition yields

$$\lambda\dot{\gamma}^{ij}(A) = \frac{\partial}{\partial A_{ij}}\gamma(\lambda A) = \lambda\dot{\gamma}^{ij}(\lambda A),$$

so the first derivatives of γ are scaling-invariant. By assumption, the set

$$\bar{\Gamma}' \cap \partial B(0,1)$$

is a compact subset of Γ. It follows that

$$\mathrm{O}\,\Gamma' \cap \{A \in \mathrm{Sym}(n) : |A| = 1\}$$

is a compact subset of $\mathrm{O}\,\Gamma$, so since γ is smooth and elliptic, the following quantity is positive:

$$c_0 := \inf\{\dot{\gamma}^{ij}(A)\xi_i\xi_j : A \in \mathrm{O}\,\Gamma',\ |A| = 1,\ \xi \in \mathbb{R}^n\}.$$

Then because of the scale-invariance of $\dot{\gamma}^{ij}$, for every non-zero $A \in \mathrm{O}\,\Gamma'$ there holds

$$\dot{\gamma}^{ij}(A) = \dot{\gamma}^{ij}(|A|^{-1}A) \ge c_0\delta_{ij}.$$

The remaining estimates are proven using the same kind of argument - differentiating the homogeneity condition k times shows that k-th derivatives scale like $|A|^{1-k}$. $\square$

In all of the cases we are interested in, uniform parabolicity is guaranteed by the maximum principle. Taking φ to be the trace in the evolution equation for curvature quantities stated above, and defining $H(x,t) := \mathrm{tr}(A(x,t))$, we get

$$(\partial_t - \Delta_\gamma)H = |A|^2_\gamma H + g^{pq}\frac{\partial^2\gamma}{\partial A_{kl}\partial A_{ij}}\nabla_p A_{ij}\nabla_q A_{kl}.$$

Combining this with the equation for G sated above, we get

$$(\partial_t - \Delta_\gamma)\frac{H}{G} = \frac{1}{G}g^{pq}\frac{\partial^2\gamma}{\partial A_{kl}\partial A_{ij}}\nabla_p A_{ij}\nabla_q A_{kl} + \frac{2}{G}\left\langle \nabla\frac{H}{G}, \nabla G\right\rangle_\gamma,$$

where we have also introduced the notation $\langle v, w\rangle_\gamma := \dot\gamma^{ij}v_i w_j$. If γ is concave, applying the maximum principle to this equation gives:

LEMMA 1.3. *Let $F : M \times [0,T) \to \mathbb{R}^{n+1}$ be a solution of* (CF)*, where γ is a concave admissible speed. Then for each $(x,t) \in M \times [0,T)$ there holds*

$$\frac{H(x,t)}{G(x,t)} \leq \max_M \frac{H(\cdot,0)}{G(\cdot,0)}.$$

For each $\alpha \in (0,\infty)$ we define a cone as follows:

$$\Gamma_\alpha := \{\lambda \in \Gamma : \mathrm{tr}(\lambda) \leq \alpha\gamma(\lambda)\}.$$

With this notation, the lemma says that if $\lambda(M\times\{0\}) \subset \Gamma_\alpha$, then $\lambda(M\times[0,T)) \subset \Gamma_\alpha$. Since γ is concave and tr is linear, the sublevel sets of $\lambda \mapsto \gamma(\lambda)^{-1}\,\mathrm{tr}(\lambda)$ are convex, so each of the cones Γ_α is convex. If γ vanishes at the boundary of Γ and is strictly concave in off-radial directions, then for every $\alpha > 0$ such that Γ_α is nonempty, there holds

$$\Gamma_\alpha \Subset \Gamma.$$

Hence the lemma guarantees uniform parabolicity of the flow with $\Gamma' = \Gamma_\alpha$ and α depending only on $F(\cdot,0)$.

3.6. Short- and long-time existence. Suppose $\gamma : \Gamma \to (0,\infty)$ is an admissible speed, and let

$$F_0 : M \to \mathbb{R}^{n+1}$$

be a compact immersed hypersurface such that $\lambda(x) \in \Gamma$ for every $x \in M$. Then there is certainly a symmetric convex cone $\Gamma' \Subset \Gamma$ such that $\lambda(x) \in \Gamma'$ for each $x \in M$. Suppose we know a priori that, if $F : M\times[0,T) \to \mathbb{R}^{n+1}$ solves (CF), then $\lambda(x,t) \in \Gamma'$ remains true for every $(x,t) \in M \times [0,T)$. Then in particular, we have an a priori estimate of the form

$$|A|(x,t) \geq c(n,\gamma,\Gamma')G(x,t)$$

with $c > 0$ which is valid on all of $M\times[0,T)$. Furthermore, by the evolution equation for G stated above,

$$G(x,t) \geq \min_M G(\cdot,0)$$

for each $(x,t) \in M \times [0,T)$, so we have a positive a priori lower bound for $|A|$. Combining this with Lemma 1.2, we obtain a positive upper and lower bounds for the first derivatives of the speed, and uniform upper and lower bounds for all of its higher derivatives, all of which are valid on $M \times [0,T)$ and depend only on F_0.

With these bounds and the evolution of G, one can establish an a priori upper bound for the growth rate of

$$\max_M |A|(\cdot, t)$$

as t increases. Since $|A|$ controls the rate of change of the normal, this means that there is a uniformly positive time up to which any potential solution remains a graph over the initial hypersurface. With this fact, and the a priori estimates for the derivatives of the speed, the problem of solving (CF) on a short time interval starting from F_0 reduces to a uniformly parabolic scalar PDE problem for the graph representation. This problem always has a short-time solution - we refer to Section 3.5 of [**Lan14**] for further details.

If in addition to admitting a preserved cone $\Gamma' \Subset \Gamma$, the speed is convex or concave, then the solution remains smooth while the curvature is bounded. To be precise, if $F : M \times [0, T) \to \mathbb{R}^{n+1}$ solves (CF) and

$$\sup_{M\times[0,T)} |A|(x, t) \leq C,$$

then all of the derivatives of F are bounded in terms of C on $M \times [0, T)$. To sketch the argument, the curvature bound means the solution is locally a graph over spacetime sets of controlled size, and each local graph representation solves a uniformly parabolic fully nonlinear equation, where the parabolic operator is concave or convex in the Hessian. The bound on $|A|$ implies a C^2 estimate for the graph representation, and from this it is possible to deduce a $C^{2,\alpha}$ estimate (this fact is due independently to Evans [**Eva82**] and Krylov [**Kry82**]). Schauder theory then implies higher regularity. The technical results required to make this argument rigourous are contained in the appendix, but we do not give the details, and refer instead to Section 4.3 of [**Lan14**]. If $n = 2$, then all of this works even if γ is not convex or concave, since the Evans-Krylov theory can be replaced by a $C^{2,\alpha}$-estimate for fully nonlinear parabolic equations due to Andrews [**And04**].

CHAPTER 2

Cylindrical estimates

Our goal in this chapter is to prove a cylindrical estimate for immersed hypersurfaces moving by an admissible speed γ which is strictly concave in off-radial directions. The statement of the estimate can be summarised as follows: on any compact uniformly k-convex solution moving with inward normal velocity $G(x,t) = \gamma(\lambda(x,t))$, there holds

$$H \leq (\alpha_{k-1} + \varepsilon)G + C_\varepsilon,$$

where α_{k-1} is the value taken by $G^{-1}H$ on a cylinder $\mathbb{R}^{k-1} \times S^{n-k+1}$ and ε can be any positive number. The constant C_ε depends only on ε, the quality of the k-convexity, and the solution at the initial time. The proof we give is directly adapted from Brendle-Huisken [**BH17**][Theorem 3.1]. Analogous estimates for the mean curvature flow and other convex admissible speed functions were established in [**HS09**][Theorem 5.3] and [**AL14**], respectively. A cylindrical estimate for speeds which are only weakly concave is established in [**LL**][Theorem 1.1].

The cylindrical estimate implies that, as G blows up, the principal curvatures of the solution must approach the cone

$$\Gamma_{\alpha_{k-1}} := \{\lambda \in \Gamma : \mathrm{tr}(\lambda) \leq \alpha_{k-1}\gamma(\lambda)\},$$

which is the smallest sub-level set of the function

$$\lambda \mapsto \gamma(\lambda)^{-1}\,\mathrm{tr}(\lambda)$$

which contains the principal curvatures of every cylinder of the form $\mathbb{R}^{k-1} \times S^{n-k+1}$. Since it is possible to construct uniformly k-convex solutions which form singularities modeled on a homothetically shrinking $\mathbb{R}^{k-1} \times S^{n-k+1}$, the constant α_{k-1} is sharp. We note that the assumption that γ is strictly concave in off-radial directions ensures that the cone $\Gamma_{\alpha_{k-1}}$ sits inside the $(k-1)$-nonnegative cone (see Lemma 2.1 below), so the cylindrical estimate implies that the second fundamental form is becoming $(k-1)$-nonnegative at a singularity.

To establish the cylindrical estimate we prove an a priori supremum estimate for the pinching quantity

$$u_\sigma := \frac{H - (\alpha_{k-1} + \varepsilon)G}{G^{1-\sigma}},$$

where σ is chosen to be small depending on ε and the initial data. Let us sketch how this works. Since u_σ scales like G^σ, its evolution equation contains a positive zeroth-order term:

$$(\partial_t - \Delta_\gamma)u_\sigma = \sigma|A|_\gamma^2 u_\sigma + \dots$$

We can try to combat this term using the gradient of curvature term in the evolution of H, which is nonpositive by the concavity of γ, but this cannot work pointwise. Hence a direct maximum principle argument does not seem to be viable.

The idea, introduced by Huisken in [**Hui84**], is to instead use integral estimates and an iteration argument to get the desired supremum bound. One of the key observations is that, by integrating some form of Simons' identity, it is possible to bound integrals of certain curvature quantities in terms of integrals of gradients of the curvature. There are different ways of implementing this idea, but a particularly simple route is taken by Brendle-Huisken in [**BH17**]. They take the square of a symmetrised version of Simons' identity, and so obtain an identity of the following form:

$$\sum_{i,j} \lambda_i^2 \lambda_j^2 (\lambda_i - \lambda_j)^2 = A * A * A * \nabla^2 A,$$

where the $*$ denotes some contraction. The quantity on the left vanishes precisely when the principal curvatures take the form

$$\lambda_1 = \cdots = \lambda_m = 0, \qquad \lambda_{m+1} = \cdots = \lambda_n$$

for some m, but the uniform k-convexity and definition of α_{k-1} imply that this cannot occur over the support of u_σ. This crucial property of u_σ means that, on its support, the norm of the second fundamental form can be controlled by some term of the form $A * A * A * \nabla^2 A$, and an integration by parts argument then allows us to estimate integrals of $|A|^2 u_\sigma$ in terms of integrals of ∇u_σ and ∇A (this is all made precise in Proposition 2.9). These terms can then be controlled using the diffusion term and good gradient-of-curvature term in the evolution of u_σ, leading to powerful L^p-estimates. Along with the Michael-Simon Sobolev inequality, these are the key ingredient needed to carry out Huisken's Stampacchia iteration argument.

The chapter is divided into two halves. First we establish some basic properties of the cones Γ_α, and then state the cylindrical estimate precisely (see Theorem 2.4) before discussing some of its consequences. We then analyse the terms in the evolution equation of the quantity $H^{-1}G$ and reduce the proof of the cylindrical estimate to an application of the Stampacchia procedure. In the second half the Stampacchia argument is presented in a general form (this is Theorem 2.6) which implies a supremum estimate for a general function uG^σ, where σ is small and u satisfies certain structural conditions. Crucially, u must be a bounded subsolution of a certain parabolic equation where the right-hand side contains a negative gradient-of-curvature term, and at no point on the support of u can the second fundamental form equal that of a cylinder.

With this general supremum estimate established, in order to prove a pinching estimate, one only needs to identify an appropriate quantity u satisfying the structural conditions. Hence Theorem 2.6 can be used to recover the cylindrical and convexity estimates for mean curvature flow, and the analogous estimates for fully nonlinear flows mentioned above. We also use it to prove the convexity estimate in the next chapter.

1. Curvature cones

In this section we establish some elementary properties of the cones

$$\Gamma_\alpha = \{\lambda \in \mathbb{R}^n : \mathrm{tr}(\lambda) \leq \alpha\gamma(\lambda)\},$$

where $\gamma : \Gamma \to (0, \infty)$ is an admissible speed function which is strictly concave in off-radial directions and vanishes at $\partial\Gamma$. These assumptions imply that for every

$\alpha \in (0, \infty)$ there holds

$$\Gamma_\alpha \Subset \Gamma.$$

For each $0 \leq k \leq n-1$, we write

$$\mathrm{Cyl}_k := \{\lambda \cdot (e_{i_1} + \cdots + e_{i_{n-k}}) : \lambda > 0,\ 1 \leq i_1 < \cdots < i_{n-k} \leq n\},$$

where the e_i are the standard basis vectors in $\mathbb{R}^n$. This is the set of possible eigenvalue n-tuples of a cylinder $\mathbb{R}^k \times S^{n-k}$. We then write

$$\mathrm{Cyl} := \bigcup_{0 \leq k \leq n-1} \mathrm{Cyl}_k\,.$$

Observe that if $\mathrm{Cyl}_k \subset \Gamma$ then the value of $\gamma(\lambda)^{-1} \operatorname{tr}(\lambda)$ is constant for $\lambda \in \mathrm{Cyl}_k$, so we may define

$$\alpha_k := \gamma(\lambda)^{-1} \operatorname{tr}(\lambda), \qquad \lambda \in \mathrm{Cyl}_k\,.$$

The cones Γ_{α_k} will play an important role. We record some of their key properties in the following lemma.

Lemma 2.1. *Let γ be strictly concave in off-radial directions and fix $0 \leq k \leq n-2$. Then the following statements are true:*

(1) *The set Cyl_k is in the convex hull of Cyl_{k+1};*
(2) *Suppose $k \geq 1$ and $\mathrm{Cyl}_k \subset \Gamma$. Then for each $0 \leq l < k$, $\mathrm{Cyl}_l \subset \operatorname{int} \Gamma_{\alpha_k}$;*
(3) *If $\mathrm{Cyl}_l \subset \Gamma$ and $l > k$ then Cyl_l is in the complement of Γ_{α_k}.*

Proof. (1) Since Cyl_{k+1} consists of rays, its convex hull is a cone. A convex cone contains every positive linear combination of any finite subset of its elements, so we only need to show that each point in Cyl_k can be expressed as a positive linear combination of points in Cyl_{k+1}. By symmetry, it suffices to check this for one element of Cyl_k, so let us take

$$\hat{\lambda} = \sum_{i=k+1}^{n} e_i.$$

For each $j \in \{k+1, \dots, n\}$ the vector

$$\lambda_j := \hat{\lambda} - e_j$$

is an element of Cyl_{k+1}, and summing over j we get

$$\sum_{j=k+1}^{n} (\hat{\lambda} - e_j) = (n-k)\hat{\lambda} - \hat{\lambda} = (n-k-1)\hat{\lambda}.$$

We have $n-k-1 \geq 1$ by assumption, so

$$\hat{\lambda} = \frac{1}{n-k-1} \sum_{j=k+1}^{n} (\hat{\lambda} - e_j),$$

which is the required decomposition of $\hat{\lambda}$.

(2) Note that since Γ_{α_k} is convex and contains Cyl_k, it also contains the convex hull of Cyl_k. Claim (1) thus implies that, since $l < k$, $\mathrm{Cyl}_l \subset \Gamma_{\alpha_k}$. Suppose now that

$$\hat{\lambda} \in \mathrm{Cyl}_l \cap \partial \Gamma_{\alpha_k},$$

and relabel indices and rescale so that

$$\hat{\lambda} = \sum_{k=l+1}^{n} e_i.$$

Let $\lambda_j := \hat{\lambda} - e_j$ for each $l+1 \le j \le n$, and write X for the set of positive linear combinations of the vectors λ_j. Then X is an open cone of dimension $n-l-1 < n$ which, as we showed in (1), has $\hat{\lambda}$ in its interior. Since $X \in \Gamma_{\alpha_k}$, and the smooth hypersurface $\Sigma := \partial\Gamma_{\alpha_k} \setminus \{0\}$ contains $\hat{\lambda}$, the dimension of $\ker(A_\Sigma)$ must therefore have dimension at least $n-l-1$ at $\hat{\lambda}$. On the other hand $l < m \le n-1$, so $l \le n-2$, and the dimension of $\ker(A_\Sigma)$ is therefore at least two at the point $\hat{\lambda}$. Since γ is strictly concave in off-radial directions, the second fundamental form of Σ vanishes precisely in the radial direction at each point, so its kernel is one-dimensional. We have thus a reached a contradiction, so the original assumption must have been false. That is, $\hat{\lambda} \in \operatorname{int}\Gamma_{\alpha_k}$.

(3) We know by claim (2) that if $\mathrm{Cyl}_l \subset \Gamma$ and $k < l$ then $\mathrm{Cyl}_k \subset \operatorname{int}\Gamma_{\alpha_l}$. The interior of Γ_{α_l} consists precisely of those points λ where $\mathrm{tr}(\lambda) < \alpha_l\gamma(\lambda)$, so we have

$$\alpha_k = \frac{\mathrm{tr}}{\gamma}\bigg|_{\mathrm{Cyl}_k} < \alpha_l,$$

which implies $\Gamma_{\alpha_k} \subset \operatorname{int}\Gamma_{\alpha_l}$. Since Cyl_l is a subset of the boundary of Γ_{α_l}, this implies that Cyl_l is in the complement of Γ_{α_k}. □

Lemma 2.2. *Fix $k \ge 1$ and suppose $\mathrm{Cyl}_k \subset \Gamma$. Then for each $\lambda \in \Gamma_{\alpha_k}$ satisfying $\lambda_1 \le \cdots \le \lambda_n$ there holds*

$$\lambda_1 + \cdots + \lambda_k \ge 0,$$

with equality if and only if

$$\lambda_1 = \cdots = \lambda_k.$$

Proof. Fix a nonzero vector $\lambda \in \Gamma_{\alpha_k}$, normalise so that $\mathrm{tr}(\lambda) = 1$, and relabel the entries so that $\lambda_1 \le \cdots \le \lambda_n$. We first prove that

$$\lambda_1 + \cdots + \lambda_k \ge 0.$$

If $\lambda_1 \ge 0$ then there is nothing to prove, so let m be the largest natural number such that $\delta := \lambda_1 + \cdots + \lambda_m < 0$.

Since Γ_{α_k} is a symmetric convex cone, we can take any permutation of the entries of λ, add the result to λ, and get back an element of Γ_{α_k}. In particular, cycling over the first m entries and taking the average gives

$$\frac{\delta}{m}(e_1 + \cdots + e_m) + \sum_{i=m+1}^{n} \lambda_i e_i \in \Gamma_{\alpha_k}.$$

Similarly, we have

$$\hat{\lambda} := \frac{\delta}{m}(e_1 + \cdots + e_m) + \frac{1-\delta}{n-m}(e_{m+1} + \cdots + e_n) \in \Gamma_{\alpha_k}.$$

Appealing again to convexity, we find that for each $s \in (0,1]$,

$$(1-s)\hat{\lambda} + s(e_1 + \cdots + e_n) \in \operatorname{int}\Gamma_{\alpha_k}.$$

Since $\delta < 0$ by assumption, we can choose $s_0 \in (0,1]$ so that $(1-s_0)\delta/m + s_0 = 0$, and so find that

$$\left(\frac{(1-s_0)(1-\delta)}{n-m} + s_0\right)(e_{m+1} + \cdots + e_n) \in \operatorname{int}\Gamma_{\alpha_k},$$

so we have found an element common to Cyl_m and $\operatorname{int}\Gamma_{\alpha_k}$. Since Γ_{α_k} is a cone, and Cyl_m consists of individual rays, it follows that $\mathrm{Cyl}_m \subset \operatorname{int}\Gamma_{\alpha_k}$. Appealing then to part (3) of Lemma 2.1, we find that $m < k$. By the definition of m, we therefore have

$$\lambda_1 + \cdots + \lambda_k \geq 0.$$

Next we consider the equality case. Suppose $\lambda \in \Gamma_{\alpha_k}$ satisfies $\lambda_1 + \cdots + \lambda_k = 0$. Then if we set

$$\lambda' := \sum_{i=k+1}^{n} e_i,$$

the line segment $\overline{\lambda\lambda'}$ lies in the hyperplane P_0, where

$$P_\varepsilon := \{\lambda \in \mathbb{R}^n : \lambda_1 + \cdots + \lambda_k = \varepsilon\}.$$

We have just shown that P_ε is disjoint from Γ_{α_k} for all $\varepsilon < 0$, so any points which are common to P_0 and Γ_{α_k} must lie in $\partial\Gamma_{\alpha_k}$. In particular,

$$\overline{\lambda\lambda'} \subset \partial\Gamma_{\alpha_k}.$$

Since the boundary of Γ_{α_k} is strictly convex in off-radial directions, $\overline{\lambda\lambda'}$ must therefore lie in a ray emanating from the origin. Since λ' is in Cyl_k, this implies that λ is also in Cyl_k. □

As an easy corollary, we get the following:

Corollary 2.3. *Fix $1 \leq k \leq n-1$ and suppose $\mathrm{Cyl}_k \subset \Gamma$. Then for every $\eta > 0$, there is a constant $\varepsilon > 0$ depending on η such that if $\lambda \in \Gamma_{\alpha_k+\varepsilon}$ and $\lambda_1 \leq \cdots \leq \lambda_n$, then*

$$\lambda_1 + \cdots + \lambda_k \geq -\eta\,\mathrm{tr}(\lambda).$$

In addition, for each $0 \leq k \leq n-1$ with $\mathrm{Cyl}_k \subset \Gamma$, there is a constant ε' depending on η such that if $\lambda \in \Gamma_{\alpha_k+\varepsilon'}$ satisfies $\lambda_1 \leq \cdots \leq \lambda_n$ and $\lambda_1 + \cdots + \lambda_k \leq \varepsilon'\,\mathrm{tr}(\lambda)$, then

$$\sum_{1\leq i\leq k} |\lambda_i| + \sum_{i=k+1}^{n} \left|\lambda_i - \frac{\mathrm{tr}(\lambda)}{n-k}\right| \leq \eta\,\mathrm{tr}(\lambda).$$

Proof. If the first claim is false, then there is a positive number η and a sequence of points $\lambda^{(j)}$ for $j \in \mathbb{N}$ such that

$$\lambda^{(j)} \in \Gamma_{\alpha_k+1/j},$$

but

$$\lambda_1^{(j)} + \cdots + \lambda_k^{(j)} \geq -\eta\,\mathrm{tr}(\lambda^{(j)}).$$

Let us normalise so that $\mathrm{tr}(\lambda^{(j)}) = 1$ for each $j \in \mathbb{N}$. Since

$$\frac{\mathrm{tr}(\lambda^{(j)})}{\gamma(\lambda^{(j)})} \leq \alpha_k + \frac{1}{j},$$

and the sets $\Gamma_\alpha \cap \{\lambda \in \mathbb{R}^n : \mathrm{tr}(\lambda) = 1\}$ are compact, we can extract a subsequence converging to some limit $\hat{\lambda} \in \Gamma_{\alpha_k}$. The lemma then implies that

$$\hat{\lambda}_1 + \cdots + \hat{\lambda}_k \geq 0,$$

contradicting our assumption. The proof of the second claim proceeds in a very similar way. $\square$

2. The cylindrical estimate and consequences

We now give a precise statement of the cylindrical estimate. The conditions on the cone Γ' are stated quite generally, but the key case to keep in mind is when Γ is the $(k+1)$-positive cone, and $\Gamma' = \Gamma_\alpha$ for some $\alpha < \infty$. Then the condition $\lambda(x,t) \in \Gamma'$ is guaranteed by Lemma 1.3.

THEOREM 2.4. *Fix $n \geq 2$ and let $\gamma : \Gamma \to (0,\infty)$ be an admissible speed function which is strictly concave in off-radial directions and vanishes at $\partial\Gamma$. Let*

$$F : M \times [0,T) \to \mathbb{R}^{n+1}$$

be a compact evolving immersion satisfying

$$\partial_t F(x,t) = -G(x,t)\nu(x,t),$$

where $G(x,t) := \gamma(\lambda(x,t))$. Suppose there is a closed symmetric convex cone $\Gamma' \Subset \Gamma$ such that

$$\lambda(x,t) \in \Gamma' \qquad \forall\ (x,t) \in M \times [0,T),$$

and let

$$k := \max\{0 \leq l \leq n-1 : \mathrm{Cyl}_l \subset \Gamma'\}.$$

Then, for each $\varepsilon > 0$, there is a constant

$$C_\varepsilon = C_\varepsilon(n, \gamma, \Gamma', \sup_M G(\cdot, 0), \mu_0(M), T)$$

such that the inequality

$$H(x,t) \leq (\alpha_k + \varepsilon)G(x,t) + C_\varepsilon$$

holds for each $(x,t) \in M \times [0,T)$.

Before proving the theorem, let us draw some consequences by comparing with Corollary 2.3. Given any $\varepsilon > 0$, the theorem says that if $G(x,t)$ is sufficiently large depending on ε and the initial data, then

$$H(x,t) \leq (\alpha_k + \varepsilon)G(x,t). \tag{2}$$

Suppose first that $k = 0$, in which case the assumptions imply that the solution is uniformly convex. Then (2) and Corollary 2.3 tell us that there is an $\eta(\varepsilon)$ such that

$$\sum_{i=1}^n \left| \lambda_i(x,t) - \frac{1}{n}H(x,t) \right| \leq \eta H(x,t),$$

and $\eta(\varepsilon) \to 0$ as $\varepsilon \to 0$. That is, points where the curvature is extremely large are approximately umbilic.

If $k \geq 1$, then (2) and Corollary 2.3 tell us that $\lambda(x,t)$ is close to being k-positive: there is an $\eta(\varepsilon)$ such that

$$\lambda_1(x,t) + \cdots + \lambda_k(x,t) \geq -\eta(\varepsilon) G(x,t),$$

and $\eta(\varepsilon) \to 0$ as $\varepsilon \to 0$. In particular, if $k = 1$ then $A(x,t)$ is approximately nonnegative - this property is what allowed Brendle-Huisken to establish their convexity estimate in [**BH17**]. In addition to this, there is an $\eta'(\varepsilon)$ such that if

$$\lambda_1(x,t) + \cdots + \lambda_k(x,t) \leq \eta'(\varepsilon) H(x,t),$$

then λ is extremely close to Cyl_k:

$$\sum_{1\leq i \leq k} |\lambda_i(x,t)| + \sum_{i=k+1}^{n} \left| \lambda_i(x,t) - \frac{1}{n-k} H(x,t) \right| \leq \eta(\varepsilon) G(x,t).$$

The value of $\eta'(\varepsilon)$ also tends to zero as $\varepsilon \to 0$.

A key step in the proof of the cylindrical estimate involves extracting a uniform gradient of curvature term from the term involving the Hessian of γ appearing in the evolution of H. This argument is contained in the following lemma.

Lemma 2.5. *Let Γ' be a closed symmetric cone such that*

$$\Gamma' \Subset \Gamma \setminus \mathrm{Cyl}_{n-1},$$

and consider a symmetric $(n \times n)$-matrix A which is diagonal and has eigenvalues $\lambda \in \Gamma'$. Then there is a positive constant $c = c(n, \gamma, \Gamma')$ such that

$$\sum_k \ddot{\gamma}^{pq,rs}(A) T_{kpq} T_{krs} \leq -c \frac{|T|^2}{\mathrm{tr}(A)}$$

for every totally symmetric T.

Proof. Let S be any $n \times n$-matrix. Since γ is one-homogeneous there holds $\ddot{\gamma}^{pq,rs}(A) A_{pq} A_{rs} = 0$, so for any symmetric S we have

$$\begin{aligned}\ddot{\gamma}^{pq,rs}(A) S_{pq} S_{rs} &= \ddot{\gamma}^{pq,rs}(A) \left(S_{pq} - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)} A_{pq} \right) \left(S_{rs} - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)} A_{rs} \right) \\ &\quad + 2 \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)} \ddot{\gamma}^{pq,rs}(A) S_{pq} A_{rs}.\end{aligned}$$

The last term on the right vanishes, since $f(t) := \ddot{\gamma}^{pq,rs}(tS_{pq} + A_{pq})(tS_{rs} + A_{rs})$ is nonpositive and vanishes at $t = 0$, giving

$$0 = f'(0) = 2\ddot{\gamma}^{pq,rs} S_{pq} A_{rs}.$$

Now let

$$c_0 = \min\{-\ddot{\gamma}^{pq,rs}(A) S_{pq} S_{rs} : A \in \mathrm{O}\,\Gamma', \ \mathrm{tr}(A) = 1, \ \mathrm{tr}(S) = 0, \ |S| = 1\},$$

which is strictly positive, since γ is strictly concave in off-radial directions and the conditions $\mathrm{tr}(A) = 1$ and $\mathrm{tr}(S) = 0$ prevent S from being proportional to A over this compact set. By scaling, we have

$$\ddot{\gamma}^{pq,rs}(A) S_{pq} S_{rs} \leq -c_0 \frac{|S|^2}{\mathrm{tr}(A)}$$

for all $A \in \mathrm{O}\,\Gamma'$, as long as S is traceless. In particular, for every symmetric S we have

$$\ddot{\gamma}^{pq,rs}(A)\left(S_{pq} - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)}A_{pq}\right)\left(S_{rs} - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)}A_{rs}\right) \leq -\frac{c_0}{\mathrm{tr}(A)}\left|S - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)}A\right|^2.$$

Collecting these facts, we obtain

$$\ddot{\gamma}^{pq,rs}(A)S_{pq}S_{rs} \leq -\frac{c_0(n,\gamma,\Gamma')}{\mathrm{tr}(A)}\left|S - \frac{\mathrm{tr}(S)}{\mathrm{tr}(A)}A\right|^2.$$

Next we observe that since

$$4\sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2 \geq \sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq} - T_{pkq} + \frac{\mathrm{tr}(T_p)}{\mathrm{tr}(A)}A_{kq}\right)^2,$$

and A is diagonal, if T is totally symmetric then there holds

$$\begin{aligned}4\sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2 &\geq \sum_{k,p,q}\left(-\frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq} + \frac{\mathrm{tr}(T_p)}{\mathrm{tr}(A)}A_{kq}\right)^2\\ &= 2\sum_k \frac{|\lambda|^2 - \lambda_k^2}{\mathrm{tr}(\lambda)^2}\,\mathrm{tr}(T_k)^2.\end{aligned}$$

Let us define

$$c_1 := \min\{|\lambda|^2 - \lambda_k^2 : \lambda \in \Gamma',\ \mathrm{tr}(\lambda) = 1,\ 1 \leq k \leq n\}.$$

Since $|\lambda|^2 - \lambda_k^2$ only vanishes if $\lambda = 0$ or $\lambda \in \mathrm{Cyl}_{n-1}$, the assumption $\Gamma' \Subset \Gamma \setminus \mathrm{Cyl}_{n-1}$ ensures that $c_1 > 0$. It follows that

$$c_1\sum_k |\,\mathrm{tr}(T_k)|^2 \leq 2\sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2.$$

On the other hand,

$$\begin{aligned}|T|^2 &= \sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq} + \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2\\ &\leq 2\sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2 + 2\frac{|A|^2}{\mathrm{tr}(A)^2}\sum_k \mathrm{tr}(T_k)^2,\end{aligned}$$

so by setting

$$C_0 := \max\{|A|^2 : A \in \mathrm{O}\,\Gamma',\ \mathrm{tr}(A) = 1\},$$

we obtain

$$\begin{aligned}|T|^2 &= \sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq} + \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2\\ &\leq 2(1 + c_1^{-1}C_0)\sum_{k,p,q}\left(T_{kpq} - \frac{\mathrm{tr}(T_k)}{\mathrm{tr}(A)}A_{pq}\right)^2.\end{aligned}$$

Combining this with the bound from the previous paragraph completes the proof. □

With this lemma in place, we can now establish the cylindrical estimate by applying Huisken's Stampacchia procedure. It will be convenient to give the arguments out of order - we establish the key supremum estimate in Theorem 2.6 just below, and in the meantime demonstrate that Theorem 2.6 implies Theorem 2.4.

Proof of Theorem 2.4 assuming Theorem 2.6. Recall that we are assuming

$$\lambda(x,t) \in \Gamma'$$

for each $(x,t) \in M \times [0,T)$, where $\Gamma' \Subset \Gamma$. Also, $\mathrm{Cyl}_l \subset \Gamma'$ if and only if $l \leq k$. Fix a small constant $\varepsilon > 0$ and consider the function

$$u(x,t) := \frac{H(x,t)}{G(x,t)} - \alpha_k - \varepsilon.$$

We claim that on the support of u, the principal curvatures of the solution are compactly contained away from Cyl. To see this, observe that if $u(x_0,t_0) > 0$, there holds

$$\lambda(x_0,t_0) \in \Gamma' \setminus \Gamma_{\alpha_k+\varepsilon} =: \Gamma''.$$

By Lemma 2.1, we know that $\mathrm{Cyl}_l \Subset \Gamma_{\alpha_k+\varepsilon}$ for every $l \leq k$. If $k < l \leq n-1$ and $\mathrm{Cyl}_l \subset \Gamma$, then by the definition of k there holds $\mathrm{Cyl}_l \Subset \Gamma \setminus \Gamma'$. Putting these statements together, we get

$$\Gamma'' \Subset \Gamma \setminus \mathrm{Cyl}\,.$$

We saw in the proof of Lemma 1.3 that u satisfies the equation

$$(\partial_t - \Delta_\gamma)u = \frac{1}{G} g^{kl}\ddot{\gamma}^{pq,rs}\nabla_k A_{pq}\nabla_l A_{rs} + \frac{2}{G}\langle \nabla u, \nabla G\rangle_\gamma.$$

In light of the previous paragraph, we can choose an orthonormal basis of principal directions and apply Lemma 2.5 at each point in $\mathrm{sp}(u)$ to obtain

$$(\partial_t - \Delta_\gamma)u \leq -c_0(n,\gamma,\Gamma',\varepsilon)\frac{|\nabla A|^2}{G^2} + \frac{2}{G}\langle \nabla u, \nabla G\rangle_\gamma.$$

Here we have also used the fact that γ and tr are comparable on Γ'. We now make use of Young's inequality to estimate the remaining gradient term, and so arrive at

$$(\partial_t - \Delta_\gamma)u \leq -c_0(n,\gamma,\Gamma',\varepsilon)\frac{|\nabla A|^2}{G^2} + r\frac{|\nabla G|_\gamma^2}{G^2} + \frac{1}{r}|\nabla u|_\gamma^2,$$

where r can be taken to be any positive number. By setting

$$C_0 := \max\{\dot{\gamma}^i(\lambda) : \lambda \in \Gamma',\ \mathrm{tr}(\lambda) = 1,\ 1 \leq i \leq n\},$$

and computing in an orthonormal basis of eigenvectors for A, we get

$$|\nabla G|_\gamma^2 = \dot{\gamma}^k\dot{\gamma}^p\dot{\gamma}^q\nabla_k A_{pp}\nabla_k A_{qq} \leq C_1(n,\gamma,\Gamma')|\nabla A|^2.$$

Therefore, by setting $r = c_0/2C_1$ we ensure that

$$(\partial_t - \Delta_\gamma)u \leq -c_1\frac{|\nabla A|^2}{G^2} + C_2|\nabla u|^2,$$

where c_1 and C_2 depend only on n, γ, Γ' and ε. Since $u \leq C_3(n,\gamma,\Gamma')$, we can pass to the weaker estimate

$$(\partial_t - \Delta_\gamma)u \leq -c_1C_3^{-1}u\frac{|\nabla A|^2}{G^2} + C_2C_3\frac{|\nabla u|^2}{u}$$

on $\operatorname{sp}(u)$. Applying Theorem 2.6 (with $k_0 = 0$), we obtain the estimate

$$u(x,t) \le \varepsilon + C_\varepsilon G(x,t)^{-1},$$

where $C_\varepsilon = C_\varepsilon(n, \gamma, \Gamma', \varepsilon, \sup G(\cdot, 0), \mu_0(M), T)$, and since ε can be taken to be any positive number, this proves the claim. $\square$

3. Huisken's Stampacchia procedure

In this section we prove a general pinching estimate for hypersurface flows by admissible speed functions. Consequences of this result include the cylindrical estimate stated in the previous section and the convexity estimate of the next chapter. The result can also be used to recover the analogous cylindrical and convexity estimates for mean curvature flow, and other flows as well.

THEOREM 2.6. *Fix $n \ge 2$ and let $\gamma : \Gamma \to (0,\infty)$ be an admissible speed. Let $F : M \times [0,T) \to \mathbb{R}^{n+1}$ be compact evolving immersion which satisfies*

$$\partial_t F(x,t) = -G(x,t)\nu(x,t)$$

for every $(x,t) \in M \times [0,T)$, where $G(x,t) := \gamma(\lambda(x,t))$. Let $u : M \times [0,T) \to \mathbb{R}$ be a smooth function satisfying $u \le C_0$, set

$$R^{-1} := \sup_M G(\cdot, 0),$$

and suppose there is a constant $k_0 > 0$ and a symmetric cone $\Gamma' \Subset \Gamma \setminus \mathrm{Cyl}$ such that

$$\lambda(x,t) \in \Gamma' \qquad \forall\ (x,t) \in \operatorname{sp}(u) \cap \operatorname{sp}(G - k_0 R^{-1}).$$

Assume also that there are positive constants C_1, C_2, C_3, C_4 and $\delta \in (0,2]$ such that

$$(\partial_t - \Delta_\gamma) u \le C_1 \frac{|\nabla u|^2}{u} - \frac{1}{C_2} u \frac{|\nabla A|^2}{G^2} + C_3 |A|^{2-\delta} + C_4 \tag{3}$$

holds at every point in $\operatorname{sp}(u) \cap \operatorname{sp}(G - k_0 R^{-1})$. Set $C' := (C_0, C_1, C_2, C_3, C_4)$. Then for every $\varepsilon > 0$ there is a constant

$$K_\varepsilon = K_\varepsilon(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T)$$

such that

$$u(x,t) \le \varepsilon + K_\varepsilon G(x,t)^{-1}$$

for each $(x,t) \in M \times [0,T)$.

REMARK 2.7. *Suppose T is the maximal time of smooth existence. By scaling, the constant K_ε can be written as $K_\varepsilon = \tilde{K}_\varepsilon R^{-1}$, where*

$$\tilde{K}_\varepsilon = \tilde{K}_\varepsilon(n, \gamma, \Gamma', k_0, C', \delta, \theta_1, \theta_2, \theta_3)$$

and

$$\theta_1 := \min_{M \times [0,T)} \frac{|A|_\gamma^2}{G^2}; \qquad \theta_2 := \frac{\inf_M G(\cdot,0)}{\sup_M G(\cdot,0)}; \qquad \theta_3 := \sup_M G(\cdot,0)^n \cdot \mu_0(M).$$

Here we have appealed to the fact that T can be bounded from above in terms of θ_1, θ_2 and R^{-1}. This is a consequence of the evolution equation

$$(\partial_t - \Delta_\gamma) G = |A|_\gamma^2 G.$$

If the derivatives of γ are bounded from above then θ_1 is automatically bounded from below by the Cauchy-Schwarz inequality.

REMARK 2.8. *The conclusion of the theorem remains true if, rather than being smooth, u is only locally Lipschitz and satisfies the differential inequality in the following weak sense: for every nonnegative Lipschitz function $\varphi : M \times [0,T) \to \mathbb{R}$ supported in $\mathrm{sp}(u) \cap \mathrm{sp}(G - k_0 R^{-1})$, the inequality*

$$\begin{aligned}\int_M \varphi \partial_t u \, d\mu_t \leq &- \int_M \langle \nabla u, \nabla \varphi \rangle_\gamma \, d\mu_t - \int_M \varphi \ddot{\gamma}^{ij,pq} \nabla_i A_{pq} \nabla_j u \, d\mu_t + C_1 \int_M \varphi \frac{|\nabla u|^2}{u} \, d\mu_t \\ &- \frac{1}{C_2} \int_M \varphi u \frac{|\nabla A|^2}{G^2} \, d\mu_t + C_3 \int_M |A|^{2-\delta} \varphi \, d\mu_t + C_4 \int_M \varphi \, d\mu_t \qquad (4)\end{aligned}$$

holds for almost every $t \in [0,T)$. If u is smooth and satisfies (3) *then this inequality is a consequence of the divergence theorem. In the proof of the theorem, we will take care to only use this weaker assumption. This will be important in the next chapter, where we need to apply the theorem to λ_1, which may not be smooth.*

3.1. Poincaré and Sobolev inequalities. The following Poincaré-type inequality first appeared in the form used here in [**BH17**] (see the proof of Theorem 3.1). Earlier incarnations were used to prove pinching estimates in [**Hui84**] and [**HS09**].

PROPOSITION 2.9. *Let M be a smooth hypersurface in $\mathbb{R}^{n+1}$ satisfying $|A| > 0$ and consider a compactly supported Lipschitz function $u : M \to \mathbb{R}$. Suppose that u is nonnegative, and that there is a symmetric cone*

$$\Gamma' \Subset \{\lambda \in \mathbb{R}^n : \mathrm{tr}(\lambda) > 0\} \setminus \mathrm{Cyl}$$

such that

$$\lambda(x) \in \Gamma' \qquad \forall\, x \in \mathrm{sp}(u).$$

Then there is a positive constant $c = c(n, \Gamma')$ with the property that for every $r > 0$,

$$c \int_M u^2 |A|^2 \, d\mu \leq \frac{1}{r} \int_M |\nabla u|^2 \, d\mu + (1+r) \int_M u^2 \frac{|\nabla A|^2}{|A|^2} \, d\mu,$$

where μ is the induced measure on M.

PROOF. Simons' identity implies that the second fundamental form of M satisfies

$$\nabla_i \nabla_j A_{kl} + \nabla_j \nabla_i A_{kl} - \nabla_k \nabla_l A_{ij} - \nabla_l \nabla_k A_{ij} = 2C_{ijkl},$$

where C is defined to be

$$C = A \otimes A^2 - A^2 \otimes A.$$

In a principal frame for A, the only non-zero components of C are $C_{iijj} = \lambda_i \lambda_j (\lambda_i - \lambda_j)$ for $i \neq j$, so

$$|C|^2 = \sum_{1 \leq i,j \leq n} \lambda_i^2 \lambda_j^2 (\lambda_i - \lambda_j)^2. \qquad (5)$$

Let $\Gamma = \{\lambda \in \mathbb{R}^n : \mathrm{tr}(\lambda) > 0\}$ and consider the function $h : \Gamma \to [0, \infty)$ defined by

$$h(\lambda) := \sum_{1 \leq i < j \leq n} \frac{\lambda_i^2 \lambda_j^2 (\lambda_i - \lambda_j)^2}{|\lambda|^6}.$$

The norm is positive on Γ, so the right-hand side is well defined. We then define c to be the constant

$$c := \inf\{h(\lambda) : \lambda \in \Gamma',\ |\lambda| = 1\},$$

which is guaranteed to be positive, since h vanishes precisely on Cyl, but by assumption $\Gamma' \Subset \Gamma \setminus \mathrm{Cyl}$. By the scaling-invariance, we then have

$$h(\lambda) = h(\lambda/|\lambda|) \geq c$$

for each $\lambda \in \Gamma'$, which is to say that

$$\sum_{1\leq i<j\leq n} \lambda_i^2\lambda_j^2(\lambda_i - \lambda_j)^2 \geq c|\lambda|^6.$$

This means that at each point in $\mathrm{sp}(u)$,

$$|C|^2 \geq c|A|^6.$$

We thus have

$$c\int_M u^2|A|^2\,d\mu \leq \int_M u^2|A|^{-4}|C|^2\,d\mu.$$

Using Simons' identity and the symmetry of A, we obtain

$$c\int_M u^2|A|^2\,d\mu \leq 8\int_M u^2|A|^{-4}C^{ijkl}(\nabla_i\nabla_j A_{kl} - \nabla_k\nabla_l A_{ij})\,d\mu.$$

Consider the first Hessian term on the right (the remaining term can be handled in the same way). Setting $T^i := u^2|A|^{-4}C^{ijkl}\nabla_j A_{kl}$, we may write

$$\begin{aligned} u^2|A|^{-4}C^{ijkl}\nabla_i\nabla_j A_{kl} &= \nabla_i T^i - 2u\nabla_i u|A|^{-4}C^{ijkl}\nabla_j A_{kl} \\ &\quad - 4u^2|A|^{-5}\nabla_i|A|C^{ijkl}\nabla_j A_{kl} - u^2|A|^{-4}\nabla_i C^{ijkl}\nabla_j A_{kl}. \end{aligned}$$

The divergence term vanishes upon integration, and there is a constant $K = K(n)$ such that

$$|C| \leq K|A|^3, \qquad |\nabla|A|| \leq K|\nabla A|, \qquad |\nabla C| \leq K|A|^2|\nabla A|,$$

so we can bound

$$\begin{aligned} &\int_M u^2|A|^{-4}C^{ijkl}\nabla_i\nabla_j A_{kl}\,d\mu \\ &\qquad \leq K\int_M \big(u|\nabla u||A|^{-1} + u^2|A|^{-2}|\nabla A|\big)\,|\nabla A|\,d\mu, \end{aligned}$$

where K is a larger constant, still depending only on n. The claim now follows from Young's inequality. $\square$

We will also make use of the Michael-Simon Sobolev inequality [**MS73**].

PROPOSITION 2.10. *Let u be a compactly supported Lipschitz function on a smooth hypersurface M. Then there holds*

$$\left(\int_M |u|^{\frac{n}{n-1}}\,d\mu\right)^{\frac{n-1}{n}} \leq C(n)\int_M |\nabla u| + |H||u|\,d\mu.$$

Let us record here a straightforward consequence of the Sobolev inequality for later use:

COROLLARY 2.11. *Let u be a compactly supported nonnegative Lipschitz function on a smooth hypersurface M. Then the inequality*

$$\Big(\int_M u^{2q}\,d\mu\Big)^{\frac{1}{q}} \le C(n)\int_M |\nabla u|^2\,d\mu + C(n)\Big(\int_{\mathrm{sp}(u)} |H|^n\,d\mu\Big)^{\frac{2}{n}}\Big(\int_M u^{2q}\,d\mu\Big)^{\frac{1}{q}}$$

holds with $q = \frac{n}{n-2}$ if $n \ge 3$. If $n = 2$, then for every $q \ge 1$,

$$\Big(\int_M u^{2q}\,d\mu\Big)^{\frac{1}{q}} \le Cq^2\mu(M)^{\frac{1}{q}}\int_M |\nabla u|^2\,d\mu + C\int_{\mathrm{sp}(u)} |H|^2\,d\mu\Big(\int_M u^{2q}\,d\mu\Big)^{\frac{1}{q}}.$$

PROOF. Let $\beta \ge 1$ be a real number to be determined later. Then u^β is a Lipschitz function, so the Michael-Simon Sobolev inequality and Hölder's inequality give

$$\begin{aligned}\Big(\int_M u^{\frac{n\beta}{n-1}}\,d\mu\Big)^{\frac{n-1}{n}} &\le C\int_M \beta u^{\beta-1}|\nabla u| + |H|u^\beta\,d\mu\\ &\le C\beta\Big(\int_M u^{2(\beta-1)}\,d\mu\Big)^{\frac{1}{2}}\Big(\int_M |\nabla u|^2\,d\mu\Big)^{\frac{1}{2}}\\ &\quad + C\Big(\int_{\mathrm{sp}(u)} |H|^n\,d\mu\Big)^{\frac{1}{n}}\Big(\int_M u^{\frac{n\beta}{n-1}}\,d\mu\Big)^{\frac{n-1}{n}}.\end{aligned}$$

Suppose for now that $n > 2$. Let us square both sides and choose $\beta = 2\frac{n-1}{n-2}$. This ensures that $2(\beta - 1) = \frac{n\beta}{n-1}$, so we obtain

$$\begin{aligned}\Big(\int_M u^{\frac{2n}{n-2}}\,d\mu\Big)^{\frac{2(n-1)}{n}} &\le C\int_M u^{\frac{2n}{n-2}}\,d\mu\cdot\int_M |\nabla u|^2\,d\mu\\ &\quad + C\Big(\int_{\mathrm{sp}(u)} |H|^n\,d\mu\Big)^{\frac{2}{n}}\Big(\int_M u^{\frac{2n}{n-2}}\,d\mu\Big)^{\frac{2(n-1)}{n}},\end{aligned}$$

which gives the desired estimate after canceling.

When $n = 2$ we proceed as before, but use the Hölder inequality again to get

$$\begin{aligned}\int_M u^{2\beta}\,d\mu &\le C\beta^2\int_M u^{2(\beta-1)}\,d\mu\cdot\int_M |\nabla u|^2\,d\mu + C\int_{\mathrm{sp}(u)} |H|^2\,d\mu\cdot\int_M u^{2\beta}\,d\mu\\ &\le C\beta^2\mu(M)^{\frac{1}{\beta}}\Big(\int_M u^{2\beta}\,d\mu\Big)^{\frac{\beta-1}{\beta}}\Big(\int_M |\nabla u|^2\,d\mu\Big)\\ &\quad + C\int_{\mathrm{sp}(u)} |H|^2\,d\mu\cdot\int_M u^{2\beta}\,d\mu,\end{aligned}$$

and then simply rearrange to obtain the desired estimate. $\square$

3.2. Modifying u. Let $u_\sigma(x,t) := u(x,t)G(x,t)^\sigma$. We want to estimate the quantity $(\partial_t - \Delta_\gamma)u_\sigma$, assuming u is smooth and satisfies (3). For v positive there holds

$$\begin{aligned}(\partial_t - \Delta_\gamma)(uv) &= u(\partial_t - \Delta_\gamma)v + v(\partial_t - \Delta_\gamma)u - 2\langle\nabla u, \nabla v\rangle_\gamma\\ &= u(\partial_t - \Delta_\gamma)v + v(\partial_t - \Delta_\gamma)u - \frac{2}{v}\langle\nabla(uv), \nabla v\rangle_\gamma + 2\frac{u}{v}|\nabla v|_\gamma^2,\end{aligned}$$

and

$$(\partial_t - \Delta_\gamma)G^\sigma = \sigma|A|_\gamma^2 G^\sigma - \sigma(\sigma-1)G^{\sigma-2}|\nabla G|_\gamma^2.$$

Combining these formulae, we arrive at

$$\begin{aligned}(\partial_t - \Delta_\gamma)u_\sigma &= u(\sigma|A|_\gamma^2 G^\sigma - \sigma(\sigma-1)G^{\sigma-2}|\nabla G|_\gamma^2) + G^\sigma(\partial_t - \Delta_\gamma)u \\ &\quad - \frac{2}{G^\sigma}\langle \nabla u_\sigma, \nabla G^\sigma\rangle_\gamma + 2\frac{u}{G^\sigma}|\nabla G^\sigma|_\gamma^2 \\ &= \sigma|A|_\gamma^2 u_\sigma + \sigma(\sigma+1)u_\sigma\frac{|\nabla G|_\gamma^2}{G^2} - 2\sigma\left\langle \nabla u_\sigma, \frac{\nabla G}{G}\right\rangle_\gamma \\ &\quad + G^\sigma(\partial_t - \Delta_\gamma)u.\end{aligned}$$

If the inequality

$$(\partial_t - \Delta_\gamma)u \leq C_1\frac{|\nabla u|^2}{u} - \frac{1}{C_2}u\frac{|\nabla A|^2}{G^2} + C_3|A|^{2-\delta} + C_4,$$

holds on $\mathrm{sp}(u)\cap\mathrm{sp}(G - k_0R^{-1})$, then on this same set,

$$\begin{aligned}(\partial_t - \Delta_\gamma)u_\sigma &\leq \sigma|A|_\gamma^2 u_\sigma + \sigma(\sigma+1)u_\sigma\frac{|\nabla G|_\gamma^2}{G^2} - 2\sigma\left\langle \nabla u_\sigma, \frac{\nabla G}{G}\right\rangle_\gamma \\ &\quad + C_1G^\sigma\frac{|\nabla u|^2}{u} - C_2^{-1}uG^\sigma\frac{|\nabla A|^2}{G^2} + C_3|A|^{2-\delta}G^\sigma + C_4G^\sigma.\end{aligned}$$

Recall that in Theorem 2.6 we assume

$$\lambda(x,t)\in\Gamma' \qquad \forall (x,t)\in \mathrm{sp}(u)\cap\mathrm{sp}(G-k_0R^{-1}),$$

where $\Gamma' \Subset \Gamma\setminus\mathrm{Cyl}$. Under this assumption there is a $C_5 = C_5(n,\gamma,\Gamma',C_1)$ such that on $\mathrm{sp}(u)\cap\mathrm{sp}(G-k_0R^{-1})$ there holds

$$C_1G^\sigma\frac{|\nabla u|^2}{u} \leq C_5G^\sigma\frac{|\nabla u|_\gamma^2}{u} = C_5\frac{|\nabla u_\sigma|_\gamma^2}{u_\sigma} - 2C_5\sigma\left\langle \nabla u_\sigma, \frac{\nabla G}{G}\right\rangle_\gamma + C_5\sigma^2u_\sigma\frac{|\nabla G|_\gamma^2}{G^2},$$

and consequently, the inequality

$$\begin{aligned}(\partial_t - \Delta_\gamma)u_\sigma &\leq \sigma|A|_\gamma^2u_\sigma + C_5\frac{|\nabla u_\sigma|_\gamma^2}{u_\sigma} + \sigma((1+C_5)\sigma+1)u_\sigma\frac{|\nabla G|_\gamma^2}{G^2} \\ &\quad - 2(1+C_5)\sigma\left\langle \nabla u_\sigma, \frac{\nabla G}{G}\right\rangle_\gamma - C_2^{-1}u_\sigma\frac{|\nabla A|^2}{G^2} + C_3|A|^{2-\delta}G^\sigma + C_4G^\sigma\end{aligned}$$

holds on $\mathrm{sp}(u)\cap\mathrm{sp}(G-k_0R^{-1})$. By Young's inequality there is a $C_6 = C_6(n,\gamma,\Gamma',C_1)$ such that on $\mathrm{sp}(u)\cap\mathrm{sp}(G-k_0R^{-1})$,

$$(\partial_t - \Delta_\gamma)u_\sigma \leq \sigma|A|_\gamma^2u_\sigma + C_6\frac{|\nabla u_\sigma|_\gamma^2}{u_\sigma} - (C_2^{-1} - C_6\sigma)u_\sigma\frac{|\nabla A|^2}{G^2} + C_3|A|^{2-\delta}G^\sigma + C_4G^\sigma.$$

Here we have used the fact that at points where $\lambda\in\Gamma'$ we have a bound of the form

$$|\nabla G|_\gamma^2 = \dot\gamma^i\gamma^j\gamma^k\nabla_iA_{jj}\nabla_iA_{kk} \leq C(n,\gamma,\Gamma')|\nabla A|^2.$$

If u is only Lipschitz and satisfies (4), then similar considerations show that for every Lipschitz test function φ supported in $\mathrm{sp}(u) \cap \mathrm{sp}(G - k_0 R^{-1})$ there holds

$$\begin{aligned}\int_M \varphi \partial_t u_\sigma \, d\mu_t \leq & -\int_M \langle \nabla u_\sigma, \nabla \varphi \rangle_\gamma \, d\mu_t - \int_M \varphi \ddot{\gamma}^{ij,pq} \nabla_i A_{pq} \nabla_j u_\sigma \, d\mu_t \\ & + C_6 \int_M \varphi \frac{|\nabla u_\sigma|_\gamma^2}{u_\sigma} \, d\mu_t - (C_2^{-1} - C_6 \sigma) \int_M \varphi u_\sigma \frac{|\nabla A|^2}{G^2} \, d\mu_t \\ & + \sigma \int_M |A|_\gamma^2 u_\sigma \varphi \, d\mu_t + C_3 \int_M |A|^{2-\delta} G^\sigma \varphi \, d\mu_t + C_4 \int_M G^\sigma \varphi \, d\mu_t. \end{aligned} \tag{6}$$

3.3. The L^p-estimate. We are now ready to prove the crucial L^p-estimate for the function

$$u_{\sigma,k} := \max\{u_\sigma - k, 0\}.$$

PROPOSITION 2.12. *Let F and u satisfy the assumptions of Theorem 2.6. Then there are positive constants p_0 and ℓ_0 which depend only n, γ, Γ', C' and δ, and a positive constant $k_1 = k_1(C_0, k_0, R)$, with the following property. For all k, σ and p satisfying*

$$k \geq C_0 \cdot \max\{1, k_0 R^{-1}\}, \qquad p \geq p_0, \qquad \sigma \leq \ell_0 p^{-\frac{1}{2}},$$

there holds

$$\sup_{t \in [0,T)} \left(\int_M u_{\sigma,k}^p \, d\mu_t \right) \leq C$$

for some

$$C = C(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T, k, \sigma, p).$$

The same inequality also holds if u is only a Lipschitz function satisfying (4).

PROOF. Consider constants $\sigma \in (0,1)$ and $k > 0$. Observe that, on $\mathrm{sp}(u_{\sigma,k})$, we have

$$k \leq u G^\sigma \leq C_0 G^\sigma,$$

so by assuming $k \geq k_1 := C_0 \cdot \max\{1, k_0 R^{-1}\}$, we ensure that

$$G \geq \max\{1, k_0 R^{-1}\}^{\frac{1}{\sigma}} \geq \max\{1, k_0 R^{-1}\}$$

holds on $\mathrm{sp}(u_{\sigma,k})$. By assumption, we then have that $\lambda(x,t) \in \Gamma'$ holds for every $(x,t) \in \mathrm{sp}(u_{\sigma,k})$, and in addition to this, the function $u_{\sigma,k}$ is a valid test function to which we can apply (4) (or more precisely (6)). For $p \geq 2$ the function $p u_{\sigma,k}^{p-1}$ is also a valid test function, and inserting this choice into (6) gives

$$\begin{aligned} p \int_M u_{\sigma,k}^p \partial_t u_\sigma \, d\mu_t \leq & -p(p-1) \int_M u_{\sigma,k}^{p-2} |\nabla u_\sigma|_\gamma^2 \, d\mu_t - p \int_M u_{\sigma,k}^{p-1} \ddot{\gamma}^{ij,pq} \nabla_i A_{pq} \nabla_j u_\sigma \, d\mu_t \\ & + C_6 p \int_M u_{\sigma,k}^{p-1} u_\sigma^{-1} |\nabla u_\sigma|_\gamma^2 \, d\mu_t - (C_2^{-1} - C_6 \sigma) p \int_M u_{\sigma,k}^{p-1} u_\sigma \frac{|\nabla A|^2}{G^2} \, d\mu_t \\ & + \sigma p \int_M |A|_\gamma^2 u_\sigma u_{\sigma,k}^{p-1} \, d\mu_t + C_3 p \int_M |A|^{2-\delta} G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t \\ & + C_4 p \int_M G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t \end{aligned}$$

for almost every $t \in [0, T)$. For $p \geq 2$ and

$$\ell_0 \leq \frac{1}{2C_2C_6},$$

the condition $\sigma \leq \ell_0 p^{-\frac{1}{2}}$ ensures that

$$-(C_2^{-1} - C_6\sigma)p \int_M u_{\sigma,k}^{p-1} u_\sigma \frac{|\nabla A|^2}{G^2}\, d\mu_t \leq -c_0 p \int_M u_{\sigma,k}^p \frac{|\nabla A|^2}{G^2}\, d\mu_t$$

where $c_0 := (2C_2)^{-1}$. Using

$$\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t = p\int_M u_{\sigma,k}^{p-1}\partial_t u_\sigma \, d\mu_t - \int_M HG u_{\sigma,k}^p \, d\mu_t$$

and estimating

$$C_6 p \int_M u_{\sigma,k}^{p-1} u_\sigma^{-1} |\nabla u_\sigma|_\gamma^2 \, d\mu_t \leq C_6 p \int_M u_{\sigma,k}^{p-2} |\nabla u_\sigma|_\gamma^2 \, d\mu_t$$

we obtain

$$\begin{aligned}
\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t \leq{}& -(p(p-1) - C_6 p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2 \, d\mu_t \\
&- p\int_M u_{\sigma,k}^{p-1}\ddot{\gamma}^{ij,pq}\nabla_i A_{pq}\nabla_j u_\sigma \, d\mu_t - c_0 p \int_M u_{\sigma,k}^p \frac{|\nabla A|^2}{G^2}\, d\mu_t \\
&+ \sigma p \int_M |A|_\gamma^2 u_\sigma u_{\sigma,k}^{p-1}\, d\mu_t + C_3 p \int_M |A|^{2-\delta} G^\sigma u_{\sigma,k}^{p-1}\, d\mu_t \\
&+ C_4 p \int_M G^\sigma u_{\sigma,k}^{p-1}\, d\mu_t.
\end{aligned}$$

Choose $C_7(n, \gamma, \Gamma')$ so large that

$$|\ddot{\gamma}^{ij,rs}(A)| \leq C_7 \gamma(A)^{-1}$$

holds for all combinations of indices and every $A \in \mathrm{O}\,\Gamma'$. Then, working in an orthonormal frame at any point in $\mathrm{sp}(u_{\sigma,k})$, and using Young's inequality, we can bound

$$\begin{aligned}
p u_{\sigma,k}^{p-1}\ddot{\gamma}^{ij,rs}\nabla_i A_{rs}\nabla_j u_\sigma &\leq C_7 p \sum_{i,j,r,s} u_{\sigma,k}^{p-1} G^{-1} |\nabla_i A_{rs}||\nabla_j u_\sigma| \\
&\leq C_7 p^{\frac{3}{2}} \sum_{i,j,r,s} u_{\sigma,k}^{p-2}|\nabla_j u_\sigma|^2 + C_7 p^{\frac{1}{2}} \sum_{i,j,r,s} u_{\sigma,k}^p \frac{|\nabla_i A_{rs}|^2}{G^2} \\
&\leq C_8 p^{\frac{3}{2}} u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2 + C_8 p^{\frac{1}{2}} u_{\sigma,k}^p \frac{|\nabla A|^2}{G^2}.
\end{aligned}$$

Here C_8 depends only on n, γ and Γ'. Hence there holds

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \le &-(p(p-1)-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t\\ &+\sigma p\int_M |A|_\gamma^2 u_{\sigma,k}^{p-1}u_\sigma\,d\mu_t + C_3p\int_M |A|^{2-\delta}G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t\\ &+C_4p\int_M G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t \end{aligned} \tag{7}$$

Next we set

$$C_9(n,\gamma,\Gamma') := \sup\{\dot\gamma^i(\lambda) : \lambda\in\Gamma',\ 1\le i\le n\}$$

so that we have $|A|_\gamma^2 \le C_9|A|^2$ on $\mathrm{sp}(u_{\sigma,k})$. Using this we bound

$$\begin{aligned}\sigma p\int_M |A|_\gamma^2 u_{\sigma,k}^{p-1}u_\sigma\,d\mu_t &= \sigma p\int_M |A|_\gamma^2 u_{\sigma,k}^{p-1}(u_{\sigma,k}+k)\,d\mu_t\\ &\le C_9\sigma p\int_M |A|^2 u_{\sigma,k}^p\,d\mu_t + C_9k\sigma p\int_M |A|^2u_{\sigma,k}^{p-1}\,d\mu_t,\end{aligned}$$

so that for almost every $t\in[0,T)$ there holds

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \le &-(p(p-1)-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t\\ &+C_9\sigma p\int_M |A|^2u_{\sigma,k}^p\,d\mu_t + C_9k\sigma p\int_M |A|^2u_{\sigma,k}^{p-1}\,d\mu_t\\ &+C_3p\int_M |A|^{2-\delta}G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t + C_4p\int_M G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t.\end{aligned}$$

Let us restrict ℓ_0 further so that

$$\ell_0 \le \min\left\{\frac{1}{2C_2C_6},\frac{\delta}{2}\right\}.$$

Then $\sigma\le \ell_0p^{-\frac{1}{2}}$ ensures that $\sigma\le\delta/2$. Since we have arranged that $G\ge 1$ on $\mathrm{sp}(u_{\sigma,k})$ we can estimate

$$\begin{aligned}C_3p\int_M &|A|^{2-\delta}G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t + C_4p\int_M G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t\\ &\le C_3p\int_M |A|^{2-\delta}G^{\delta/2}u_{\sigma,k}^{p-1}\,d\mu_t + C_4p\int_M Gu_{\sigma,k}^{p-1}\,d\mu_t\\ &\le C_{10}p\int_M |A|^{2-\delta/2}u_{\sigma,k}^{p-1}\,d\mu_t + C_{10}p\int_M |A|u_{\sigma,k}^{p-1}\,d\mu_t\end{aligned}$$

where C_{10} depends on n, γ, Γ', C_3, C_4 and δ. Using Young's inequality we obtain

$$C_3 p\int_M |A|^{2-\delta}G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t + C_4 p\int_M G^\sigma u_{\sigma,k}^{p-1}\,d\mu_t \leq \sigma p\int_M |A|^2 u_{\sigma,k}^{p-1}\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^{p-1}\,d\mu_t,$$

where C_{11} depends only on C_{10}, δ and σ. Hence for almost every $t\in[0,T)$ we have

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \leq{}& -(p(p-1)-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t + C_9\sigma p\int_M |A|^2u_{\sigma,k}^p\,d\mu_t\\ &+(C_9k+1)\sigma p\int_M |A|^2 u_{\sigma,k}^{p-1}\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^{p-1}\,d\mu_t.\end{aligned}$$

We now use Young's inequality,

$$ab\leq \frac{p-1}{p}a^{\frac{p}{p-1}}+\frac{1}{p}b^p,$$

to estimate

$$(C_9k+1)\sigma p\int_M |A|^2u_{\sigma,k}^{p-1}\,d\mu_t \leq \sigma(p-1)\int_M |A|^2u_{\sigma,k}^p\,d\mu_t + (C_9k+1)^p\sigma\int_{\mathrm{sp}(u_{\sigma,k})}|A|^2\,d\mu_t$$

and

$$C_{11}p\int_M u_{\sigma,k}^{p-1}\,d\mu_t\leq C_{11}\int_M (p-1)u_{\sigma,k}^p+1\,d\mu_t\leq C_{11}p\int_M u_{\sigma,k}^p+1\,d\mu_t.$$

Hence for almost every $t\in[0,T)$ we have

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \leq{}& -(p(p-1)-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t + (C_9+1)\sigma p\int_M |A|^2u_{\sigma,k}^p\,d\mu_t\\ &+(C_9k+1)^p\sigma\int_{\mathrm{sp}(u_{\sigma,k})}|A|^2\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^p+1\,d\mu_t.\end{aligned}$$

The two zeroth-order terms on the last line are of lower order and will be dealt with later. The remaining zeroth-order term can be absorbed by the good gradient terms via the Poincaré inequality, as follows. Since $\Gamma'\Subset\Gamma\setminus\mathrm{Cyl}$ and $\lambda\in\Gamma'$ holds on $\mathrm{sp}(u_{\sigma,k})$, we can apply Proposition 2.9 with $u=u_{\sigma,k}^{\frac{p}{2}}$. This provides us with a constant $c_1=c_1(n,\gamma,\Gamma')$ such that

$$c_1\int_M |A|^2u_{\sigma,k}^p\,d\mu_t\leq s(p/2-1)^2\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t + (1+s^{-1})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t,$$

where s can be any positive number. From this we obtain

$$(C_9+1)\sigma p\int_M |A|_\gamma^2 u_{\sigma,k}^p\,d\mu_t \le c_1^{-1}(C_9+1)s\sigma p^3\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t + c_1^{-1}(C_9+1)(1+s^{-1})\sigma p\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t,$$

where to ease notation we have used $p\ge 2$ to estimate $(p/2-1)^2\le p^2$. Setting $s=p^{-\frac{1}{2}}$ and $C_{12}:=c_1^{-1}(C_9+1)$, we therefore have

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \le{}& -(p(p-1)-C_{12}\sigma p^{\frac{5}{2}}-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p-C_{12}\sigma p^{\frac{3}{2}}-C_{12}\sigma p-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t\\ &+(C_9k+1)^p\sigma\int_{\mathrm{sp}(u_{\sigma,k})}|A|^2\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^p+1\,d\mu_t.\end{aligned}$$

We now insert the assumption $\sigma\le \ell_0 p^{-\frac{1}{2}}$ and make ℓ_0 a bit smaller so that

$$\ell_0\le\min\left\{\frac{1}{2C_2C_6},\frac{\delta}{2},\frac{1}{2C_{12}},\frac{c_0}{2C_{12}}\right\}.$$

This ensures that for almost every $t\in[0,T)$,

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t \le{}& -(p^2/2-p-C_8p^{\frac{3}{2}}-C_6p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2\,d\mu_t\\ &-(c_0p/2-C_{12}\ell_0p^{\frac{1}{2}}-C_8p^{\frac{1}{2}})\int_M u_{\sigma,k}^p\frac{|\nabla A|^2}{G^2}\,d\mu_t\\ &+(C_9k+1)^p\sigma\int_{\mathrm{sp}(u_{\sigma,k})}|A|^2\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^p+1\,d\mu_t.\end{aligned}$$

We now take p_0 so large that the inequalities

$$-(p^2/2-p-C_8p^{\frac{3}{2}}-C_6p)\le -p^2/4$$

and

$$-(c_0p/2-C_{12}\ell_0p^{\frac{1}{2}}-C_8p^{\frac{1}{2}})\le -c_0p/4$$

both hold for every $p\ge p_0$. The constants c_0, C_6, C_8, C_{12} and ℓ_0 depend only on n, γ, Γ', C' and δ, so we can choose p_0 having only these same dependencies. In particular, with this choice of parameters, for almost every $t\in[0,T)$ we have

$$\frac{d}{dt}\int_M u_{\sigma,k}^p\,d\mu_t\le (C_9k+1)^p\sigma\int_{\mathrm{sp}(u_{\sigma,k})}|A|^2\,d\mu_t + C_{11}p\int_M u_{\sigma,k}^p+1\,d\mu_t.$$

Let $C_{13} := C_{13}(n, \gamma, \Gamma')$ be such that $|A|^2 \leq C_{13}HG$ holds on $\mathrm{sp}(u)$. Then by the last inequality, for almost every $t \in [0, T)$ we have

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p + 1 + C_{13}(C_9k+1)^p\sigma\, d\mu_t \\ &\leq (C_9k+1)^p\sigma \int_{\mathrm{sp}(u_{\sigma,k})} |A|^2\, d\mu_t + C_{11}p\int_M u_{\sigma,k}^p + 1\, d\mu_t \\ &- (1 + C_{13}(C_9k+1)^p\sigma)\int_M HG\, d\mu_t \\ &\leq C_{11}p\int_M u_{\sigma,k}^p + 1\, d\mu_t \\ &\leq C_{11}p\int_M u_{\sigma,k}^p + 1 + C_{13}(C_9k+1)^p\sigma\, d\mu_t.\end{aligned}$$

In short, the function

$$\zeta(t) := \int_M u_{\sigma,k}^p + 1 + C_{13}(C_9k+1)^p\sigma\, d\mu_t,$$

which is Lipschitz-continuous, satisfies

$$\zeta'(t) \leq C_{11}p\zeta(t)$$

for almost every $t \in [0, T)$. Hence

$$\int_M u_{\sigma,k}^p\, d\mu_t \leq \zeta(t) \leq \zeta(0)\exp(C_{11}pT)$$

for each $t \in [0, T)$, and this gives the desired estimate. □

3.4. The supremum estimate. With the L^p-estimate in hand, we are ready to carry out the Stampacchia iteration procedure to derive a supremum estimate for the function u_σ. The argument closely follows the proof of Theorem 5.1 in [**Hui84**].

PROPOSITION 2.13. *Let F and u be as in the statement of Theorem 2.6. Then there are constants p_1 and ℓ_1, each of which may depend only on n, γ, Γ', C' and δ, with the following property. Suppose $p \geq p_1$ and $\sigma \leq \ell_1 p^{-\frac{1}{2}}$, and define*

$$A(k) := \mathrm{sp}(u_{\sigma,k}), \qquad |A(k)| := \int_0^T\int_{\mathrm{sp}(u_{\sigma,k})} d\mu_t dt.$$

Then there is a constant $C = C(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T, \sigma, p)$ such that

$$|A(h)| \leq \frac{C}{(h-k)^p}|A(k)|^\theta$$

for all $h > k > k_2$, where $k_2 = k_2(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T, \sigma, p)$ and $\theta > 1$ depends only on n.

PROOF. Suppose

$$p_1 \geq p_0, \qquad \ell_1 \leq \ell_0,$$

where p_0 and ℓ_0 are the constants from Proposition 2.12. Then the assumptions $k \geq k_1$, $p \geq p_1$ and $\sigma \leq \ell_1 p^{-\frac{1}{2}}$ ensure that (7) holds: for almost every $t \in [0, T)$,

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t \leq &-(p(p-1) - C_8 p^{\frac{3}{2}} - C_6 p)\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2 \, d\mu_t \\ &- (c_0 p - C_8 p^{\frac{1}{2}})\int_M u_{\sigma,k}^p \frac{|\nabla A|^2}{G^2} \, d\mu_t \\ &+ \sigma p \int_M |A|_\gamma^2 u_{\sigma,k}^{p-1} u_\sigma \, d\mu_t + C_3 p \int_M |A|^{2-\delta} G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t \\ &+ C_4 p \int_M G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t.\end{aligned}$$

By our choice of p_0, we can estimate

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t \leq &-\frac{p^2}{4}\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2 \, d\mu_t + \sigma p \int_M |A|_\gamma^2 u_{\sigma,k}^{p-1} u_\sigma \, d\mu_t \\ &+ C_3 p \int_M |A|^{2-\delta} G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t + C_4 p \int_M G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t,\end{aligned}$$

and as before

$$\begin{aligned}C_3 p \int_M |A|^{2-\delta} G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t &+ C_4 p \int_M G^\sigma u_{\sigma,k}^{p-1} \, d\mu_t \\ &\leq \sigma p \int_M |A|^2 u_{\sigma,k}^{p-1} \, d\mu_t + C_{11} p \int_M u_{\sigma,k}^{p-1} \, d\mu_t,\end{aligned}$$

hence

$$\begin{aligned}\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t \leq &-\frac{p^2}{4}\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|_\gamma^2 \, d\mu_t + \sigma p \int_M |A|_\gamma^2 u_{\sigma,k}^{p-1} u_\sigma \, d\mu_t \\ &+ \sigma p \int_M |A|^2 u_{\sigma,k}^{p-1} \, d\mu_t + C_{11} p \int_M u_{\sigma,k}^{p-1} \, d\mu_t\end{aligned}$$

for almost every $t \in [0, T)$. Let $A(k, t) := \mathrm{sp}(u_{\sigma,k}(\cdot, t))$. Without loss of generality we may assume $k \geq 1$ so that

$$\sigma p \int_M |A|_\gamma^2 u_{\sigma,k}^{p-1} u_\sigma \, d\mu_t + \sigma p \int_M |A|^2 u_{\sigma,k}^{p-1} \, d\mu_t \leq C(n, \gamma, \Gamma')\sigma p \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t.$$

We chose k_1 to ensure that $G \geq 1$ on $\mathrm{sp}(u_{\sigma,k})$ for all $k \geq k_1$, so we can also bound

$$C_{11} p \int_M u_{\sigma,k}^{p-1} \, d\mu_t \leq C_{11} p \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t.$$

We chose C_{11} depending only on n, γ, Γ', δ and σ, and may therefore conclude that there is a constant

$$B_0 = B_0(n, \gamma, \Gamma', C', \delta, \sigma, p)$$

such that for almost every $t \in [0, T)$,

$$\frac{d}{dt}\int_M u_{\sigma,k}^p \, d\mu_t \leq -\frac{1}{B_0}\frac{p^2}{4}\int_M u_{\sigma,k}^{p-2}|\nabla u_\sigma|^2 \, d\mu_t + B_0 \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t.$$

Let $v_k = u_{\sigma,k}^{\frac{p}{2}}$. We then compute that

$$|\nabla v_k|^2 = \frac{p^2}{4} u_{\sigma,k}^{p-2} |\nabla u_\sigma|^2,$$

which upon substitution into the previous inequality yields

$$\frac{d}{dt}\int_M v_k^2 \, d\mu_t + \frac{1}{B_0}\int_M |\nabla v_k|^2 \, d\mu_t \leq B_0 \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t. \tag{8}$$

Define

$$\alpha := \sup\{\gamma(\lambda)^{-1} \operatorname{tr}(\lambda) : \lambda \in \Gamma'\}.$$

Then since $\gamma \in \Gamma'$ holds on $\operatorname{sp}(u_{\sigma,k})$ (as a consequence of $k \geq k_1$), we have

$$\begin{aligned}\int_{A(k,t)} H^n \, d\mu_t &\leq \alpha^n \int_{A(k,t)} G^n \, d\mu_t \\ &\leq \frac{\alpha^n}{k^p}\int_{A(k,t)} G^n u_\sigma^p \, d\mu_t \\ &= \frac{\alpha^n}{k^p}\int_{A(k,t)} u_{\sigma'}^p \, d\mu_t,\end{aligned}$$

where $\sigma' := \sigma + n/p$. From this it follows that

$$\begin{aligned}\int_{A(k,t)} H^n \, d\mu_t &\leq \frac{2^{p-1}\alpha^n}{k^p}\int_{A(k,t)} (u_{\sigma'} - k_1)^p + k_1^p \, d\mu_t \\ &\leq \frac{\alpha^n}{2}\left(\frac{2}{k}\right)^p \int_M u_{\sigma',k_1}^p \, d\mu_t + \frac{\alpha^n}{2}\left(\frac{2k_1}{k}\right)^p \mu_0(M).\end{aligned}$$

We can apply Proposition 2.12 to to the first term as long as $\sigma' \leq \ell_0 p^{-\frac{1}{2}}$. To arrange this, we choose p_1 and ℓ_1 a bit smaller so that

$$p_1 \geq \max\left\{p_0, \frac{4n^2}{\ell_0^2}\right\}, \qquad \ell_1 \leq \frac{\ell_0}{2}.$$

Then for $p \geq p_1$ and $\sigma \leq \ell_1 p^{-\frac{1}{2}}$ we have

$$\sigma' = \sigma + \frac{n}{p} \leq \left(\frac{\ell_0}{2} + \frac{n}{p^{\frac{1}{2}}}\right)\frac{1}{p^{\frac{1}{2}}} \leq \frac{\ell_0}{p^{\frac{1}{2}}}.$$

Invoking Proposition 2.12 we get

$$\int_{A(k,t)} H^n \, d\mu_t \leq B_1 k^{-p}$$

for some $B_1 = B_1(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T, \sigma, p)$.

The Sobolev inequality from Corollary 2.11 tells us that

$$\left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} \leq K_0 \int_M |\nabla v_k|^2 \, d\mu_t + K_1 \left(\int_{A(k,t)} H^n \, d\mu_t\right)^{\frac{2}{n}} \left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}},$$

where $K_i = K_i(n)$ and $q = \frac{n}{n-2}$ if $n \geq 3$. If $n = 2$, the same estimate holds for $q = 2$ (this is an arbitrary choice) but K_0 picks up an additional dependence and $\mu_0(M)$. Inserting

$$\int_{A(k,t)} H^n \, d\mu_t \leq B_1 k^{-p}$$

gives

$$\left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} \leq K_0 \int_M |\nabla v_k|^2 \, d\mu_t + K_1 B_1^{\frac{2}{n}} k^{-\frac{2p}{n}} \left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}}.$$

We increase p_1 if necessary so that

$$p_1 \geq \max\left\{p_0, \frac{4n^2}{\ell_0^2}, n\right\}.$$

Then $p \geq p_1$ ensures that

$$K_1 B_1^{\frac{2}{n}} k^{-\frac{2p}{n}} \leq K_1 B_1^{\frac{2}{n}} k^{-2}$$

and for every

$$k \geq \max\{k_1, 2K_1^{\frac{1}{2}} B_1^{\frac{1}{n}}\}$$

there holds

$$K_1 B_1^{\frac{2}{n}} k^{-\frac{2p}{n}} \leq 1/4,$$

and consequently

$$\left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} \leq 4K_0 \int_M |\nabla v_k|^2 \, d\mu_t.$$

Inserting this into (8) gives

$$\frac{d}{dt}\int_M v_k^2 \, d\mu_t + \frac{1}{4K_0 B_0}\left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} \leq B_0 \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t \tag{9}$$

for almost every $t \in [0, T)$.

Observe that

$$u_\sigma \leq C_0 G^\sigma \leq C_0 \max\{1, R^{-1}\}$$

on M_0. Therefore, by choosing k a bit larger so that

$$k \geq \max\{k_1, 2K_1^{\frac{1}{2}} B_1^{\frac{1}{n}}, C_0, C_0 R^{-1}\} =: k_2,$$

we ensure $u_{\sigma,k}^p = v_k^2 \equiv 0$ on M_0. Integrating (9) in time then gives

$$\int_M v_k^2 \, d\mu_\tau + \frac{1}{4K_0 B_0} \int_0^\tau \left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} dt \leq B_0 \int_0^\tau \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t dt$$

for each $\tau \in [0, T)$. Throwing away terms on the left yields

$$\sup_{t \in [0,T)} \int_M v_k^2 \, d\mu_t \leq B_0 \int_0^T \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t \, dt$$

and

$$\int_0^T \left(\int_M v_k^{2q} \, d\mu_t\right)^{\frac{1}{q}} dt \leq 4K_0 B_0^2 \int_0^T \int_{A(k,t)} G^2 u_\sigma^p \, d\mu_t \, dt,$$

which we recombine to get

$$\sup_{t\in[0,T)}\int_M v_k^2\,d\mu_t+\int_0^T\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{q}}dt\le B_1\int_0^T\int_{A(k,t)}G^2u_\sigma^p\,d\mu_t\,dt,$$

where $B_1:=B_0+4K_0B_0^2$.

Fix a constant $q_0\in(1,q)$. To exploit the second term on the left above, we use the following interpolation inequality for L^p spaces:

$$\|f\|_{q_0}\le\|f\|_r^{1-\theta}\|f\|_q^{\theta},$$

where $\theta\in(0,1)$ and $\frac{1}{q_0}=\frac{\theta}{q}+\frac{1-\theta}{r}$. We set $r=1$ and $\theta=\frac{1}{q_0}$. This gives

$$\left(\int_M v_k^{2q_0}\,d\mu_t\right)^{\frac{1}{q_0}}\le\left(\int_M v_k^2\,d\mu_t\right)^{\frac{q_0-1}{q_0}}\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{qq_0}}.$$

Raising both sides to q_0, integrating in time and using Young's inequality we have

$$\begin{aligned}\int_0^T\int_M v_k^{2q_0}\,d\mu_t dt&\le\int_0^T\left(\int_M v_k^2\,d\mu_t\right)^{q_0-1}\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{q}}dt\\&\le\left(\sup_{t\in[0,T)}\int_M v_k^2\,d\mu_t dt\right)^{q_0-1}\int_0^T\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{q}}dt\\&\le\frac{q_0-1}{q_0}\left(\sup_{t\in[0,T)}\int_M v_k^2\,d\mu_t\right)^{q_0}+\frac{1}{q_0}\left(\int_0^T\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{q}}dt\right)^{q_0}\\&\le\left(\sup_{t\in[0,T)}\int_M v_k^2\,d\mu_t+\int_0^T\left(\int_M v_k^{2q}\,d\mu_t\right)^{\frac{1}{q}}dt\right)^{q_0},\end{aligned}$$

hence

$$\left(\int_0^T\int_M v_k^{2q_0}\,d\mu_t dt\right)^{\frac{1}{q_0}}\le B_1\int_0^T\int_{A(k,t)}G^2u_\sigma^p\,d\mu_t\,dt.$$

Turning now to the right-hand side, let r be a large constant depending only on n whose value we will fix later. By Hölder's inequality,

$$B_1\int_0^T\int_{A(k,t)}G^2u_\sigma^p\,d\mu_t\,dt\le B_1|A(k)|^{1-\frac{1}{r}}\left(\int_0^T\int_{A(k,t)}G^{2r}u_\sigma^{pr}\,d\mu_t\right)^{\frac{1}{r}}.$$

We estimate the integral on the right-hand side by

$$\begin{aligned}\int_{A(k,t)}G^{2r}u_\sigma^{pr}\,d\mu_t&=\int_{A(k,t)}u_{\sigma+2/p}^{pr}\,d\mu_t\\&\le\int_{A(k,t)}(u_{\sigma+2/p}-k_1+k_1)^{pr}\,d\mu_t\\&\le 2^{pr-1}\int_M u_{\sigma+2/p,k_1}^{pr}\,d\mu_t+2^{pr-1}\mu_0(M)k_1^{pr}.\end{aligned}$$

We would like to bound the first term on the last line using Proposition 2.12. For this to work, we need

$$\sigma + \frac{2}{p} \le \frac{\ell_0}{(pr)^{\frac{1}{2}}},$$

which can be achieved by further restricting p_1 and ℓ_1 as follows:

$$p_1 \ge \max\left\{p_0, \frac{4n^2}{\ell_0^2}, \frac{16r}{\ell_0^2}\right\}, \qquad \ell_1 \le \frac{\ell_0}{2r^{\frac{1}{2}}}.$$

Then by Proposition 2.12 there holds

$$\int_{A(k,t)} G^{2r} u_\sigma^{pr}\, d\mu_t \le B_2$$

for some

$$B_2 = B_2(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T, \sigma, p, r).$$

Substituting back in, we obtain

$$\left(\int_0^T \int_M v_k^{2q_0}\, d\mu_t dt\right)^{\frac{1}{q_0}} \le B_1 T^{\frac{1}{r}} B_2^{\frac{1}{r}} |A(k)|^{1-\frac{1}{r}} =: B_3 |A(k)|^{1-\frac{1}{r}}. \tag{10}$$

For every $h > k \ge k_2$ there holds

$$\begin{aligned}
|A(h)| &= \int_0^T \int_{\mathrm{sp}(u_{\sigma,h}(\cdot,t))} \frac{(h-k)^p}{(h-k)^p}\, d\mu_t\, dt \\
&\le \int_0^T \int_{\mathrm{sp}(u_{\sigma,h}(\cdot,t))} \frac{v_k^2}{(h-k)^p}\, d\mu_t\, dt \\
&\le \frac{1}{(h-k)^p} \int_0^T \int_M v_k^2\, d\mu_t\, dt,
\end{aligned}$$

and

$$\int_0^T \int_M v_k^2\, d\mu_t\, dt \le |A(k)|^{1-\frac{1}{q_0}} \left(\int_0^T \int_M v_k^{2q_0}\, d\mu_t\, dt\right)^{\frac{1}{q_0}}.$$

Combining these two estimates with (10), we now arrive at

$$|A(h)| \le \frac{B_3}{(h-k)^p} |A(k)|^{2-\frac{1}{q_0}-\frac{1}{r}}.$$

Fixing $r = r(n)$ large enough so that

$$\theta := 2 - \frac{1}{q_0} - \frac{1}{r} > 1,$$

the proof is complete. □

The iteration inequality just proven implies that $|A(k)| = 0$ for large k, by Stampacchia's lemma (for a proof see Lemma B.1. in [**KS80a**]).

Lemma 2.14. *Let $\varphi : [\bar{k}, \infty) \to [0, \infty)$ be a nonincreasing function such that*

$$\varphi(h) \le \frac{C}{(h-k)^\alpha} \varphi(k)^\theta$$

for all $h > k > \bar{k}$, where $C, \alpha > 0$ and $\theta > 1$ are constants. Then $\varphi(\bar{k}+d) = 0$, where

$$d^\alpha = C\varphi(\bar{k})2^{\frac{\alpha\theta}{\theta-1}}.$$

Proof of Theorem 2.6. Fix $p = p_1$ and $\sigma = \ell_1 p_1^{-\frac{1}{2}}$ and set $\varphi(k) = |A(k)|$. Then by Proposition 2.13 there is a

$$C = C(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T)$$

such that

$$\varphi(h) \leq \frac{C}{(h-k)^{p_1}}\varphi(k)^\theta$$

for every $h > k > \bar{k}$, where $\bar{k} = \bar{k}(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T)$. Since $\theta > 1$, Stampacchia's Lemma now implies that

$$\varphi(\bar{k} + d) = 0,$$

where $d^{p_1} = C\varphi(\bar{k})2^{\frac{\theta p_1}{p_1-1}}$. Since $\varphi(\bar{k}) = |A(\bar{k})| \leq T\mu_0(M)$, and $\theta = \theta(n)$, the constant $k' := \bar{k} + d$ depends only on n, C, p_1, $\mu_0(M)$ and T. Unpacking the dependencies of C and p_1, we see that

$$k' = k'(n, \gamma, \Gamma', k_0, C', \delta, R, \mu_0(M), T).$$

By the definition of $A(k)$, the inequality

$$u \leq k' G^{-\sigma_1}$$

holds on all of $M \times [0, T)$. Appealing to Young's inequality, we obtain

$$u \leq \varepsilon + \sigma\left(\frac{k' G^{-\sigma_1}}{\varepsilon^{1-\sigma_1}}\right)^{\frac{1}{\sigma_1}} =: \varepsilon + K_\varepsilon G^{-1}.$$

The constant K_ε only depends on k', σ_1 and ε, so the theorem is proven. $\square$

CHAPTER 3

A convexity estimate

Our goal in this chapter is to establish a convexity estimate for certain fully nonlinear flows of k-convex hypersurfaces. The first result of this kind was established by Huisken and Sinestrari for mean-convex mean curvature flow in [**HS99b**] and [**HS99a**]. They use Stampacchia iteration and induction to prove that on a compact mean-convex solution of dimension $n \geq 2$, for each $2 \leq k \leq n$, the elementary symmetric polynomial S_k of degree k applied to λ satisfies an estimate of the form

$$S_k(\lambda(x,t)) \geq -\varepsilon H(x,t)^k - C_{\varepsilon,k}.$$

Here ε can be any positive constant, and $C_{\varepsilon,k}$ depends only on ε, k, and the solution at the initial time. Hence this estimate implies that, at points where the curvature is sufficiently large, the quantity $S_k(\lambda)$ is almost positive.

Combining these estimates one concludes that the smallest principal curvature satisfies an estimate of the same form:

$$\lambda_1(x,t) \geq -\varepsilon H(x,t) - C_\varepsilon.$$

In particular, the rescaled second fundamental form $H^{-1}A$ is approximately nonnegative at points of extremely large curvature, and the limit of any smoothly converging blow-up sequence must have nonnegative second fundamental form. We note that the Huisken-Sinestrari convexity estimate plays the same role in the study of mean curvature flow as the Hamilton-Ivey estimate (see [**Ham93**][Theorem 24.4] and [**Ive93**]) in three-dimensional Ricci flow, which in turn was an essential ingredient in Perelman's proof of the geometrisation conjecture.

Andrews, Langford and McCoy generalised the convexity estimate to flows by convex admissible speed functions in [**ALM14**]. Rather than working with elementary symmetric polynomials, the authors apply the Stampacchia procedure to a single cleverly chosen curvature quantity, which is in a sense a smooth approximation of the smallest principal curvature. In this way, their proof is technically simpler than that of Huisken-Sinestrari, even in the mean curvature flow case. Using an idea of Brendle [**Bre15**], Langford then gave another proof of the convexity estimate for mean curvature flow in [**Lan17**] by applying the Stampacchia procedure directly to the nonsmooth quantity λ_1.

For flows where the speed function is instead concave in the curvature, the evolution of the second fundamental form picks up a nonpositive gradient term, which in general makes it much more difficult to control the second fundamental form 'from below'. It is possible that for certain concave speeds with some special algebraic structure, the pinching quantities utilised by Huisken-Sinestrari and Andrews-Langford-McCoy could be used to prove a convexity estimate, but this approach will not work for concave speeds in general. Despite these new hurdles in the concave case, it is possible to prove a convexity estimate if the speed is supported in the two-positive cone. Then, as we saw in the last chapter, one can prove a cylindrical estimate which implies a convexity estimate (see the discussion following the statement of Theorem 2.4).

This is no longer the case if we look at flows even of three-convex hypersurfaces. The problem is that, in general, a uniformly k-convex solution can form singularities modeled on a shrinking $\mathbb{R}^{k-1} \times S^{n-k+1}$. This means that if we want to prove an a priori estimate showing that the principal curvatures pinch onto some convex cone $\Gamma' \subset \overline{\Gamma}_+$ at a singularity, this cone must contain the convex hull of Cyl_{k-1}. When $k = 2$, the convex hull of $\mathrm{Cyl}_{k-1} = \mathrm{Cyl}_1$ intersected with $\partial\Gamma_+$ is a set of finitely many rays, so Γ' needs to contain these rays, but this is not very restrictive. In particular, the boundary of Γ' can have this property and still be strictly convex. The general case is not so favourable, since the convex hull of Cyl_{k-1} intersects $\partial\Gamma_+$ in a set of finitely many cones of dimension $k-1$. This means that, in order for Γ' to sit in $\overline{\Gamma}_+$ and contain the convex hull of Cyl_{k-1}, its boundary must contain regions with at least $k-1$ flat directions. On the other hand, we need the boundary of Γ' to have a certain amount of convexity in order to overcome the nonpositive gradient term appearing in the evolution of A. This is essentially the reason why, at present, we are only able to establish a convexity estimate for some very special speeds.

We note that similar issues arise when one tries to identify which speed functions give rise to a flow that preserves convexity. It turns out that concavity of the speed is not enough to ensure this, and in fact, the speed needs to satisfy a weak kind of convexity condition known as inverse-concavity. Andrews proved that inverse-concave speeds preserve convexity in [**And07**] (see also Corollary 3.7 below) and in [**AMZ13**] the authors construct compact solutions moving by a concave admissible speed which start off convex and become non-convex in finite time.

We introduce here the first examples of nonlinear concave speed functions which are defined on the k-positive cone in $\mathbb{R}^n$ for $n \geq 4$ and $k \geq 3$, and for which it is possible to prove a convexity estimate. For each $k \geq 3$ we work with a family of speeds that interpolates (in a nonlinear fashion) between the k-harmonic mean and the mean:

$$\gamma_\rho(\lambda) := \left(\sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\rho}{\lambda_{i_1} + \cdots + \lambda_{i_k}} + \frac{1-\rho}{\lambda_1 + \cdots + \lambda_n} \right)^{-1},$$

where $\rho \in (0,1]$. For every $\rho > 0$ the function γ_ρ is an admissible speed which vanishes at the boundary of the k-positive cone and is strictly concave in off-radial directions. As ρ goes to zero, γ_ρ becomes steeper and more concave near the boundary, but approaches the trace function on the interior of the k-positive cone. This gives us very precise control over the size of the Hessian of γ_ρ, and hence over the size of the troublesome term in the evolution of A, over a large region in curvature space. It is this property and the cylindrical estimate of Chapter 2 that allow us to prove the convexity estimate:

THEOREM 3.1. *Fix $n \geq 4$ and $k \in \{3, \ldots, n-1\}$. Let Γ denote the k-positive cone in $\mathbb{R}^n$,*

$$\Gamma := \{\lambda \in \mathbb{R}^n : \lambda_{i_1} + \cdots + \lambda_{i_k} > 0 \ \ \forall \ \ 1 \leq i_1 < \cdots < i_k \leq n\},$$

and let $\gamma_\rho : \Gamma \to (0,\infty)$ be defined as above. Consider a compact evolving immersion $F : M \times [0,T) \to \mathbb{R}^{n+1}$ which satisfies

$$\partial_t F(x,t) = -G_\rho(x,t)\nu(x,t)$$

for each $(x,t) \in M \times [0,T)$, where $G_\rho(x,t) := \gamma_\rho(\lambda(x,t))$. Then there is a positive constant $\rho_0 = \rho_0(n,k)$ with the property that if $\rho \in (0,\rho_0]$, for every $\varepsilon > 0$,

$$\lambda_1(x,t) \geq -\varepsilon G_\rho(x,t) - C_\varepsilon$$

holds for every $(x,t) \in M \times [0,T)$. The constant C_ε has the dependencies

$$C_\varepsilon = C_\varepsilon(n,k,\rho,\bar{\alpha},R,\mu_0(M),T),$$

where

$$\bar{\alpha} := \max_M \frac{H(\cdot,0)}{G_\rho(\cdot,0)}, \qquad R^{-1} := \sup_M G(\cdot,0).$$

The outline of the chapter is as follows. We first derive an evolution equation for the smallest principal curvature (interpreted in an appropriate weak sense) which holds for a general hypersurface flow by an admissible speed. Following this we discuss the inverse-concavity condition and some of its implications, which include a strong maximum principle for the smallest principal curvature. This result plays a key role in Chapter 4. We then establish some algebraic properties of the family of speeds γ_ρ and begin studying compact immersions moving with inward normal velocity $\gamma_\rho(\lambda(x,t))$. By applying the cylindrical estimate from Chapter 2 we show that the gradient terms in the evolution of λ_1 have a favourable structure at high curvature scales, provided ρ is small. This allows us to apply the Stampacchia principle to (a modified quantity built from) λ_1 and prove the estimate. Following [**Bre15**] and [**Lan17**], we estimate the nonsmooth function λ_1 directly, rather than working with a smooth approximation.

The arguments are structured so that little extra work is needed to prove an analogous theorem when the background space is a Riemannian manifold satisfying a certain curvature condition. In the final section of the chapter we sketch the extra arguments needed to prove the convexity estimate in this more general setting.

Let us remark here that our construction also works for curvature conditions other than k-convexity. Indeed, given any convex curvature cone Γ and a concave admissible speed $\gamma : \Gamma \to (0,\infty)$ which vanishes at $\partial\Gamma$, compact hypersurfaces moving by the speed

$$\lambda \mapsto (\rho\gamma(\lambda)^{-1} + (1-\rho)\operatorname{tr}(\lambda)^{-1})^{-1}$$

satisfy a convexity estimate provided that $\rho > 0$ is sufficiently small.

1. An equation for λ_1

For the next result we consider a general admissible speed $\gamma : \Gamma \to (0,\infty)$. Let $F : M \times [0,T) \to \mathbb{R}^{n+1}$ be a smooth immersion evolving by (CF). We first recall the evolution equation for the second fundamental form: choosing a basis at any point in spacetime, we have

$$(\partial_t - \Delta_\gamma)A_{ij} = |A|^2_\gamma A_{ij} - 2Gg^{kl}A_{ik}A_{lj} + \ddot{\gamma}^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs}.$$

For each $(x,t) \in M \times [0,T)$, let λ_1 denote the smallest principal curvature of $A(x,t)$. This is a Lipschitz function on $M \times [0,T)$, which can be written as

$$\lambda_1(x,t) = \min_{v \in T_xM \setminus \{0\}} |v|^{-2} A(x,t)(v,v).$$

In the following proposition, we use the evolution of the second fundamental form to derive an evolution equation for λ_1, which holds in the sense of viscosities. This kind of computation was first carried out by Andrews in [**And07**] (see Theorem 3.2) to prove a stronger version of Hamilton's tensor maximum principle. Further refinements were made by Langford in [**Lan14**] (see Theorem 4.18). We use the following terminology:

DEFINITION 3.2. *Consider a function $f : M \times [0,T) \to \mathbb{R}$ and a point $(x_0,t_0) \in M \times (0,T)$. We call φ a lower support for f at (x_0,t_0) if there is a positive constant r such that*

$$\varphi \in C^2(B_{g(t_0)}(x_0,r) \times [-r^2 + t_0, t_0])$$

and there holds

$$\varphi(x,t) \le f(x,t)$$

on $B_{g(t_0)}(x_0,r) \times [-r^2+t_0,t_0]$, with equality at (x_0,t_0). If the inequality is reversed φ is called an upper support for f at (x_0,t_0).

For any point $(x_0,t_0) \in M \times [0,T)$, we say that $\{e_i\}_{i=1}^n \subset T_{x_0}M$ is a principal frame at (x_0,t_0) if the e_i are unit-length eigenvectors of $A(x_0,t_0)$ such that

$$A(x_0,t_0)(e_i,e_i) = \lambda_i, \qquad \lambda_1 \le \cdots \le \lambda_n.$$

PROPOSITION 3.3. *Fix a point $(x_0,t_0) \in M \times (0,T)$ and let φ be a lower support for λ_1 at (x_0,t_0). Then, in a principal frame at (x_0,t_0), there holds*

$$(\partial_t - \Delta_\gamma)\varphi \ge |A|^2_\gamma \varphi + \ddot{\gamma}^{pq,rs} \nabla_1 A_{pq} \nabla_1 A_{rs} + 2\dot{\gamma}^k \sum_{\lambda_p > \lambda_1} \frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1}.$$

PROOF. Let us choose coordinates $\{x^i\}$ on M in a neighbourhood of x_0 and denote the corresponding coordinate tangent vectors by $\{e_i\}$. We may assume that at (x_0,t_0), the $\{e_i\}$ form a principal frame satisfying

$$A(e_i,e_i) = \lambda_i, \qquad \lambda_1 \le \cdots \le \lambda_n,$$

and that $\nabla_i e_k = 0$ for all i and k. These coordinates give rise to a local system of $2n$ coordinates for the tangent bundle TM, since we can write any smooth vectorfield v defined near x_0 as $v^i e_i$.

Define $Z(x,t,v) := |v|^{-2} A(x,t)(v,v)$ for all $(x,t) \in M \times [0,T)$ and nonzero $v \in T_xM$. This is a smooth function on $TM \times [0,T)$ with the property that

$$Z(x,t,v) - \lambda_1(x,t) \ge 0,$$

and this holds with equality at (x_0,t_0,e_1). Since φ is a lower support for λ_1, we get that $Z(x,t,v) - \varphi(x,t) \ge 0$ with equality at (x_0,t_0,e_1). This means that all $2n$ of the first-order coordinate derivatives of $Z(x,t,v) - \varphi(x,t)$ vanish at (x_0,t_0,e_1), and

$$\partial_t (Z - \varphi)(x_0,t_0) \le 0.$$

In addition, the second-order coordinate derivatives form a nonnegative $2n \times 2n$-matrix. We will use these facts to derive the desired inequality for φ.

With respect to coordinates, we have

$$Z(x,t,v) = (g_{pq}(x,t)v^p v^q)^{-1} A_{pq}(x,t)v^p v^q.$$

Therefore, at the point (x_0, t_0, e_1) there holds

$$\begin{aligned}
\frac{\partial \varphi}{\partial t}(x_0,t_0) &\geq \frac{\partial Z}{\partial t}(x_0,t,e_1)\\
&= -\frac{\partial g}{\partial t}(e_1,e_1)A_{11} + \frac{\partial A}{\partial t}(e_1,e_1)\\
&= 2GA_{11}^2 + \Delta_\gamma A_{11} + |A|_\gamma^2 A_{11} - 2GA_{11}^2 + \ddot{\gamma}^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs}\\
&= \Delta_\gamma A_{11} + |A|_\gamma^2 A_{11} + \ddot{\gamma}^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs}.
\end{aligned}$$

The remaining first derivatives of Z are given by:

$$\begin{aligned}
\frac{\partial Z}{\partial x^l} &= -\frac{\partial g_{pq}}{\partial x^l}v^p v^q \cdot |v|^{-4}A(v,v) + |v|^{-2}\frac{\partial A_{pq}}{\partial x^l}v^p v^q;\\
\frac{\partial Z}{\partial v^q} &= -2g_{pq}v^p|v|^{-4}A(v,v) + 2|v|^{-2}A_{pq}v^p,
\end{aligned}$$

and the second derivatives of Z at (x_0, t_0, v) are:

$$\begin{aligned}
\frac{\partial Z}{\partial x^k \partial x^l} &= -\frac{\partial^2 g_{pq}}{\partial x^k \partial x^l}v^p v^q \cdot |v|^{-4}A(v,v) + |v|^{-2}\frac{\partial^2 A_{pq}}{\partial x^k \partial x^l}v^p v^q;\\
\frac{\partial Z}{\partial x^k \partial v^p} &= -2g_{pq}v^q|v|^{-4}\frac{\partial A_{rs}}{\partial x^k}v^r v^s + 2|v|^{-2}\frac{\partial A_{pq}}{\partial x^k}v^q;\\
\frac{\partial Z}{\partial v^p \partial v^q} &= -2g_{pq}|v|^{-4}A(v,v) + 8g_{iq}v^i g_{jp}v^j|v|^{-6}A(v,v)\\
&\quad - 4g_{iq}v^i|v|^{-4}A_{pr}v^r - 4g_{ip}v^i|v|^{-4}A_{rq}v^r + 2|v|^{-2}A_{pq}.
\end{aligned}$$

Evaluating at (x_0, t_0, e_1) gives

$$\frac{\partial \varphi}{\partial x^i}(x,t) = \frac{\partial Z}{\partial x^i}(x_0,t_0,e_1) = \nabla_i A_{11},$$

and

$$\begin{aligned}
\frac{\partial^2 Z}{\partial x^k \partial x^l} &= -\frac{\partial^2 g_{11}}{\partial x^k \partial x^l}A_{11} + \frac{\partial^2 A_{11}}{\partial x^k \partial x^l};\\
\frac{\partial^2 Z}{\partial x^k \partial v^p} &= -2\delta_{p1}\frac{\partial A_{11}}{\partial x^k} + 2\frac{\partial A_{p1}}{\partial x^k};\\
\frac{\partial^2 Z}{\partial v^p \partial v^q} &= -2g_{pq}A_{11} + 8g_{p1}g_{1q}A_{11} - 4g_{p1}A_{1q} - 4g_{1q}A_{p1} + 2A_{pq}\\
&= -2\delta_{pq}A_{11} + 2A_{pq}.
\end{aligned}$$

Since the Hessian of $Z - \varphi$ is nonnegative at (x_0, t_0, e_1), for every $n \times n$-matrix Γ_k^p we have

$$\begin{aligned}
\frac{\partial^2 \varphi}{\partial x^k \partial x^l} &\leq \frac{\partial^2 Z}{\partial x^k \partial x^l} + \Gamma_k^p \frac{\partial^2 Z}{\partial x^l \partial v^p} + \Gamma_l^q \frac{\partial^2 Z}{\partial x^k \partial v^q} + \Gamma_k^p \Gamma_l^q \frac{\partial^2 Z}{\partial v^p v^q}\\
&= \frac{\partial^2 A_{11}}{\partial x^k \partial x^l} - \frac{\partial^2 g_{11}}{\partial x^k \partial x^l}A_{11} - 2\Gamma_k^1 \frac{\partial A_{11}}{\partial x^l} + 2\Gamma_k^p \frac{\partial A_{p1}}{\partial x^l}\\
&\quad - 2\Gamma_l^1 \frac{\partial A_{11}}{\partial x^k} + 2\Gamma_l^q \frac{\partial A_{q1}}{\partial x^k} - 2\delta_{pq}\Gamma_k^p \Gamma_l^q A_{11} + 2\Gamma_k^p \Gamma_l^q A_{pq}.
\end{aligned}$$

From this it follows that at (x_0, t_0, e_1),

$$\dot{\gamma}^{kl}\nabla_k\nabla_l\varphi \le \dot{\gamma}^{kl}\frac{\partial^2 A_{11}}{\partial x^k \partial x^l} - \dot{\gamma}^{kl}\frac{\partial^2 g_{11}}{\partial x^k \partial x^l}A_{11} + 4\dot{\gamma}^{kl}\sum_{p>1}\Gamma_k^p\nabla_l A_{p1}$$
$$+ 2\dot{\gamma}^{kl}\sum_{p>1}\Gamma_k^p\Gamma_l^p(\lambda_p - \lambda_1).$$

Next we observe that

$$\nabla_k\nabla_l A_{11} = \frac{\partial^2 A_{11}}{\partial x^k \partial x^l} - 2\frac{\partial \Gamma_{k1}^1}{\partial x^l}A_{11} = \frac{\partial^2 A_{11}}{\partial x^k \partial x^l} - \frac{\partial^2 g_{11}}{\partial x^k \partial x^l}A_{11},$$

and so obtain

$$\dot{\gamma}^{kl}\nabla_k\nabla_l\varphi \le \dot{\gamma}^{kl}\nabla_k\nabla_l A_{11} + 2\dot{\gamma}^{kl}\sum_{p>1}\Big(2\Gamma_k^p\nabla_l A_{p1} + \Gamma_k^p\Gamma_l^p(\lambda_p - \lambda_1)\Big).$$

Combining this last inequality with the time-derivative condition above, we find that at (x_0, t_0) there holds

$$(\partial_t - \Delta_\gamma)\varphi \ge |A|_\gamma^2\varphi + \ddot{\gamma}^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs}$$
$$- 2\dot{\gamma}^{kl}\sum_{p>1}\Big(2\Gamma_k^p\nabla_l A_{p1} + \Gamma_k^p\Gamma_l^p(\lambda_p - \lambda_1)\Big).$$

Let us choose $\Gamma_k^p = 0$ for all indices p such that $\lambda_p = \lambda_1$, and

$$\Gamma_k^p = -\frac{\nabla_k A_{p1}}{\lambda_p - \lambda_1}$$

whenever $\lambda_p > \lambda_1$. Using the fact that in a principal frame $\dot{\gamma}^{kl} = \dot{\gamma}^k\delta^{kl}$, this gives

$$(\partial_t - \Delta_\gamma)\varphi \ge |A|_\gamma^2\varphi + \ddot{\gamma}^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + 2\dot{\gamma}^k\sum_{\lambda_p>\lambda_1}\frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1}.$$

□

The useful output of these careful computations is the nonnegative gradient term on the right-hand side. When exploiting this term, we use the following elementary lemma.

LEMMA 3.4. *Fix $(x_0, t_0) \in M \times (0, T)$ and suppose λ_1 admits a lower support φ at (x_0, t_0). Then if e_1 and e_2 are two orthonormal vectors in $T_{x_0}M$ which are such that*

$$A(x_0, t_0)(e_1, e_1) = A(x_0, t_0)(e_2, e_2) = \lambda_1(x_0, t_0),$$

there holds $\nabla A(e_1, e_2) = 0$ at (x_0, t_0).

PROOF. Extend e_1 and e_2 to an orthonormal basis $\{e_i\}$, and then to a local orthonormal frame on a spatial neighbourhood of x_0, using parallel transport with respect to the Levi-Civita connection. Then, computing at x_0, we have

$$\nabla_k A(e_1, e_2) = e_k(A_{12})$$
$$= \frac{1}{4}e_k(A(e_1 + e_2, e_1 + e_2) - A(e_1 - e_2, e_1 - e_2)).$$

On the other hand, since $|e_1 + e_2| \equiv \sqrt{2}$, there holds

$$A(e_1 + e_2, e_1 + e_2) \geq \sqrt{2}\lambda_1 \geq \sqrt{2}\varphi,$$

and by assumption, this inequality becomes an equality at x_0. From this we conclude that, at x_0,

$$e_k(A(e_1 + e_2, e_1 + e_2)) = \sqrt{2}\nabla_k\varphi,$$

but the same argument shows that

$$e_k(A(e_1 - e_2, e_1 - e_2)) = \sqrt{2}\nabla_k\varphi,$$

also holds at x_0. Combining these two equalities with the computation above gives the result. $\square$

One easy consequence of the lemma just proven is that our choice of Γ_k^p in the proof of Proposition 3.3 was optimal. Indeed, in light of the lemma, at (x_0, t_0) we have

$$\begin{aligned} -2\dot{\gamma}^{kl} \sum_{p>1} &\Big(2\Gamma_k^p \nabla_l A_{p1} + \Gamma_k^p \Gamma_l^p (\lambda_p - \lambda_1)\Big) \\ &= -2\dot{\gamma}^{kl} \sum_{\lambda_p > \lambda_1} \Big(2\Gamma_k^p \nabla_l A_{p1} + \Gamma_k^p \Gamma_l^p (\lambda_p - \lambda_1)\Big). \end{aligned}$$

Using the fact that in a principal frame $\dot{\gamma}^{kl} = \dot{\gamma}^k \delta^{kl}$, pulling out a factor of $\lambda_p - \lambda_1$ and completing the square, we obtain

$$\begin{aligned} -2\dot{\gamma}^{kl} \sum_{p>1} &\Big(2\Gamma_k^p \nabla_l A_{p1} + \Gamma_k^p \Gamma_l^p (\lambda_p - \lambda_1)\Big) \\ &= 2\dot{\gamma}^k \sum_{\lambda_p > \lambda_1} (\lambda_p - \lambda_1) \left[\frac{|\nabla_k A_{p1}|^2}{(\lambda_p - \lambda_1)^2} - \left(\Gamma_k^p + \frac{\nabla_k A_{p1}}{\lambda_p - \lambda_1}\right)^2\right], \end{aligned}$$

In order to maximise the right-hand side, it is clear that we should choose Γ_k^p so that

$$\Gamma_k^p = -\frac{\nabla_k A_{p1}}{\lambda_p - \lambda_1}$$

for each p such that $\lambda_p(x_0, t_0) > \lambda_1(x_0, t_0)$.

1.1. Inverse-concavity. If the speed γ is inverse-concave on the positive cone, meaning the function $\gamma_*(A) := \gamma(A^{-1})^{-1}$ is concave in A, then the gradient terms appearing in the equation for λ_1 have a favourable structure at points where $\lambda_1 \geq 0$. This observation seems to have been made by Huisken, and communicated to Urbas [**Urb91**] and Andrews [**And07**]. In particular, one finds that for flows by inverse-concave speeds, if a compact solution is strictly convex at the initial time, then this remains true up to the maximal time of smooth existence. The proof of this fact uses the following characterisation of inverse-concavity in terms of second derivatives.

LEMMA 3.5. *Let $\gamma : \mathrm{O}\,\Gamma_+ \to (0,\infty)$ be a smooth symmetric function, and for each matrix $A \in \mathrm{O}\,\Gamma_+$, define $\gamma_*(A) = \gamma(A^{-1})^{-1}$. Then there holds*

$$
\begin{aligned}
&-\ddot{\gamma}_*^{pq,rs}(A^{-1})T^*_{pq}T^*_{rs}\\
&\qquad = \frac{1}{\gamma(A)^2}\left(\ddot{\gamma}^{pq,rs}(A) + 2\dot{\gamma}^{pr}(A)(A^{-1})_{qs} - \frac{2}{\gamma(A)}\dot{\gamma}^{pq}(A)\dot{\gamma}^{rs}\right)T_{pq}T_{rs},
\end{aligned}
$$

where

$$T^*_{pq} := (A^{-1})_{pk}(A^{-1})_{ql}T_{kl}.$$

In particular, γ is inverse-concave if and only if

$$\ddot{\gamma}^{pq,rs}(A) + 2\dot{\gamma}^{pr}(A)(A^{-1})_{qs} - \frac{2}{\gamma(A)}\dot{\gamma}^{pq}(A)\dot{\gamma}^{rs}(A) \geq 0$$

for every $A > 0$.

PROOF. Fix a matrix $B \in \mathrm{O}\,\Gamma_+$, and let A denote B^{-1}. The first derivatives of γ_* at B are given by

$$\dot{\gamma}_*^{pq}(B) = -\gamma(A)^{-2}\dot{\gamma}^{kl}(A)A_{pk}A_{ql},$$

and differentiating again gives

$$
\begin{aligned}
\ddot{\gamma}_*^{pq,rs}(B) &= 2\gamma(A)^{-3}\dot{\gamma}^{kl}(A)\dot{\gamma}^{ij}(A)A_{kp}A_{lq}A_{ir}A_{js} - \gamma(A)^{-2}\ddot{\gamma}^{kl,ij}(A)A_{kp}A_{lq}A_{ir}A_{js}\\
&\quad + \gamma(A)^{-2}\dot{\gamma}^{kl}(A)A_{pi}A_{jk}A_{ql} + \gamma(A)^{-2}\dot{\gamma}^{kl}(A)A_{pk}A_{qi}A_{jl}.
\end{aligned}
$$

Contracting this against $T^*_{pq} := B_{pk}B_{ql}T_{kl}$ gives

$$
\begin{aligned}
\ddot{\gamma}_*^{pq,rs}(B)T^*_{pq}T^*_{rs} &= 2\gamma(A)^{-3}\dot{\gamma}^{kl}(A)\dot{\gamma}^{ij}(A)T_{kl}T_{ij} - \gamma(A)^{-2}\ddot{\gamma}^{kl,ij}(A)T_{kl}T_{ij}\\
&\quad + 2\gamma(A)^{-2}\dot{\gamma}^{kl}(A)B_{ij}T_{ik}T_{jl}\\
&= -\gamma(A)^{-2}\left(\ddot{\gamma}^{ij,kl}(A) + 2\dot{\gamma}^{ik}(A)B_{jl} - \frac{2}{\gamma(A)}\dot{\gamma}^{ij}(A)\dot{\gamma}^{kl}(A)\right)T_{ij}T_{kl}.
\end{aligned}
$$

□

Combining this computation with the evolution of λ_1 gives the following inequality. Crucially, the right-hand side is nonnegative except for terms which contain $|\nabla\varphi|$ as a factor and thus vanish at a spatial minimum of λ_1. We note that the analysis carried out here is slightly more detailed than in [**And07**], where Andrews computes only at a nonnegative spacetime minimum of λ_1. The main difference here is that we use Lemma 3.4 to get an equation for λ_1 which holds everywhere. This will allow us to apply the strong maximum principle to prove Corollary 3.8 below.

PROPOSITION 3.6. *Let $\gamma : \Gamma \to (0,\infty)$ be such that $\Gamma_+ \subset \Gamma$ and suppose the restriction of γ to Γ_+ is inverse-concave. Fix a point $(x_0,t_0) \in M \times (0,T)$ and let φ be a lower support for λ_1 at (x_0,t_0). Then if $\lambda_1(x_0,t_0) \in \Gamma' \cap \overline{\Gamma}_+$ for some $\Gamma' \Subset \Gamma$, in a principal frame at (x_0,t_0) there holds*

$$(\partial_t - \Delta_\gamma)\varphi \geq |A|^2_\gamma\varphi - C\frac{|\nabla\varphi|^2}{G} - C\frac{|\nabla\varphi||\nabla_1 A|}{G} + 2\frac{|\nabla_1 G|^2}{G},$$

where $C = C(n,\gamma,\Gamma')$.

PROOF. At the point (x_0, t_0), we have the decomposition

$$\nabla_1 A = S + T,$$

where

$$S := \nabla_1 A_{11} e^1 \otimes e^1 + \sum_{p\geq 2} \nabla_1 A_{p1} e^p \otimes e^1 + \sum_{q\geq 2} \nabla_1 A_{1q} e^1 \otimes e^q$$
$$T := \sum_{p,q\geq 2} \nabla_1 A_{pq} e^p \otimes e^q.$$

Suppose that the dimension of the eigenspace of λ_1 at (x_0, t_0) is m. By Lemma 3.4, if $1 < p \leq m$, then $\nabla_q A_{p1}(x_0, t_0) = 0$ for all indices q. Therefore, by the Codazzi equations, at (x_0, t_0) we have

$$T = \sum_{p,q\geq m+1} \nabla_1 A_{pq} e^p \otimes e^q.$$

We may write

$$\ddot{\gamma}^{pq,rs} \nabla_1 A_{pq} \nabla_1 A_{rs} = \ddot{\gamma}^{pq,rs} S_{pq} S_{rs} + 2\ddot{\gamma}^{pq,rs} S_{pq} T_{rs} + \ddot{\gamma}^{pq,rs} T_{pq} T_{rs}.$$

Recall that at (x_0, t_0) there holds

$$(\partial_t - \Delta_\gamma)\varphi \geq |A|^2_\gamma \varphi + \ddot{\gamma}^{pq,rs} \nabla_1 A_{pq} \nabla_1 A_{rs} + 2\dot{\gamma}^k \sum_{\lambda_p > \lambda_1} \frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1}.$$

Since we are assuming that $\lambda_1(x_0, t_0) \geq 0$, we can use the Codazzi equations to estimate

$$\begin{aligned} 2\dot{\gamma}^k \sum_{\lambda_p > \lambda_1} \frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1} &= 2\dot{\gamma}^k \sum_{p\geq m+1} \frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1} \\ &= 2\dot{\gamma}^1 \sum_{p\geq m+1} \frac{|\nabla_1 A_{p1}|^2}{\lambda_p - \lambda_1} + 2 \sum_{k,p\geq m+1} \dot{\gamma}^k \frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1} \\ &\geq 2 \sum_{k,p\geq m+1} \dot{\gamma}^k \frac{|T_{kp}|^2}{\lambda_p}, \end{aligned}$$

and putting this all together we find that at (x_0, t_0),

$$\begin{aligned} (\partial_t - \Delta_\gamma)\varphi \geq{}& |A|^2_\gamma \varphi + \ddot{\gamma}^{pq,rs} S_{pq} S_{rs} + 2\ddot{\gamma}^{pq,rs} S_{pq} T_{rs} \\ &+ \sum_{p,q,r,s\geq m+1} (\ddot{\gamma}^{pq,rs} + 2\dot{\gamma}^{pr} \lambda_q^{-1} \delta^{qs}) T_{pq} T_{rs}. \end{aligned}$$

If B is any positive diagonal matrix with eigenvalues μ, we can write

$$\begin{aligned} &\sum_{p,q,r,s\geq m+1} (\ddot{\gamma}^{pq,rs}(B) + 2\dot{\gamma}^{pr}(B) \mu_q^{-1} \delta^{qs}) T_{pq} T_{rs} \\ &\qquad = (\ddot{\gamma}^{pq,rs}(B) + 2\dot{\gamma}^{pr}(B)(B^{-1})^{qs}) T_{pq} T_{rs}, \end{aligned}$$

and conclude using the inverse-concavity and Lemma 3.5 that

$$\sum_{p,q,r,s\geq m+1} (\ddot{\gamma}^{pq,rs}(B) + 2\dot{\gamma}^{pr}(B) \mu_q^{-1} \delta^{qs}) T_{pq} T_{rs} \geq \frac{2}{\gamma(B)} (\dot{\gamma}^{pq}(B) T_{pq})^2.$$

Since $T_{pq} = 0$ if $p \leq m$ or $q \leq m$, by approximation, the same inequality also holds in the case that

$$0 = \mu_1 = \cdots = \mu_m, \qquad 0 < \mu_{m+1} \leq \cdots \leq \mu_n,$$

as long as γ is smooth in a neighbourhood of μ. Therefore, at the point (x_0, t_0), we have

$$\begin{aligned}\sum_{p,q,r,s \geq m+1} &(\ddot{\gamma}^{pq,rs}(A) + 2\dot{\gamma}^{pr}(A)\lambda_q^{-1}\delta^{qs})T_{pq}T_{rs} \\ &\geq \frac{2}{G}(\dot{\gamma}^{pq}(A)T_{pq})^2 \\ &= \frac{2}{G}(\nabla_1 G - \dot{\gamma}^{pq}(A)S_{pq})^2 \\ &= \frac{2}{G}|\nabla_1 G|^2 - \frac{4}{G}\nabla_1 G(\dot{\gamma}^{pq}(A)S_{pq}) + \frac{2}{G}(\dot{\gamma}^{pq}(A)S_{pq})^2,\end{aligned}$$

and consequently,

$$\begin{aligned}(\partial_t - \Delta_\gamma)\varphi \geq{}& |A|_\gamma^2\varphi + \ddot{\gamma}^{pq,rs}S_{pq}S_{rs} + 2\ddot{\gamma}^{pq,rs}S_{pq}T_{rs} \\ &+ \frac{2}{G}|\nabla_1 G|^2 - \frac{4}{G}\nabla_1 G(\dot{\gamma}^{pq}(A)S_{pq}) + \frac{2}{G}(\dot{\gamma}^{pq}(A)S_{pq})^2.\end{aligned}$$

To estimate the remaining terms, we note that since $\lambda(x_0, t_0) \in \Gamma'$,

$$\ddot{\gamma}^{pq,rs}S_{pq}S_{rs} + 2\ddot{\gamma}^{pq,rs}S_{pq}T_{rs} \geq -C(n, \gamma, \Gamma')\left(\frac{|S|^2}{G} + \frac{|S||T|}{G}\right),$$

and

$$-4\nabla_1 G(\dot{\gamma}^{pq}(A)S_{pq}) + 2(\dot{\gamma}^{pq}(A)S_{pq})^2 \geq -C(n, \gamma, \Gamma')(|\nabla_1 A||S| + |S|^2).$$

By the Codazzi equations,

$$|S|^2 = |\nabla_1 A_{11}|^2 + 2\sum_{p \geq 2}|\nabla_1 A_{p1}|^2 \leq 2|\nabla A_{11}|^2,$$

and $\nabla_p A_{11}(x_0, t_0) = \nabla_p \varphi(x_0, t_0)$, so we have

$$|S|^2 \leq 2|\nabla\varphi|^2.$$

Also, $|T|^2 \leq |\nabla_1 A|^2$, so at the point (x_0, t_0) there holds

$$(\partial_t - \Delta_\gamma)\varphi \geq |A|_\gamma^2\varphi - C\frac{|\nabla\varphi|^2}{G} - C\frac{|\nabla\varphi||\nabla_1 A|}{G} + 2\frac{|\nabla_1 G|^2}{G},$$

where $C = C(n, \gamma, \Gamma')$. $\square$

Now we can apply the maximum principle to draw some important conclusions from these computations. First, by an elementary argument we have:

COROLLARY 3.7. *Let $\gamma : \Gamma \to (0, \infty)$ be an admissible speed with $\Gamma_+ \subset \Gamma$ and suppose the restriction of γ to Γ_+ is inverse-concave. Let $F : M \times [0, T) \to \mathbb{R}^{n+1}$ be a compact solution of* (CF) *which is such that*

$$\lambda(x, t) \in \Gamma' \qquad \forall\ (x, t) \in M \times [0, T),$$

where $\Gamma' \Subset \Gamma$ *is a symmetric cone, and*

$$\min_{x\in M} \lambda_1(x,0) > 0.$$

Then λ_1 *is positive on* $M \times [0,T)$.

Next, by the strong maximum principle for viscosity solutions of parabolic equations (see for example [**DL04**]), we can conclude that if λ_1 is nonnegative and vanishes at an interior point, then it vanishes everywhere backwards in time. This result will play a key role when we derive curvature derivative estimates in the next chapter (see Step 5 of the proof of Theorem 4.11). Note that we do not require M to be complete.

A similar result for two-convex solutions can be found in [**BL16**][Theorem A.1]. The two-convex case is made somewhat simpler by the fact that no zero eigenvalue of A can occur with multiplicity.

Corollary 3.8. *Let* $\gamma : \Gamma \to (0,\infty)$ *be an admissible speed such that* $\Gamma_+ \subset \Gamma$ *and suppose the restriction of* γ *to* Γ_+ *is inverse-concave. Let* M *be a connected smooth manifold and consider an evolving immersion* $F : M \times [-T+t_0, t_0] \to \mathbb{R}^{n+1}$ *which solves* (CF)*. Suppose there is a symmetric cone* $\Gamma' \Subset \Gamma$ *such that* $\lambda(x,t) \in \Gamma' \cap \overline{\Gamma}_+$ *for each* $(x,t) \in M \times [-T+t_0, t_0]$*. If in addition* $\lambda_1(x_0,t_0) = 0$ *for some* $x_0 \in M$*, then* $\lambda_1 \equiv 0$ *on* $M \times [-T+t_0,t_0]$*. Furthermore, if* $v \in T_xM$ *is in* $\ker(A(x,t))$ *then there holds*

$$\nabla_v G(x,t) = 0.$$

Proof. The fact that λ_1 vanishes identically follows from Proposition 3.6 and the strong maximum principle in [**DL04**]. If $v \in T_xM$ is as in the kernel of $A(x,t)$, then we can normalise so that $|v| = 1$ and extend v to a principal frame at (x,t). Since the constant function $\varphi \equiv 0$ is a lower support for λ_1 at (x,t), Proposition 3.6 implies that

$$0 = |\nabla_v G|^2(x,t).$$

$\square$

2. A speed for k-convex hypersurfaces

We consider a fixed dimension $n \geq 4$, and a fixed $3 \leq k \leq n-1$. Let

$$\gamma_1(\lambda) = \left(\sum_{1\leq i_1 < \cdots < i_k \leq n} \frac{1}{\lambda_{i_1} + \cdots + \lambda_{i_k}}\right)^{-1},$$

for each λ in the cone

$$\Gamma := \{\lambda \in \mathbb{R}^n : \lambda_{i_1} + \cdots + \lambda_{i_k} > 0 \ \ \forall \ \ 1 \leq i_1 < \cdots < i_k \leq n\}.$$

For the rest of this chapter we will be concerned with the following family of speed functions: for each $\rho \in (0,1]$, we define $\gamma_\rho : \Gamma \to (0,\infty)$ by

$$\gamma_\rho(\lambda) = (\rho\gamma_1(\gamma)^{-1} + (1-\rho)\operatorname{tr}(\lambda)^{-1})^{-1}.$$

As the parameter ρ ranges over $(0,1]$, γ_ρ interpolates between γ_1 and the trace, and for each $\lambda \in \Gamma$ there holds

$$\min_{1\leq i_1 < \cdots < i_k \leq n} \frac{\lambda_{i_1} + \cdots + \lambda_{i_k}}{\rho} > \gamma_\rho(\lambda).$$

This property ensures that a family of hypersurfaces evolving with inward normal velocity given by γ_ρ applied to the principal curvatures will remain strictly k-convex for all time, even in a curved background space (we elaborate on this in Section 5 below). If ρ is small, then γ_ρ is approximately linear away from $\partial\Gamma$, and it is essentially this property that will allow us to prove the convexity estimate (Theorem 3.1).

For now, we observe that γ_ρ is concave for every $\rho \in (0,1]$, is strictly concave in off-radial directions, and is also inverse-concave on the positive cone. Indeed, for $\lambda \in \Gamma_+$, there holds

$$\begin{aligned}&\gamma_\rho(\lambda_1^{-1},\dots,\lambda_n^{-1})^{-1}\\ &\quad= \rho \sum_{1\le i_1<\dots<i_k\le n}\left(\frac{1}{\lambda_{i_1}}+\dots+\frac{1}{\lambda_{i_k}}\right)^{-1} + (1-\rho)\left(\frac{1}{\lambda_1}+\dots+\frac{1}{\lambda_n}\right)^{-1},\end{aligned}$$

and each of the summands on the right is a concave function of λ.

2.1. Estimates for the derivatives of γ_ρ. As the parameter ρ tends to zero, γ_ρ converges smoothly to the trace function on Γ, but since γ_ρ always vanishes at the boundary of Γ, this convergence clearly cannot be uniform. Our aim now is to establish estimates which control how the first and second derivatives of γ_ρ differ from those of the trace at a fixed distance from $\partial\Gamma$.

Let us define

$$h(x_1,x_2) = \left(\frac{1}{x_1}+\frac{1}{x_2}\right)^{-1}, \qquad \forall\, x_1,\, x_2 > 0,$$

so that we may write

$$\gamma_\rho(\lambda) = h\left(\frac{\gamma_1(\lambda)}{\rho}, \frac{\operatorname{tr}(\lambda)}{1-\rho}\right).$$

Using the fact that

$$\dot h^p(x) = \frac{h(x)^2}{x_p^2},$$

we compute

$$\dot\gamma_\rho^{pq}(A) = \rho\frac{\gamma_\rho(\lambda)^2}{\gamma_1(\lambda)^2}\dot\gamma_1^{pq}(A) + (1-\rho)\frac{\gamma_\rho(\lambda)^2}{\operatorname{tr}(\lambda)^2}\delta^{pq}.$$

Furthermore, since for each $\xi \in \mathbb{R}^2$, the Hessian of h acts by

$$\ddot h^{pq}(x)\xi_p\xi_q = -2\frac{h(x)^3}{x_1x_2}\left(\frac{\xi_1}{x_1}-\frac{\xi_2}{x_2}\right)^2,$$

for $A \in \mathrm{O}\,\Gamma$ we have

$$\begin{aligned}\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{pq} &= \rho\frac{\gamma_\rho(A)^2}{\gamma_1(A)^2}\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs}\\ &\quad - 2\rho(1-\rho)\frac{\gamma_\rho(A)^3}{\gamma_1(A)\operatorname{tr}(A)}\left(\frac{\dot\gamma_1^{pq}(A)T_{pq}}{\gamma_1(A)} - \frac{\operatorname{tr}(T)}{\operatorname{tr}(A)}\right)^2.\end{aligned} \tag{11}$$

LEMMA 3.9. *For each $\rho \in (0,1]$ and $\lambda \in \Gamma$, there holds*

$$\gamma_1(\lambda) \le \gamma_\rho(\lambda) \le \min\{\operatorname{tr}(\lambda), \rho^{-1}\gamma_1(\lambda)\}.$$

PROOF. We first observe that

$$\gamma_\rho(\lambda)^{-1} = \rho\gamma_1(\lambda)^{-1} + (1-\rho)\operatorname{tr}(\lambda)^{-1} \geq \rho\gamma_1(\lambda)^{-1},$$

which is one of the desired upper bounds. Also, since

$$\gamma_1(\lambda) \leq \min_{1\leq i_1<\cdots<i_k\leq n} \lambda_{i_1} + \cdots + \lambda_{i_k} \leq \operatorname{tr}(\lambda),$$

there holds

$$\operatorname{tr}(\lambda)^{-1} \leq \rho\gamma_1(\lambda)^{-1} + (1-\rho)\operatorname{tr}(\lambda)^{-1} \leq \gamma_1(\lambda)^{-1}.$$

Inverting this gives the remaining two inequalities. □

Inserting these estimates into the expression for the derivatives of γ_ρ from above, we immediately obtain the following:

LEMMA 3.10. *For each $\rho \in (0,1]$ and $A \in \mathrm{O}\,\Gamma$, there holds:*

$$\dot\gamma_\rho^{kl}(A) \leq \min\left\{\frac{1}{\rho}, \frac{\operatorname{tr}(A)^2}{\gamma_1(A)^2}\right\}\dot\gamma_1^{kl}(A) + \delta^{kl};$$
$$\dot\gamma_\rho^{kl}(A) \geq \rho\dot\gamma_1^{kl}(A) + (1-\rho)\frac{\gamma_1(A)^2}{\operatorname{tr}(A)^2}\delta^{kl}.$$

For the second derivatives, we have:

LEMMA 3.11. *For each $\rho \in (0,1]$ and symmetric $A \in \mathrm{O}\,\Gamma$, there holds*

$$\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{rs} \leq \rho\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs},$$

and

$$-\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{rs}$$
$$\leq \min\left\{\rho^{-2}, \rho\frac{\operatorname{tr}(A)^3}{\gamma_1(A)^3}\right\}\left(-\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs} + 4\frac{(\dot\gamma_1^{pq}T_{pq})^2}{\operatorname{tr}(A)} + 4\frac{\operatorname{tr}(T)^2}{\operatorname{tr}(A)}\right).$$

PROOF. The first estimate follows immediately from (11) and Lemma 3.9. To obtain the second, we first bound

$$2\rho(1-\rho)\frac{\gamma_\rho(\lambda)^3}{\gamma_1(\lambda)\operatorname{tr}(\lambda)}\left(\frac{\dot\gamma_1^{pq}(A)T_{pq}}{\gamma_1(\lambda)} - \frac{\operatorname{tr}(T)}{\operatorname{tr}(\lambda)}\right)^2$$
$$\leq 4\rho\frac{\gamma_\rho(\lambda)^3}{\gamma_1(\lambda)\operatorname{tr}(\lambda)}\left(\frac{(\dot\gamma_1^{pq}(A)T_{pq})^2}{\gamma_1(\lambda)^2} + \frac{\operatorname{tr}(T)^2}{\operatorname{tr}(\lambda)^2}\right)$$
$$\leq 4\rho\frac{\gamma_\rho(\lambda)^3}{\gamma_1(\lambda)^3}\left(\frac{(\dot\gamma_1^{pq}(A)T_{pq})^2}{\operatorname{tr}(\lambda)} + \frac{\operatorname{tr}(T)^2}{\operatorname{tr}(\lambda)}\right).$$

Using the fact that $\gamma_\rho \geq \gamma_1$, we obtain

$$-\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{rs} \leq \rho\frac{\gamma_\rho(\lambda)^3}{\gamma_1(\lambda)^3}\left(-\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs} + 4\frac{(\dot\gamma_1^{pq}T_{pq})^2}{\operatorname{tr}(\lambda)} + 4\frac{\operatorname{tr}(T)^2}{\operatorname{tr}(\lambda)}\right),$$

and the claim then follows from Lemma 3.9. □

At points which are at a controlled distance from the boundary of Γ, we can say more. We introduce the notation

$$\Gamma_\alpha = \{\lambda \in \mathbb{R}^n : \operatorname{tr}(\lambda) \le \alpha\gamma_1(\lambda)\},$$

and observe that since γ_1 is strictly concave in off-radial directions and vanishes at $\partial\Gamma$, the set

$$\Gamma_\alpha \cap \{\lambda \in \mathbb{R}^n : \operatorname{tr}(\lambda) = 1\}$$

is a compact subset of Γ.

LEMMA 3.12. *For every $\alpha < \infty$, there is a positive constant $C = C(n, k, \alpha)$ with the following properties. If $A \in \mathrm{O}\,\Gamma_\alpha$, then*

$$C^{-1}\delta^{ij} \le \dot\gamma_\rho^{ij}(A) \le C\delta^{ij},$$

and

$$-\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{rs} \le C\rho\frac{|T|^2}{\operatorname{tr}(A)}.$$

PROOF. Consider the set

$$S := \{A \in \mathrm{O}\,\Gamma_\alpha : \operatorname{tr}(A) = 1\}.$$

As noted above, S is a compact subset of $\mathrm{O}\,\Gamma$. Therefore, since $\dot\gamma_1^{ij}$ is smooth and positive-definite on Γ, the quantity

$$c_0 := \min\{\dot\gamma_1^{ij}(A)\xi_i\xi_j : A \in S,\ \xi \in \mathbb{R}^n,\ |\xi| = 1\}$$

is strictly positive, and depends only on n, k, and α. If $A \in \mathrm{O}\,\Gamma_\alpha$, and $\xi \in \mathbb{R}^n$, then since $\dot\gamma_1^{ij}$ is scaling-invariant there holds

$$\dot\gamma_1^{ij}(A)\xi_i\xi_j = \dot\gamma_1^{ij}(\operatorname{tr}(A)^{-1}A)\xi_i\xi_j \ge c_0|\xi|^2.$$

For $A \in \mathrm{O}\,\Gamma_\alpha$, Lemma 3.10 tells us that

$$\rho\dot\gamma_1^{ij}(A) + (1-\rho)\alpha^{-2}\delta^{ij} \le \dot\gamma_\rho^{ij}(A) \le \alpha^2\dot\gamma_1^{ij}(A) + \delta^{ij}.$$

We also have $\dot\gamma_1^{ij} \le C(n,k)\delta^{ij}$, so appealing to the lower bound just derived, we find that

$$\min\{c_0, \alpha^{-2}\}\delta^{ij} \le \dot\gamma_\rho^{ij}(A) \le (\alpha^2 C(n,k) + 1)\delta^{ij}.$$

Next we define $C_0 = C_0(n,k,\alpha)$ by

$$C_0 := \max\{-\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs} : A \in S,\ T \in \operatorname{Sym}(n),\ |T| = 1\},$$

and observe that C_0 is nonnegative by the concavity of, since γ_1. Then if A is any matrix in $\mathrm{O}\,\Gamma_\alpha$ and $T \in \operatorname{Sym}(n)$, there holds

$$-\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs} = -\operatorname{tr}(A)^{-1}\ddot\gamma_1^{pq,rs}(\operatorname{tr}(A)^{-1}A)T_{pq}T_{rs} \le C_0\operatorname{tr}(A)^{-1}|T|^2.$$

Combining this inequality with the second bound in Lemma 3.11, we obtain

$$\begin{aligned}
&-\ddot\gamma_\rho^{pq,rs}(A)T_{pq}T_{rs} \\
&\qquad \le \rho\alpha^3\bigg(-\ddot\gamma_1^{pq,rs}(A)T_{pq}T_{rs} + 4\frac{(\dot\gamma_1^{pq}T_{pq})^2}{\operatorname{tr}(A)} + 4\frac{\operatorname{tr}(T)^2}{\operatorname{tr}(A)}\bigg) \\
&\qquad \le C(n,k,\alpha)\rho\operatorname{tr}(A)^{-1}|T|^2.
\end{aligned}$$

$\square$

3. Consequences of the cylindrical estimate

We continue to consider a fixed $n \geq 4$ and $k \in \{3, \dots, n-1\}$. Let M be a compact smooth n-manifold and suppose $F : M \times [0, T) \to \mathbb{R}^{n+1}$ is an evolving immersion satisfying the equation

$$\partial_t F(x, t) = -G_\rho(x, t)\nu(x, t),$$

where $G_\rho(x, t) := \gamma_\rho(x, t)$ and $\rho > 0$. It follows from Lemma 1.3 that since γ_ρ is concave, we have the estimate

$$\frac{H(x,t)}{G_\rho(x,t)} \leq \max_M \frac{H(\cdot, 0)}{G_\rho(\cdot, 0)}$$

for all $(x, t) \in M \times [0, T)$. Inserting the definition of γ_ρ, we find that

$$\rho \frac{H(x,t)}{G_1(x,t)} + 1 - \rho \leq \rho \max_M \frac{H(\cdot, 0)}{G_1(\cdot, 0)} + 1 - \rho,$$

which we may simplify to

$$\frac{H(x,t)}{G_1(x,t)} \leq \max_M \frac{H(\cdot, 0)}{G_1(\cdot, 0)}.$$

Interestingly, this estimate no longer contains the parameter ρ.

This same algebraic property lets us draw a powerful consequence from the cylindrical estimate derived in the previous chapter. Let us define $\alpha_j^{(\rho)}$ to be the value attained by the function $\lambda \mapsto \gamma_\rho(\lambda)^{-1} \operatorname{tr}(\lambda)$ on a cylinder of the form $\mathbb{R}^j \times S^{n-j}$. Then, applied to the speed γ_ρ, the cylindrical estimate can be stated as follows.

THEOREM 3.13. *Fix $\rho > 0$, consider $F : M \times [0, T) \to \mathbb{R}^{n+1}$ as above, and define*

$$R^{-1} := \sup_M G_\rho(\cdot, 0), \qquad \bar{\alpha} := \max_M \frac{H(\cdot, 0)}{G_\rho(\cdot, 0)}.$$

Then, for each $\varepsilon > 0$, there is a positive $C_\varepsilon = C_\varepsilon(n, k, \rho, \bar{\alpha}, R, \mu_0(M), T)$ with the property that

$$\frac{H(x,t)}{G_\rho(x,t)} \leq \alpha_{k-1}^{(\rho)} + \varepsilon + C_\varepsilon G_\rho(x,t)^{-1}$$

for each $(x, t) \in M \times [0, T)$.

Using the definition of γ_ρ as before, we can simplify the statement of the estimate. The parameter ρ does not completely disappear, but in the end, appears only in the lower-order term.

COROLLARY 3.14. *For each $\varepsilon > 0$, there is a positive $C_\varepsilon = C_\varepsilon(n, k, \rho, \bar{\alpha}, R, \mu_0(M), T)$ with the property that*

$$\frac{H(x,t)}{G_1(x,t)} \leq \alpha_{k-1}^{(1)} + \varepsilon + C_\varepsilon G_\rho(x,t)^{-1}$$

for each $(x, t) \in M \times [0, T)$.

PROOF. Inserting the definition of γ_ρ into the cylindrical estimate, we find that for each $\varepsilon > 0$,

$$\rho\frac{H(x,t)}{G_1(x,t)} + 1 - \rho \leq \rho\alpha^{(1)}_{k-1} + 1 - \rho + \varepsilon + C_\varepsilon G_\rho(x,t)^{-1},$$

which we rearrange to obtain

$$\frac{H(x,t)}{G_1(x,t)} \leq \alpha^{(1)}_{k-1} + \rho^{-1}\varepsilon + \rho^{-1}C_\varepsilon G_\rho(x,t)^{-1}.$$

Since ε was arbitrary, this implies the desired estimate. □

As a consequence of this estimate, we see that for any positive ρ, if $G_\rho(x,t)$ is sufficiently large, then $\lambda(x,t)$ is extremely close to the cone $\Gamma_{\alpha^{(1)}_{k-1}}$. The smaller ρ is, the larger $G_\rho(x,t)$ will need to be to ensure that this is the case, but the conclusion on the position of λ in curvature space is independent of ρ.

4. Proof of the convexity estimate

The proof of Theorem 3.1 will proceed by an application the Stampacchia principle from Chapter 2 to a certain curvature quantity, which we now construct and analyse. Throughout this section, $n \geq 4$ and $k \in \{3, \dots, n-1\}$ are fixed, M is a smooth compact n-manifold, and $F : M \times [0,T) \to \mathbb{R}^{n+1}$ is an evolving immersion satisfying

$$\partial_t F(x,t) = -G_\rho(x,t)\nu(x,t),$$

with $\rho \in (0,1]$. We define

$$R^{-1} := \max_M G_\rho(\cdot,0), \qquad \bar{\alpha} := \max_M \frac{H(\cdot,0)}{G_\rho(\cdot,0)}.$$

LEMMA 3.15. *Fix a spacetime point $(x_0,t_0) \in M \times (0,T)$, and suppose that $\lambda(x_0,t_0) \in \Gamma_\alpha$, where $\alpha > 0$. Then if φ is a lower support for λ_1 at (x_0,t_0), in a principal frame at (x_0,t_0) there holds*

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \geq |A|^2_{\gamma_\rho}\varphi - C\rho\frac{|\nabla_1\varphi|^2}{H} + (C^{-1} - C\rho)\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H},$$

where $C = C(n,k,\alpha)$.

PROOF. We let m be the dimension of the kernel of $A(x_0,t_0) - \lambda_1(x_0,t_0)g(x_0,t_0)$, so that $\lambda_p > \lambda_1$ if and only if $p \geq m+1$. By Proposition 3.3, we know that at (x_0,t_0) there holds

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \geq |A|^2_{\gamma_\rho}\varphi + \ddot{\gamma}^{pq,rs}_\rho\nabla_1 A_{pq}\nabla_1 A_{rs} + 2\dot{\gamma}^i_\rho\sum_{\lambda_p>\lambda_1}\frac{|\nabla_i A_{p1}|^2}{\lambda_p - \lambda_1},$$

and since $\lambda(x_0,t_0) \in \Gamma_\alpha$, by Lemma 3.12 we can estimate $\dot{\gamma}^i_\rho(\lambda(x_0,t_0)) \geq c_0(n,k,\alpha)$ and so obtain

$$2\dot{\gamma}^i_\rho\sum_{\lambda_p>\lambda_1}\frac{|\nabla_i A_{p1}|^2}{\lambda_p-\lambda_1} \geq 2c_0\sum_i\sum_{p\geq m+1}\frac{|\nabla_i A_{p1}|^2}{\lambda_p-\lambda_1}.$$

Since $\lambda_1 + \cdots + \lambda_k > 0$ and $k \leq n-1$ we have $\lambda_n < \operatorname{tr}(\lambda)$, and consequently

$$\lambda_p - \lambda_1 < \lambda_p + \lambda_2 + \cdots + \lambda_k < k\operatorname{tr}(\lambda)$$

for each $p \geq m+1$. Substituting this in, we find that at (x_0, t_0) there holds

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \geq |A|^2_{\gamma_\rho}\varphi + \ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + 2c_0 k^{-1}\sum_i \sum_{p\geq m+1}\frac{|\nabla_i A_{p1}|^2}{H}.$$

By Lemma 3.4, the definition of m, and the Codazzi equations, the tensor $\nabla_1 A$ has the following structure at (x_0, t_0):

$$\nabla_1 A = \nabla_1 A_{11} e^1 \otimes e^1 + \sum_{p\geq m+1}\nabla_1 A_{p1} e^p \otimes e^1 + \sum_{q\geq m+1}\nabla_1 A_{1q} e^1 \otimes e^q$$
$$+ \sum_{p,q\geq m+1}\nabla_1 A_{pq} e^p \otimes e^q.$$

Using the Codazzi equations again, we find that at (x_0, t_0),

$$\begin{aligned}
\sum_i \sum_{p\geq m+1}|\nabla_i A_{p1}|^2 &= \sum_i \sum_{p\geq m+1}|\nabla_1 A_{pi}|^2 \\
&= \sum_{p\geq m+1}|\nabla_1 A_{p1}|^2 + \sum_{i,p\geq m+1}|\nabla_1 A_{pi}|^2 \\
&= \frac{1}{2}\sum_{p\geq m+1}|\nabla_1 A_{p1}|^2 + \frac{1}{2}\sum_{q\geq m+1}|\nabla_1 A_{1q}|^2 + \sum_{p,q\geq m+1}|\nabla_1 A_{pq}|^2 \\
&\geq \frac{1}{2}(|\nabla_1 A|^2 - |\nabla_1 A_{11}|^2).
\end{aligned}$$

Hence at (x_0, t_0) we have

$$\begin{aligned}
(\partial_t - \Delta_{\gamma_\rho})\varphi &\geq |A|^2_{\gamma_\rho}\varphi + \ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + c_0 k^{-1}\left(\frac{|\nabla_1 A|^2}{H} - \frac{|\nabla_1 A_{11}|^2}{H}\right) \\
&= |A|^2_{\gamma_\rho}\varphi + \ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + c_0 k^{-1}\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H}.
\end{aligned}$$

To finish the proof, we use Lemma 3.12 and the assumption $\lambda(x_0, t_0) \in \Gamma_\alpha$ to conclude that at (x_0, t_0),

$$\begin{aligned}
\ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} &\geq -C(n,k,\alpha)\rho\frac{|\nabla_1 A|^2}{H} \\
&= -C(n,k,\alpha)\rho\frac{|\nabla_1 \varphi|^2}{H} - C(n,k,\alpha)\rho\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H}.
\end{aligned}$$

Here we have also used $\nabla_1 A_{11}(x_0, t_0) = \nabla_1\varphi(x_0, t_0)$. □

Next we rearrange the version of the cylindrical estimate from Corollary 3.14 to find that

$$0 \leq G_1(x,t) - \frac{1}{\alpha^{(1)}_{k-1} + \varepsilon}H(x,t) + \frac{C_\varepsilon}{\alpha^{(1)}_{k-1} + \varepsilon}\frac{G_1(x,t)}{G_\rho(x,t)}$$

for each $(x,t) \in M \times [0,T)$. We (somewhat arbitrarily) set the parameter ε equal to $\varepsilon_0 := 100^{-1}\alpha^{(1)}_{k-1}$ in this estimate, and use $G_1(x,t) \leq G_\rho(x,t)$ (this was proven in Lemma 3.9) to obtain

$$0 \leq G_\rho(x,t) - 2\mu H(x,t) + K,$$

where

$$\mu := \frac{1}{2(1+100^{-1})\alpha^{(1)}_{k-1}}, \qquad K := \frac{C_{\varepsilon_0}}{(1+100^{-1})\alpha^{(1)}_{k-1}}$$

We will make use of the function $h(x,t) := G_\rho(x,t) - \mu H(x,t) + K$, which by construction satisfies

$$\mu H(x,t) \le h(x,t) \le G_\rho(x,t) + K$$

for every $(x,t) \in M \times [0,T)$. The constant μ depends only on n and k, and

$$K = K(n,k,\rho,\bar{\alpha},R,\mu_0(M),T).$$

The function h evolves according to

$$(\partial_t - \Delta_{\gamma_\rho})h = |A|^2_{\gamma_\rho}(h-K) - \mu g^{kl}\ddot{\gamma}_\rho^{pq,rs}\nabla_k A_{pq}\nabla_l A_{rs}.$$

We are going to make use of the good gradient term on the right to control the gradient terms appearing in the evolution of λ_1. Here it will be crucial that the coefficient μ depends only on n and k, since we will have to choose ρ small depending on μ. In principle, we could carry out the entire construction of this section with μ taken to be any positive value larger than $\alpha^{(1)}_{k-1}$, so our choice of ρ is not canonical.

For each $\eta \in (0,1]$, we define

$$f_\eta(x,t) = \frac{-\lambda_1(x,t) - \eta G_\rho(x,t)}{h(x,t)}.$$

This is the function to which we are eventually going to apply the Stampacchia procedure. Our immediate goal is to derive an evolution equation for f_η and analyse the gradient terms appearing on the right-hand side. To do so, we employ the following elementary lemma.

LEMMA 3.16. *Let $\Gamma \subset \mathbb{R}^n$ be an open convex symmetric cone containing $(1,\dots,1)$, and let $\gamma : \Gamma \to \mathbb{R}$ be a smooth symmetric function which is one-homogeneous, concave, and satisfies*

$$\gamma(1,\dots,1) > 0.$$

Then if $\lambda \in \Gamma$ is such that $\lambda_1 \le \dots \le \lambda_n$, there holds $\dot{\gamma}^1(\lambda) \ge 0$.

PROOF. Fix $\lambda \in \Gamma$ satisfying $\lambda_1 \le \dots \le \lambda_n$. Since γ is concave, the super-level set

$$S := \{z \in \Gamma : \gamma(z) \ge \gamma(\lambda)\}$$

is convex, and since γ is symmetric, each of the vectors

$$(\lambda_m, \dots, \lambda_{n-1}, \lambda_n, \lambda_1, \dots, \lambda_{m-1})$$

is in S. Taking the average, we get $\bar{\lambda} \in S$, where

$$\bar{\lambda} := \frac{\operatorname{tr}(\lambda)}{n}(1,\dots,1).$$

Since Γ is open, convex, symmetric, and contains $(1,\dots,1)$,

$$\Gamma \subset \{z \in \mathbb{R}^n : \operatorname{tr}(z) > 0\},$$

so all of the entries of $\bar{\lambda}$ are positive. Therefore, since $\gamma(1,\dots,1) > 0$, for every $s \ge 1$ there holds

$$\gamma(s\bar{\lambda}) = s\gamma(\bar{\lambda}) \ge \gamma(\bar{\lambda}) \ge \gamma(\lambda),$$

which means $s\bar{\lambda} \in S$. Appealing again to the convexity of S, we find that for each $s \geq 1$, the line segment connecting λ with $s\bar{\lambda}$ is contained in S. It follows that

$$\{\lambda + s\bar{\lambda} : s \in [0, \infty)\} \subset S.$$

Another way to say this is that $\gamma(\lambda) \leq \gamma(\lambda + s\bar{\lambda})$ for all $s \geq 0$, so we have

$$0 \leq \frac{d}{ds}\bigg|_{s=0} \gamma(\lambda + s\bar{\lambda}) = \dot{\gamma}^i(\lambda)\bar{\lambda}_i = \frac{\operatorname{tr}(\lambda)}{n}\sum_{i=1}^n \dot{\gamma}^i(\lambda).$$

Without loss of generality, we may assume that $\lambda_1 < \cdots < \lambda_n$, since the general case then follows by approximation. Then since γ is concave, $\dot{\gamma}^j(\lambda) \leq \dot{\gamma}^i(\lambda)$ is true whenever $i < j$. Substituting this fact into the inequality above, we get

$$0 \leq \operatorname{tr}(\lambda)\dot{\gamma}^1(\lambda),$$

and the claim follows. $\square$

With the lemma in hand, we can establish the following estimate for the gradient terms in the evolution of f_η.

Proposition 3.17. *Let $(x_0, t_0) \in M \times (0, T)$ be such that $\lambda(x_0, t_0) \in \Gamma_\alpha$, where $\alpha > 0$, and let φ be an upper support function for f_η at the point (x_0, t_0). Suppose in addition that $f_\eta(x_0, t_0) \geq 0$. Then at the point (x_0, t_0) there holds*

$$\begin{aligned}(\partial_t - \Delta_{\gamma_\rho})\varphi \leq K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + \mu\frac{\varphi}{h}g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h \\ + C\rho\frac{h}{H}|\nabla_1\varphi|^2 - (C^{-1} - C\rho)\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{hH},\end{aligned}$$

where $C = C(n, k, \alpha)$.

Proof. We first observe that the smooth function

$$\tilde{\varphi}(x, t) := -h(x, t)\varphi(x, t) - \eta G_\rho(x, t).$$

is a lower support for λ_1 at (x_0, t_0), and

$$\varphi(x, t) = \frac{-\tilde{\varphi}(x, t) - \eta G_\rho(x, t)}{h(x, t)}.$$

For smooth functions u and v, with $v > 0$, there holds

$$(\partial_t - \Delta_\gamma)\frac{u}{v} = \frac{1}{v}(\partial_t - \Delta_\gamma)u - \frac{u}{v^2}(\partial_t - \Delta_\gamma)v + \frac{2}{v}\dot{\gamma}^{ij}\nabla_i\left(\frac{u}{v}\right)\nabla_j v,$$

so we have

$$\begin{aligned}(\partial_t - \Delta_{\gamma_\rho})\varphi &= -\frac{1}{h}(\partial_t - \Delta_{\gamma_\rho})(\tilde{\varphi} + \eta G_\rho) + \frac{\tilde{\varphi} + \eta G_\rho}{h^2}|A|^2_{\gamma_\rho}(h - K) \\ &\quad - \mu\frac{\tilde{\varphi} + \eta G_\rho}{h^2}g^{kl}\ddot{\gamma}_\rho^{pq,rs}\nabla_k A_{pq}\nabla_l A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h \\ &= K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + |A|^2_{\gamma_\rho}\frac{\tilde{\varphi}}{h} - \frac{1}{h}(\partial_t - \Delta_{\gamma_\rho})\tilde{\varphi} \\ &\quad + \mu\frac{\varphi}{h}g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h.\end{aligned}$$

Applying Lemma 3.15 to $\tilde{\varphi}$, we find that at the point (x_0, t_0),

$$(\partial_t - \Delta_{\gamma_\rho})\tilde{\varphi} \geq |A|^2_{\gamma_\rho}\tilde{\varphi} - C\rho\frac{|\nabla_1\tilde{\varphi}|^2}{H} + (C^{-1} - C\rho)\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H},$$

where $C = C(n, k, \alpha)$, and therefore,

$$\begin{aligned}(\partial_t - \Delta_{\gamma_\rho})\varphi \leq{}& K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + \mu\frac{\varphi}{h}g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h\\ &+\frac{1}{h}\bigg(C\rho\frac{|\nabla_1\tilde{\varphi}|^2}{H} - (C^{-1} - C\rho)\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H}\bigg).\end{aligned}$$

We are going to decompose and then absorb part of the term $\nabla_1\tilde{\varphi}$. At (x_0, t_0) there holds

$$\begin{aligned}\nabla_1\tilde{\varphi} &= -h\nabla_1\varphi - \varphi\nabla_1 h - \eta\nabla_1 G_\rho\\ &= -h\nabla_1\varphi - \varphi(\dot{\gamma}_\rho^i - \mu)\nabla_1 A_{ii} - \eta\dot{\gamma}_\rho^i\nabla_1 A_{ii}\\ &= -h\nabla_1\varphi - (\eta\dot{\gamma}_\rho^1 + \varphi(\dot{\gamma}_\rho^1 - \mu))\nabla_1\tilde{\varphi} - \sum_{i\geq 2}^n(\eta\dot{\gamma}_\rho^i + \varphi(\dot{\gamma}_\rho^i - \mu))\nabla_1 A_{ii},\end{aligned}$$

which we rearrange to obtain

$$(1 + \eta\dot{\gamma}_\rho^1 + \varphi(\dot{\gamma}_\rho^1 - \mu))\nabla_1\tilde{\varphi} = -h\nabla_1\varphi - \sum_{i\geq 2}^n(\eta\dot{\gamma}_\rho^i + \varphi(\dot{\gamma}_\rho^i - \mu))\nabla_1 A_{ii}.$$

The function

$$\lambda \mapsto \gamma_\rho(\lambda) - \mu\,\mathrm{tr}(\lambda)$$

is concave and one-homogeneous, and positive for each

$$\lambda \in \mathrm{int}\,\Gamma_{\mu^{-1}}.$$

Recalling that $\mu^{-1} := 2(1 + 100^{-1})\alpha^{(1)}_{k-1}$, we see that $\mathrm{Cyl}_{k-1} \subset \Gamma_{\mu^{-1}}$ and consequently $\mathrm{Cyl}_0 \subset \mathrm{int}\,\Gamma_{\mu^{-1}}$ (see Lemma 2.1), so there holds

$$\gamma_\rho(1, \dots, 1) - \mu\,\mathrm{tr}(1, \dots, 1) > 0.$$

We may therefore apply Lemma 4 to conclude that the quantity $\dot{\gamma}_\rho^1(\lambda(x_0, t_0)) - \mu$ is nonnegative. We are assuming that

$$\varphi(x_0, t_0) = f_\eta(x_0, t_0) \geq 0,$$

so we have

$$1 + \eta\dot{\gamma}_\rho^1 + \varphi(\dot{\gamma}_\rho^1 - \mu) \geq 1.$$

In particular, at the point (x_0, t_0) there holds

$$\nabla_1\tilde{\varphi} = \frac{1}{1 + \eta\dot{\gamma}_\rho^1 + \varphi(\dot{\gamma}_\rho^1 - \mu)}\bigg(-h\nabla_1\varphi - \sum_{i\geq 2}^n(\eta\dot{\gamma}_\rho^i + \varphi(\dot{\gamma}_\rho^i - \mu))\nabla_1 A_{ii}\bigg).$$

Let us introduce the abbreviation $\xi^i := \eta\dot{\gamma}_\rho^i + \varphi(\dot{\gamma}_\rho^i - \mu)$, so that we may write the last identity as

$$\nabla_1\tilde{\varphi} = -\frac{1}{1 + \xi^1}h\nabla_1\varphi - \sum_{i\geq 2}^n\frac{\xi^i}{1 + \xi^1}\nabla_1 A_{ii}.$$

Then since $\xi^1 \geq 0$ we can bound

$$|\nabla_1 \tilde{\varphi}|^2 \leq 2h^2 |\nabla_1 \varphi|^2 + C(n) \sum_{i \geq 2}^n |\xi^i|^2 |\nabla_1 A_{ii}|^2.$$

There holds

$$|\xi^i|^2 \leq 2\eta^2 |\dot{\gamma}_\rho^i|^2 + 4\varphi^2 (|\dot{\gamma}_\rho^i|^2 + \mu^2),$$

and $\eta \in (0, 1]$ by definition. Since $\lambda(x_0, t_0) \in \Gamma_\alpha$, we can bound $\dot{\gamma}_\rho^i(\lambda(x_0, t_0))$ purely in terms of n, k and α using Lemma 3.12, and at (x_0, t_0),

$$0 \leq \varphi = \frac{-\lambda_1 - \eta G_\rho}{h} \leq \frac{\lambda_2 + \cdots + \lambda_k}{h} \leq (k-1)\frac{H}{h} \leq \frac{k-1}{\mu}.$$

Putting these facts together, we can bound $|\xi^i|$ purely in terms of n, k and α, hence

$$|\nabla_1 \tilde{\varphi}|^2 \leq 2h^2 |\nabla_1 \varphi|^2 + C(n, k, \alpha) \sum_{i \geq 2}^n |\nabla_1 A_{ii}|^2.$$

Substituting this estimate back in, we find that at (x_0, t_0) there holds

$$\begin{aligned}(\partial_t - \Delta_{\gamma_\rho})\varphi \leq{}& K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + \mu\frac{\varphi}{h} g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h \\ &+ C\rho\frac{h}{H}|\nabla_1\varphi|^2 + \frac{1}{h}\bigg(C\rho\sum_{i\geq 2}^n \frac{|\nabla_1 A_{ii}|^2}{H} - (C^{-1} - C\rho)\sum_{p+q>2}\frac{|\nabla_1 A_{pq}|^2}{H}\bigg),\end{aligned}$$

which completes the proof. □

Applying the proposition with $\alpha = (1 + 100^{-1})\alpha^{(1)}_{k-1}$, we obtain the following corollary.

Corollary 3.18. *There is a positive constant*

$$\rho_0 = \rho_0(n, k)$$

with the following property. Fix $(x_0, t_0) \in M \times (0, T)$, let φ be an upper support function for f_η at the point (x_0, t_0), and suppose that

$$\lambda(x_0, t_0) \in \Gamma_{(1+100^{-1})\alpha^{(1)}_{k-1}}, \qquad f_\eta(x_0, t_0) \geq 0.$$

Then at the point (x_0, t_0) there holds

$$\begin{aligned}(\partial_t - \Delta_{\gamma_\rho})\varphi \leq{}& K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + \mu\frac{\varphi}{h} g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i\varphi\nabla_j h \\ &+ C\rho\frac{h}{H}|\nabla_1\varphi|^2,\end{aligned}$$

where $C = C(n, k)$.

4.1. Applying Stampacchia. We need to carry out a few more steps before the Stampacchia procedure can be applied to prove the convexity estimate. Let us introduce the notation

$$\Gamma' := \Gamma_{(1+100^{-1})\alpha^{(1)}_{k-1}}.$$

LEMMA 3.19. *Suppose $\rho \leq \rho_0$. Then if $(x_0, t_0) \in M \times (0, T)$ is such that*

$$\lambda(x_0, t_0) \in \Gamma', \qquad f_\eta(x_0, t_0) > 0,$$

and φ is an upper support for f_η at (x_0, t_0), at (x_0, t_0) there holds

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \leq CK|A|\varphi - C^{-1}\rho\varphi\frac{|\nabla A|^2}{hH} + C(\rho^{-1} + KH^{-1})\frac{|\nabla\varphi|^2}{\varphi},$$

where $C = C(n, k)$.

PROOF. By Corollary 3.18 we have that at (x_0, t_0),

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \leq K|A|^2_{\gamma_\rho}\frac{\varphi}{h} + \mu\frac{\varphi}{h}g^{ij}\ddot{\gamma}^{pq,rs}_\rho\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}^{ij}_\rho\nabla_i\varphi\nabla_j h + C\rho\frac{h}{H}|\nabla_1\varphi|^2,$$

where $C = C(n, k)$. Since $\lambda(x_0, t_0) \in \Gamma'$, by Lemma 3.12 we can estimate

$$h^{-1}|A|^2_{\gamma_\rho} \leq C(n, k)\mu^{-1}H^{-1}|A|^2 \leq C(n, k)|A|,$$

so at (x_0, t_0),

$$K|A|^2_{\gamma_\rho}\frac{\varphi}{h} \leq C(n, k)K|A|\varphi.$$

Next, by Lemma 3.12, we know that

$$g^{ij}\ddot{\gamma}^{pq,rs}_\rho\nabla_i A_{pq}\nabla_j A_{rs} \leq \rho g^{ij}\ddot{\gamma}^{pq,rs}_1\nabla_i A_{pq}\nabla_j A_{rs}.$$

Therefore, since γ_1 is strictly concave in off-radial directions, we can invoke Lemma 2.5 to bound

$$g^{ij}\ddot{\gamma}^{pq,rs}_\rho\nabla_i A_{pq}\nabla_j A_{rs} \leq -c_0(n, k)\rho\frac{|\nabla A|^2}{H}.$$

This estimate also relies on the fact that $\Gamma' \Subset \Gamma \setminus \mathrm{Cyl}_{n-1}$, which holds since Cyl_{n-1} lies outside of the k-positive cone for each $k \geq n - 1$.

Since $\varphi(x_0, t_0) > 0$ and $\dot{\gamma}^i_\rho(\lambda(x_0, t_0)) \leq C(n, k)$, at the point (x_0, t_0) we can use Young's inequality to estimate

$$\frac{2}{h}\dot{\gamma}^{ij}_\rho\nabla_i\varphi\nabla_j h \leq s^{-1}C(n, k)\frac{H}{h}\frac{|\nabla\varphi|^2}{\varphi} + sC(n, k)\varphi\frac{|\nabla h|^2}{hH},$$

where s can be any positive number. Since at (x_0, t_0) we have

$$|\nabla h|^2 \leq 2|\nabla G_\rho|^2 + 2\mu^2|\nabla A|^2 \leq C(n, k)|\nabla A|^2,$$

this leads to

$$\frac{2}{h}\dot{\gamma}^{ij}_\rho\nabla_i\varphi\nabla_j h \leq s^{-1}C_0(n, k)\frac{H}{h}\frac{|\nabla\varphi|^2}{\varphi} + sC_0(n, k)\varphi\frac{|\nabla A|^2}{hH}.$$

Setting $s = \frac{c_0\mu}{2C_0}\rho$ and putting all of this together, we get

$$(\partial_t - \Delta_{\gamma_\rho})\varphi \leq CK|A|\varphi - \frac{c_0\mu}{2}\rho\varphi\frac{|\nabla A|^2}{hH} + C\rho^{-1}\frac{H}{h}\frac{|\nabla\varphi|^2}{\varphi} + C\rho\frac{h}{H}|\nabla_1\varphi|^2.$$

Now, using Lemma 3.9 to estimate

$$\frac{H}{h} \le \mu^{-1}, \qquad \frac{h}{H} \le \frac{G_\rho}{H} + \frac{K}{H} \le 1 + \frac{K}{H},$$

and using

$$\varphi(x_0, t_0) \le f_\eta(x_0, t_0) \le k\mu^{-1},$$

we find that at (x_0, t_0),

$$C\rho^{-1}\frac{H}{h}\frac{|\nabla\varphi|^2}{\varphi} + C\rho\frac{h}{H}|\nabla_1\varphi|^2 \le C(n,k)(\rho^{-1} + KH^{-1})\frac{|\nabla\varphi|^2}{\varphi}.$$

□

Next we need to translate the viscosity inequality for f_η into an integral inequality. Following [**Bre15**], we verify that f_η is a locally semiconvex function. By this we mean that around any point in spacetime there is a small neighbourhood where f_η can be written as the sum of a smooth function and a convex function. It suffices to show that:

LEMMA 3.20. *Let $F : M \times [0, T) \to \mathbb{R}^{n+1}$ be a smooth family of smooth immersions. Then λ_1 is locally semiconcave on $M \times (0, T)$.*

PROOF. Fix a point $(x_0, t_0) \in M \times (0, T)$. It suffices to show that λ_1 is the sum of a smooth and a concave function on a small neighbourhood of the form

$$Q(x_0, t_0, r) = B_{g(t_0)}(x_0, r) \times [-r^2 + t_0, t_0 + r^2].$$

Observe that if r is sufficiently small, the function

$$\tilde{k}(x, t) := d_{g(t)}(x, x_0)^2$$

satisfies $\nabla_i\nabla_j\tilde{k} \ge g_{ij}$ on each spatial slice of $Q(x_0, t_0, r)$. Hence the function

$$k(x,t) := \frac{1}{2}t^2 + \tilde{k}(x,t)$$

is uniformly convex on $Q(x_0, t_0, r)$.

Making r a bit smaller if necessary, we can express λ_1 as the infimum over a family of smooth functions with uniformly bounded C^2-norms. For each unit vector v in $T_{x_0}M_{t_0}$, let $\tilde{X}_v$ be the vector field obtained by extending v by parallel transport on M_{t_0}, and then using ∂_t to extend the resulting vectorfield in time directions. For r small enough, $\tilde{X}_v$ has positive length on $Q(x_0, t_0, r)$, so we can set $X_v := \tilde{X}_v/X_v$, and then define

$$Z_v(x,t) := A(x,t)(X_v(x,t), X_v(x,t))$$

for each $(x, t) \in Q(x_0, t_0, r)$. For each $(x, t) \in Q(x_0, t_0, r)$ there holds

$$\lambda_1(x,t) := \inf\{Z_v(x,t) : v \in T_{x_0}M_{t_0},\ |v| = 1\}.$$

The C^2-norm of each of the functions Z_v over $Q(x_0, t_0, r)$ is bounded independently of v, so there is a constant $\Lambda > 0$ with the property that $Z_v - \Lambda k$ is concave on $Q(x_0, t_0, r)$ for every choice of v. Since the infimum of a family of concave functions is again concave,

$$\lambda_1(x,t) - \Lambda k(x,t) = \inf\{Z_v(x,t) - \Lambda k(x,t) : v \in T_{x_0}M_{t_0},\ |v| = 1\}$$

is concave. This completes the proof. □

In particular, Alexandrov's theorem (see Lemma A.1 in the appendix) now tells us that λ_1 (and hence f_η) is twice differentiable on a set of full measure in $M \times [0, T)$. Using this property we now verify that f_η satisfies the hypotheses of the Stampacchia procedure as stated in Theorem 2.6.

PROPOSITION 3.21. *For each $\rho \leq \rho_0$ there is a positive constant k_0 depending only on n, k, ρ, $\bar{\alpha}$, $R^{-n}\mu_0(M)$ and $R^{-2}T$ with the following properties. For each $(x, t) \in \mathrm{sp}(G_\rho - k_0 R^{-1})$ there holds $\lambda(x, t) \in \Gamma'$, and if φ is a nonnegative Lipschitz test function satisfying*

$$\mathrm{sp}(\varphi) \subset \mathrm{sp}(f_\eta) \cap \mathrm{sp}(G_\rho - k_0 R^{-1}),$$

then for almost every $t \in [0, T)$,

$$\begin{aligned}\int_M \varphi \partial_t f_\eta \, d\mu_t \leq &- \int_M \langle \nabla\varphi, \nabla f_\eta \rangle_{\gamma_\rho} \, d\mu_t - \int_M \varphi \ddot{\gamma}_\rho^{ij,pq} \nabla_i A_{pq} \nabla_j f_\eta \, d\mu_t \\ &- C^{-1} \int_M \varphi f_\eta \frac{|\nabla A|^2}{H^2} \, d\mu_t + C \int_M \varphi \frac{|\nabla f_\eta|^2}{f_\eta} \, d\mu_t + CK \int_M |A| \varphi \, d\mu_t\end{aligned}$$

where $C = C(n, k, \rho)$.

PROOF. First recall that we defined

$$K = \frac{C_{\varepsilon_0}}{(1 + 100^{-1})\alpha_{k-1}^{(1)}},$$

where $\varepsilon_0 = 100^{-1}\alpha_{k-1}^{(1)}$ and C_{ε_0} is the constant coming from the cylindrical estimate in Corollary 3.14. By the remark following Theorem 2.6, for each $\varepsilon > 0$, there is a constant $\tilde{C}_\varepsilon = \tilde{C}_\varepsilon(n, k, \rho, \bar{\alpha}, R^{-n}\mu_0(M), R^{-2}T)$ such that $C_\varepsilon = \tilde{C}_\varepsilon R^{-1}$. We may therefore write $K = \tilde{K} R^{-1}$, where $\tilde{K} = \tilde{K}(n, k, \rho, \bar{\alpha}, R^{-n}\mu_0(M), R^{-2}T)$. Recall that the statement

$$\lambda(x, t) \in \Gamma' = \Gamma_{\alpha_{k-1}^{(1)} + \varepsilon_0}$$

is equivalent to saying

$$G_1(x, t)^{-1} H(x, t) \leq \alpha_{k-1}^{(1)} + \varepsilon_0.$$

To ensure that this holds whenever $G_\rho(x, t) \geq k_0 R^{-1}$, by the cylindrical estimate

$$G_1(x, t)^{-1} H(x, t) \leq \alpha_{k-1}^{(1)} + \varepsilon_0/2 + \tilde{C}_{\varepsilon_0/2} R^{-1} G_\rho(x, t)^{-1},$$

it suffices to take $k_0 \geq 2\varepsilon_0^{-1} \tilde{C}_{\varepsilon_0/2}$.

By the last lemma f_η is a semiconvex function on $M \times [0, T)$, so by Alexandrov's theorem there is a set Q of full measure in $M \times [0, T)$ where f_η has two spatial derivatives and two time derivatives. At any point in Q, f_η admits an upper support φ, and at the point of contact there holds

$$(\partial_t - \Delta_{\gamma_\rho}) f_\eta \leq (\partial_t - \Delta_{\gamma_\rho}) \varphi.$$

We chose k_0 to ensure that $\lambda \in \Gamma'$ on $\mathrm{sp}(G_\rho - k_0 R^{-1})$, so by Lemma 3.19, inside the set $Q \cap \mathrm{sp}(f_\eta) \cap \mathrm{sp}(G_\rho - k_0 R^{-1})$ there holds

$$(\partial_t - \Delta_{\gamma_\rho}) f_\eta \leq CK|A| f_\eta - C^{-1} \rho f_\eta \frac{|\nabla A|^2}{hH} + C(\rho^{-1} + KH^{-1}) \frac{|\nabla f_\eta|^2}{f_\eta},$$

where $C = C(n,k)$. We now impose the further restriction $k_0 \geq \tilde{K}$, so that

$$KH(x,t)^{-1} \leq \tilde{K}R^{-1}G_\rho(x,t)^{-1} \leq 1$$

whenever $G(x,t) \geq k_0 R^{-1}$. This ensures that at on $Q \cap \mathrm{sp}(f_\eta) \cap \mathrm{sp}(G_\rho - k_0 R^{-1})$ there holds

$$(\partial_t - \Delta_{\gamma_\rho}) f_\eta \leq C_1 K|A| - C_1^{-1} f_\eta \frac{|\nabla A|^2}{H^2} + C_1 \frac{|\nabla f_\eta|^2}{f_\eta},$$

where $C_1 = C_1(n,k,\rho)$. Here we have also used the fact that f_η is bounded in terms of n and k.

If φ is a nonnegative Lipschitz function on $M \times [0,T)$ with the property

$$\mathrm{sp}(\varphi) \subset \mathrm{sp}(f_\eta) \cap \mathrm{sp}(G_\rho - k_0 R^{-1}),$$

we can multiply the last inequality by φ and integrate to get

$$\begin{aligned}\int_M \varphi \partial_t f_\eta \, d\mu_t \leq{}& \int_M \varphi \Delta_{\gamma_\rho} f_\eta \, d\mu_t - C_1^{-1} \int_M \varphi f_\eta \frac{|\nabla A|^2}{H^2} \, d\mu_t \\ &+ C_1 \int_M \varphi \frac{|\nabla f_\eta|^2}{f_\eta} \, d\mu_t + \int_M C_1 K |A| \varphi \, d\mu_t\end{aligned}$$

for almost every $t \in [0,T)$. To finish we appeal to Lemma A.3 from the appendix, which says that since f_η is semiconvex in space and φ is nonnegative,

$$\int_M \varphi \Delta_{\gamma_\rho} f_\eta \, d\mu_t \leq -\int_M \langle \nabla \varphi, \nabla f_\eta \rangle_{\gamma_\rho} \, d\mu_t - \int_M \varphi \ddot{\gamma}_\rho^{ij,pq} \nabla_i A_{pq} \nabla_j f_\eta \, d\mu_t.$$

□

All that is left now is to apply the Stampacchia theorem to f_η and verify that this gives the desired estimate.

Proof of Theorem 3.1. We have already observed that f_η is bounded:

$$f_\eta \leq \frac{\lambda_2 + \cdots + \lambda_k}{h} \leq k\mu^{-1}.$$

Also, if $(x,t) \in \mathrm{sp}(u)$ then by the definition of f_η there holds

$$\lambda(x,t) \in \{\lambda \in \Gamma : \min_{1 \leq i \leq n} \lambda_i \leq -\eta \gamma_\rho(\lambda)\},$$

so for k_0 as above,

$$\lambda(x,t) \in \Gamma' \cap \{\lambda \in \Gamma : \min_{1 \leq i \leq n} \lambda_i \leq -\eta \gamma_\rho(\lambda)\} =: \Gamma''$$

for every $(x,t) \in \mathrm{sp}(u) \cap \mathrm{sp}(G_\rho - k_0 R^{-1})$. Since $\mathrm{Cyl} \subset \Gamma_+$, we have $\Gamma'' \Subset \Gamma \setminus \mathrm{Cyl}$.

Combining these facts with Proposition 3.21, we see that for every $\rho \leq \rho_0$ and $\eta \in (0,1]$, the function f_η satisfies all of the hypotheses of Theorem 2.6. We conclude that, for each $\eta \in (0,1]$, there is a constant

$$C_\eta = C_\eta(n,k,\rho,\bar{\alpha},R,\mu_0(M),T,K)$$

such that

$$f_\eta(x,t) \leq \eta + C_\eta G_\rho(x,t)^{-1}$$

holds for every $(x,t) \in M \times [0,T)$. Since we chose K depending only on

$$n, k, \rho, \bar{\alpha}, R, \mu_0(M), T,$$

the constant C_η depends only on these quantities and η.

Using the definition of f_η we obtain

$$\begin{aligned}-\lambda_1(x,t) - \eta G_\rho(x,t) &\leq (\eta + C_\eta G_\rho(x,t)^{-1})h(x,t) \\ &\leq \eta G_\rho(x,t) + C_\eta + K\eta + KG_\rho(x,t)^{-1}.\end{aligned}$$

Since η can be made arbitrarily small, this gives an estimate of the desired form at all points where $G_\rho \geq 1$. On the other hand, at points where $G_\rho(x,t) \leq 1$ we have

$$\lambda_1(x,t) \geq -C(n,k,\rho,\bar{\alpha}),$$

so there is nothing to prove. □

5. Curved ambient spaces

In contrast to the Euclidean case, in a general Riemannian background, a family of hypersurfaces moving by mean curvature flow which is k-convex initially may fail to be k-convex after a finite amount of time. On the other hand, if the background geometry is bounded, then it is possible to construct fully nonlinear flows which do preserve k-convexity, by using a speed which vanishes at the boundary of the k-positive cone. The $k = 1$ and $k = 2$ cases were considered in [**And94b**] and [**BH17**], respectively. Each of the speeds γ_ρ gives rise to a flow preserving k-convexity in a Riemannian background, and if ρ is small relative to n and k, then compact solutions satisfy a convexity estimate. The argument is very similar to the Euclidean case, so we only sketch the details.

Consider a fixed $n \geq 4$ and $3 \leq k \leq n-1$. Let $(N, \bar{g})$ be a Riemannian manifold of dimension $n+1$, and suppose there is a constant C_0 such that

$$\sup_N |\bar{R}|_{\bar{g}} + |\bar{\nabla}\bar{R}|_{\bar{g}} \leq C_0,$$

where $\bar{\nabla}$ and $\bar{R}$ are the Levi-Civita connection and Riemann curvature tensor of the metric $\bar{g}$. Let $F : M \times [0,T) \to (N, \bar{g})$ be a solution of the equation

$$\partial_t F(x,t) = -\gamma_\rho(\lambda(x,t))\nu(x,t),$$

where M is a compact smooth n-manifold. Then, in an orthonormal basis, the Weingarten map satisfies

$$\partial_t A_q^p = \nabla_p \nabla_q G_\rho + A_{pr}A_{rq}G_\rho + \bar{R}(e_p, \nu, e_q, \nu)G_\rho,$$

where $G_\rho(x,t) := \gamma_\rho(\lambda(x,t))$. Taking the trace with $\dot{\gamma}_\rho^{pq}$ then gives

$$(\partial_t - \Delta_{\gamma_\rho})G_\rho = |A|^2_{\gamma_\rho} G_\rho + \dot{\gamma}_\rho^{pq}\bar{R}(e_p, \nu, e_q, \nu)G_\rho,$$

so since $\dot{\gamma}_\rho^{pq} \leq C_1(n,k,\rho)g^{pq}$, the maximum principle implies that

$$G_\rho(x,t) \geq \exp(-Ct) \cdot \min_M G_\rho(\cdot, 0),$$

where C depends only on C_0 and C_1. On the other hand, on the k-positive cone

$$\lambda_1 + \cdots + \lambda_k > \rho \cdot G_\rho,$$

so this estimate shows that the solution remains strictly k-convex for as long as it exists. If in addition the curvature tensor of N satisfies

$$\bar{R}(e_2, e_1, e_2, e_1) + \cdots + \bar{R}(e_{k+1}, e_1, e_{k+1}, e_1) > 0$$

for every set of orthonormal vectors $\{e_i\}_{i=1}^{k+1}$, then

$$\dot{\gamma}_\rho^{pq}\bar{R}(e_p,\nu,e_q,\nu) > 0,$$

so we even have

$$(\partial_t - \Delta_{\gamma_\rho})G_\rho \geq |A|^2_{\gamma_\rho}G_\rho \geq C(n,k,\rho)^{-1}G_\rho^3,$$

which is sufficient to conclude that G_ρ becomes unbounded in finite time. We may assume T is the maximal time.

Taking the trace of the evolution of the Weingarten map and using Simons' identity, one finds that the mean curvature of the hypersurface satisfies

$$(\partial_t - \Delta_{\gamma_\rho})H \leq |A|^2_{\gamma_\rho}H + g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + CH + C,$$

where C depends only on C_0 and C_1. Hence

$$(\partial_t - \Delta_{\gamma_\rho})\frac{H}{G_\rho} \leq \frac{1}{G_\rho}g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{G_\rho}\left\langle \nabla\left(\frac{H}{G_\rho}\right), \nabla G_\rho\right\rangle_{\gamma_\rho} + C\frac{H}{G_\rho},$$

where C may now depend additionally on the global spactime minimum of G_ρ. It follows that $G_\rho^{-1}H$ can grow at most exponentially, so there is a constant $\bar{\alpha}$ depending on T with the property that

$$\max_{M\times[0,T)}\frac{H}{G_\rho} \leq \bar{\alpha}.$$

With minor modifications, the proof of Lemma 2.5 allows us to prove that

$$g^{kl}\ddot{\gamma}_\rho^{pq,rs}\nabla_k A_{pq}\nabla_l A_{rs} \leq -c(n,k,\rho,\bar{\alpha})\frac{|\nabla A|^2}{H} + \frac{C}{H}.$$

The constant C comes from the curvature term in the Codazzi equations, and so depends only on C_0. The proof of Huisken's Stampacchia principle also goes through with minor modifications (see for example the proof of Theorem 3.1 in [**BH17**]), so with the good gradient term in hand, we obtain a cylindrical estimate:

$$H(x,t) \leq (\alpha_{k-1}^{(\rho)} + \varepsilon)G_\rho(x,t) + C_\varepsilon,$$

where C_ε has all of the same dependencies as in the Euclidean case, and now depends additionally on C_0. As before, if $G_\rho(x,t)$ is above a certain threshold depending only on n, k, ρ, $\bar{\alpha}$, M_0 and N, then

$$\lambda(x,t) \in \Gamma_{\alpha_{k-1}^{(1)}+\varepsilon_0}, \qquad \varepsilon_0 := 100^{-1}\alpha_{k-1}^{(1)}.$$

The second fundamental form satisfies

$$(\partial_t - \Delta_{\gamma_\rho})A_{ij} \geq |A|^2_{\gamma_\rho}A_{ij} - 2g^{kl}A_{ik}A_{lj}G_\rho + \ddot{\gamma}^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} - C|A| - C,$$

where C depends only on C_0 and C_1. On a set Q of full measure in $M\times[0,T)$, with respect to a principle frame, we have

$$(\partial_t - \Delta_{\gamma_\rho})\lambda_1 \geq |A|^2_{\gamma_\rho}\lambda_1 + \ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + 2\dot{\gamma}_\rho^k\sum_{\lambda_p>\lambda_1}\frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1} - C|A| - C.$$

We now form the same pinching function as before,

$$f_\eta := \frac{-\lambda_1 - \eta G_\rho}{h},$$

where $h = G_\rho - \mu H + K$ and

$$\mu := \frac{1}{2(1+100^{-1})\alpha_{k-1}^{(1)}}, \qquad K := \frac{C_{\varepsilon_0}}{(1+100^{-1})\alpha_{k-1}^{(1)}}.$$

On $Q \cap \mathrm{sp}(f_\eta)$ there holds

$$(\partial_t - \Delta_{\gamma_\rho}) f_\eta \le K|A|^2_{\gamma_\rho}\frac{f_\eta}{h} - \frac{1}{h}\Bigg(\ddot{\gamma}_\rho^{pq,rs}\nabla_1 A_{pq}\nabla_1 A_{rs} + 2\dot{\gamma}_\rho^k \sum_{\lambda_p > \lambda_1}\frac{|\nabla_k A_{p1}|^2}{\lambda_p - \lambda_1}\Bigg)$$
$$+ \mu\frac{f_\eta}{h} g^{ij}\ddot{\gamma}_\rho^{pq,rs}\nabla_i A_{pq}\nabla_j A_{rs} + \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i f_\eta \nabla_j h + C\frac{|A|}{h} + \frac{C}{h}.$$

As before, it is possible to choose $\rho_0 = \rho_0(n,k)$ so small that, at points in Q where

$$\lambda(x,t) \in \Gamma_{\alpha_{k-1}^{(1)}+\varepsilon_0}, \qquad f_\eta(x_0,t_0) > 0,$$

there holds

$$(\partial_t - \Delta_{\gamma_\rho}) f_\eta \le K|A|^2_{\gamma_\rho}\frac{f_\eta}{h} - \frac{1}{C} f_\eta \frac{|\nabla A|^2}{hH} + C(1+H^{-1})\frac{|\nabla f_\eta|^2}{f_\eta}$$
$$+ \frac{2}{h}\dot{\gamma}_\rho^{ij}\nabla_i f_\eta \nabla_j h + C\frac{f_\eta}{hH} + C\frac{|A|}{H} + \frac{C}{H}.$$

The constant C depends only on n, k, ρ, $\bar{\alpha}$, M_0 and C_0. This inequality has exactly the same form as in the Euclidean case, except for the presence of the the last three terms, which are of lower order and can all be absorbed in the Stampacchia argument. Hence, as before, we get

$$\lambda_1(x,t) \ge -\varepsilon H(x,t) - C_\varepsilon,$$

where C_ε depends only on ε, n, k, $\bar{\alpha}$, ρ, M_0 and the geometry of N via C_0.

CHAPTER 4

Harnack and gradient estimates

In this chapter we establish some regularity results for families of compact embedded hypersurfaces moving by a concave admissible speed. In addition to concavity, we assume the speed is inverse-concave on the positive cone and such that solutions satisfy a convexity estimate (see Definition 4.9 for the exact class of solutions we consider). In particular, the results here apply to the speeds γ_ρ considered in the previous chapter provided $\rho \leq \rho_0$, to the two-harmonic mean, and also to mean curvature flow (although in this last case, we mostly recover existing results). By combining the interior noncollapsing estimate from [**ALM13**] with the convexity estimate and a generalisation of the induction on scales argument in [**BH17**][Theorem 6.2], we are able to establish a global Harnack inequality for the curvature, which can be described as follows (see Theorem 4.17 for the precise statement). We show that for every large Λ there is a curvature threshold C depending on Λ and the initial data such that if the value of G at some point (x_0, t_0) exceeds C, then G is controlled from above and below over a backward parabolic neighbourhood of size $\Lambda G(x_0, t_0)^{-1}$ about (x_0, t_0).

The Harnack inequality implies a pointwise gradient estimate, which says that at points where the curvature is sufficiently large relative to the initial data, the estimate

$$|\nabla A|^2 \leq CG^4$$

holds for some universal C. Estimates of this kind first appeared in Huisken's work on convex solutions of mean curvature flow [**Hui84**] and in the work of Huisken-Sinestrari in [**HS09**] on immersed two-convex solutions of mean curvature flow of dimension $n \geq 3$. In the two-convex case, the proof by Huisken-Sinestrari uses a cylindrical estimate, but otherwise consists only of an application of the maximum principle applied to $|\nabla A|^2$. We note that this kind of gradient estimate cannot hold for mean-convex solutions of dimension $n = 2$ - indeed, the gradient estimate fails on the grim reaper, and it is not difficult to construct solutions which form a singularity modeled on the product of a grim reaper with $\mathbb{R}$ (take a thin torus over a solution of curve-shortening flow with self intersections as in [**Ang91**]). The gradient estimate proven in [**HS09**] holds as long as the solution is k-convex with $3k < 2n + 1$ (see Theorem 5.4 in [**HS15**]).

In the fully nonlinear case it is not clear clear whether a pointiwise gradient estimate can be obtained via the maximum principle. One difficulty seems to be that, whereas for the mean curvature flow the worst reaction term in the evolution equation for $|\nabla A|^2$ is linear in $|\nabla A|^2$, for nonlinear speeds there is an additional term which is quadratic in $|\nabla A|^2$. A different approach is taken by Brendle and Huisken in [**BH17**], who prove a pointwise gradient estimate for embedded solutions moving by the two-harmonic mean curvature. The key step is an induction on scales argument (inspired in part by Section 12 of [**Per02**]) which is combined with the cylindrical estimate to show that, locally about any point where the curvature is sufficiently large, the solution is a radial graph over an interior sphere of controlled size. The existence of the interior sphere is guaranteed by the interior noncollapsing estimate. The radial graph representation then lets one express the solution locally as a scalar solution of a fully nonlinear parabolic PDE, and the desired gradient estimate follows from the regularity theory of Evans and Krylov.

A pointwise gradient estimate of the kind we have been discussing can be integrated along geodesics to obtain a local Harnack inequality for the curvature. This works on an intrinsic spacetime neighbourhood, but the size of this neighbourhood is bounded from above on the scale of the curvature. Using this fact it is possible to establish, for example, that blow-up sequences converge locally in spacetime. To get global convergence, it is necessary to establish curvature bounds which become valid over an arbitrarily large spacetime set about any sequence where the curvature is blowing up. This kind of result is what we refer to as a global Harnack inequality.

We note that differential Harnack inequalities have been established for strictly convex solutions of mean curvature flow [**Ham95**], and flows by convex and inverse-concave admissible speeds [**And94c**]. These inequalities can be integrated to get curvature bounds over arbitrarily large spacetime regions. The analysis carried out in [**HS09**] to perform surgery also implies curvature bounds at bounded distances near a singularity. For embedded mean-convex solutions of mean curvature flow, a global Harnack inequality has been established by Haslhofer-Kleiner in [**HK17a**][Corollary 3.8]. These authors also make use of ideas from Section 12 of [**Per02**]. The method employed by Haslhofer-Kleiner makes use of Huisken's monotonicity formula [**Hui90**] via White's ε-regularity theorem [**Whi05**], and exterior noncollapsing estimates, neither of which is available for flows by a concave nonlinear speed function.

The structure of the chapter is as follows. We first look at solutions which can be written as a radial graph over a sphere in some spacetime neighbourhood, and use the maximum principle to show that on a smaller neighbourhood, any such solution satisfies an a priori upper bound for the curvature (this is Theorem 4.3). This bound depends on how steep the graph is, and the size of the spacetime neighbourhood on which the solution is graphical. This result is very similar to a theorem of Ecker and Huisken for solutions of mean curvature flow which can locally be written as a graph over a hyperplane [**EH91**][Theorem 3.1]. Their result was adapted to the case of radial graphs moving by nonlinear curvature functions in [**BH17**][Proposition 5.1]. We apply the maximum principle to a slightly different quantity to the one in [**BH17**], and in doing so prove an estimate which is stronger on very large spacetime sets. We then recall the notion of a pseudocone from [**BH17**], and establish some technical results needed to prove Theorem 4.11, which is the most difficult step in establishing the Harnack inequality. This result contains the induction on scales argument adapted from [**BH17**], which is used to show that around any point where the curvature is sufficiently large, a connected component of the solution is locally a radial graph. It is at this step that we use the inverse-concavity of the speed, which implies that the smallest principal curvature satisfies a strong maximum principle (see Corollary 3.8).

By Theorem 4.3, the local graph property gives a local scaling-invariant upper bound for the curvature in a backward neighbourhood of any point where the curvature is sufficiently large. With some further argumentation, we use this upper bound to derive an analogous lower bound in Theorem 4.15. Following this we prove the main result of the chapter, the global Harnack inequality, in Theorem 4.17. With the Harnack inequality in hand, scaling-invariant estimates for all of the derivatives of A quickly follow.

The results proven here can be used to show that, for the class of flows under consideration, every blow-up sequence subconverges smoothly to a complete convex ancient solution. Identifying an appropriate notion of convergence and establishing the necessary compactness theorems are interesting problems in their own right, which we will address elsewhere.

1. Curvature bounds for radial graphs

We first state some evolution equations which will be used just below. In this section, $\gamma : \Gamma \to (0, \infty)$ can be any admissible speed - no concavity or convexity properties are required.

LEMMA 4.1. *Let $F : M \times [0, T) \to \mathbb{R}^{n+1}$ be a solution of* (CF)*, where $G(x, t) := \gamma(\lambda(x, t))$ and γ is an admissible speed. Suppose $|F| > 0$ on $M \times [0, T)$ and set*

$$f := \frac{(F \cdot \nu)^2}{|F|^2}.$$

Then the following evolution equations hold:

$$(\partial_t - \Delta_\gamma)|F|^2 = -2\dot{\gamma}^{ij} g_{ij};$$

$$(\partial_t - \Delta_\gamma)F \cdot \nu = |A|_\gamma^2 F \cdot \nu - 2G;$$

$$\begin{aligned}(\partial_t - \Delta_\gamma)f = {} & 2(|A|_\gamma^2 - 2|F|^{-1}G)f - \frac{|F|^2}{2(F \cdot \nu)^2}|\nabla f|_\gamma^2 + \frac{1}{|F|^2}\langle \nabla f, \nabla |F|^2 \rangle_\gamma \\ & + 2\frac{(F \cdot \nu)^2}{|F|^4}\left(\dot{\gamma}^{ij} g_{ij} - \frac{|\nabla |F|^2|_\gamma^2}{4|F|^2}\right).\end{aligned}$$

PROOF. For the first equation, we compute in normal coordinates

$$\begin{aligned}(\partial_t - \Delta_\gamma)|F|^2 &= 2F \cdot \partial_t F - 2F \cdot \dot{\gamma}^{ij}\frac{\partial^2 F}{\partial x^i \partial x^j} - 2\dot{\gamma}^{ij}\frac{\partial F}{\partial x^i} \cdot \frac{\partial F}{\partial x^j} \\ &= -2GF \cdot \nu + 2\dot{\gamma}^{ij} A_{ij} F \cdot \nu - 2\dot{\gamma}^{ij} g_{ij},\end{aligned}$$

and since γ is a one-homogeneous function, we have the result.

It is well known that since the vectorfield $F + 2tG\nu$ generates a parabolic rescaling of the solution, its normal component satisfies

$$(\partial_t - \Delta_\gamma)(F + 2tG\nu) \cdot \nu = |A|_\gamma^2 (F + 2tG\nu) \cdot \nu.$$

A detailed derivation of this fact can be found, for example, in [**Lan14**][Lemma 3.4]. Inserting the evolution of the speed, we obtain

$$(\partial_t - \Delta_\gamma)F \cdot \nu = |A|_\gamma^2 F \cdot \nu - 2G.$$

From this, we readily obtain

$$\begin{aligned}(\partial_t - \Delta_\gamma)(F \cdot \nu)^2 &= 2|A|_\gamma^2 (F \cdot \nu)^2 - 2|\nabla (F \cdot \nu)|_\gamma^2 - 4GF \cdot \nu \\ &= 2|A|_\gamma^2 (F \cdot \nu)^2 - \frac{1}{2(F \cdot \nu)^2}|\nabla (F \cdot \nu)^2|_\gamma^2 - 4GF \cdot \nu.\end{aligned}$$

Applying the parabolic operator to a quotient gives

$$(\partial_t - \Delta_\gamma)\frac{u}{v} = \frac{1}{v}(\partial_t - \Delta_\gamma)u - \frac{u}{v^2}(\partial_t - \Delta_\gamma)v + \frac{2}{v}\left\langle \nabla\left(\frac{u}{v}\right), \nabla v\right\rangle_\gamma,$$

so for

$$f := \frac{(F\cdot\nu)^2}{|F|^2},$$

we have

$$\begin{aligned}(\partial_t - \Delta_\gamma)f &= \frac{1}{|F|^2}(\partial_t - \Delta_\gamma)(F\cdot\nu)^2 - \frac{(F\cdot\nu)^2}{|F|^4}(\partial_t - \Delta_\gamma)|F|^2 \\ &\quad + \frac{2}{|F|^2}\langle \nabla f, \nabla |F|^2\rangle_\gamma \\ &= 2|A|^2_\gamma f - \frac{1}{2|F|^2(F\cdot\nu)^2}|\nabla (F\cdot\nu)^2|^2_\gamma - 4G\frac{(F\cdot\nu)}{|F|^2} \\ &\quad + 2\frac{(F\cdot\nu)^2}{|F|^4}\dot\gamma^{ij}g_{ij} + \frac{2}{|F|^2}\langle \nabla f, \nabla|F|^2\rangle_\gamma.\end{aligned}$$

We may rewrite the first of the gradient terms as follows:

$$\begin{aligned}-&\frac{1}{2|F|^2(F\cdot\nu)^2}|\nabla(F\cdot\nu)^2|^2_\gamma \\ &= -\frac{|F|^2}{2(F\cdot\nu)^2}\left|\frac{\nabla(F\cdot\nu)^2}{|F|^2}\right|^2_\gamma \\ &= -\frac{|F|^2}{2(F\cdot\nu)^2}\left|\nabla f + \frac{(F\cdot\nu)^2}{|F|^4}\nabla|F|^2\right|^2_\gamma \\ &= -\frac{|F|^2}{2(F\cdot\nu)^2}|\nabla f|^2_\gamma - \frac{1}{|F|^2}\langle\nabla f, \nabla|F|^2\rangle_\gamma - \frac{(F\cdot\nu)^2}{2|F|^6}|\nabla|F|^2|^2_\gamma.\end{aligned}$$

Substituting back in now gives the result. □

REMARK 4.2. *We observe that the final term in the evolution of f is always positive. Indeed, since $\nabla|F|^2$ equals $2F^\top$, where the $\top$ denotes projection onto the tangent space of M, we are done if $F^\top$ is zero. Otherwise, we can choose an orthonormal frame where $e_1 = |F^\top|^{-1}F^\top$ and observe that*

$$\dot\gamma^{ij}g_{ij} - \frac{|\nabla|F|^2|^2_\gamma}{4|F|^2} \geq \sum_{i=1}^n \dot\gamma_{ii} - \frac{1}{|F^\top|^2}\dot\gamma(F^\top, F^\top) = \sum_{i=2}^n \dot\gamma_{ii} > 0.$$

We do not make use of this observation here, but it may be useful in other contexts.

We now establish the curvature bound for local radial graph solutions. The only difference between our result and Proposition 5.1 in [**BH17**] lies in the choice of the function w - we work with the scaling-invariant quantity f defined above, as opposed to $F\cdot\nu$, which scales like distance. This modification gives rise to an extra reaction term in the evolution of ψ, but this term is only quadratic in ψ and can be combated in the same way as the other reaction terms - using the good cubic term which arises when we divide by w. By working with the modified quantity we gain an extra factor of L on the left-hand side of the final estimate, which makes it more powerful at large distances.

THEOREM 4.3. *Fix constants $\theta > 0$, $r > 0$, $K > 0$ and $L > 1$, and a point $p \in \mathbb{R}^{n+1}$. Let*

$$F : M \times [-K^2r^2, 0] \to \mathbb{R}^{n+1}$$

be a solution of (CF) *such that for each $t \in [-K^2r^2, 0]$ the hypersurface $F(M, t)$ is properly embedded and bounds a smooth domain Ω_t. Suppose there is a symmetric cone $\Gamma' \Subset \Gamma$ such that $\lambda(x, t) \in \Gamma'$ for every $(x, t) \in M \times [-K^2r^2, 0]$, and that $B(p, r) \subset \Omega_0$. For each $t \in [-K^2r^2, 0]$, let U_t denote the connected component of $\Omega_t \cap B(p, Lr)$ which contains $B(p, r)$, set $N_t := \partial U_t \cap B(p, Lr)$, and define*

$$\mathbf{N} := \{(x, t) \in B(p, Lr) \times [-K^2r^2, 0] : x \in N_t\}.$$

Finally, suppose that for each $(x, t) \in \mathbf{N}$ there holds

$$\nu(x, t) \cdot \frac{F(x, t) - p}{|F(x, t) - p|} \geq \theta.$$

Then there is a constant $C = C(n, \gamma, \Gamma')$ such that

$$\left(\frac{L^2r^2}{4} - |x|^2\right)(t + K^2r^2)^{\frac{1}{2}} G(x, t) \leq C \max\{1, K\} L^2\theta^{-2}r^2$$

for all $(x, t) \in \mathbf{N} \cap (B(0, Lr/2) \times [-K^2r^2, 0])$.

PROOF. It will be convenient to identify $M \times [-K^2r^2, 0]$ with the set

$$\mathbf{M} = \{(x, t) \in \mathbb{R}^{n+1} \times [-K^2r^2, 0] : x \in M_t\}$$

via the evolving immersion F. In particular, functions defined on the solution may equivalently be viewed as functions on $M \times [-K^2r^2, 0]$ or on $\mathbf{M}$. Let us shift the solution in space if necessary so that $p = 0$. We define functions

$$\eta(x, t) = \frac{L^2r^2}{4} - |F(x, t)|^2, \qquad w(x, t) = f(x, t) - \theta^2/2, \qquad v(x, t) = w(x, t)^{-\frac{1}{2}},$$

and set

$$\psi(x, t) = \eta(x, t)v(x, t)G(x, t)$$

for each $(x, t) \in \mathbf{N}$. Note that the assumption on the normal of F says exactly that $w \geq \theta^2/2$ on $\mathbf{N}$. Since η is negative at points which lie outside the ball $B(0, Lr/2)$, on each timeslice the support of ψ is compactly contained in N_t.

We compute at an arbitrary point in the support of ψ, writing C for a large constant which depends only on n, γ and Γ'. From Lemma 4.1, we know that

$$\begin{aligned}
(\partial_t - \Delta_\gamma)w &= 2(|A|^2_\gamma - 2|F|^{-1}G)f - \frac{1}{2f}|\nabla f|^2_\gamma + \frac{1}{|F|^2}\langle \nabla f, \nabla |F|^2\rangle_\gamma \\
&\quad + 2\frac{f}{|F|^2}\left(\dot\gamma^{ij}g_{ij} - \frac{|\nabla |F|^2|^2_\gamma}{4|F|^2}\right) \\
&\geq 2(|A|^2_\gamma - 2|F|^{-1}G)(w + \theta^2/2) - \frac{1}{2(w + \theta^2/2)}|\nabla w|^2_\gamma \\
&\quad + \frac{1}{|F|^2}\langle \nabla w, \nabla |F|^2\rangle_\gamma - (w + \theta^2/2)\frac{|\nabla |F|^2|^2_\gamma}{2|F|^4}
\end{aligned}$$

Let ε be a small positive constant to be chosen later, and estimate

$$\frac{1}{|F|^2}\langle \nabla w, \nabla |F|^2\rangle_\gamma \geq -\frac{\varepsilon}{2(w+\theta^2/2)}|\nabla w|_\gamma^2 - C\varepsilon^{-1}(w+\theta^2/2)\frac{|\nabla |F|^2|_\gamma^2}{|F|^4},$$

so that we obtain

$$\begin{aligned}(\partial_t - \Delta_\gamma)w \geq 2(|A|_\gamma^2 - 2|F|^{-1}G)(w+\theta^2/2) - \frac{1+\varepsilon}{2(w+\theta^2/2)}|\nabla w|_\gamma^2 \\ - C\varepsilon^{-1}(w+\theta^2/2)\frac{|\nabla |F|^2|_\gamma^2}{|F|^4}.\end{aligned}$$

Using this inequality, we get

$$\begin{aligned}(\partial_t - \Delta_\gamma)v &= -\frac{1}{2}w^{-\frac{3}{2}}(\partial_t - \Delta_\gamma)w - \frac{3}{4}w^{-\frac{5}{2}}|\nabla w|_\gamma^2 \\ &\leq -w^{-\frac{3}{2}}(|A|_\gamma^2 - 2|F|^{-1}G)(w+\theta^2/2) + \frac{1+\varepsilon}{4}\frac{w^{-\frac{3}{2}}}{w+\theta^2/2}|\nabla w|_\gamma^2 \\ &\quad + C\varepsilon^{-1}w^{-\frac{3}{2}}(w+\theta^2/2)\frac{|\nabla |F|^2|_\gamma^2}{|F|^4} - \frac{3}{4}w^{-\frac{5}{2}}|\nabla w|_\gamma^2.\end{aligned}$$

Rewriting ∇w in terms of ∇v gives

$$\frac{1+\varepsilon}{4}\frac{w^{-\frac{3}{2}}}{w+\theta^2/2}|\nabla w|_\gamma^2 = (1+\varepsilon)\frac{w^{-\frac{3}{2}}v^{-6}}{w+\theta^2/2}|\nabla v|_\gamma^2 = (1+\varepsilon)\frac{v^{-1}}{1+\theta^2 v^2/2}|\nabla v|_\gamma^2,$$

and

$$-\frac{3}{4}w^{-\frac{5}{2}}|\nabla w|_\gamma^2 = -3w^{\frac{1}{2}}|\nabla v|_\gamma^2 = -3v^{-1}|\nabla v|_\gamma^2,$$

so we have

$$\begin{aligned}(\partial_t - \Delta_\gamma)v \leq -(|A|_\gamma^2 - 2|F|^{-1}G)(v+\theta^2 v^3/2) - \left(3 - \frac{1+\varepsilon}{1+\theta^2 v^2/2}\right)v^{-1}|\nabla v|_\gamma^2 \\ + C\varepsilon^{-1}(v+\theta^2 v^3/2)\frac{|\nabla |F|^2|_\gamma^2}{|F|^4}.\end{aligned}$$

We will use the good terms on the right-hand side of the last inequality to control various other terms appearing in the evolution of η and G. The evolution of η is given by

$$(\partial_t - \Delta_\gamma)\eta = -(\partial_t - \Delta_\gamma)|x|^2 = 2\dot{\gamma}^{ij}g_{ij},$$

and combining this with the equation for v gives

$$\begin{aligned}(\partial_t - \Delta_\gamma)(\eta v) \leq -(|A|_\gamma^2 - 2|F|^{-1}G)(\eta v + \theta^2\eta v^3/2) - \left(3 - \frac{1+\varepsilon}{1+\theta^2 v^2/2}\right)\eta v^{-1}|\nabla v|_\gamma^2 \\ + C\varepsilon^{-1}(\eta v + \theta^2\eta v^3/2)\frac{|\nabla |F|^2|_\gamma^2}{|F|^4} + 2\dot{\gamma}^{ij}g_{ij}v - 2\langle\nabla\eta, \nabla v\rangle_\gamma.\end{aligned}$$

Incorporating now the evolution of G, we obtain

$$\begin{aligned}&(\partial_t-\Delta_\gamma)\psi\\&=-(|A|^2_\gamma-2|F|^{-1}G)(\eta vG+\theta^2\eta v^3G/2)-\left(3-\frac{1+\varepsilon}{1+\theta^2v^2/2}\right)\eta v^{-1}G|\nabla v|^2_\gamma\\&+C\varepsilon^{-1}(\eta vG+\theta^2\eta v^3G/2)\frac{|\nabla|F|^2|^2_\gamma}{|F|^4}+2\dot\gamma^{ij}g_{ij}vG-2G\langle\nabla\eta,\nabla v\rangle_\gamma\\&+|A|^2_\gamma\eta vG-2v\langle\nabla\eta,\nabla G\rangle_\gamma-2\eta\langle\nabla v,\nabla G\rangle_\gamma.\end{aligned}$$

We simplify

$$\begin{aligned}-(|A|^2_\gamma-2|F|^{-1}G)(\eta vG+\theta^2\eta v^3G/2)+|A|^2_\gamma\eta vG&\\=-(|A|^2_\gamma-2|F|^{-1}G)\theta^2\eta v^3G/2+2|F|^{-1}G^2\eta v&\end{aligned}$$

an cancel some of the gradient terms by expanding

$$\begin{aligned}-2v^{-1}\eta G|\nabla v|^2_G&=-2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma+2\langle\nabla v,\nabla(\eta G)\rangle_\gamma\\&=-2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma+2G\langle\nabla v,\nabla\eta\rangle_\gamma+2\eta\langle\nabla v,\nabla G\rangle_\gamma,\end{aligned}$$

and substituting back in:

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi&=-(|A|^2_\gamma-2|F|^{-1}G)\theta^2\eta v^3G/2+2|F|^{-1}G^2\eta v\\&-\left(1-\frac{1+\varepsilon}{1+\theta^2v^2/2}\right)\eta v^{-1}G|\nabla v|^2_\gamma\\&+C\varepsilon^{-1}(\eta vG+\theta^2\eta v^3G/2)\frac{|\nabla|F|^2|^2_\gamma}{|F|^4}+2\dot\gamma^{ij}g_{ij}vG\\&-2v\langle\nabla\eta,\nabla G\rangle_\gamma-2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

Now we write ∇G in terms of the gradients of ψ, v and η,

$$\begin{aligned}-2v\langle\nabla\eta,\nabla G\rangle_\gamma&=-2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma+2\eta^{-1}G\langle\nabla\eta,\nabla(\eta v)\rangle_\gamma\\&=-2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma+2G\langle\nabla\eta,\nabla v\rangle_\gamma+2\eta^{-1}vG|\nabla\eta|^2_\gamma\end{aligned}$$

and use

$$-\left(1-\frac{1+\varepsilon}{1+\theta^2v^2/2}\right)\eta v^{-1}G|\nabla v|^2_\gamma=-\frac{\theta^2v^2/2-\varepsilon}{1+\theta^2v^2/2}\eta v^{-1}G|\nabla v|^2_\gamma$$

to arrive at

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi&=-(|A|^2_\gamma-2|F|^{-1}G)\theta^2\eta v^3G/2+2|F|^{-1}G^2\eta v\\&-\frac{\theta^2v^2/2-\varepsilon}{1+\theta^2v^2/2}\eta v^{-1}G|\nabla v|^2_\gamma+C\varepsilon^{-1}(\eta vG+\theta^2\eta v^3G/2)\frac{|\nabla|F|^2|^2_\gamma}{|F|^4}\\&+2\dot\gamma^{ij}g_{ij}vG-2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma+2G\langle\nabla\eta,\nabla v\rangle_\gamma\\&+2\eta^{-1}vG|\nabla\eta|^2_\gamma-2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

Since $(F\cdot\nu)^2\le|F|^2$, we have

$$v^2=w^{-1}=(f-\theta^2/2)^{-1}\ge1.$$

Therefore, taking $\varepsilon=\theta^2/4$ ensures that

$$\theta^2v^2/2-\varepsilon\ge\theta^2v^2/4+\theta^2/4-\varepsilon\ge\theta^2v^2/4.$$

We can use Young's inequality to estimate,

$$2G\langle\nabla\eta,\nabla v\rangle_\gamma \le \frac{\theta^2v^2/4}{1+\theta^2v^2/2}\eta v^{-1}G|\nabla v|^2_\gamma + \frac{1+\theta^2v^2/2}{\theta^2v^2/4}\eta^{-1}vG|\nabla\eta|^2_\gamma,$$

and inserting the last two inequalities now gives

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi \le& -(|A|^2_\gamma - 2|F|^{-1}G)\theta^2\eta v^3G/2 + 2|F|^{-1}G^2\eta v\\ &+ C(\theta^{-2}\eta vG + \eta v^3G/2)\frac{|\nabla|F|^2|^2_\gamma}{|F|^4}\\ &+ 2\dot\gamma^{ij}g_{ij}vG + \frac{1+\theta^2v^2/2}{\theta^2v^2/4}\eta^{-1}vG|\nabla\eta|^2_\gamma\\ &+ 2\eta^{-1}vG|\nabla\eta|^2_\gamma - 2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma - 2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

The condition $\lambda(x,t)\in\Gamma'$ implies the bounds

$$G^2 \le C|A|^2_\gamma, \qquad \dot\gamma^{ij}g_{ij} \le C,$$

and $|F|\ge r$ by assumption, so by Young's inequality we have

$$|A|^2_\gamma - 2|F|^{-1}G \ge C^{-1}G^2 - 2r^{-1}G \ge C^{-1}G^2 - Cr^{-2}.$$

Subsituting in these bounds, and writing v in terms of ψ, we obtain

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi \le& -C^{-1}\theta^2\eta^{-2}\psi^3 + C\theta^2r^{-2}v^2\psi + 2|F|^{-1}\eta^{-1}v^{-1}\psi^2\\ &+ C(\theta^{-2}\psi + v^2\psi/2)\frac{|\nabla|F|^2|^2_\gamma}{|F|^4} + 2C\eta^{-1}\psi + \frac{1+\theta^2v^2/2}{\theta^2v^2/4}\eta^{-2}\psi|\nabla\eta|^2_\gamma\\ &+ 2\eta^{-2}\psi|\nabla\eta|^2_\gamma - 2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma - 2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

We have the bounds $1\le v^2\le 2\theta^{-2}$,

$$|\nabla\eta|^2_\gamma \le C|F|^2 \le CL^2r^2,$$

and

$$\frac{|\nabla|F|^2|^2_\gamma}{|F|^4} = 4\frac{|F^\top|^2}{|F|^2} \le 4|F|^{-2} \le Cr^{-2},$$

so

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi \le& -C^{-1}\theta^2\eta^{-2}\psi^3 + C\theta^{-2}r^{-2}\psi + 2r^{-1}\eta^{-1}\psi^2 + 2C\eta^{-1}\psi\\ &+ CL^2\theta^{-2}r^2\eta^{-2}\psi + CL^2r^2\eta^{-2}\psi - 2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma - 2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

Finally, using $\eta^2\le L^4r^4$, we arrive at

$$\begin{aligned}(\partial_t-\Delta_\gamma)\psi \le& -C^{-1}\theta^2\eta^{-2}\psi^3/2 + CL^4\theta^{-2}r^2\eta^{-2}\psi + CL^2r\eta^{-2}\psi^2\\ &- 2\eta^{-1}\langle\nabla\eta,\nabla\psi\rangle_\gamma - 2v^{-1}\langle\nabla v,\nabla\psi\rangle_\gamma.\end{aligned}$$

Let us define

$$\Theta := \sup\{(t+K^2r^2)^{\frac12}\psi(x,t) : (x,t)\in\mathbf{N}\},$$

let Λ be a large positive constant, and assume $\Theta>\Lambda$. We are going to choose Λ so that this gives a contradiction.

Define

$$\bar t := \inf\left\{t\in[-K^2r^2,0] : \sup_{N_t}(t+K^2r^2)^{\frac12}\psi(\cdot,t) > \Lambda\right\}.$$

Then there is a sequence $t_j > \bar{t}$ such that $t_j \to \bar{t}$ and

$$(t_j + K^2r^2)^{\frac{1}{2}}\psi(x_j, t_j) > \Lambda$$

for some $x_j \in N_{t_j} \cap B(0, Lr/2)$. Passing to a convergent subsequence, we may assume that the x_j converge to a point $\bar{x} \in M_{\bar{t}} \cap \overline{B}(0, Lr/2)$. To conclude that $\bar{x} \in N_{\bar{t}}$ we need to find a path from the origin to $\bar{x}$ inside $\overline{\Omega}_{\bar{t}} \cap B(0, Lr)$, but this is easily achieved. Indeed, for each j there is a continuous path $\alpha_j : [0, 1] \to \overline{\Omega}_{t_j} \cap B(0, Lr)$ such that $\alpha_j(0) = 0$ and $\alpha_j(1) = x_j$. Since $\bar{t} < t_j$ each of the paths α_j maps into $\Omega_{\bar{t}} \cap B(0, Lr)$, so by smoothness of Ω_t and the fact that $\bar{x} \in \overline{B}(0, Lr/2)$, if j is large enough the path α_j can be extended to a continuous a path from the origin to $\bar{x}$ which stays inside $\overline{\Omega}_{\bar{t}} \cap B(0, Lr)$.

Arguing again using smoothness, we conclude that there is a small $\delta > 0$ such that

$$B_{g(\bar{t})}(\bar{x}, \delta) \times [\bar{t}, \bar{t} + \delta^2]$$

is a subset of $\mathbf{N}$ and contains (x_j, t_j) for large j. Consequently,

$$(t_j + K^2r^2)^{\frac{1}{2}}\psi(x_j, t_j) \to (\bar{t} + K^2r^2)^{\frac{1}{2}}\psi(\bar{x}, \bar{t}),$$

hence by the definitions of the (x_j, t_j) and $(\bar{x}, \bar{t})$ it must be the case that

$$(\bar{t} + K^2r^2)^{\frac{1}{2}}\psi(\bar{x}, \bar{t}) = \Lambda.$$

In particular, since $\Lambda > 0$ and ψ vanishes on $\partial B(0, Lr/2)$ and at $t = 0$, we have $\bar{x} \in B(0, Lr/2)$ and $\bar{t} > 0$. Therefore, making $\delta > 0$ smaller if necessary, we can ensure that

$$Q(\delta) := B_{g(\bar{t})}(\bar{x}, \delta) \times [-\delta^2 + \bar{t}, \bar{t}]$$

is contained in $\mathbf{N}$.

The function

$$(x, t) \mapsto (t + K^2r^2)^{\frac{1}{2}}\psi(x, t)$$

restricted to the set $Q(\delta)$ attains its spacetime maximum at $(\bar{x}, \bar{t})$, so by the computation above

$$\begin{aligned}
0 &\leq \frac{1}{2}\frac{1}{(\bar{t} + K^2r^2)^{\frac{1}{2}}}\psi(\bar{x}, \bar{t}) + (\bar{t} + K^2r^2)^{\frac{1}{2}}\partial_t\psi(\bar{x}, \bar{t}) \\
&\leq \frac{1}{2}\frac{1}{(\bar{t} + K^2r^2)^{\frac{1}{2}}}\psi(\bar{x}, \bar{t}) - C^{-1}(\bar{t} + K^2r^2)^{\frac{1}{2}}\theta^2\eta(\bar{x}, \bar{t})^{-2}\psi(\bar{x}, \bar{t})^3 \\
&\quad + C(\bar{t} + K^2r^2)^{\frac{1}{2}}\theta^2\eta(\bar{x}, \bar{t})^{-2}\psi(\bar{x}, \bar{t})^2(L^2\theta^{-2}r + L^4\theta^{-4}r^2\psi(\bar{x}, \bar{t})^{-1}).
\end{aligned}$$

If $\psi(\bar{x}, \bar{t}) \geq 10CL^2\theta^{-2}r$, then rearranging the last inequality gives

$$C^{-1}\theta^2\eta(\bar{x}, \bar{t})^{-2}(\bar{t} + K^2r^2)\psi^3(\bar{x}, \bar{t}) \leq \psi(\bar{x}, \bar{t}),$$

so there holds

$$\Lambda = (\bar{t} + K^2r^2)\psi(\bar{x}, \bar{t})^2 \leq C\theta^{-2}\eta^2 \leq CL^4\theta^{-2}r^4.$$

If on the other hand $\psi(\bar{x}, \bar{t}) \leq 10C^2L^2\theta^{-2}r$, then

$$\Lambda = (\bar{t} + K^2r^2)^{\frac{1}{2}}\psi(\bar{x}, \bar{t}) \leq CKL^2\theta^{-2}r^2.$$

Therefore, we can force a contradiction by choosing

$$\Lambda = C\max\{1, K\}L^2\theta^{-2}r^2$$

for some sufficiently large C, in which case our original assumption $\Theta > \Lambda$ must have been false. Combining $\Theta \leq \Lambda$ with $v \geq 1$ we conclude that

$$\begin{aligned}\eta(x,t)(t+K^2r^2)^{\frac{1}{2}}G(x,t) &\leq C\max\{1,K\}L^2\theta^{-2}r^2v(x,t)^{-1}\\ &\leq C\max\{1,K\}L^2\theta^{-2}r^2\end{aligned}$$

for each $(x,t) \in \mathbf{N}$. □

2. Pseudocones

Following Brendle-Huisken [**BH17**][Section 6], we introduce the notion of a pseudocone. This is the name we give to a piece of cone whose boundary has been bowed outward slightly, so that parallel to its axis of rotation, the boundary has a small but definite amount of negative curvature. We define a pseudocone $C(x,p,r)$ for each pair of points x and p in $\mathbb{R}^{n+1}$, and each positive r, as follows: set

$$\varphi(s) = \frac{1}{2}(s+s^2), \qquad s \in [0,1],$$

and define

$$C(x,p,r) = \{(1-s)x + sp + \tau v : s \in (0,1),\, 0 < \tau < r\varphi(s),\, \langle v, p-x\rangle = 0\}.$$

The point x is the vertex, and p is the center of the base, which is a solid n-ball of radius r. This ball sits in the hyperplane which is orthogonal to $p - x$ and passes through p. It will also be convenient to give a name to the smooth, outwardly curved part of the boundary of $C(x,p,r)$, so we set

$$S(x,p,r) = \{(1-s)x + sp + \tau v : s \in (0,1),\, \tau = r\varphi(s),\, \langle v, p-x\rangle = 0\}.$$

Observe that

$$\partial C(x,p,r) \subset \{x\} \cup S(x,p,r) \cup \bar{B}(x,r),$$

and that near the point x, the region $C(x,p,r)$ is asymptotic to a cone of aperture

$$2\tan\left(\frac{r}{2|p-x|}\right).$$

LEMMA 4.4. *Suppose $|p - x| > r$. Then the smallest principal curvature of $S(x,p,r)$ is at most*

$$-\frac{1}{10}\frac{r}{|p-x|^2}.$$

PROOF. We can apply a rigid motion taking $S(x,p,r)$ to the hypersurface $S(0,de_1,r)$, where $d := |x-p|$. Since $S(de_1,0,r)$ is given by rotating the graph of the function

$$f(s) := r\varphi(d^{-1}s), \qquad s \in (0,d)$$

about the e_1 axis, at each point of its boundary, there are $n-1$ positive principal curvatures, and the remaining principal curvature is equal to the curvature of the graph of f. For the latter we have the estimate

$$\begin{aligned}\frac{-f''(s)}{(1+|f'(s)|^2)^{\frac{3}{2}}} &= \frac{-rd^{-2}\varphi''(d^{-1}s)}{(1+r^2d^{-2}|\varphi'(d^{-1}s)|^2)^{\frac{3}{2}}}\\ &= \frac{-rd^{-2}}{(1+r^2d^{-2}(1/2+d^{-1}s)^2)^{\frac{3}{2}}}\\ &\leq -(4/13)^{\frac{3}{2}}rd^{-2}.\end{aligned}$$

Here we have used $d^{-1}s \leq 1$ and $d^{-1}r \leq 1$. Since $(4/13)^{\frac{3}{2}} \geq 1/10$, this completes the proof. □

Consider positive constants $r > 0$ and $\Lambda > 1$, and let Ω be an open subset of $B(p, \Lambda r)$ which contains $B(p, r)$. It is clear that if for each $x \in \Omega$, the pseudocone $C(x, p, r)$ is contained in Ω, then Ω is starshaped about the point x. If Ω is also smooth, the smooth part of if its boundary has a globally defined outward-pointing unit normal vectorfield ν, and we have the following estimate:

Lemma 4.5. *Let Ω be a smooth open subset of $B(p, \Lambda r)$ with the property that $C(x, p, r) \subset \Omega$ for each $x \in \Omega$. Then for each $y \in \partial\Omega \cap B(x, \Lambda r)$ there holds*

$$\nu(y) \cdot \frac{y-p}{|y-p|} \geq \frac{1}{\sqrt{5}}\frac{1}{\Lambda},$$

where ν is the outward-pointing unit normal to $\partial\Omega \cap B(p, \Lambda r)$.

Proof. Fix $y \in \partial\Omega \cap B(p, \Lambda r)$. Approximating y by a sequence of points in $\Omega \cap B(p, \Lambda r)$, we find that

$$C(y, x, r) \subset \bar{\Omega}.$$

The aperture of $C(y, x, r)$ is at least

$$\phi := 2\tan\left(\frac{1}{2\Lambda}\right),$$

and the hyperplane tangent to $\partial\Omega$ at y lies outside $C(y, p, r)$, so we have

$$\nu(y) \cdot \frac{(y-p)}{|y-p|} \geq \cos\left(\frac{\pi-\phi}{2}\right) = \frac{r/2}{\sqrt{r^2/4+\Lambda^2r^2}}.$$

Inserting $\Lambda > 1$ now gives the result. □

The following is another technical result which we make use of below. It gives us a way of characterising whether or not a smooth domain is starshaped in terms of the curvature of its boundary.

Lemma 4.6. *Let Ω be a smooth, connected open subset of $B(p, \Lambda r)$ and suppose that $B(p, r)$ is contained in Ω. Suppose also that, for some $y \in \Omega$, the pseudocone $C(y, p, r)$ is not contained in Ω. Then there exists a point $x \in \Omega \cap B(p, \Lambda r)$ such that $C(x, p, r) \subset \Omega$, but the hypersurface $S(x, p, r)$ makes interior contact with $\partial\Omega$ at some $\tilde{x} \in B(p, \Lambda r)$. In particular, if λ_1 denotes the smallest principal curvature of the hypersurface $\partial\Omega \cap B(p, \Lambda r)$, then there holds*

$$\lambda_1(\tilde{x}) \leq -\frac{1}{10\Lambda^2}r^{-1}.$$

PROOF. Let us define

$$B := \{x \in \Omega : C(x, p, r) \subset \Omega\}.$$

By assumption, the set $\Omega \setminus B$ is nonempty. Therefore, since B is relatively closed in Ω, and Ω is connected, we conclude that B cannot be relatively open in Ω. That is, there is a point $x \in B$ and a sequence $x_i \in \Omega \setminus B$ such that $x_i \to x$. If the boundary of $C(x_i, p, r)$ is in $\bar{\Omega}$, then $C(x_i, p, r) \subset \Omega$, so there is a sequence $\tilde{x}_i \in \partial C(x_i, p, r)$ which is such that $\tilde{x}_i \in B(p, \Lambda r) \setminus \overline{\Omega}$. Passing to a subsequence, we may assume the $\tilde{x}_i$ converge to a limit $\tilde{x} \in \partial\Omega \cap B(p, \Lambda r)$.

We recall that

$$\partial C(x_i, p, r) \subset \{x_i\} \cup S(x_i, p, r) \cup \bar{B}(p, r),$$

and by assumption $x_i \in \Omega$ and $\bar{B}(p, r) \subset \bar{\Omega}$. It follows then that $\tilde{x}_i$ remains at a uniformly positive distance from both the vertex and base of $C(x_i, p, r)$ as $i \to \infty$, and in the limit we get

$$\tilde{x} \in S(x, p, r) \cap \partial\Omega.$$

The upper bound for $\lambda_1(\tilde{x})$ follows from Lemma 4.4. $\square$

3. A class of solutions

Let $\gamma : \Gamma \to (0, \infty)$ be an admissible speed and consider an evolving embedding of a compact manifold M,

$$F : M \times [0, T) \to \mathbb{R}^{n+1},$$

which satisfies (CF), where the normal velocity is $G(x, t) := \gamma(\lambda(x, t))$. We write Ω_t for the smooth open domain bounded by $M_t := F(M, t)$, and $\mathbf{\Omega}$ for the subset of $\mathbb{R}^{n+1} \times [0, T)$ given by

$$\mathbf{\Omega} := \{\Omega_t : t \in [0, T)\}.$$

Similarly, we set

$$\mathbf{M} := \{M_t : t \in [0, T)\}.$$

Where there is no chance of confusion we forget about the embedding F and refer to $\mathbf{\Omega}$ and $\mathbf{M}$ as solutions of (CF). We may view objects such as the principal curvatures as being defined on $M \times [0, T)$ or on $\mathbf{M}$.

The results in the rest of this chapter all apply to embedded solutions which are κ-noncollapsed, in the following sense.

DEFINITION 4.7. *Fix an admissible speed γ and let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be a solution of* (CF)*. Given $t \in [0, T)$, we say that Ω_t is κ-noncollapsed if for each $x \in M_t$, there is a ball of radius $\kappa G(x, t)^{-1}$ inside Ω_t which is tangent to M_t at x. To be precise, if we set $x' := x - \kappa G(x, t)^{-1}\nu(x, t)$, then*

$$B(x', \kappa G(x, t)^{-1}) \subset \Omega_t.$$

The evolving embedding F, or the family of domains $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$, is said to be κ-noncollapsed if this is true for each $t \in [0, T)$.

By work of Andrews in the mean curvature flow case [**And12**], or Andrews-Langford-McCoy for flows by a concave admissible speed function [**ALM13**], a solution $\mathbf{\Omega}$ is κ-noncollapsed as long as this is true of the initial domain Ω_0. Hence every embedded compact solution moving by a concave admissible speed is κ-noncollaped for some $\kappa > 0$ depending only on Ω_0.

In addition to noncollapsing, the results we prove here all assume that high-curvature regions are becoming convex. We capture this by assuming the solution is φ-almost-convex, in the following sense. Note that φ-almost-convexity follows if a convexity estimate of the form established in Chapter 3 holds. In this case the function φ depends on the speed γ and the solution at the initial time.

Definition 4.8. *Fix an admissible speed γ and let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be a smooth solution of* (CF)*. Let $\varphi : [0, \infty) \to (0, \infty)$ be a non-increasing function such that*

$$\lim_{x \to \infty} \frac{\varphi(x)}{x} = 0.$$

We say that $\mathbf{\Omega}$ is φ-almost-convex if for each $(x, t) \in \mathbf{M}$ there holds

$$\lambda_1(x, t) \geq -\varphi(G(x, t)).$$

Since we will repeatedly want to refer back to the same class of solutions with the properties introduced here, let us give this class a name. We recall that inverse-concavity means the function

$$\lambda \mapsto \gamma(\lambda_1^{-1}, \dots, \lambda_n^{-1})^{-1}$$

is concave on Γ_+.

Definition 4.9. *Let $\gamma : \Gamma \to (0, \infty)$ be a concave admissible speed which is such that $\Gamma_+ \subset \Gamma$, and assume the restriction of γ to Γ_+ is inverse-concave. Let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be a smooth solution of* (CF) *with precompact timeslices. We say that $\mathbf{\Omega}$ is admissible if it is κ-noncollapsed, φ-almost-convex, and there is a symmetric cone $\Gamma' \Subset \Gamma$ such that*

$$\lambda(x, t) \in \Gamma', \qquad \forall\ (x, t) \in \mathbf{M}.$$

4. High curvature regions are locally starshaped

To ease the language somewhat in some of the proofs below, we introduce the following terminology. We implicitly use the fact that $\Omega_t + h \subset \Omega_t$ for every $t \in [0, T)$ and $h > 0$.

Definition 4.10. *Fix an admissible speed γ and let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be a solution of* (CF)*. Fix positive constants $\kappa > 0$ and $\Lambda > \kappa$, let (x_0, t_0) be a point in $\mathbf{M}$, and define $r_0^{-1} := G(x_0, t_0)$. Suppose $[-\Lambda^2 r_0^2 + t_0, t_0] \subset [0, T)$ and that the ball $B(x_0', \kappa r_0)$ is contained in Ω_{t_0}, where*

$$x_0' := x_0 - \kappa r_0 \nu(x_0, t_0).$$

For each $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$ let U_t denote the connected component of $\Omega_t \cap B(x_0', \Lambda r_0)$ which contains $B(x_0', \kappa r_0)$. We say that $\mathbf{\Omega}$ is (κ, Λ)-starshaped about (x_0, t_0) if for every $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$ and $p \in U_t$ the pseudocone $C(p, x_0', \kappa r_0)$ is contained in U_t.

The next theorem is the key technical result of this chapter. It tells us that an admissible solution $\mathbf{\Omega}$ is (κ, Λ)-starshaped about any point in $\mathbf{M}$ where the curvature is sufficiently large.

THEOREM 4.11. *Let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be an admissible solution in the sense of Definition 4.9. Then for every $\Lambda > \kappa$ there is a constant $K = K(n, \gamma, \Gamma', \kappa, \varphi, \Lambda)$ with the property that if $G(x_0, t_0) \geq K$, then Ω_t is (κ, Λ)-starshaped about (x_0, t_0).*

We make use of the following technical lemma. This is proven using general regularity results for parabolic PDE, so we defer most of the details to the appendix.

LEMMA 4.12. *Fix $r > 0$ and $\Lambda > 0$. Let $\mathbf{\Omega} = \{\Omega_t : t \in [-\Lambda^2 r^2 + t_0, t_0]\}$ be a solution of* (CF) *such that*

$$\lambda(x, t) \in \Gamma' \Subset \Gamma \qquad \forall\ (x, t) \in \mathbf{M}.$$

Suppose there is a point $p \in \mathbb{R}^{n+1}$ such that $B(p, r) \subset \Omega_{t_0}$ and for each $t \in [-\Lambda^2 r^2 + t_0, t_0]$ let U_t denote the connected component of $\Omega_t \cap B(p, \Lambda r)$ which contains $B(p, r)$. Suppose in addition that there is a positive K_0 such that

$$G(x, t) \leq K_0 r^{-1}$$

for each $x \in \partial U_t \cap B(p, \Lambda r)$ and $t \in [-\Lambda^2 r^2 + t_0, t_0]$, and let $x_0 \in \partial U_{t_0} \cap B(p, \Lambda r/2)$ be such that

$$G(x_0, t_0) \geq k_0 r^{-1},$$

where k_0 is some positive constant. Then there is a positive constant

$$\delta = \delta(n, \gamma, \Gamma', \Lambda, k_0, K_0)$$

and a smooth function

$$u : B(0, \delta r) \cap T_{x_0} M_{t_0} \times [-\delta^2 r^2 + t_0, t_0] \to \mathbb{R}$$

such that $u(x_0, t_0) = 0$ and the mapping

$$X(\cdot, t) : x \mapsto x_0 + x + u(x, t)\nu(x_0, t_0)$$

is a local parameterisation of $\partial U_t \cap B(p, \Lambda r)$ for each $t \in [-\delta^2 r^2 + t_0, t_0]$. Furthermore,

$$G(x, t) \geq \frac{k_0}{2} \qquad \forall\ (x, t) \in B(0, \delta r) \cap T_{x_0} M_{t_0} \times [-\delta^2 r^2 + t_0, t_0],$$

and for each $m \in \mathbb{N}$ the spatial derivatives of u satisfy

$$|D^m u|^2 \leq C(n, m, \gamma, \Gamma', \Lambda, k_0, K_0) r^{-2m+2}.$$

PROOF. With the upper bound on curvature, it is straightforward to find a δ with the right dependencies such that, locally about (x_0, t_0), the solution is a graph over $B(0, \delta r) \cap T_{x_0} M_{t_0} \times [-\delta^2 r^2 + t_0, t_0]$ with uniformly bounded first derivatives. Making δ a bit smaller if necessary, the desired lower bound for G follows from Lemma B.5, and the derivative bounds are then a consequence of Proposition B.4. $\square$

With the lemma in place, we are set to prove Theorem 4.11. The technique we use was introduced in [**BH17**] to prove curvature derivative estimates for a two-convex embedding evolving by the two-harmonic mean of its principal curvatures. Let us give our argument and then afterwards discuss differences with the proof of Theorem 6.2 in [**BH17**].

Proof of Theorem 4.11. **Step 1: Point-picking.** Suppose towards a contradiction that there is a large $\Lambda > \kappa$ for which the statement is false. This means we can find a sequence of spacetime points $\tilde{x}_k \in M_{\tilde{t}_k}$ with the property that $G(\tilde{x}_k, \tilde{t}_k)$ tends to infinity, but Ω_t fails to be (κ, Λ)-starshaped about $(\tilde{x}_k, \tilde{t}_k)$. We are going to modify this sequence using Perelman's point-picking trick: Let $Q(\kappa, \Lambda)$ be the set of points $(x, t) \in \mathbf{M}$ with the property that Ω_t is (κ, Λ)-starshaped about (x, t). For each pair of postive integers k and j, let

$$U_{k,j} := \{(x,t) \in \mathbf{M} : t \le \tilde{t}_k,\ (x,t) \notin Q(\kappa,\Lambda),\ G(x,t) \ge 2^j r_k^{-1}\}.$$

By assumption, $U_{k,0}$ is nonempty, since it contains $(\tilde{x}_k, \tilde{t}_k)$. On the other hand, $U_{k,j}$ is always empty if j is sufficiently large, since G is bounded on the time interval $[0, \tilde{t}_k]$. Therefore, for each k, there is a largest value $j_k \in \mathbb{N}$ such that U_{k,j_k} is nonempty. Let (x_k, t_k) be a spacetime point in U_{k,j_k}. By this process we have made sure that if $t \le t_k$ and $G(x,t) \ge 2G(x_k, t_k)$, then $(x,t) \in Q(\kappa, \Lambda)$.

Step 2: Separating curvature scales. Set $r_k^{-1} := G(x_k, t_k)$ and $x'_k := x_k - \kappa r_k \nu(x_k, t_k)$. For each $t \in [-\Lambda^2 r_k^2 + t_k, t_k]$, let U_t^k be the connected component of $\Omega_t \cap B(x'_k, \Lambda r_k)$ containing $B(x'_k, \kappa r_k)$. Since $r_k \to 0$ and $t_k \to T$, we may assume that $-\Lambda^2 r_k^2 + t_k \ge 0$. By definition, since $(x_k, t_k) \notin Q(\kappa, \Lambda)$, there must be some $\tau_k \in [-\Lambda^2 r_k^2 + t_k, t_k]$ and a point $z_k \in U_{\tau_k}$ such that

$$C(z_k, x'_k, \kappa r_k) \not\subset U_{\tau_k}.$$

Therefore, by Lemma 4.6, there is a point $\tilde{y}_k \in U_{\tau_k}$ such that

$$C(\tilde{y}_k, x'_k, \kappa r_k) \subset U_{\tau_k},$$

and the hypersurface $S(\tilde{y}_k, x'_k, \kappa r_k)$ makes interior contact with $\partial U_{\tau_k} \cap B(x'_k, \Lambda r_k)$ at some point, which we denote by y_k. Lemma 4.4 tells us that the smallest principal curvature of $S(\tilde{y}_k, x'_k, \kappa r_k)$ is at most $-\frac{1}{10}\kappa r_k |x'_k - \tilde{y}_k|^{-2}$, so we have

$$\lambda_1(y_k, \tau_k) \le -\frac{1}{10}\kappa r_k |x'_k - \tilde{y}_k|^{-2} \le -\frac{\kappa}{10\Lambda^2} r_k^{-1}.$$

On the other hand, by the φ-almost convexity property there holds

$$\frac{\lambda_1(y_k, \tau_k)}{G(y_k, \tau_k)} \ge -\frac{\varphi(G(y_k, \tau_k))}{G(y_k, \tau_k)}.$$

There is a potentially large positive constant $C = C(n, \gamma, \Gamma')$ such that

$$\lambda_1(x,t) \ge -CG(x,t)$$

holds on $\mathbf{M}$, so we have

$$-\frac{\kappa}{10\Lambda^2} r_k^{-1} \ge -CG(y_k, \tau_k),$$

which implies that $G(y_k, \tau_k) \to \infty$. Hence by the almost convexity,

$$-\frac{\kappa}{10\Lambda^2}\frac{r_k^{-1}}{G(y_k, \tau_k)} \ge \frac{\lambda_1(y_k, \tau_k)}{G(y_k, \tau_k)} \to 0,$$

from which we conclude that

$$\frac{G(y_k, \tau_k)}{G(x_k, t_k)} \to \infty.$$

In particular, passing to a subsequence if necessary, the point-picking construction ensures that Ω_t is (κ, Λ)-starshaped about (y_k, τ_k) for each index k.

Step 3: Extracting a local limit. Let $s_k := \kappa G(y_k, \tau_k)^{-1}$, and consider the sequence of rescaled solutions $\mathbf{\Omega}^k = \{\Omega_t^k : t \in [-\kappa^{-2}\Lambda^2, 0]\}$ defined as follows:

$$\Omega_t^k := s_k^{-1}(\Omega_{\tau_k + s_k^2 t} - y_k).$$

We write $G^{(k)}$ for $\gamma(\lambda^{(k)})$, and $\lambda^{(k)}$ for the principal curvatures of $\mathbf{\Omega}^k$. There holds $G^{(k)}(0,0) = \kappa^{-1}$, and applying a rotation in the ambient space if necessary, we may assume that

$$\nu^{(k)}(0,0) = e_1,$$

in which case the κ-noncollapsing says that the ball $\partial B(-e_1, 1)$ makes interior contact with $M_0^k := \partial\Omega_0^k$ at the origin. Here $\nu^{(k)}$ is the unit normal to

$$\mathbf{M}^k := \{M_t^k : t \in [-\kappa^{-2}\Lambda^2, 0]\}.$$

Since Ω_t is (κ, Λ)-starshaped about (y_k, τ_k), applying the rescaling, we find that Ω_t^k is $(1, \kappa^{-1}\Lambda)$-starshaped about $(0,0)$. That is, if we define U_t^k to be the connected component of $\Omega_t^k \cap B(-e_1, \kappa^{-1}\Lambda)$ which contains $B(-e_1, 1)$, then

$$C(-e_1, p, 1) \subset U_t^k \qquad \forall\, p \in U_t^k,\ t \in [-\kappa^{-2}\Lambda^2, 0].$$

Therefore, by Lemma 4.5, $\mathbf{\Omega}^k$ satisfies the hypotheses of Theorem 4.3 in

$$B(-e_1, \kappa^{-1}\Lambda) \times [-\kappa^{-2}\Lambda^2, 0],$$

with $p = -e_1$, $r = \kappa^{-1}$, $L = K = \Lambda$, and $\theta = \theta(\kappa, \Lambda)$. We may assume that Λ is much larger than one. Thus, by Theorem 4.3, we have a curvature bound

$$G^{(k)}(x,t) \le K_0(n, \gamma, \Gamma', \kappa, \Lambda),$$

valid for each

$$x \in \partial U_t^k \cap B(-e_1, \kappa^{-1}\Lambda/4), \qquad t \in [-\kappa^{-2}\Lambda^2/2, 0].$$

Since $G^{(k)}(0,0) = \kappa^{-1}$, $\mathbf{\Omega}^k$ satisfies all of the assumptions of Lemma 4.12 at $(0,0)$. Hence there is a positive $\delta = \delta(n, \gamma, \Gamma', \kappa, \Lambda)$ and a sequence of smooth functions

$$u^{(k)} : B(0,\delta) \cap e_1^\perp \times [-\delta^2, 0] \to \mathbb{R}$$

such that $u^{(k)}(0,0) = 0$ and the mapping

$$X^{(k)}(\cdot, t) : x \mapsto x + u^{(k)}(x,t)e_1$$

is a local parameterisation of M_t^k for each $t \in [-\delta^2, 0]$. Furthermore, for each $m \in \mathbb{N}$, the spatial derivatives of u satisfy

$$|D^m u^{(k)}|^2 \le C r^{-2m+2},$$

where C is independent of k, and the value of G on the graph of $u^{(k)}$ is bounded from below by $(2\kappa)^{-1}$.

Now we can apply the Arzela-Ascoli theorem to extract a subsequence of the $u^{(k)}$ converging smoothly to some $\hat{u}$. The embedding

$$\hat{X}(x,t) := x + \hat{u}(x,t)e_1$$

satisfies

$$(\partial_t \hat{X}(x,t))^\perp = -\hat{G}(x,t),$$

and as a consequence of the φ-almost-convexity, has nonnegative second fundamental form. On the other hand, we had $\lambda_1(y_k, \tau_k) \le 0$ for every index k, so $\lambda_1^{(k)}(0,0) \le 0$.

It follows that $\hat{\lambda}_1$ vanishes at $(0,0)$, and since we are assuming that γ is inverse-concave on Γ_+, we can invoke the strong maximum principle of Corollary 3.8. This tells us that $\hat{\lambda}_1$ vanishes identically, and if $v \in \ker(\hat{A})$, then $\nabla_v \hat{G} = 0$. Let us define $\hat{M}_t = \hat{X}(B(0,\delta) \cap e_1^\perp, t)$ for each $t \in [-\delta^2, 0]$.

Step 4: The rescaled pseudocone. We now want to track the pseudocone $S(\tilde{y}_k, x_k', r_k)$ under the rescaling. Recall that $S(\tilde{y}_k, x_k', r_k)$ makes interior contact with M_{τ_k} at y_k. Observe also that since $G(y_k, \tau_k)$ blows up much more quickly than $r_k^{-1} = G(x_k, t_k)$, as $k \to \infty$, the sequence y_k must be approaching the vertex of the pseudocone - otherwise, the interior contact would imply an upper bound for $G(y_k, \tau_k)$ on the scale of r_k^{-1}.

Let us write

$$d_k := s_k^{-1} |\tilde{y}_k - y_k|$$

for the rescaled distance from y_k to the vertex of $C(\tilde{y}_k, x_k', \kappa r_k)$. If $d_k \to \infty$, then for large k, the rescaled pseudocone

$$C_k := s_k^{-1}(C(\tilde{y}_k, x_k', \kappa r_k) - y_k)$$

is very close to a halfspace, and since $\nu^{(k)}(0,0) = e_1$, we get that

$$C_k \to \{x \in \mathbb{R}^{n+1} : \langle x, e_1 \rangle < 0\},$$

and this convergence is smooth on compact subsets of the ambient space. Since $G^{(k)}(0,0) = \kappa^{-1}$ for every k, for large enough k this contradicts the fact that C_k makes interior contact with M_0^k at $(0,0)$. Therefore, the sequence d_k must be uniformly bounded from above, and passing to a subsequence in k, we may assume that the d_k approach some limit $\hat{d} > 0$.

To recap, whilst C_k may move around as k varies, its boundary always passes through zero, where its normal agrees with e_1, and the distance from the origin to the vertex is positive, uniformly bounded, and converges to $\hat{d} \geq 0$. Applying a further rotation (leaving e_1 fixed) for each k, we may arrange that the vertex of C_k is approaching $-\hat{d}e_2$. Then, since $s_k^{-1} \to \infty$, the negative curvature in the boundary of C_k is being scaled away, and the C_k must converge to a round round cone $\hat{K}$, which has positive aperture $\phi > 0$ depending only on κ and Λ. In addition, $\hat{K}$ has the following properties:

- The vertex of $\hat{K}$ is at $-\hat{d}e_1$;
- The ray $E_2 := \{se_2 : s \in (-\hat{d}, \infty)\}$ is in $\partial\hat{K}$;
- The vector e_1 is normal to $\partial\hat{K}$ at each point in E_2.

Step 5: Extending the local limit. The presence of the cone $\hat{K}$ allows us to extend the local limiting solution $\hat{M}_t$ defined in Step 3. For each $\sigma < \infty$, let us define a strip

$$S_\sigma := \{x \in e_1^\perp : |x - \langle x, e_2 \rangle e_2| \leq \delta/2, \ \langle x, e_2 \rangle \in [0, \sigma]\},$$

where δ is the constant appearing in Step 3, which comes from the technical lemma. We claim that for each $\sigma < \infty$, there is a sequence of smooth functions

$$u^{(k,\sigma)} : S_\sigma \times [-\delta^2/4, 0] \to \mathbb{R}$$

which satisfy $u^{(k,\sigma)}(0,0) = 0$, are such that the maps

$$X^{(k,\sigma)}(x,t) := x + u^{(k,\sigma)}(x,t)e_1$$

locally parameterise M_t^k, and which converge in C^∞ to a limit $\hat{u}^{(\sigma)}$. Moreover, $\hat{G}^{(\sigma)} \geq (2\kappa)^{-1}$ on $\operatorname{graph}(\hat{u}^{(\sigma)}(\cdot,t))$ for each $t \in [-\delta^2/4, 0]$. We use an upper (k,σ) to denote quantities on the graph of $u^{(k,\sigma)}$, while quantities with a hat and upper (σ) are defined on the graph of $\hat{u}^{(\sigma)}$.

We have already proven the claim for $\sigma \leq \delta/2$ in Step 3, so suppose for a contradiction that $\sigma_0 > \delta/2$ is the largest constant for which the claim holds. The family of hypersurfaces $\hat{M}_t^{\sigma_0} := \operatorname{graph}(\hat{u}^{(\sigma_0)}(\cdot,t))$ solves

$$\partial_t \hat{u}^{(\sigma_0)}(x,t)(e_1 \cdot \hat{\nu}^{(\sigma_0)}(x,t)) = -\hat{G}^{(\sigma_0)}(x,t),$$

and the lower bound for the curvature $\hat{G}^{(\sigma)} \geq (2\kappa)^{-1}$ and convexity estimate ensure that the second fundamental form of $\hat{M}_t^{\sigma_0}$ is nonnegative. On the other hand, $\hat{M}_0^{\sigma_0}$ is exterior to $\hat{K}$; the boundary of $\hat{K}$ contains the ray E_2 parallel to e_2; and e_2 is in $T_0\hat{M}_0^{\sigma_0}$. These properties mean that the function

$$f(s) := \hat{u}^{(\sigma_0)}(se_2, 0)$$

is nonnegative and satisfies $f'(0)$. Since f is also convex, we have $f \equiv 0$, which is the same as saying that $\hat{M}_0^{\sigma_0}$ contains the line segment $E_2^{\sigma_0} := \{se_2 : s \in [0,\sigma_0]\}$. In particular, e_2 is tangent to $\hat{M}_0^{\sigma_0}$ at each point $x \in E_2^{\sigma_0}$, and lies in the kernel of the second fundamental form. Hence, by the strong maximum principle of Corollary 3.8, we have $\nabla_{e_2}\hat{G}^{(\sigma_0)}(\cdot,0)$ on $E_2^{\sigma_0}$, and consequently

$$\hat{G}^{(\sigma_0)}(\cdot,0) \equiv \hat{G}^{(\sigma_0)}(0,0) = \kappa^{-1}$$

on $E_2^{\sigma_0}$.

In particular, $\hat{G}^{(\sigma_0)}(\sigma_0 e_2, 0) = \kappa^{-1}$, so there is a sequence of points

$$z_k \in \operatorname{graph}(u^{(k,\sigma_0)}(\cdot,0))$$

which converge to $\sigma_0 e_2$, and are such that $G^{(k,\sigma)}(z_k,0) \to \kappa^{-1}$. Consequently, for large k, the curvature of M_t^k at $(z_k,0)$ is approximately κ^{-1}, and by the point-picking, Ω_t^k is $(1,\kappa^{-1}\Lambda)$-starshaped about $(z_k,0)$. Now we can proceed almost exactly as in Step 3 to get a local limit near $(\sigma_0 e_2, 0)$. Indeed, the technical lemma (Lemma 4.12) shows that near $(z_k,0)$, M_t^k can be expressed as an evolving graph over the set

$$B(0,\delta) \cap \nu^{(k)}(z_k,0)^\perp \times [-\delta^2, 0].$$

Therefore, since $z_k \to \sigma_0 e_2$ and $\nu^{(k)}(z_k,0) \to e_1$, for sufficiently large k there is a smooth function

$$\tilde{u}^{(k,\sigma_0)} : B(0,\delta/2) \cap e_1^\perp \times [-\delta^2/4, 0]$$

which gives a local graph representation of M_t^k, and satisfies $\tilde{u}^{(k,\sigma_0)}(z_k,0) \to 0$. Moreover, the technical lemma says that all of the derivatives of the $\tilde{u}^{(k,\sigma_0)}$ are bounded independently of k, and the value of G on the graph of $\tilde{u}^{(k,\sigma_0)}$ is bounded from below by $(2\kappa)^{-1}$.

Now, where their domains of definition overlap, the functions $u^{(k,\sigma_0)}$ and $\tilde{u}^{(k,\sigma_0)}$ agree, so for $\sigma' := \sigma_0 + \delta/2$, we can define

$$u^{(k,\sigma')}(x,t) := \begin{cases} u^{(k,\sigma_0)}(x,t) & (x,t) \in S_{\sigma_0} \times [-\delta^2/4, 0] \\ \tilde{u}^{(k,\sigma_0)}(x,t) & (x,t) \in B(\sigma_0 e_2, \delta/2) \cap e_1^\perp \times [-\delta^2/4, 0]. \end{cases}$$

We have bounds on all of the derivatives of $u^{(k,\sigma')}$, and G is bounded from below by $(2\kappa)^{-1}$ on its graph, so passing to a subsequence in k, we get a contradiction to the maximality of σ_0.

Step 6: Drawing a contradiction. Let $\sigma > 0$ be a large constant to be chosen in a moment, and let

$$\hat{M}_t^\sigma := \operatorname{graph}(\hat{u}^{(\sigma)}(\cdot, t))$$

for each $t \in [-\delta^2/4, 0]$, where $\hat{u}^{(\sigma)}$ is the smooth limiting function constructed in the previous step. We saw that the zero timeslice $\hat{M}_0^\sigma$ contains $E_2 \cap S_\sigma$, and $\hat{M}_0^\sigma$ lies in the complement of $\hat{K}$. In particular, this means that at the point $z := \frac{\sigma}{2} e_2$, $\hat{M}_0^\sigma$ touches $\partial\hat{K}$ from the outside. On the other hand, there is a purely geometric constant C such that the principal curvatures of the hypersurface

$$\{x \in \partial\hat{K} : \langle x, e_2\rangle = \sigma/2\}$$

are all bounded from above by $C\sigma^{-1}$, so we have

$$\hat{G}^{(\sigma)}(z, 0) \leq C'(n, \gamma, \Gamma', C)\sigma^{-1}.$$

Since we showed above that $\hat{G}^{(\sigma)}(\cdot, 0) \equiv \kappa^{-1}$ on $E_2 \cap S_\sigma$, choosing σ sufficiently large depending only on κ and C', we obtain a contradiction. Hence, our original assumption must have been false - there is no such sequence $(\tilde{x}_k, \tilde{t}_k)$. $\square$

The preceding proof diverges from the proof of Theorem 6.2 in [**BH17**] at Step 3. In both cases a local limiting solution is obtained by rescaling about y_k, and λ_1 can be shown to vanish identically on this limit by a strong maximum principle argument. However, the special structure of the cylindrical estimate in the two-convex case then implies that the remaining principal curvatures are all equal, so the limit is a piece of a cylinder with one flat direction. In particular, its curvature is constant, and by induction, the limit can be extended arbitrarily far out in space. On the other hand, the limiting cylinder is supposed to contain a cone of positive aperture by the same argument as in Step 4, which is absurd.

The cylindrical estimate for a general k-convex solution is not strong enough to conclude from $\lambda_1 \equiv 0$ that the local limit has constant curvature. By tracking more carefully the rescaled pseudocone tangent to y_k, and using the inverse-concavity of the speed, we were able to extend the limit along a ray in the boundary of $\hat{K}$, which turned out to be sufficient. In the end, the contradiction comes from the fact that the final timeslice of the local limiting solution cannot make exterior contact with the boundary of a round cone. This is interesting, since the same kind of property is used by Haslhofer-Kleiner to prove their global convergence theorem (which implies a global Harnack inequality) in [**HK17a**][Theorem 1.12], and by Perelman to obtain a similar result for compact three-dimensional Ricci flow [**Per02**][Section 12]. In each of these two arguments and ours, noncollapsing and almost-convexity are combined in different ways, but at a crucial step the strong maximum principle is used to say that the final timeslice of a solution with 'nonnegative curvature' cannot coincide with (or in our case make exterior contact with) a round cone.

5. A Harnack inequality for the curvature

We now know that on an admissible solution, about any point (x_0, t_0) of sufficiently large curvature, there is a large connected component of the solution which is starshaped. The curvature bound for radial graphs then tells us that over this connected component, we have an upper bound for the curvature in terms of its value at (x_0, t_0).

LEMMA 4.13. *Let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be an admissible solution in the sense of Definition 4.9. Then for every $\Lambda > \kappa$ there is a positive $K = K(n, \gamma, \Gamma', \kappa, \varphi, \Lambda)$ with the following property. Suppose $r_0^{-1} := G(x_0, t_0) \geq K$, and for each*

$$t \in [-\Lambda^2 r_0^2 + t_0, t_0]$$

let U_t denote the connected component of $\Omega_t \cap B(x_0, \Lambda r_0)$ which contains $B(x_0', \kappa r_0)$, where $x_0' = x_0 - \kappa r_0 \nu(x_0, t_0)$. Then for each $x \in \partial U_t \cap B(0, \Lambda r_0)$ and $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$, there holds

$$G(x, t) \leq C(n, \gamma, \Gamma', \kappa)\Lambda^2 r_0^{-1}.$$

PROOF. We may assume $\Lambda \geq 100\kappa$. By Theorem 4.11, we can choose K so that if $r_0^{-1} \geq K$, then $\mathbf{\Omega}$ is $(\kappa, 4\Lambda)$-starshaped about (x_0, t_0). By Lemma 4.5, this means that

$$\nu(x, t) \cdot \frac{x - x_0'}{|x - x_0'|} \geq \frac{1}{4\sqrt{5}}\Lambda^{-1}$$

for each

$$x \in \partial U_t \cap B(x_0', 4\Lambda r_0), \qquad t \in [-16\Lambda^2 r_0^2 + t_0, t_0].$$

Therefore, by Theorem 4.3, if $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$ and $x \in \partial U_t \cap B(x_0', \Lambda r_0)$ then there holds

$$G(x, t) \leq C(n, \gamma, \Gamma', \kappa)\Lambda^2 r_0^{-1}.$$

□

Thus the technical lemma implies a pointwise estimate for all of the derivatives of A which is valid at points of large curvature:

THEOREM 4.14. *Let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be an admissible solution in the sense of Definition 4.9. Then there is a constant $K = K(n, \gamma, \Gamma', \kappa, \varphi)$ with the property that if $G(x_0, t_0) \geq K$, the estimate*

$$|\nabla^k A|^2(x_0, t_0) \leq C(n, k, \gamma, \Gamma', \kappa) G(x_0, t_0)^{-2k-2}$$

holds for each $k \in \mathbb{N}$.

PROOF. Let K be the constant from Lemma 4.13 with $\Lambda = 100\kappa$. For each

$$t \in [-\Lambda^2 r_0^2 + t_0, t_0]$$

let U_t denote the connected component of $\Omega_t \cap B(x_0, \Lambda r_0)$ which contains $B(x_0', \kappa r_0)$, where $x_0' = x_0 - \kappa r_0 \nu(x_0, t_0)$. Then for each $x \in \partial U_t \cap B(0, \Lambda r_0)$ and $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$ there holds

$$G(x_0, t_0) \leq C(n, \gamma, \Gamma', \kappa) r_0^{-1}.$$

Applying Lemma 4.12, we conclude that there is a small positive $\delta = \delta(n, \gamma, \Gamma', \kappa)$ and a smooth function

$$u : B(0, \delta r_0) \cap T_{x_0} M_{t_0} \times [-\delta^2 r_0^2 + t_0, t_0]$$

which represents the solution as a graph near (x_0, t_0). In particular, $u(0,0) = 0$, the map

$$x \mapsto x_0 + x + u(x,t)\nu(x_0, t_0)$$

is a local parameterisation of M_t for each $t \in [-\delta^2 r_0^2 + t_0, t_0]$, and the derivatives of u at $(0,0)$ satisfy

$$|D^k u|^2(0,0) \le C(n,k,\gamma,\Gamma',\kappa) r_0^{-2k+2}.$$

The quantity $|\nabla^k A|_g^2(x_0,t_0)$ is bounded in terms of the derivatives of u of order at most $k+2$ at $(0,0)$, so we obtain

$$|\nabla^k A|_g^2(x_0,t_0) \le C(n,k,\gamma,\Gamma',\kappa) r_0^{-2k-2},$$

as required. □

We would now like to establish that, in a large backwards neighbourhood around any point where the curvature is large, we also have control on the curvature from below. This is the key to establishing derivative estimates which are valid on increasingly large subsets of spacetime, and not just at a point or on a small neighbourhood.

THEOREM 4.15. *Let $\mathbf{\Omega} = \{\Omega_t : t \in [0,T)\}$ be an admissible solution in the sense of Definition 4.9. Then for every $\Lambda > \kappa$ there is a positive $K = K(n,\gamma,\Gamma',\kappa,\varphi,\Lambda)$ with the following property. Suppose $r_0^{-1} := G(x_0,t_0) \ge K$, and for each*

$$t \in [-\Lambda^2 r_0^2 + t_0, t_0]$$

let U_t denote the connected component of $\Omega_t \cap B(x_0, \Lambda r_0)$ which contains $B(x_0', \kappa r_0)$, where $x_0' = x_0 - \kappa r_0 \nu(x_0,t_0)$. Then for each $x \in \partial U_t \cap B(0,\Lambda r_0)$ and $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$ there holds

$$G(x,t) \ge c(n,\gamma,\Gamma',\kappa)\Lambda^{-1} r_0^{-1}.$$

PROOF. Fix $\Lambda \ge 100\kappa$. Let μ be a small positive constant. Suppose there are sequences of spacetime points (x_k,t_k) and (y_k,τ_k) in $\mathbf{M}$ with the following properties:

- $r_k^{-1} := G(x_k,t_k) \to \infty$;
- $s_k^{-1} := G(y_k,\tau_k) = \mu G(x_k,t_k)$;
- $\tau_k \in [-\Lambda^2 r_k^2 + t_0, t_0]$ and $y_k \in \partial U^k_{\tau_k} \cap B(x_k', \Lambda r_k)$.

For each $t \in [-\Lambda^2 r_k^2 + t_0, t_0]$, U_t^k denotes the connected component of $\Omega_t \cap B(x_k', \Lambda r_k)$ which contains $B(x_k', \kappa r_k)$, where $x_k' := x_k - \kappa r_k \nu(x_k,t_k)$. Using these assumptions, we are going to prove a lower bound for μ of the form

$$\mu \ge c(n,\gamma,\Gamma',\kappa,\varphi)\Lambda^{-1}.$$

This proves the theorem since, for any μ violating this inequality, no sequence satisfying the above properties can exist.

We first observe that

$$\begin{aligned}
|y_k' - x_k'| &\le |y_k' - y_k| + |y_k - x_k| + |x_k - x_k'| \\
&\le \kappa s_k + (\Lambda + \kappa) r_k \\
&= (\kappa(1+\mu^{-1}) + \Lambda) r_k,
\end{aligned}$$

and $B(y_k', \kappa s_k) = B(y_k', \kappa\mu^{-1} r_k)$, so for $\Lambda_0 := 10(\kappa(1+2\mu^{-1}) + \Lambda)$ there holds

$$B(y_k', \kappa s_k) \subset B(x_k', \Lambda_0 r_k).$$

For each $t \in [-\Lambda_0^2 r_k^2 + t_k, t_k]$, let $U_t^{0,k}$ denote the connected component of $\Omega_t \cap B(x_k', \Lambda_0 r_k)$ which contains $B(x_k', \kappa r_k)$. Notice that U_t^k is always a subset of $U_t^{0,k}$, since $\Lambda_0 > \Lambda$. Also, since r_k^{-1} blows up as $k \to \infty$, if k is sufficiently large then Theorem 4.11 tells us that $C(x, x_k', \kappa r_k) \subset U_t^{0,k}$ for every $x \in U_t^{0,k}$. In particular, since

$$y_k \in \partial U_{\tau_k}^k \cap B(x_k', \Lambda r_k) \subset \partial U_{\tau_k}^{0,k} \cap B(x_k', \Lambda r_k),$$

the line segment $\overline{x_k' y_k}$ is in $\overline{U_{\tau_k}^{0,k}}$. It follows that each point $y \in B(y_k', \kappa s_k)$ can be connected to x_k' by a path which lies in $\Omega_{\tau_k} \cap B(x_k', \Lambda_0 r_k)$, namely, the path

$$\overline{x_k' y_k} \cup \overline{y_k y_k'} \cup \overline{y_k' y}.$$

This means that $B(y_k', \kappa s_k) \subset U_{\tau_k}^{0,k}$.

The idea now is to let $B(y_k', \kappa s_k)$ evolve by the flow, so that it remains inside the solution. Then if μ is small, there will be a large ball B_k inside Ω_{t_k} at a controlled distance from x_k. If k is sufficiently large then the solution is locally a graph over the boundary of this large ball, which means that on a large spacetime neighbourhood around B_k, the curvature is bounded from above on the order of the inverse of the radius of B_k. If B_k is too large relative to r_k (i.e. μ is too small), we get a contradiction to the fact that $G(x_k, t_k) = r_k^{-1}$.

Let

$$R_k(t)^2 := \kappa^2 s_k^2 - 2\gamma(1, \ldots, 1)(t - \tau_k), \qquad t \leq T_k := \tau_k + \frac{\kappa^2 s_k^2}{2\gamma(1, \ldots, 1)},$$

so that the boundary of $B(y_k', R_k(t))$ is a solution of the flow, and by the last paragraph and the avoidance principle, $B(y_k', R_k(t)) \subset U_t^{0,k}$ for all $\tau_k \leq t < T_k$. We have

$$T_k \geq t_k - \Lambda^2 r_k^2 + \frac{\kappa^2 s_k^2}{2\gamma(1, \ldots, 1)} = t_k + \left(\frac{\kappa^2}{2\gamma(1, \ldots, 1)} - \Lambda^2 \mu^2\right) s_k^2,$$

so let us assume μ is small enough to ensure

$$\frac{\kappa^2}{2\gamma(1, \ldots, 1)} - \Lambda^2 \mu^2 > 0.$$

Then $T_k > t_k$, so $B(y_k', R_k(t))$ still has positive radius at time $t = t_k$. In particular, $B(y_k', R_k(t_k)) \subset U_{t_k}^{0,k}$.

We have the following lower bound for $R_k(t)$ at time $t = t_k$:

$$\begin{aligned} R_k(t_k)^2 &= \mu^{-2} \kappa^2 r_k^2 - 2\gamma(1, \ldots, 1)(t_k - \tau_k) \\ &\geq (\mu^{-2}\kappa^2 - 2\gamma(1, \ldots, 1)\Lambda^2) r_k^2 \\ &=: \eta^2 r_k^2. \end{aligned}$$

We recall from above that $|x_k - y_k'| \leq \Lambda_0 r_k / 10$, so if we set

$$\Lambda_1 = \eta^{-1} \Lambda_0,$$

then there certainly holds

$$x_k \in B(y_k', \Lambda_1 R_k(t_k)/10).$$

Also, since

$$B(y_k', \Lambda_1 R_k(t_k)) \subset B(x_k', \Lambda_0 r_k + \Lambda_1 \kappa s_k),$$

if we set $\Lambda_2 := 10(\Lambda_0 + \mu^{-1}\Lambda_1\kappa)$, then

$$B(y_k', \Lambda_1 R_k(t_k)) \subset B(x_k', \Lambda_2 r_k).$$

For each $t \in [-\Lambda_2^2 r_k^2 + t_0, t_0]$ let $U_t^{2,k}$ be the connected component of $\Omega_t \cap B(x_k', \Lambda_2 r_k)$ which contains $B(x_k', \kappa r_k)$. By the last paragraph

$$B(y_k', R_k(t_k)) \subset U_{t_k}^{0,k} \subset U_{t_k}^{2,k}.$$

We are going to show that if k is large then $\partial U_t^{2,k}$ is a radial graph over $B(y_k', R_k(t_k))$ inside

$$B(y_k', \Lambda_1 R_k(t_k)) \times [-\Lambda_1^2 R_k(t_k)^2 + t_k, t_k].$$

Since $x_k \in B(y_k', \Lambda_1 R_k(t_k)/10)$, this will imply an upper bound for the curvature at (x_k, t_k).

We know by Lemma 4.13 that if k is sufficiently large,

$$G(x,t) \leq C(n, \gamma, \Gamma', \kappa)\Lambda_2^2 r_k^{-1}.$$

for each $x \in \partial U_t^{2,k} \cap B(x_k', \Lambda_2 r_k)$ and $t \in [-\Lambda_2^2 r_k^2 + t_k, t_k]$. Suppose for a contradiction that there is a time $\tilde{t}_k \in [-\Lambda_2^2 r_k^2 + t_k, t_k]$ and a point $z_k \in U_{\tilde{t}_k}^{2,k}$ such that $C(z_k, y_k', R_k(t_k))$ is not contained in $U_{\tilde{t}_k}^{2,k}$. Then by Lemma 4.6 there is a point $\tilde{z}_k \in U_{\tilde{t}_k}^{2,k}$ such that the hypersurface $S(\tilde{z}_k, y_k', R_k(t_k))$ makes interior contact with $\partial U_{\tilde{t}_k}^{2,k} \cap B(x_k', \Lambda_2 r_k)$ at some point, which we denote by $\tilde{a}_k$. In particular, by Lemma 4.6 there holds

$$\lambda_1(\tilde{a}_k, \tilde{t}_k) \leq -\frac{1}{10} R_k(t_k)|\tilde{z}_k - y_k'|^{-2}.$$

Since $\tilde{z}_k$ and y_k' are both in $B(x_k', \Lambda_2 r_k)$ we can bound

$$|\tilde{z}_k - y_k'| \leq 2\Lambda_2 r_k,$$

and $R_k(t_k) \geq \eta r_k$, so there holds

$$\lambda_1(\tilde{a}_k, \tilde{t}_k) \leq -\frac{\eta}{40\Lambda_2^2} r_k^{-1}.$$

It follows that $\lambda_1(\tilde{a}_k, \tilde{t}_k)$ tends to $-\infty$ as $k \to \infty$, and since

$$G(\tilde{a}_k, \tilde{t}_k) \geq c(n, \gamma, \Gamma')|A|(\tilde{a}_k, \tilde{t}_k) \geq c(n, \gamma, \Gamma')|\lambda_1|(\tilde{a}_k, \tilde{t}_k)$$

it must be the case that $G(\tilde{a}_k, \tilde{t}_k) \to \infty$. Therefore, by the almost-convexity property, given any $\varepsilon > 0$ there holds

$$\frac{\lambda_1(\tilde{a}_k, \tilde{t}_k)}{G(\tilde{a}_k, \tilde{t}_k)} \geq -\frac{\varphi(G(\tilde{a}_k, \tilde{t}_k))}{G(\tilde{a}_k, \tilde{t}_k)} \geq -\varepsilon$$

for every sufficiently large k. We conclude that

$$-\varepsilon r_k^{-1} \leq -\frac{\eta}{40\Lambda_2^2} r_k^{-1}$$

for all sufficiently large k, but this is a contradiction if ε is too small. To recap, we have shown that for each $t \in [-\Lambda_2^2 r_k^2 + t_k, t_k]$ and $z \in U_t^{2,k}$, the pseudocone $C(z, y_k', R_k(t_k))$ is contained in $U_t^{2,k}$.

Since $B(y_k', \Lambda_1 R_k(t_k)) \subset B(x_k', \Lambda_2 r_k)$ and

$$\Lambda_1 R_k(t_k) \leq \Lambda_1 \kappa s_k \leq \Lambda_2 r_k$$

we can say in particular that if k is sufficiently large, then for each $z \in U_t^{2,k} \cap B(y_k', \Lambda_1 R_k(t_k))$ and $t \in [-\Lambda_1^2 R_k(t_k)^2 + t_k, t_k]$ there holds

$$C(z, y_k', R_k(t_k)) \subset U_t^{2,k} \cap B(y_k', \Lambda_1 R_k(t_k)).$$

Invoking Lemma 4.5 we find that for each

$$x \in \partial U_t^{2,k} \cap B(y_k', \Lambda_1 R_k(t_k)), \qquad t \in [-\Lambda_1^2 R_k(t_k)^2 + t_k, t_k]$$

there holds

$$\nu(x,t) \cdot \frac{x - y_k'}{|x - y_k'|} \geq \frac{1}{\sqrt{5}} \Lambda_1^{-1}.$$

It follows then from Theorem 4.3 that for each

$$x \in \partial U_t^{2,k} \cap B(y_k', \Lambda_1 R_k(t_k)/4), \qquad t \in [-\Lambda_1^2 R_k(t_k)^2/2 + t_k, t_k]$$

there holds $G(x,t) \leq C(n, \gamma, \Gamma')\Lambda_1^2 R_k(t_k)^{-1}$. We chose Λ_1 large enough to ensure that $x_k \in B(y_k', \Lambda_1 R_k(t_k)/10)$, so in particular,

$$r_k^{-1} = G(x_k, t_k) \leq C(n, \gamma, \Gamma')\Lambda_1^2 R_k(t_k)^{-1}.$$

By definition

$$\Lambda_1^2 = \eta^{-2}\Lambda_0^2,$$

and we know that $R_k(t_k) \geq \eta r_k$, so

$$\Lambda_1^2 R_k(t_k)^{-1} \leq \eta^{-3}\Lambda_0^2 r_k^{-1},$$

and there holds

$$r_k^{-1} \leq C(n, \gamma, \Gamma')\eta^{-3}\Lambda_0^2 r_k^{-1}.$$

Let us write

$$L = \frac{1}{\mu\Lambda}.$$

We recall the definition of η and insert the definition of L:

$$\eta = \mu^{-2}\kappa^2 - 2\gamma(1, \ldots, 1)\Lambda^2 = (L^2\kappa^2 - 2\gamma(1, \ldots, 1))\Lambda^2.$$

Suppose $L \geq 10\kappa^{-1}\gamma(1, \ldots, 1)^{\frac{1}{2}}$ so that we can estimate

$$\eta \geq \frac{\kappa^2}{2} L^2 \Lambda^2.$$

Similarly, using the definition of Λ_0 we can estimate

$$\begin{aligned}\Lambda_0 &= 10(\kappa(1 + 2\mu^{-1}) + \Lambda) \\ &= 10(\kappa(1 + 2L\Lambda) + \Lambda) \\ &\leq 100\kappa L\Lambda\end{aligned}$$

as long as $L \geq \kappa^{-1}$ and $\Lambda \geq 1$. Substituting these bounds back in we get

$$\begin{aligned}1 &\leq C(n, \gamma, \Gamma')\eta^{-3}\Lambda_0^2 \\ &\leq C(n, \gamma, \Gamma', \kappa)L^{-6}\Lambda^{-6} \cdot L^2\Lambda^2 \\ &= C(n, \gamma, \Gamma', \kappa)\mu^4.\end{aligned}$$

Hence there are two cases: either

$$L \geq \max\{10\kappa^{-1}\gamma(1, \ldots, 1)^{\frac{1}{2}}, \kappa^{-1}\}$$

and $\mu \geq c(n, \gamma, \Gamma', \kappa)$, or else

$$L \leq \max\{10\kappa^{-1}\gamma(1, \dots, 1)^{\frac{1}{2}}, \kappa^{-1}\}$$

and by the definition of L there holds $\mu \geq c(n, \gamma, \kappa)\Lambda^{-1}$. Putting the two cases together we get the desired estimate,

$$\mu \geq c(n, \gamma, \Gamma', \kappa)\Lambda^{-1}.$$

□

As a corollary we obtain the following result, which in some sense serves as a replacement for the exterior noncollapsing estimate available for mean curvature flow.

PROPOSITION 4.16. *Let* $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ *be an admissible solution in the sense of Definition 4.9. Then for each* $\Lambda > \kappa$ *there is a constant* $K = K(n, \gamma, \Gamma', \kappa, \varphi, \Lambda)$ *with the following property. If* $(x_0, t_0) \in \mathbf{M}$ *is such that* $r_0^{-1} := G(x_0, t_0) \geq K$ *then for each* $t \in [-\Lambda^2 r_0^2 + t_0, t_0]$*, the set*

$$\Omega_t \cap B(x_0, \Lambda r_0)$$

is connected.

PROOF. Fix $\Lambda > 1$. It suffices to show that, for an arbitrary sequence $(x_j, t_j) \in \mathbf{M}$ satisfying $r_j^{-1} := G(x_j, t_j) \to \infty$, if j is sufficiently large then

$$\Omega_t^j \cap B(x_j, \Lambda r_j)$$

is connected for every $t \in [-\Lambda^2 r_j^2 + t_j, t_j]$.

Fix $\tilde{\Lambda} > \Lambda$ to be chosen in the course of the proof. We first rescale:

$$\Omega_t^j := r_j^{-1}(\Omega_{r_j^2 t + t_j} - x_j), \qquad t \in [-r_j^{-2} t_j, 0].$$

We may assume j is so large that $[-\tilde{\Lambda}^2, 0]$ is contained in $[-r_j^{-2} t_j, 0]$. For each $t \in [-\tilde{\Lambda}^2, 0]$, let U_t^j denote the connected component of $\Omega_t^j \cap B(0, \tilde{\Lambda})$ which contains $r_j^{-1}(B(x_j', \kappa r_j) - x_j)$, where $x_j' := x_j - \kappa r_j \nu(x_j, t_j)$. Then, if j is sufficiently large relative to $\tilde{\Lambda}$, Theorem 4.15 tells us that

$$G(x, t) \geq c_0(n, \gamma, \Gamma', \kappa)\tilde{\Lambda}^{-1}$$

for each $x \in U_t^j$ and $t \in [-\tilde{\Lambda}^2, 0]$. Let us define

$$f_j(t) := \mathrm{dist}(\partial U_t^j, 0),$$

and observe that this function Lipschitz continuous on $[-\tilde{\Lambda}^2, 0]$. At a time of differentiability $\bar{t}$ for f_j, if $\bar{x} \in \partial U^j \cap B(0, \tilde{\Lambda})$ is such that $f_j(\bar{t}) = |\bar{x}|$ then there holds

$$f_j'(\bar{t}) \leq -G(\bar{x}, \bar{t}) \leq -c_0 \tilde{\Lambda}^{-1}.$$

Since $f_j(t) \to 0$ as $t \to 0$, we conclude that at time $t = -\tilde{\Lambda}^2$, U_t^j contains the ball

$$B' := B(0, c_0 \tilde{\Lambda}).$$

Let us define y_j to be the image of x_j' under the rescaling, i.e.,

$$y_j := r_j^{-1}(x_j' - x_j).$$

We may assume $\tilde{\Lambda} \geq 2\kappa$ so that each of the balls $B(y_j, \kappa)$ is contained in $B(0, \tilde{\Lambda})$. We claim that if j is sufficiently large and $x \in \Omega^j_\tau \cap B'$ for some $\tau \in [-\tilde{\Lambda}^2, 0]$ then $\overline{y_j x} \subset \Omega^j_\tau \cap B'$. In fact we prove the stronger statement that

$$C(x, y_j, \kappa/2) \subset \Omega^j_\tau \cap B'.$$

Observe that if $x \in \Omega^j_\tau \cap B'$ then there is a small positive ρ such that

$$B(x, \rho) \subset \Omega^j_\tau \cap B'.$$

We know by the previous paragraph that $B' \subset \Omega_{-\tilde{\Lambda}^2}$, so if it is the case that

$$C(x, y_j, \kappa/2) \not\subset \Omega^j_\tau \cap B',$$

then there must be some final time $\tilde{\tau} \leq \tau$ such that

$$C(x, y_j, \kappa/2) \subset \Omega^j_{\tilde{\tau}} \cap B'.$$

In particular, $\partial C(x, y_j, \kappa/2)$ touches $\partial \Omega^j_{\tilde{\tau}} \cap B'$ from the inside. Let the point of contact be denoted by y.

Since $B(x, \rho)$ and $B(y_j, \kappa)$ are both contained in $\Omega^j_{\tilde{\tau}}$ the point y must lie in $S(x, y_j, \kappa/2)$. Furthermore, there is a path from y_j to y which stays inside the pseudocone, so y and y_j are in the same connected component of $\Omega^j_{\tilde{\tau}} \cap B'$, which means that $y \in \partial U^j_{\tilde{\tau}}$. On the other hand, by Lemma 4.13 we have an upper bound for the curvature on $\partial U^j_{\tilde{\tau}}$ which is independent of j, so the almost-convexity property implies that if j is large then the second fundamental form of the solution is becoming nonnegative at $(y, \tilde{\tau})$. Since the smallest principal curvature of $S(x, y_j, \kappa/2)$ is bounded from above by a negative constant independent of j, if j is large then we have a contradiction to the interior contact at y.

To recap, we have shown that

$$\overline{x y_j} \subset C(x, y_j, \kappa/2) \subset \Omega^j_\tau \cap B'$$

for every $x \in \Omega^j_\tau$ and $\tau \in [-\tilde{\Lambda}^2, 0]$. In particular, this means that $\Omega^j_t \cap B'$ has a single connected component for each $t \in [-\tilde{\Lambda}^2, 0]$, provided that j is sufficiently large. Therefore, if we take

$$\tilde{\Lambda} \geq c_0^{-1} \Lambda,$$

then $B' = B(0, \Lambda)$ and we have that $\Omega^j_t \cap B(0, \Lambda)$ has a single connected component for each $t \in [-\Lambda^2, 0]$, provided that j is sufficiently large. This completes the proof. $\square$

Combining the last result with the upper and lower curvature estimates in Lemma 4.13 and Theorem 4.15, we finally obtain the following Harnack inequality. For the mean curvature flow there is a result of this kind in [**HK17a**], but the method used there does not yield any information about how the constants depend on Λ. It would be interesting to know whether the dependence of the upper bound on Λ can be improved.

THEOREM 4.17. *Let $\mathbf{\Omega} = \{\Omega_t : t \in [0, T)\}$ be an admissible solution in the sense of Definition 4.9. Then for every $\Lambda > \kappa$ there is a constant $K = K(n, \gamma, \Gamma', \kappa, \varphi, \Lambda)$ with the following property. If $r_0^{-1} := G(x_0, t_0) \geq K$ then for each*

$$(x, t) \in \mathbf{M} \cap \Big(B(x_0, \Lambda r_0) \times [-\Lambda^2 r_0^2 + t_0, t_0]\Big),$$

there holds

$$C^{-1}\Lambda^{-1}r_0^{-1} \leq G(x,t) \leq C\Lambda^2 r_0^{-1},$$

where $C = C(n, \gamma, \Gamma', \kappa)$.

As a consequence, we obtain higher derivative bounds by essentially the same argument used to prove Theorem 4.14.

COROLLARY 4.18. *Let* $\mathbf{\Omega} = \{\Omega_t : t \in [0,T)\}$ *be an admissible solution in the sense of Definition 4.9. Then for every* $\Lambda > \kappa$ *there is a constant* $K = K(n, \gamma, \Gamma', \kappa, \varphi, \Lambda)$ *with the following property. If* $r_0^{-1} := G(x_0, t_0) \geq K$ *then for each* $k \in \mathbb{N}$ *and*

$$(x,t) \in \mathbf{M} \cap \Big(B(x_0, \Lambda r_0) \times [-\Lambda^2 r_0^2 + t_0, t_0]\Big),$$

there holds

$$|\nabla^k A|^2(x,t) \in C(n,k,\gamma,\Gamma',\kappa) r_0^{-2k-2}.$$

APPENDIX A

Semiconvex functions

Let (M, g) be a smooth Riemannian manifold, and $U \subset M$ an open subset. We say that $f : U \to \mathbb{R}$ is semiconvex if there is a smooth function $p : U \to \mathbb{R}$ such that $f + p$ is convex. We say that f is locally semiconvex if for each $x \in U$ there is a small ball around x on which f is semiconvex.

Alexandrov showed that a convex function on an open subset of Euclidean space is almost-everywhere twice differentiable (see Section 6.4 of [**EG15**]). Since a semiconvex function is the difference between a convex and a smooth function, semiconvex functions also have this property. We also have the following result, which is proven in Section 6.3 of [**EG15**].

LEMMA A.1. *Let U be an open subset of $\mathbb{R}^n$ and suppose $f : U \to \mathbb{R}$ is convex. For each pair of indices i and j, there is a Radon measure μ_{ij} on U such that for every $\varphi \in C_0^2(U)$,*

$$\int_U f \frac{\partial^2 \varphi}{\partial x^i \partial x^j}\, dx = \int_U \varphi \, d\mu_{ij}.$$

Moreover, the density of the absolutely continuous part of μ_{ij} is $\frac{\partial^2 f}{\partial x^i \partial x^j}$.

The hypothesis on f can be weakened to local semiconvexity using a covering argument. Hence we can prove the following:

LEMMA A.2. *Let (M, g) be a Riemannian manifold and consider a compactly supported locally semiconvex function f defined on M. Then for each smooth vectorfield V on M, there is a Radon measure μ_V with the property that*

$$\int_M \varphi \nabla^2 f(V, V)\, d\mu_g + \int_M \varphi \, d\mu_V = -\int_M \nabla_k(\varphi \cdot V \otimes V)^{kl} \nabla_l f \, d\mu_g$$

holds for every $\varphi \in C_0^2(M)$.

PROOF. Let $\{U_\alpha\}$ be an open cover of the support of f such that each U_α admits a coordinate chart (so we can identify U_α with an open subset of $\mathbb{R}^n$). Let $\{\zeta_\alpha\}$ be a family of nonnegative smooth functions such that ζ_α is compactly supported in U_α, and the ζ_α are a partition of unity on $\mathrm{sp}(f)$. Composing f with a coordinate chart yields a locally semiconvex function on Euclidean space, so by Alexandrov's theorem, f is almost-everywhere twice differentiable. In addition, for each pair of indices k and l, the previous lemma tells us that there is a (singular) Radon measure μ_{kl}^α on U_α with the property that if $\zeta \in C_0^2(U_\alpha)$ then

$$\int_{U_\alpha} \zeta \frac{\partial^2 f}{\partial x^k \partial x^l}\, dx + \int_{U_\alpha} \zeta \, d\mu_{kl}^\alpha = \int_{U_\alpha} \frac{\partial^2 \zeta}{\partial x^k \partial x^l} f \, dx.$$

Since f is locally Lipschitz, we may also write this as

$$\int_{U_\alpha} \zeta \frac{\partial^2 f}{\partial x^k \partial x^l}\, dx + \int_{U_\alpha} \zeta \, d\mu^\alpha_{kl} = -\int_{U_\alpha} \frac{\partial \zeta}{\partial x^k} \frac{\partial f}{\partial x^l}\, dx.$$

If V is a smooth vectorfield defined in U_α, we apply this formula with

$$\zeta = \zeta_\alpha \varphi V^k V^l \sqrt{\det g}$$

to obtain:

$$\begin{aligned}
\int_{U_\alpha} \zeta_\alpha \varphi \frac{\partial^2 f}{\partial x^k \partial x^l} V^k V^l \sqrt{\det g}\, dx &+ \int_{U_\alpha} \zeta_\alpha \varphi V^k V^l \sqrt{\det g}\, d\mu^\alpha_{kl} \\
&= -\int_{U_\alpha} \frac{\partial \zeta_\alpha}{\partial x^k} \cdot \varphi V^k V^l \frac{\partial f}{\partial x^l} \sqrt{\det g}\, dx \\
&\quad - \int_{U_\alpha} \zeta_\alpha \frac{1}{\sqrt{\det g}} \frac{\partial}{\partial x^k} (\varphi V^k \sqrt{\det g}) V^l \frac{\partial f}{\partial x^l} \sqrt{\det g}\, dx \\
&\quad - \int_{U_\alpha} \zeta_\alpha \varphi V^k \frac{\partial V^l}{\partial x^k} \frac{\partial f}{\partial x^l} \sqrt{\det g}\, dx \\
&= \int_{U_\alpha} \varphi V^k V^l \nabla_k f \nabla_l \zeta_\alpha \, d\mu - \int_{U_\alpha} \zeta_\alpha \operatorname{div}(\varphi V) V^l \nabla_l f \, d\mu \\
&\quad - \int_{U_\alpha} \zeta_\alpha \varphi V^k \nabla_k V^l \nabla_l f \, d\mu + \int_{U_\alpha} \zeta_\alpha \varphi \Gamma^l_{km} \frac{\partial f}{\partial x^l} V^k V^m \sqrt{\det g}\, dx,
\end{aligned}$$

We expand the divergence as $\operatorname{div}(\varphi V) = V^k \nabla_k \varphi + \varphi \nabla_k V^k$ and rearrange to get

$$\begin{aligned}
\int_{U_\alpha} \zeta_\alpha \varphi \nabla^2 f(V, V)\, d\mu &+ \int_{U_\alpha} \zeta_\alpha \varphi V^k V^l \sqrt{\det g}\, d\mu^\alpha_{kl} \\
&= \int_{U_\alpha} \varphi V^k V^l \nabla_k f \nabla_l \zeta_\alpha \, d\mu - \int_{U_\alpha} \zeta_\alpha V^k V^l \nabla_k \varphi \nabla_l f \, d\mu \\
&\quad - \int_{U_\alpha} \zeta_\alpha \varphi (\nabla_k V^k V^l + V^k \nabla_k V^l) \nabla_l f \, d\mu \\
&= \int_{U_\alpha} \varphi V^k V^l \nabla_k f \nabla_l \zeta_\alpha \, d\mu - \int_{U_\alpha} \zeta_\alpha \nabla_k (\varphi \cdot V \otimes V)^{kl} \nabla_l f \, d\mu.
\end{aligned}$$

Summing over α, we obtain

$$\begin{aligned}
\int_M \varphi \nabla^2 f(V, V)\, d\mu &+ \sum_\alpha \int_{U_\alpha} \zeta_\alpha \varphi V^k V^l \sqrt{\det g}\, d\mu^\alpha_{kl} \\
&= -\int_M \nabla_k (\varphi \cdot V \otimes V)^{kl} \nabla_l f \, d\mu,
\end{aligned}$$

so it suffices to define

$$\mu_V(U) = \sum_\alpha \int_{U \cap U_\alpha} \zeta_\alpha V^k V^l \sqrt{\det g} \cdot \mu^\alpha_{kl}$$

for each measurable $U \subset M$. □

This integration-by-parts formula implies the following inequality, which is used in the proof of the convexity estimate in Chapter 3. Here $\gamma : \Gamma \to \mathbb{R}$ can be any smooth symmetric function which is elliptic (increasing in each of its arguments).

LEMMA A.3. *Let M be a smooth hypersurface with principal curvatures in Γ. Suppose $f : M \to \mathbb{R}$ is locally semiconvex and $\varphi : M \to \mathbb{R}$ is smooth, nonnegative, Lipschitz continuous and compactly supported. Then there holds*

$$\int_M \varphi \Delta_\gamma f \, d\mu \leq -\int_M \nabla_i \dot{\gamma}^{ij} \nabla_j f \, d\mu - \int_M \dot{\gamma}^{ij} \nabla_i \varphi \nabla_j f \, d\mu,$$

where μ denotes the induced measure.

PROOF. It suffices to consider $\varphi \in C_0^2(M)$, since the general case then follows by approximation. Let $\{U_\alpha\}$ and $\{\zeta_\alpha\}$ be as in the last proof, but where the ζ_α now form a partition of unity on $\mathrm{sp}(\varphi)$. Consider a fixed U_α equipped with coordinates $\{x^i\}$ and let $\{e^i\}$ denote the associated smooth frame of coordinate one-forms. Since γ is increasing in each of its arguments the mapping

$$(\omega, \omega) \mapsto \dot{\gamma}(\omega, \omega) := \dot{\gamma}^{ij} \omega_i \omega_j$$

defines a smooth inner product of one-forms, so we can apply the Gram-Schmidt algorithm to produce from $\{e^i\}$ a smooth frame $\{\tilde{e}^i\}$ of one-forms such that

$$\dot{\gamma}(\tilde{e}^i, \tilde{e}^j) = \delta^{ij}.$$

We now define a local frame $\{\tilde{e}_i\}$ for the tangent bundle by the condition

$$\tilde{e}^i(\tilde{e}_j) = \delta^i_j.$$

With respect to this basis,

$$\dot{\gamma} = \dot{\gamma}(\tilde{e}^i, \tilde{e}^j) \tilde{e}_i \otimes \tilde{e}_j = \sum_i \tilde{e}_i \otimes \tilde{e}_i,$$

so we can express

$$\Delta_\gamma f = \dot{\gamma}^{ij} \nabla_i \nabla_j f = \sum_i \nabla^2 f(\tilde{e}_i, \tilde{e}_i).$$

Since $\varphi\zeta_\alpha$ is nonnegative, we can now use the result from above to estimate

$$\begin{aligned}
\int_M \varphi\zeta_\alpha \Delta_\gamma f \, d\mu &= \sum_i \int_M \varphi\zeta_\alpha \nabla^2 f(\tilde{e}_i, \tilde{e}_i) \, d\mu \\
&\leq \sum_i \int_M \varphi\zeta_\alpha \nabla^2 f(\tilde{e}_i, \tilde{e}_i) \, d\mu + \sum_i \int_M \varphi\zeta_\alpha \, d\mu_{\tilde{e}_i} \\
&= -\sum_i \int_M \nabla_k (\varphi\zeta_\alpha \cdot \tilde{e}_i \otimes \tilde{e}_i)^{kl} \nabla_l f \, d\mu \\
&= -\int_M \nabla_k (\varphi\zeta_\alpha \cdot \dot{\gamma})^{kl} \nabla_l f \, d\mu.
\end{aligned}$$

We thus have

$$\int_M \varphi\zeta_\alpha \Delta_\gamma f \, d\mu \leq -\int_M \varphi\zeta_\alpha \nabla_i \dot{\gamma}^{ij} \nabla_j f \, d\mu - \int_M \dot{\gamma}^{ij} \nabla_i (\varphi\zeta_\alpha) \nabla_j f \, d\mu,$$

and since this is true for every α we can sum up to obtain

$$\int_M \varphi \Delta_\gamma f \, d\mu \leq -\int_M \nabla_i \dot{\gamma}^{ij} \nabla_j f \, d\mu - \int_M \dot{\gamma}^{ij} \nabla_i \varphi \nabla_j f \, d\mu.$$

$\square$

APPENDIX B

Regularity of graphical solutions

In this section, we establish a number of technical results for graphical hypersurfaces moving by a concave function of the principal curvatures. The graph condition allows us to represent the solution purely in terms of a scalar function, which can be shown to satisfy a fully nonlinear parabolic equation. Using the regularity theory for such equations, it is then possible to establish bounds for derivatives of the curvature. The results here form the technical basis needed to prove the curvature derivative estimates in Chapter 4.

What follows is the famous Krylov-Safonov estimate for solutions of linear parabolic equations with bounded coefficients [**KS80b**][Section 4].

Theorem B.1. *Let $u : B(0,1) \times [-1,0] \to \mathbb{R}$ be a $C^{2;1}$ solution of the equation*

$$\partial_t u = a^{ij}(x,t) D_i D_j u + b^i(x,t) D_i u + f(x,t)$$

with a^{ij}, b^i and f bounded. Suppose the equation is uniformly parabolic, meaning that

$$\lambda |\xi|^2 \leq a^{ij}(x,t)\xi_i \xi_j \leq \Lambda |\xi|^2,$$

and that

$$\sup_{B(0,1)\times[-1,0]} |b(x,t)| \leq K.$$

Then for each $\theta \in (0,1)$ there holds

$$\sup_{(x,s)\neq(y,t)\in B(0,\theta)\times[-\theta,0]} \frac{|u(x,s)-u(y,t)|}{|x-y|^\alpha + |s-t|^{\frac{\alpha}{2}}} \leq C \sup_{B(0,1)\times[-T,0]} (|u| + |f|)$$

where α depends only on n, λ and Λ, and C depends on all of these quantities, and additionally on K and θ.

The next theorem was proven independently by Krylov [**Kry82**][Theorem 2.1] and Evans [**Eva82**][Theorem 1]. Both of these works make essential use of the Krylov-Safonov Hölder-estimate. We have not tried to state the theorem in the fullest possible generality - the following version suffices for our purposes.

Theorem B.2. *Let $u \in C^4(B(0,1)\times[-1,0])$ be a solution of the equation*

$$\partial_t u = \Phi(D^2 u, Du, u, x),$$

where $\Phi \in C^2(U \times B(0,1))$, and U is an open subset of $\mathrm{Sym}(n)\times\mathbb{R}^n\times\mathbb{R}$. We assume that the equation is uniformly parabolic, in the sense that there are positive constants λ and Λ for which

$$\lambda|\xi|^2 \leq \frac{\partial\Phi}{\partial A_{ij}}(A,p,z,x)\xi_i\xi_j \leq \Lambda|\xi|^2$$

for every $(A, p, z, x) \in U \times B(0,1)$ *and* $\xi \in \mathbb{R}^n$. *We also assume that* F *is convex in its first argument, by which we mean that*

$$\frac{\partial^2 \Phi}{\partial A_{ij} \partial A_{kl}}(A, p, z, x) T_{ij} T_{kl} \geq 0$$

holds for all $(A, p, z, x) \in U \times B(0,1)$ *and* $T \in \mathrm{Sym}(n)$, *and finally, that there is a* K *such that*

$$\|\Phi\|_{C^2(U \times B(0,1))} \leq K.$$

Then for each $\theta \in (0,1)$, *there holds*

$$\sup_{(x,s)\neq(y,t)\in B(0,\theta)\times[-\theta,0]} \left(\frac{|\partial_t u(x,s) - \partial_t u(y,t)|}{|x-y|^\alpha + |s-t|^{\frac{\alpha}{2}}} + \frac{|D^2 u(x,s) - D^2 u(y,t)|}{|x-y|^\alpha + |s-t|^{\frac{\alpha}{2}}} \right) \leq C,$$

where α *depends only on* n, λ, Λ, K, *and*

$$\sup_{B(0,1)\times[-1,0]} |\partial_t u| + |D^2 u| + |Du|,$$

and C *depends on all of these same quantities, and on* θ.

A standard bootstrap argument employing the Schauder estimates for linear parabolic equations (see for example [**Lie96**][Theorem 4.9]) now yields higher regularity:

Corollary B.3. *Let* u *and* Φ *now be smooth functions, which otherwise satisfy all of the same assumptions as in the theorem. Suppose in addition that for each* $k \in \mathbb{N}$, *the* C^k*-norm of* Φ *over the set* $U \times B(0,1)$ *is bounded by a constant* K_k. *Then, for each* $k \geq 3$ *there is a constant* L_k *depending only on* n, λ, Λ, K_{k-1} *and*

$$\sup_{B(0,1)\times[-1,0]} |\partial_t u| + |D^2 u| + |Du|$$

such that the C^k*-norm of* u *over* $B(0,1/2) \times [-1/2, 0]$ *is bounded by* L_k.

Now we will apply the general results stated above to graphical hypersurfaces moving by curvature. Let $\gamma : \Gamma \to (0,\infty)$ be an admissible speed, which we also assume to be concave. If M is a smooth hypersurface with principal curvatures $\lambda \in \Gamma$ then we write $G(x)$ for $\gamma(\lambda(x))$.

Proposition B.4. *Fix a positive constant* r *and consider a smooth function*

$$u : B(0,r) \times [-r^2, 0] \to \mathbb{R},$$

where $B(0,r) \subset \mathbb{R}^n$. *Identify* $\mathbb{R}^n$ *with* $e_{n+1}^\perp$ *in* $\mathbb{R}^{n+1}$ *and suppose the family of hypersurfaces defined by*

$$F(x,t) := x + u(x,t) e_{n+1}, \qquad (x,t) \in B(0,r) \times [-r^2, 0]$$

solves the equation

$$(\partial_t F(x,t))^\perp = -G(x,t).$$

Suppose further that

$$\nu(x,t) \cdot e_{n+1} \geq \theta, \qquad |u| \leq Lr,$$

where ν *is the upward-pointing unit normal. Finally, suppose that for each* $(x,t) \in B(0,r) \times [-r^2, 0]$ *we have upper and lower curvature bounds and curvature pinching:*

$$K_0^{-1} r^{-1} \leq G(x,t) \leq K_1 r^{-1}, \qquad \lambda(x,t) \in \Gamma' \Subset \Gamma.$$

Then for each $k \in \mathbb{N}$ and $(x,t) \in B(0,r/2) \times [-r^2/4, 0]$ there holds

$$|D^k u|^2(x,t) \le C(n,k,\gamma,\Gamma',\theta,L,K_0,K_1) r^{-2k+2}$$

Equivalently, for each $k \in \mathbb{N}$ and $(x,t) \in B(0,r/2) \times [-r^2/4, 0]$ there holds

$$|\nabla^k A|_g^2(x,t) \le C(n,k,\gamma,\Gamma',\theta,L,K_0,K_1) r^{-2k-2},$$

where g denotes the induced metric on the graph of u.

Proof. If we can prove the estimates for $r = 1$, then the general case follows by scaling, so let us assume $r = 1$. In order to apply the regularity theory for scalar parabolic equations, we need to convert the evolution equation for F into an equation for u. For this we need expressions for the metric, inverse-metric, and normal in terms of u:

$$\begin{aligned} g_{ij} &= \delta_{ij} + D_i u D_j u; \\ g^{ij} &= \delta_{ij} - \frac{D_i u D_j u}{1 + |Du|^2}; \\ \nu &= \frac{-Du + e_{n+1}}{\sqrt{1 + |Du|^2}}. \end{aligned}$$

It follows that the coefficients of the second fundamental form and Weingarten map can be expressed as:

$$A_{ij} = -\frac{D_i D_j u}{\sqrt{1 + |Du|^2}}; \qquad g^{ik} A_{kj} = -\left(\delta_{ik} - \frac{D_i u D_k u}{1 + |Du|^2}\right) \frac{D_k D_j u}{\sqrt{1 + |Du|^2}}.$$

From these formulae, we see that

$$(\partial_t F(x,t))^\perp = -G(x,t)$$

holds if and only if

$$\partial_t u = -\sqrt{1 + |Du|^2}\,\gamma\left(-\left(\delta_{ik} - \frac{D_i u D_k u}{1 + |Du|^2}\right) \frac{D_k D_j u}{\sqrt{1 + |Du|^2}}\right).$$

Note that if A is any diagonalisable matrix with eigenvalues λ in Γ, we define $\gamma(A)$ to be $\gamma(\lambda)$. The matrix $g^{ik} A_{kj}$ is not, in general, symmetric, and it is convenient to replace it with a symmetric matrix having the same eigenvalues. To achieve this we borrow a trick from Urbas [**Urb91**][Equation 2.21] and define

$$P_{ij} = \delta_{ij} - \frac{D_i u D_j u}{\sqrt{1 + |Du|^2}(1 + \sqrt{1 + |Du|^2})}, \qquad \hat{A}_{ij} = P_{ik} A_{kl} P_{lj}.$$

Then $\hat{A}_{ij}$ is symmetric, and since $P_{ij}^2 = g^{ij}$, one can show that v is an eigenvector of $g^{ik} A_{kj}$ with eigenvalue λ if and only if $P^{-1} v$ is an eigenvector of $\hat{A}$ with eigenvalue λ. Since the value of γ depends only on the eigenvalues of its argument, we may write

$$\partial_t u = -\sqrt{1 + |Du|^2}\,\gamma\left(- P_{ik} \frac{D_k D_l u}{\sqrt{1 + |Du|^2}} P_{lj}\right).$$

The uniform graph property implies that

$$\sqrt{1 + |Du|^2} \le \theta^{-1},$$

so there is a constant C depending only on θ such that

$$C^{-1}\delta_{ij} \leq g_{ij} \leq C\delta_{ij}, \qquad C^{-1}\delta_{ij} \leq g^{ij} \leq C\delta_{ij}.$$

These bounds imply that

$$|D^2u|^2 \leq C(\theta)|A|^2 \leq C(n,\gamma,\Gamma',\theta)K_1^2,$$

and by assumption

$$|\partial_t u| = \sqrt{1+|Du|^2}|G| \leq \theta^{-1}K_1,$$

so we have

$$|\partial_t u| + |u| + |Du| + |D^2u| \leq C(n,\gamma,\Gamma',\theta,L,K_1).$$

Let us define functions

$$\begin{aligned} P : B(0,2\sqrt{\theta^{-2}-1}) &\to \mathrm{Sym}(n) \\ p &\mapsto \delta_{ij} - \frac{p_ip_j}{\sqrt{1+|p|^2}(1+\sqrt{1+|p|^2})} \end{aligned}$$

and

$$\begin{aligned} \Psi : \mathrm{Sym}(n) \times B(0,2\sqrt{\theta^{-2}-1}) &\to \mathrm{Sym}(n) \\ (A,p) &\mapsto P_{ik}(p)\frac{A_{kl}}{\sqrt{1+|p|^2}}P_{lj}(p). \end{aligned}$$

Let U be the preimage of the set

$$\{S \in \mathrm{O}\,\Gamma' : (2K_0)^{-1} < \gamma(S) < 2K_1\}$$

under Ψ. Then on the set $\Psi(U)$, the derivatives of γ up to all orders can be bounded purely in terms of n, K_0, K_1 and Γ', and this implies that all of the derivatives of

$$\begin{aligned} \Phi : \Psi^{-1}(U) &\to (0,\infty) \\ (A,p) &\mapsto -\gamma(-\Psi(A,p)) \end{aligned}$$

are bounded as well. Also, Φ is convex in A, since Ψ is linear in A and γ is concave. Next we want to verify that the operator

$$\Phi(D^2u, Du) := -\gamma(-\Psi(D^2u,Du))$$

is uniformly elliptic. To this end, we fix A and p and compute

$$\begin{aligned} \frac{\partial\Phi}{\partial A_{pq}}(A,p) &= \frac{\partial}{\partial A_{pq}}\left(P_{ik}\frac{A_{kl}}{\sqrt{1+|p|^2}}P_{lj}\right)\dot\gamma^{ij}(-\Psi(A,p)) \\ &= \frac{1}{\sqrt{1+|p|^2}}P_{ip}P_{qj}\dot\gamma^{ij}(-\Psi(A,p)), \end{aligned}$$

and so find that for each $\xi \in \mathbb{R}^n$,

$$\frac{\partial\Phi}{\partial A_{pq}}(A,p)\xi_p\xi_q = \frac{1}{\sqrt{1+|p|^2}}\gamma^{ij}(-\Psi(A,p))(P\xi)_i(P\xi)_j \geq c(n,\gamma,\Gamma',\theta)|\xi|^2.$$

Hence u and Φ satisfy all of the hypotheses of Corollary B.3 in $B(0,1)\times[-1,0]$. We conclude that for each $k \in \mathbb{N}$ there holds

$$|u|_{C^k(B(0,1/2)\times[-1/2,0])} \leq C(n,k,\gamma,\Gamma',\theta,L,K_0,K_1),$$

and this implies that for each $k \in \mathbb{N}$, on the graph of u restricted to $B(0,1/2) \times [-1/2,0]$ there holds

$$|\nabla^k A|_g^2 \leq C(n,k,\gamma,\Gamma',\theta,L,K_0,K_1).$$

$\square$

The last result assumed a lower bound for G over the domain of u. If we only assume a lower bound for G at the spacetime origin, the result still holds on a small neighbourhood.

LEMMA B.5. *Let u and F be as in Proposition B.4, except that we only assume the lower bound*

$$G \geq K_0^{-1} r^{-1}$$

to be true at the point $(x,t) = (0,0)$. Then there is a positive

$$\delta = \delta(n,\gamma,\Gamma',\theta,K_0,K_1)$$

such that

$$\inf_{B(0,\delta r)\times[-\delta^2 r^2,0]} G \geq (2K_0)^{-1} r^{-1}.$$

PROOF. We can assume $r = 1$, since the general case then follows by scaling. We are going to apply the Krylov-Safonov estimate to the function $v := -\partial_t u$. As before, we write

$$G(x,t) = \gamma\bigg(-P_{ik}\frac{D_k D_l u}{\sqrt{1+|Du|^2}} P_{lj}\bigg).$$

Differentiating the equation for u in time gives

$$\begin{aligned}
\partial_t v &= \partial_t(\sqrt{1+|Du|^2}G)\\
&= G\frac{Du\cdot D\partial_t u}{\sqrt{1+|Du|^2}} - \dot\gamma^{ij}P_{ik}P_{lj}D_kD_l\partial_t u - 2\dot\gamma^{ij}\partial_t P_{ik}P_{lj}D_kD_l u\\
&\quad - \sqrt{1+|Du|^2}\dot\gamma^{ij}P_{ik}P_{lj}D_kD_lu\cdot\partial_t\frac{1}{\sqrt{1+|Du|^2}}.
\end{aligned}$$

Using the one-homogeneity of γ we find that the last term is equal to

$$\sqrt{1+|Du|^2}\dot\gamma^{ij}P_{ik}P_{lj}D_kD_lu\frac{Du\cdot D\partial_t u}{(1+|Du|^2)^{\frac{3}{2}}} = G\frac{Du\cdot D\partial_t u}{\sqrt{1+|Du|^2}},$$

so we have

$$\partial_t v = \dot\gamma^{ij}P_{ik}P_{lj}D_kD_l v - 2\dot\gamma^{ij}\partial_t P_{ik}P_{lj}D_kD_lu - 2G\frac{Du\cdot Dv}{\sqrt{1+|Du|^2}}.$$

Differentiating P_{ik} in time yields

$$\begin{aligned}
\partial_t P_{ik} &= -\frac{D_iuD_k\partial_t u}{\sqrt{1+|Du|^2}(1+\sqrt{1+|Du|^2})}\\
&\quad + \frac{D_iuD_ku}{(1+|Du|^2)(1+\sqrt{1+|Du|^2})^2}\partial_t(\sqrt{1+|Du|^2}(1+\sqrt{1+|Du|^2})),
\end{aligned}$$

which we can simply write as

$$\partial_t P_{ik} =: \zeta_i D_k v + \tilde{b}^r_{ik} D_r v,$$

where $|\zeta_i| \leq C(\theta)$ and $|\tilde{b}^r_{ik}| \leq C(n, \gamma, \Gamma', \theta, K_0)$. Hence

$$\begin{aligned}\partial_t v &= \dot{\gamma}^{ij} P_{ik} P_{lj} D_k D_l v - 2\dot{\gamma}^{ij} P_{lj} D_k D_l u(\zeta_i D_k v + \tilde{b}^r_{ik} D_r v) - 2G \frac{Du \cdot Dv}{\sqrt{1 + |Du|^2}} \\ &= a^{kl} D_k D_l v + b^k D_k v,\end{aligned}$$

where we have defined:

$$\begin{aligned}a^{kl} &:= \dot{\gamma}^{ij} P_{ik} P_{lj}; \\ b^k &= -2\dot{\gamma}^{ij} P_{lj} D_k D_l u \cdot \zeta_i - 2\dot{\gamma}^{ij} P_{lj} D_r D_l u \cdot \tilde{b}^k_{ir} - 2\frac{G}{\sqrt{1 + |Du|^2}} D_k u.\end{aligned}$$

The assumption that λ takes values in Γ' and the uniform graph property imply uniform ellipticity,

$$C(n, \gamma, \Gamma', \theta)^{-1} \delta_{kl} \leq a^{kl} \leq C(n, \gamma, \Gamma', \theta) \delta_{kl},$$

and we also have an estimate of the form

$$|b^k| \leq C(n, \gamma, \Gamma', \theta, K_1).$$

Thus the equation solved by v satisfies all of the hypotheses of Theorem B.1, and v itself is bounded in terms of θ and K_1, so there holds

$$\sup_{(x,s) \neq (y,t) \in B(0,1/2) \times [-1/2, 0]} \frac{|v(x, s) - v(y, t)|}{|x - y|^\alpha + |s - t|^{\frac{\alpha}{2}}} \leq C(n, \gamma, \Gamma', \theta, K_1).$$

In particular, we have

$$|v(x, t) - v(0, 0)| \leq C(|x|^\alpha + |t|^{\frac{\alpha}{2}}).$$

Since $v = \sqrt{1 + |Du|^2} \cdot G$ and we have a C^2-bound for u this implies

$$|G(x, t) - G(0, 0)| \leq C(|x|^\alpha + |t|^{\frac{\alpha}{2}}).$$

Since $G(0, 0) \geq K_0^{-1}$, we can find a uniform spacetime set where $G(x, t) \geq (2K_0)^{-1}$, as required. $\square$

Bibliography

[ADS18] Sigurd B Angenent, Panagiota Daskalopoulos, and Natasa Sesum. Uniqueness of two-convex closed ancient solutions to the mean curvature flow. *arXiv preprint arXiv:1804.07230*, 2018.

[AL14] Ben Andrews and Mat Langford. Cylindrical estimates for hypersurfaces moving by convex curvature functions. *Anal. PDE*, 7(5):1091–1107, 2014.

[AL16] Ben Andrews and Mat Langford. Two-sided non-collapsing curvature flows. *Annali della Scuola Normale Superiore di Pisa. Classe di scienze*, 15(1):543–560, 2016.

[ALM13] Ben Andrews, Mat Langford, and James McCoy. Non-collapsing in fully non-linear curvature flows. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 30(1):23–32, 2013.

[ALM14] Ben Andrews, Mathew Langford, and James McCoy. Convexity estimates for hypersurfaces moving by convex curvature functions. *Anal. PDE*, 7(2):407–433, 2014.

[ALM15] Ben Andrews, Mat Langford, and James McCoy. Convexity estimates for surfaces moving by curvature functions. *Journal of Differential Geometry*, 99(1):47–75, 2015.

[AMZ13] Ben Andrews, James McCoy, and Yu Zheng. Contracting convex hypersurfaces by curvature. *Calc. Var. Partial Differential Equations*, 47(3-4):611–665, 2013.

[And94a] Ben Andrews. Contraction of convex hypersurfaces in euclidean space. *Calculus of Variations and Partial Differential Equations*, 2(2):151–171, 1994.

[And94b] Ben Andrews. Contraction of convex hypersurfaces in riemannian spaces. *Journal of Differential Geometry*, 39(2):407–431, 1994.

[And94c] Ben Andrews. Harnack inequalities for evolving hypersurfaces. *Mathematische Zeitschrift*, 217(1):179–197, 1994.

[And03] Ben Andrews. Positively curved surfaces in the three-sphere. *arXiv preprint math/0304257*, 2003.

[And04] Ben Andrews. Fully nonlinear parabolic equations in two space variables. *arXiv preprint math/0402235*, 2004.

[And07] Ben Andrews. Pinching estimates and motion of hypersurfaces by curvature functions. *J. Reine Angew. Math.*, 608:17–33, 2007.

[And12] Ben Andrews. Noncollapsing in mean-convex mean curvature flow. *Geometry & Topology*, 16(3):1413–1418, 2012.

[Ang91] Sigurd Angenent. On the formation of singularities in the curve shortening flow. *J. Differential Geom.*, 33(3):601–633, 1991.

[BC18] Simon Brendle and Kyeongsu Choi. Uniqueness of convex ancient solutions to mean curvature flow in higher dimensions. *arXiv preprint arXiv:1804.00018*, 2018.

[BC19] Simon Brendle and Kyeongsu Choi. Uniqueness of convex ancient solutions to mean curvature flow in $\mathbb{R}^3$. *Inventiones mathematicae*, 217(1):35–76, 2019.

[BH16] Simon Brendle and Gerhard Huisken. Mean curvature flow with surgery of mean convex surfaces in $\mathbb{R}^3$. *Invent. Math.*, 203(2):615–654, 2016.

[BH17] Simon Brendle and Gerhard Huisken. A fully nonlinear flow for two-convex hypersurfaces in Riemannian manifolds. *Invent. Math.*, 210(2):559–613, 2017.

[BH18] Simon Brendle and Gerhard Huisken. Mean curvature flow with surgery of mean convex surfaces in three-manifolds. *J. Eur. Math. Soc. (JEMS)*, 20(9):2239–2257, 2018.

[BL16] Theodora Bourni and Mat Langford. Type-II singularities of two-convex immersed mean curvature flow. *Geometric Flows*, 2(1), 2016.

[Bra78] Kenneth A. Brakke. *The Motion of a Surface by Its Mean Curvature. (MN-20).* Princeton University Press, 1978.

[Bre15] Simon Brendle. A sharp bound for the inscribed radius under mean curvature flow. *Inventiones mathematicae*, 202(1):217–237, 2015.

[Bre18] Simon Brendle. Ricci flow with surgery in higher dimensions. *Annals of Mathematics*, 187(1):263–299, 2018.

[BS08] Simon Brendle and Richard M. Schoen. Classification of manifolds with weakly 1/4-pinched curvatures. *Acta Math.*, 200(1):1–13, 2008.

[BS09] Simon Brendle and Richard Schoen. Manifolds with 1/4-pinched curvature are space forms. *Journal of the American Mathematical Society*, 22(1):287–307, 2009.

[CGG91] Yun Gang Chen, Yoshikazu Giga, and Shun'ichi Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *Journal of differential geometry*, 33(3):749–786, 1991.

[Cho85] Bennett Chow. Deforming convex hypersurfaces by the n-th root of the gaussian curvature. *Journal of Differential Geometry*, 22(1):117–138, 1985.

[DG60] Ennio De Giorgi. Sulla differenziabilita e l'analiticita delle estremali degli integrali multipli regolari. *Matematika*, 4(6):23–38, 1960.

[DL04] Francesca Da Lio. Remarks on the strong maximum principle for viscosity solutions to fully nonlinear parabolic equations. *Communications on Pure & Applied Analysis*, 3(3):395, 2004.

[Eck12] Klaus Ecker. *Regularity theory for mean curvature flow*, volume 57. Springer Science & Business Media, 2012.

[EG15] Lawrence Craig Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*. CRC press, 2015.

[EH91] Klaus Ecker and Gerhard Huisken. Interior estimates for hypersurfaces moving by mean curvature. *Invent. Math.*, 105(3):547–569, 1991.

[ES64] James Eells and J. H. Sampson. Harmonic mappings of riemannian manifolds. *American Journal of Mathematics*, 86(1):109–160, 1964.

[ES91] L. C. Evans and J Spruck. Motion of level sets by mean curvature. I. *Journal of Differential Geometry*, 33(3):635–681, 1991.

[Eva82] Lawrence C. Evans. Classical solutions of fully nonlinear, convex, second-order elliptic equations. *Comm. Pure Appl. Math.*, 35(3):333–363, 1982.

[Gag84] Michael E Gage. Curve shortening makes convex curves circular. *Inventiones mathematicae*, 76(2):357–364, 1984.

[GH86] Michael Gage and Richard S Hamilton. The heat equation shrinking convex plane curves. *Journal of Differential Geometry*, 23(1):69–96, 1986.

[Gla63] Georges Glaeser. Fonctions composees differentiables. *Annals of Mathematics*, 77(1):193–209, 1963.

[Gra87] Matthew A Grayson. The heat equation shrinks embedded plane curves to round points. *Journal of Differential geometry*, 26(2):285–314, 1987.

[Ham82] Richard S. Hamilton. Three-manifolds with positive ricci curvature. *Journal of Differential Geometry*, 17(2):255–306, 1982.

[Ham93] Richard Hamilton. The formation of singularities in the ricci flow. *Surveys in differential geometry*, 2(1):7–136, 1993.

[Ham95] Richard S. Hamilton. Harnack estimate for the mean curvature flow. *Journal of Differential Geometry*, 41(1):215–226, 1995.

[HK17a] Robert Haslhofer and Bruce Kleiner. Mean curvature flow of mean convex hypersurfaces. *Comm. Pure Appl. Math.*, 70(3):511–546, 2017.

[HK17b] Robert Haslhofer and Bruce Kleiner. Mean curvature flow with surgery. *Duke Math. J.*, 166(9):1591–1626, 2017.

[HS99a] Gerhard Huisken and Carlo Sinestrari. Convexity estimates for mean curvature flow and singularities of mean convex surfaces. *Acta mathematica*, 183(1):45–70, 1999.

[HS99b] Gerhard Huisken and Carlo Sinestrari. Mean curvature flow singularities for mean convex surfaces. *Calculus of Variations and Partial Differential Equations*, 8(1):1–14, 1999.

[HS09] Gerhard Huisken and Carlo Sinestrari. Mean curvature flow with surgeries of two-convex hypersurfaces. *Invent. Math.*, 175(1):137–221, 2009.

[HS15] Gerhard Huisken and Carlo Sinestrari. Convex ancient solutions of the mean curvature flow. *J. Differential Geom.*, 101(2):267–287, 2015.

[Hui84] Gerhard Huisken. Flow by mean curvature of convex surfaces into spheres. *J. Differential Geom.*, 20(1):237–266, 1984.

[Hui90] Gerhard Huisken. Asymptotic behavior for singularities of the mean curvature flow. *J. Differential Geom.*, 31(1):285–299, 1990.

[Hui98] Gerhard Huisken. A distance comparison principle for evolving curves. *Asian Journal of Mathematics*, 2(1):127–133, 1998.

[Ive93] Thomas Ivey. Ricci solitons on compact three-manifolds. *Differential Geometry and its Applications*, 3(4):301–307, 1993.

[Kry82] N. V. Krylov. Boundedly inhomogeneous elliptic and parabolic equations. *Izv. Akad. Nauk SSSR Ser. Mat.*, 46(3):487–523, 670, 1982.

[KS80a] David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 31. Siam, 1980.

[KS80b] N. V. Krylov and M. V. Safonov. A property of the solutions of parabolic equations with measurable coefficients. *Izv. Akad. Nauk SSSR Ser. Mat.*, 44(1):161–175, 239, 1980.

[Lan14] Mat Langford. *Motion of hypersurfaces by curvature*. PhD thesis, 2014.

[Lan17] Mat Langford. A general pinching principle for mean curvature flow and applications. *Calculus of Variations and Partial Differential Equations*, 56(4):107, 2017.

[Lie96] Gary M. Lieberman. *Second order parabolic differential equations*. World Scientific Publishing Co., Inc., River Edge, NJ, 1996.

[LL] Mat Langford and Stephen Lynch. Sharp one-sided curvature estimates for fully nonlinear curvature flows and applications to ancient solutions. *Journal fr die reine und angewandte Mathematik*, (to appear).

[MS73] J. H. Michael and L. M. Simon. Sobolev and mean-value inequalities on generalized submanifolds of $\mathbb{R}^n$. *Communications on Pure and Applied Mathematics*, 26(3):361–379, 1973.

[Nas58] John Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958.

[Per02] Grisha Perelman. The entropy formula for the ricci flow and its geometric applications. *arXiv preprint math/0211159*, 2002.

[Per03] Grisha Perelman. Ricci flow with surgery on three-manifolds. *arXiv preprint math/0303109*, 2003.

[Sch75] Gerald W Schwarz. Smooth functions invariant under the action of a compact lie group. *Topology*, 14(1):63–68, 1975.

[SW09] Weimin Sheng and Xu-Jia Wang. Singularity profile in the mean curvature flow. *Methods and Applications of Analysis*, 16(2):139–156, 2009.

[Urb91] John Urbas. An expansion of convex hypersurfaces. *Journal of Differential Geometry*, 33(1):91–125, 1991.

[Whi00] Brian White. The size of the singular set in mean curvature flow of mean-convex sets. *Journal of the American Mathematical Society*, 13(3):665–695, 2000.

[Whi03] Brian White. The nature of singularities in mean curvature flow of mean-convex sets. *Journal of the American Mathematical Society*, 16(1):123–138, 2003.

[Whi05] Brian White. A local regularity theorem for mean curvature flow. *Ann. of Math. (2)*, 161(3):1487–1519, 2005.