CONVEXITY ESTIMATES FOR HIGH CODIMENSION MEAN CURVATURE FLOW

STEPHEN LYNCH AND HUY THE NGUYEN

ABSTRACT. We consider the evolution by mean curvature of smooth \(n\)-dimensional submanifolds in \(\mathbb{R}^{n+k}\) which are compact and quadratically pinched. We will be primarily interested in flows of high codimension, the case \(k \geq 2\). We prove that our submanifold is asymptotically convex, that is the first eigenvalue of the second fundamental form in the principal mean curvature direction blows up at a strictly slower rate than the mean curvature vector. We use this convexity estimate to show that at a singular time of the flow, there exists a rescaling that converges to a smooth codimension-one limiting flow which is convex and moves by translation.

1. INTRODUCTION

Let us consider a compact smooth \(n\)-manifold \(M\), and a family of immersions

\[
F : M \times [0, T) \rightarrow \mathbb{R}^{n+k}
\]

which move by mean curvature flow, that is

\[
\partial_t F(x, t) = H(x, t)
\]

for each \((x, t) \in M \times [0, T)\) where \(H\) is the mean curvature vector. The mean curvature flow constitutes a system of quasilinear weakly parabolic partial differential equations for \(F\), and since \(M\) is compact the flow must form a singularity in finite time. Singularity formation may be characterised analytically as follows: if we let \(A\) denote the second fundamental form and take \(T\) to be the maximal time then there holds

\[
\limsup_{t \rightarrow T} \sup_M |A| \langle \cdot, t \rangle = \infty.
\]

A profound and challenging problem is characterising and classifying the geometry of singularities, whose formation depends on the initial submanifold \(M_0\) where \(M_t := F(M, t)\). In the seminal work of Huisken-Sinestrari [HS99], the singularities formed by mean convex codimension one solutions were shown to be weakly convex (White obtained a similar result for embedded mean convex solutions in [Whi03]). It is natural to seek a corresponding theorem for solutions of higher codimension. However, we encounter a number of new difficulties, the foremost being that the second fundamental form and the mean curvature are vector-valued and consequently there is no direct corresponding notion of mean convexity. We instead use a different but related condition introduced by Andrews-Baker [AB10]. They showed that when \(n \geq 2\), the quadratic pinching condition

\[
|A|^2 - c|H|^2 + a \leq 0
\]

is preserved for each \(c < \frac{4}{3n}\) and \(a > 0\). That is, if the condition is satisfied at the initial time, then it is satisfied by \(M_t\) for every \(t \in [0, T)\). We note that for compact hypersurfaces, positive mean curvature implies this condition but for all sufficiently large \(c\). We will refer to submanifolds satisfying (1.1) as being quadratically pinched.

Andrews-Baker showed that if \(c < \left\{\frac{4}{3n}, \frac{1}{n-1}\right\}\) the flow contracts quadratically pinched solutions to round points. This result is a high codimension generalisation of Huisken’s work on convex
solutions of codimension one [Hui84]. Recently the second-named author has constructed a flow with surgeries for solutions of dimension $n \geq 5$ which are quadratically pinched with

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

This generalises the surgery construction for two-convex hypersurface flows due to Huisken-Sinestrari [HS09]. An important ingredient in [Ngu20] (and in the present work) is the codimension estimate due to Naff, which implies the singularities formed by a quadratically pinched solution are codimension one if $c < c_n$.

In this paper we show that a quadratically pinched mean curvature flow with $c < c_n$ is asymptotically convex in a quantifiable manner. As the second fundamental form is vector-valued, we denote by $\lambda_1 \leq \cdots \leq \lambda_n$ the eigenvalues of the second fundamental form in the principal normal direction, that is $\nu_1 = \frac{H}{|H|}$ (see Section 2 for precise definitions).

**Theorem 1.1.** Let $F : M \times [0, T) \to \mathbb{R}^{n+k}$ be a mean curvature flow of dimension $n \geq 5$ which is quadratically pinched with $c < c_n$. Then for any $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ which depends only on $n$, $\varepsilon$ and $M_0$ such that

$$\lambda_1 \geq -\varepsilon |H| - C_\varepsilon$$

on $M_t$, for each $t \in [0, T)$.

As $\varepsilon > 0$ is arbitrary, this shows that the negative part of the first eigenvalue in the principal normal direction does not grow as fast as $|H|$. We then use this estimate to characterise type II singularities of the quadratically pinched mean curvature flow near the maximum of the curvature:

**Theorem 1.2.** Suppose a type II singularity forms at time $T$, that is

$$\limsup_{t \to T} [(T - t) \max_{M_t} |A|^2] = \infty.$$ 

Then there exists a sequence of rescalings of $F$ that subconverges smoothly to a codimension one limiting flow which is either: a strictly convex translating solution; or the isometric product of $\mathbb{R}^m$ with a strictly convex translating solution of dimension $n - m$.

The paper is set out as follows. In Section 2 we gather together the necessary evolution equations and technical tools. In particular, $\lambda_1$ is not smooth but is locally Lipschitz and semiconvex, and its evolution equation must be understood in a distributional sense. In Section 3 we obtain a Poincaré-type inequality which requires the Simons’ identity for high codimension submanifolds. In Section 4, we complete the proof of the convexity estimate by applying Huisken’s Stampacchia iteration. Finally, in Section 5 we study singularity formation and prove Theorem 1.2.

**Acknowledgements.** The second named author was supported by the EPSRC grant EP/S012907/1.

## 2. Evolution Equations

Let $F : M \times [0, T) \to \mathbb{R}^{n+k}$ solve mean curvature flow and write $M_t := F(M, t)$. We recall from the work of Andrews-Baker [AB10] the following evolution equations for the second fundamental form and mean curvature vector. With respect to local orthonormal frames $\{e_i\}$ and $\{\nu_\alpha\}$ for the tangent and normal bundles,

$$\nabla_{\partial_t} A_{ij\alpha} = \Delta A_{ij\alpha} + A_{ij\beta} A_{pq\beta} A_{pq\alpha} + A_{iq\beta} A_{pq\beta} A_{pj\alpha} + A_{jq\beta} A_{pq\beta} A_{pi\alpha} - 2 A_{ip\beta} A_{jq\beta} A_{pq\alpha},$$

and

$$\nabla_{\partial_t} H_\alpha = \Delta H_\alpha + H_\beta A_{pq\beta} A_{pq\alpha}.$$
From these equations we can compute that
\[
(\partial_t - \Delta)|A|^2 = -2|\nabla A|^2 + 2|(A, A)|^2 + 2|R^\perp|^2.
\]
\[
(\partial_t - \Delta)|H|^2 = -2|\nabla H|^2 + 2|(A, H)|^2.
\]
We use \(R^\perp\) to denote the normal curvature, which is given by
\[
R^\perp_{ij\alpha\beta} = A_{ip\alpha}A_{jp\beta} - A_{jp\alpha}A_{ip\beta}.
\]
Under the quadratic pinching assumption we have \(|H| > 0\), so at any point in \(M_t\) we can choose a local orthonormal frame for the normal bundle which is such that
\[
\nu_1 = \frac{H}{|H|}.
\]
We also use the notation
\[
\hat{A} = A - \frac{1}{|H|}(A, H)\frac{H}{|H|} = A - A_1\nu_1
\]
to denote the components of the second fundamental form orthogonal to the mean curvature vector, and write
\[
h = \frac{1}{|H|}(A, H) = A_1
\]
for the scalar part of the mean curvature component of \(A\). Hence \(A\) admits the decomposition
\[
A = h\nu_1 + \hat{A}.
\]

2.1. **Pinching is preserved.** With this notation in place we can state the estimate proven by Andrews-Baker showing that quadratic pinching is preserved by the flow.

**Lemma 2.1** ([AB10], Section 3). Fix constants \(0 < c < \frac{4}{3n}\) and \(a > 0\) and let
\[
Q := |A|^2 - c|H|^2 + a.
\]
At every point in \(M \times [0, T)\) where \(Q \leq 0\) there holds
\[
(\partial_t - \Delta)Q \leq -2(|\nabla A|^2 - c|\nabla H|^2) + 2|h|^2Q - 2a|h|^2 - \frac{2a}{n(c - 1/n)}|\hat{A}|^2
\]
\[
+ \frac{1}{n(c - 1/n)}|\hat{A}|^2Q + \left(6 - \frac{2}{n(c - 1/n)}\right)|\hat{A}|^2|\hat{A}|^2 + \left(3 - \frac{2}{n(c - 1/n)}\right)|\hat{A}|^4. \tag{2.1}
\]
Note that by Proposition 6 in [AB10] we have
\[
|\nabla A|^2 \geq \frac{3}{n + 2}|\nabla H|^2 \tag{2.2}
\]
so the gradient term on the right-hand side is nonpositive. At points where \(Q \leq 0\), each of the zeroth-order reaction terms is also nonpositive. From now on we suppose the initial submanifold \(M_0\), and hence \(M_t\) for all \(t \in [0, T)\), is quadratically pinched with
\[
c \leq \frac{4}{3n} - \varepsilon_0, \quad \varepsilon > 0.
\]
For ease of notation let us define
\[
W := \left(\frac{4}{3n} - \frac{\varepsilon_0}{2}\right)|H|^2 - |A|^2, \quad w := W^{\frac{1}{2}},
\]
and observe that by the quadratic pinching \(W \geq \frac{\varepsilon_0}{2}|H|^2\) on \(M_t\).
Lemma 2.2. At each point in $M \times [0, T)$ we have the inequalities
\[
(\partial_t - \Delta)W \geq 2|h|^2 W + \frac{(n+2)}{3} \varepsilon_0 |\nabla A|^2.
\]
and
\[
(\partial_t - \Delta)w \geq |h|^2 w + \delta_0 \frac{|\nabla A|^2}{|H|},
\]
where $\delta_0 > 0$ depends only on $n$ and $\varepsilon_0$.

Proof. From Lemma 2.1 we obtain
\[
(\partial_t - \Delta)W \geq 2|h|^2 W + 2|\nabla A|^2 - 2 \left( \frac{4}{3n} - \frac{\varepsilon_0}{2} \right) |\nabla H|^2.
\]
Using (2.2) we estimate
\[
|\nabla A|^2 - \left( \frac{4}{3n} - \frac{\varepsilon_0}{2} \right) |\nabla H|^2 = \left( 1 - \frac{(n+2)}{3} \left( \frac{4}{3n} - \frac{\varepsilon_0}{2} \right) \right) |\nabla A|^2 + \left( \frac{4}{3n} - \frac{\varepsilon_0}{2} \right) \left( n - \frac{2}{3} |\nabla A|^2 - |\nabla H|^2 \right)
\geq \left( 1 - \frac{4(n+2)}{9n} \right) |\nabla A|^2 + \frac{(n+2)}{6} \varepsilon_0 |\nabla A|^2
\geq \frac{(n+2)}{6} \varepsilon_0 |\nabla A|^2,
\]
which gives the desired inequality for $W$. It follows that
\[
(\partial_t - \Delta)W^\frac{1}{2} = \frac{1}{4W^{3/2}} |\nabla W|^2 + \frac{1}{2W^{1/2}} (\partial_t - \Delta)W \\
\geq |h|^2 W^\frac{1}{2} + \frac{(n+2)}{6} \varepsilon_0 \frac{|\nabla A|^2}{W^{1/2}},
\]
and since $W \leq \frac{4}{3n} |H|^2$ we have
\[
(\partial_t - \Delta)w \geq |h|^2 w + \frac{(3n)\frac{1}{2} (n+2)}{6} \varepsilon_0 \frac{|\nabla A|^2}{|H|}.
\]
Thus it suffices to take
\[
\delta_0 = \frac{(3n)\frac{1}{2} (n+2)}{6} \varepsilon_0.
\]
\[
\square
\]

2.2. The evolution of $h$. From the equations for $A$ and $H$, we readily compute the projection $\langle A, H \rangle$ satisfies
\[
(\partial_t - \Delta)A_{ij\alpha} H_{\alpha} = -2\nabla_p A_{ij\alpha} \nabla_p H_{\alpha} + 2H_{\alpha} A_{ij\beta} A_{pq\beta} A_{pq\alpha}
+ H_{\alpha} (A_{iq\beta} A_{pq\beta} A_{pj\alpha} + A_{jq\beta} A_{qp\beta} A_{pj\alpha} - 2A_{ip\beta} A_{jq\beta} A_{pq\alpha}).
\]
The first of the reaction terms can be split into a hypersurface and a codimension component, as follows:

\[
2H_n A_{ij\beta} A_{pq\beta} A_{pq\alpha} = 2A_{ij\beta} A_{kl\beta} h_{kl} H_1 = 2h_{ij} h_{pq} h_{pq} H_1 + 2 \sum_{\beta \neq 1} A_{ij\beta} A_{pq\beta} h_{pq} H_1 = 2h_{ij} H_1 |h|^2 + 2 \sum_{\beta \neq 1} \hat{A}_{ij\beta} \hat{A}_{pq\beta} h_{pq} H_1.
\]

Similarly, the remaining reaction terms can be written as

\[
H_n (A_{ij\beta} A_{pq\beta} A_{pq\alpha} + A_{jq\beta} A_{pq\beta} A_{pq\alpha} - 2A_{ip\beta} A_{jq\beta} A_{pq\alpha}) = h_{ip} h_{pq} h_{qj} H_1 + h_{jp} h_{pq} h_{qi} H_1 - 2h_{ip} h_{jq} h_{pq} H_1 + \sum_{\beta \neq 1} A_{ip\beta} A_{pq\beta} h_{qj} H_1 + \sum_{\beta \neq 1} A_{jp\beta} A_{pq\beta} h_{qi} H_1 - 2 \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{pq\beta} h_{pq} H_1.
\]

Therefore,

\[
(\partial_t - \Delta) A_{ij\alpha} H_\alpha = -2\nabla_p A_{ij\alpha} \nabla_p H_n + 2|h|^2 h_{ij} H_1 + 2 \sum_{\beta \neq 1} \hat{A}_{ij\beta} \hat{A}_{pq\beta} h_{pq} H_1
\]

\[
+ \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{pq\beta} h_{qj} H_1 + \sum_{\beta \neq 1} \hat{A}_{jp\beta} \hat{A}_{pq\beta} h_{qi} H_1 - 2 \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{jq\beta} h_{pq} H_1.
\]

For a positive function \( f \), we have

\[
(\partial_t - \Delta) \sqrt{f} = \frac{1}{4f^{3/2}} |\nabla f|^2 + \frac{1}{2\sqrt{f}} (\partial_t - \Delta) f,
\]

hence the quantity \(|H|\) satisfies

\[
(\partial_t - \Delta)|H| = \frac{1}{4|H|^3} |\nabla |H||^2 + \frac{1}{2|H|} (-2|\nabla H|^2 + 2|\langle A, H \rangle|^2) = \frac{|\langle A, H \rangle|^2}{|H|^2} - \frac{|\nabla H|^2}{|H|} + \frac{1}{|H|^3} \langle H, \nabla_i H \rangle \langle H, \nabla_i H \rangle.
\]

There holds

\[
\frac{|\langle A, H \rangle|^2}{|H|^2} = |\langle A, |H|^{-1} H \rangle|^2 |H| = |h|^2 H_1
\]

and

\[
-\frac{|\nabla H|^2}{|H|} + \frac{1}{|H|^3} \langle H, \nabla_i H \rangle \langle H, \nabla_i H \rangle = -\frac{H_1}{H_1} (H^2_1 |\nabla H|^2 + |\nabla H_1|^2) + \frac{1}{H_1} (\nabla H_1 \nabla \nabla_1 + \nabla H_1 \nabla_1)
\]

\[
= -H_1 |\nabla H|^2
\]

so we have

\[
(2.3) \quad (\partial_t - \Delta)|H| = |h|^2 H_1 - H_1 |\nabla \nu_1|^2.
\]
For a tensor $B_{ij}$ divided by a positive scalar function $f$ there holds

$$
(\nabla \partial_t - \Delta) \frac{B_{ij}}{f} = \frac{1}{f}(\nabla \partial_t - \Delta)B_{ij} - \frac{B_{ij}}{f^2}(\partial_t - \Delta)f + \frac{2}{f^2} \langle \nabla B_{ij}, \nabla f \rangle,
$$

Therefore, dividing $\langle A, H \rangle$ by $|H|$, we obtain

$$
(\nabla \partial_t - \Delta)h_{ij} = |h|^2h_{ij} + 2H_1^{-1} \sum_{\beta \neq 1} \hat{A}_{ij\beta} \hat{A}_{pq\beta} h_{pq} H_1 + H_1^{-1} \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{jq\beta} h_{qi} H_1
$$

$$
+ H_1^{-1} \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{jq\beta} h_{qi} H_1 - 2H_1^{-1} \langle \nabla A_{ij}, \nabla H \rangle + h_{ij} |\nabla \nu_1|^2 + 2H_1^{-1} \langle \nabla h_{ij}, \nabla H_1 \rangle.
$$

Let us introduce the abbreviation

$$
T_{ij} := 2 \sum_{\beta \neq 1} \hat{A}_{ij\beta} \hat{A}_{pq\beta} h_{pq} + \sum_{\beta \neq 1} \hat{A}_{ip\beta} \hat{A}_{jq\beta} h_{qi}
$$

so that we may write

$$
(\nabla \partial_t - \Delta)h_{ij} = |h|^2h_{ij} + T_{ij} - 2H_1^{-1} \langle \nabla A_{ij}, \nabla H \rangle + h_{ij} |\nabla \nu_1|^2 + 2H_1^{-1} \langle \nabla h_{ij}, \nabla H_1 \rangle.
$$

We simplify the gradient terms by decomposing

$$
-2\langle \nabla A_{ij}, \nabla H \rangle = -2\langle \nabla h_{ij} \nu_1 + h_{ij} \nabla \nu_1 + \nabla \hat{A}_{ij}, \nabla H_1 \nu_1 + 2H_1 \nabla \nu_1 \rangle
$$

$$
= -2\langle \nabla h_{ij}, \nabla H_1 \rangle - 2H_1 h_{ij} |\nabla \nu_1|^2 - 2\langle \nabla \hat{A}_{ij}, \nabla H_1 \nu_1 \rangle
$$

$$
- 2H_1 \langle \nabla \hat{A}_{ij}, \nabla \nu_1 \rangle,
$$

and so obtain:

**Lemma 2.3.** At each point in $M \times [0, T)$ there holds

$$
(\nabla \partial_t - \Delta)h_{ij} = |h|^2h_{ij} + T_{ij} - h_{ij} |\nabla \nu_1|^2 - 2H_1^{-1} \langle \nabla \hat{A}_{ij}, \nabla H_1 \nu_1 \rangle
$$

$$
- 2\langle \nabla \hat{A}_{ij}, \nabla \nu_1 \rangle,
$$

where $T_{ij}$ is the quantity defined in (2.4).

Since $h$ is a symmetric bilinear form it has $n$ real eigenvalues, which we denote by

$$
\lambda_1 \leq \cdots \leq \lambda_n.
$$

The smallest eigenvalue can be written as

$$
\lambda_1(x, t) = \min_{|v|=1} h(x, t)(v, v),
$$

and is therefore a Lipschitz continuous function on $M \times [0, T)$. We will use the evolution equation for $h$ to estimate $(\partial_t - \Delta)\lambda_1$, interpreted in an appropriate weak sense (cf. [Whi03] and [Lan17]).

**Definition 2.4.** Let $f : M \times [0, T) \to \mathbb{R}$ be locally Lipschitz continuous and fix a point $(x_0, t_0) \in M \times (0, T)$. We say that a function $\varphi$ is a lower support for $f$ at $(x_0, t_0)$ if $\varphi$ is $C^2$ on the set $B_{\delta}(x_0, r) \times [-r^2 + t_0, t_0]$ for some $r > 0$ and there holds

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

with equality at $(x_0, t_0)$. If the inequality is reversed then $\varphi$ is an upper support for $f$ at $(x_0, t_0)$.
With this definition in place we have the following estimate:

**Lemma 2.5.** Fix \((x_0, t_0) \in M \times (0, T)\) and suppose \(\varphi\) is a lower support for \(\lambda_1\) at \((x_0, t_0)\). Then at \((x_0, t_0)\) there holds

\[
(\partial_t - \Delta) \varphi \geq |h|^2 \varphi + T_{11} - \varphi |\nabla \nu_1|^2 - 2H_1^{-1} \langle \nabla \hat{A}_{11}, \nabla H_1 \nu_1 \rangle
\]

- \(2 \langle \nabla \hat{A}_{11}, \nabla \nu_1 \rangle\)

**Proof.** We choose at the point \((x_0, t_0)\) an orthonormal basis of tangent vectors \(\{e_i\}\) which are such that

\[
h(x_0, t_0)(e_i, e_i) = \lambda_i,
\]

and extend the \(\{e_i\}\) to a spatial neighbourhood of \(x_0\) by parallel transport with respect to \(g(t_0)\). We then extend to an orthonormal frame on a backward spacetime neighbourhood of \((x_0, t_0)\) by parallel transport with respect to the connection \(\nabla_{\partial_t}\). On this neighbourhood we can define a smooth function

\[
\eta(x, t) := h(x, t)(e_1(x, t), e_1(x, t))
\]

Observe that by the definition of \(\lambda_1\) there holds

\[
\eta(x, t) \geq \lambda_1(x, t) \geq \varphi(x, t),
\]

with equality at \((x_0, t_0)\).

It follows that at the point \((x_0, t_0)\) we have

\[
\partial_t \eta \leq \partial_t \varphi, \quad \Delta \eta \geq \Delta \varphi,
\]

hence

\[
(\partial_t - \Delta) \varphi \geq (\partial_t - \Delta) \eta.
\]

At \((x_0, t_0)\) we compute

\[
\partial_t \eta = \nabla_{\partial_t} h(e_1, e_1) + 2h(e_1, \nabla_{\partial_t} e_1)
\]

\[
= \Delta h_{11} + |h|^2 h_{11} + T_{11} - h_{11} |\nabla \nu_1|^2 - 2H_1^{-1} \langle \nabla \hat{A}_{11}, \nabla H_1 \nu_1 \rangle
\]

\[
- 2 \langle \nabla \hat{A}_{11}, \nabla \nu_1 \rangle
\]

and

\[
\Delta h_{11} = \nabla_i (\nabla_i (h_{11}) - 2h(e_1, \nabla_i e_1))
\]

\[
= \Delta (h_{11}) - 2h(e_1, \Delta e_1)
\]

\[
= \Delta \eta - 2h(e_1, \Delta e_1).
\]

On the other hand, at \((x_0, t_0)\) there holds

\[
\langle e_1, \Delta e_1 \rangle = \nabla_k \langle e_1, \nabla_k e_1 \rangle = 0,
\]

which shows that \(\Delta e_1\) is orthogonal to \(e_1\), so since \(h\) is diagonal at \((x_0, t_0)\) we obtain

\[
\Delta h_{11} = \Delta \eta,
\]

and consequently

\[
\partial_t \eta = \Delta \eta + |h|^2 \varphi + T_{11} - \varphi |\nabla \nu_1|^2 - 2H_1^{-1} \langle \nabla \hat{A}_{11}, \nabla H_1 \nu_1 \rangle
\]

\[
- 2 \langle \nabla \hat{A}_{11}, \nabla \nu_1 \rangle.
\]
It follows then
\[(\partial_t - \Delta)\phi \geq (\partial_t - \Delta)\eta = |h|^2\phi + T_{11} - \phi|\nabla\nu_1|^2 - 2H_{11}^{-1}\langle \nabla\hat{A}_{11}, \nabla H_1\nu_1 \rangle - 2\langle \nabla\hat{A}_{11}, \nabla\nu_1 \rangle\]
at \((x_0, t_0)\) as required. □

Eventually we will want to prove integral estimates for the function \(\lambda_1\). To do so we appeal to Alexandrov’s theorem, following Brendle [Bre15] (see also [Lan17]). We call a function \(f : M \times [0, T) \to \mathbb{R}\) locally semiconvex (resp. semiconcave) if about every \((x_0, t_0)\) there is a small open neighbourhood on which \(f\) can be expressed as the sum of a smooth and a convex (resp. concave) function.

**Lemma 2.6.** Let \(f : M \times [0, T) \to \mathbb{R}\) be locally semiconvex. Then \(f\) is twice differentiable almost everywhere in \(M \times [0, T)\), and if \(\varphi\) is a nonnegative Lipschitz function on \(M\) then for each \(t \in [0, T)\) there holds
\[
\int_M \Delta f \cdot \varphi \, d\mu_t \leq -\int_M \langle \nabla f, \nabla \varphi \rangle \, d\mu_t.
\]
Here \(\mu_t\) is the measure induced by the immersion \(F(\cdot, t)\).

**Proof.** Choosing local coordinates and applying Alexandrov’s theorem [EG15, Section 6.4], we see that \(f\) has two derivatives at a.e. point in \(M \times [0, T)\). Furthermore, by [EG15, Section 6.3], for each \(t \in [0, T)\) there is a singular Radon measure \(\chi\) on \(M\) with the property that
\[
\int_M \Delta f \cdot \varphi \, d\mu_t + \int_M \varphi \, d\chi = -\int_M \langle \nabla f, \nabla \varphi \rangle \, d\mu_t
\]
for every \(\varphi \in C^2(M)\). Hence if \(\varphi \geq 0\) there holds
\[
\int_M \Delta f \cdot \varphi \, d\mu_t \leq -\int_M \langle \nabla f, \nabla \varphi \rangle \, d\mu_t.
\]
By approximation, the same inequality also holds if \(\varphi\) is only Lipschitz continuous. □

Since \(h\) is smooth, on every small enough set in spacetime, \(\lambda_1\) can be expressed as the minimum over a set of smooth functions which is compact in \(C^2\). This is sufficient to ensure that \(\lambda_1\) is locally semiconcave on \(M \times [0, T)\), so by the lemma we conclude that there is a set of full measure \(Q \subset M \times [0, T)\) where \(\lambda_1\) is twice differentiable.

**Lemma 2.7.** At each point in \(Q\) there holds
\[
(\partial_t - \Delta)\lambda_1 \geq |h|^2\lambda_1 + T_{11} - \lambda_1|\nabla\nu_1|^2 - 2H_{11}^{-1}\langle \nabla\hat{A}_{11}, \nabla H_1\nu_1 \rangle - 2\langle \nabla\hat{A}_{11}, \nabla\nu_1 \rangle.
\]

**Proof.** Fix a point \((x_0, t_0)\) \(\in\) \(Q\). Then \(\lambda_1\) admits a lower support \(\varphi\) at \((x_0, t_0)\), to which we can apply Lemma 2.5. Since \(\varphi(x_0, t_0) = \lambda_1(x_0, t_0)\), this gives the desired inequality. □

**Remark 2.8.** Notice that the first of the gradient terms is nonnegative whenever \(\lambda_1 \leq 0\), whereas the remaining gradient terms both contain \(\nabla\hat{A}\) as a factor. It is this structure of the gradient terms which allows us to prove the convexity estimate.
2.3. The evolution of $|\hat{A}|^2$. The following evolution equation for $|\hat{A}|^2$ was derived by Naff [Naf19a]:

$$(\partial_t - \Delta)|\hat{A}|^2 = 2|\langle \hat{A}, \hat{A} \rangle|^2 + 2 \sum_{i,j,k} [\hat{A}_{ik} \otimes \hat{A}_{jk} - \hat{A}_{jk} \otimes \hat{A}_{ik}]^2 + 2 \sum_{\alpha} |R_{ij1\alpha}^1|^2$$

$$- 2|\nabla \hat{A}|^2 + 4 \sum_{i,j,k} (\langle \nabla_k \hat{h}_{ij}, \nu_1 \rangle - H_1^{-1} \hat{h}_{ij} \nabla_k H_1) \langle \hat{A}_{ij}, \nabla_k \nu_1 \rangle.$$  

We make use of the quantity

$$v := \frac{|\hat{A}|^2}{|H|}.$$  

**Lemma 2.9.** There is a positive constant $C = C(n)$ such that

$$(\partial_t - \Delta)v \leq C |A|^2 |\hat{A}| + C \frac{|\hat{A}| |\nabla A|^2}{H_1} - 2 \frac{|\nabla \hat{A}|^2}{H_1}$$

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

**Proof.** We use the formula

$$(\partial_t - \Delta)f_1 f_2 = \frac{1}{f_2} (\partial_t - \Delta) f_1 - \frac{f_1}{f_2} (\partial_t - \Delta) f_2 + \frac{2}{f_2} \left\langle \nabla f_1, \nabla f_2 \right\rangle$$

to derive

$$(\partial_t - \Delta)v = \frac{1}{|H|} (\partial_t - \Delta)|\hat{A}|^2 - \frac{|\hat{A}|^2}{|H|^2} (\partial_t - \Delta)|H| + \frac{2}{|H|} \left\langle \nabla |\hat{A}|^2, \nabla |H| \right\rangle.$$  

Let us estimate

$$\frac{2}{|H|} \left\langle \nabla \frac{|\hat{A}|^2}{|H|}, \nabla |H| \right\rangle = \frac{2}{H_1} \left\langle \frac{1}{H_1} \nabla |\hat{A}|^2 - \frac{|\hat{A}|^2}{H_1^2} \nabla H_1, \nabla |H| \right\rangle$$

$$= \frac{4}{H_1^2} \hat{A}_{ij} \langle \nabla \hat{A}_{ij}, \nabla H_1 \rangle - 2 \frac{|\hat{A}|^2 |\nabla H_1|^2}{H_1}$$

$$\leq C(n) \frac{|\hat{A}| |\nabla A|^2}{H_1},$$

where in the last line we have used $|A|^2 \leq \frac{4}{3n} |H|^2$. By (2.3) we have

$$- \frac{|\hat{A}|^2}{|H|^2} (\partial_t - \Delta)|H| = - \frac{|\hat{A}|^2}{H_1^2} (|\hat{h}|^2 H_1 - H_1 |\nabla \nu_1|^2)$$

$$\leq \frac{|\hat{A}|^2}{H_1} \left| \frac{\nabla H}{H_1} \right|^2$$

$$\leq C(n) \frac{|\hat{A}| |\nabla A|^2}{H_1},$$

so there holds

$$(\partial_t - \Delta)v \leq \frac{1}{H_1} (\partial_t - \Delta)|\hat{A}|^2 + C(n) \frac{|\hat{A}| |\nabla A|^2}{H_1}.$$  

We recall

$$(\partial_t - \Delta)|\hat{A}|^2 = 2|\langle \hat{A}, \hat{A} \rangle|^2 + 2 \sum_{i,j,k} \left| \sum_k (\hat{A}_{ik} \otimes \hat{A}_{jk} - \hat{A}_{jk} \otimes \hat{A}_{ik}) \right|^2 + 2 \sum_{\alpha} |R_{ij1\alpha}^1|^2$$

$$- 2|\nabla \hat{A}|^2 + 4 \sum_{i,j,k} (\langle \nabla_k \hat{h}_{ij}, \nu_1 \rangle - H_1^{-1} \hat{h}_{ij} \nabla_k H_1) \langle \hat{A}_{ij}, \nabla_k \nu_1 \rangle,$$
and estimate

\[(\partial_t - \Delta)|\dot{A}|^2 \leq C(n)|\dot{A}|^4 + 2 \sum_{\alpha} |R_{ij\alpha}^{\perp}|^2 - 2|\nabla \dot{A}|^2 + C(n)\frac{|\dot{A}|}{H_1}|\nabla A|^2.\]

Then since

\[R_{ij\alpha}^{\perp} = A_{ip\alpha}A_{jp\beta} - A_{ip\beta}A_{jp\alpha},\]

we can write

\[2 \sum_{\alpha} |R_{ij\alpha}^{\perp}|^2 = 2 \sum_{\alpha \geq 2} |h_{ip\dot{A}_{j\alpha}} - \dot{A}_{ip\alpha}h_{jp\alpha}|^2\]

and use this to bound

\[(\partial_t - \Delta)|\dot{A}|^2 \leq C(n)|A|^2|\dot{A}|^2 + C(n)\frac{|\dot{A}|}{H_1}|\nabla A|^2 - 2|\nabla \dot{A}|^2.\]

Substituting this inequality into (2.5) and using the quadratic pinching gives the desired estimate. \(\square\)

2.4. Modifying \(\lambda_1\). We now form the quantity

\[f(x, t) := -\lambda_1(x, t) - \varepsilon w(x, t) + \Lambda v(x, t),\]

where \(\varepsilon\) and \(\Lambda\) are positive constants to be chosen later. Combining the evolution equations for the three components we obtain:

**Lemma 2.10.** At each point in \(Q\) there holds

\[(\partial_t - \Delta)f \leq |h|^2f + C(1 + \Lambda)|A|^2|\dot{A}| - (f + \varepsilon w)|\nabla \nu_1|^2 - \left(\frac{\varepsilon \delta_0}{2} - C\frac{|\dot{A}|}{H_1}\right)|\nabla A|^2 - \left(2\Lambda - \frac{C}{\varepsilon \delta_0}\right)|\nabla \dot{A}|^2,\]

where \(C = C(n)\).

**Proof.** At any point in \(Q\) we compute

\[(\partial_t - \Delta)f = -(\partial_t - \Delta)\lambda_1 - \varepsilon(\partial_t - \Delta)w + \Lambda(\partial_t - \Delta)\nu,\]

so by Lemma 2.7,

\[(\partial_t - \Delta)f \leq -|h|^2\lambda_1 - T_{11} + \lambda_1|\nabla \nu_1|^2 + 2H_1^{-1}\langle \nabla \dot{A}_{11}, \nabla H_1\nu_1 \rangle + 2\langle \nabla \dot{A}_{11}, \nabla \nu_1 \rangle - \varepsilon(\partial_t - \Delta)w + \Lambda(\partial_t - \Delta)\nu.\]

Inserting the estimates from Lemma 2.2 and Lemma 2.9 we find

\[(\partial_t - \Delta)f \leq |h|^2(-\lambda_1 - \varepsilon w) - T_{11} + \lambda_1|\nabla \nu_1|^2 + 2H_1^{-1}\langle \nabla \dot{A}_{11}, \nabla H_1\nu_1 \rangle + 2\langle \nabla \dot{A}_{11}, \nabla \nu_1 \rangle - \varepsilon\delta_0\frac{|\nabla A|^2}{H_1} + \Lambda \left(C|A|^2|\dot{A}| + C\frac{|\dot{A}|}{H_1}\frac{|\nabla A|^2}{H_1} - 2\frac{|\nabla \dot{A}|^2}{H_1}\right),\]

where \(C = C(n)\). Using the definition of \(f\) and rearranging we obtain

\[(\partial_t - \Delta)f \leq |h|^2(f - \Lambda v) - T_{11} + CA|A|^2|\dot{A}| + (-f - \varepsilon w + \Lambda v)|\nabla \nu_1|^2 + 2H_1^{-1}\langle \nabla \dot{A}_{11}, \nabla H_1\nu_1 \rangle + 2\langle \nabla \dot{A}_{11}, \nabla \nu_1 \rangle + CA\frac{|\dot{A}|}{H_1}\frac{|\nabla A|^2}{H_1} - \varepsilon\delta_0\frac{|\nabla A|^2}{H_1} - 2\Lambda \frac{|\nabla \dot{A}|^2}{H_1}.\]
Next we estimate
\[
-T_{11} = 2 \sum_{\beta \neq 1}^n \hat{A}_{ij\beta} \hat{A}_{pq\beta} \hat{h}_{pq} + \sum_{\beta \neq 1}^n \hat{A}_{ip\beta} \hat{A}_{pq\beta} \hat{h}_{qj} \\
+ \sum_{\beta \neq 1}^n \hat{A}_{jp\beta} \hat{A}_{pq\beta} \hat{h}_{pq} - 2 \sum_{\beta \neq 1}^n \hat{A}_{ip\beta} \hat{A}_{jq\beta} \hat{h}_{pq}
\]
\[
\leq C(n)|A|^2|\hat{A}|
\]
and
\[
2H_1^{-1}\langle \nabla \hat{A}_{11}, \nabla H_1 \nu_1 \rangle + 2\langle \nabla \hat{A}_{11}, \nabla \nu_1 \rangle \leq C(n)H_1^{-1}|\nabla \hat{A}| |\nabla A|
\]
\[
\leq \frac{\varepsilon \delta_0}{2} |\nabla A|^2 + \frac{C(n) |\nabla \hat{A}|^2}{H_1}
\]
in order to obtain
\[
(\partial_t - \Delta)f \leq |h|^2(f - \Lambda v) + C(1 + \Lambda)|A|^2|\hat{A}| - (f + \varepsilon w - \Lambda v)|\nabla \nu_1|^2
\]
\[
- \frac{\varepsilon \delta_0}{2} |\nabla A|^2 + C\Lambda \frac{|\hat{A}| |\nabla A|^2}{H_1} + \left( \frac{C}{\varepsilon \delta_0} - 2\Lambda \right) |\nabla \hat{A}|^2.
\]
Finally, by bounding
\[
|\Lambda v| |\nabla \nu_1|^2 \leq C(n)\Lambda \frac{|\hat{A}| |\nabla A|^2}{H_1}
\]
we find
\[
(\partial_t - \Delta)f \leq |h|^2(f - \Lambda v) + C(1 + \Lambda)|A|^2|\hat{A}| - (f + \varepsilon w)|\nabla \nu_1|^2
\]
\[
- \left( \frac{\varepsilon \delta_0}{2} - C\Lambda \frac{|\hat{A}|}{H_1} \right) |\nabla A|^2 - \left( 2\Lambda - \frac{C}{\varepsilon \delta_0} \right) |\nabla \hat{A}|^2.
\]
\[
\square
\]

3. A Poincaré Inequality

In this section we establish a Poincaré-type inequality for the high codimension solution $M_t$. The proof loosely follows [Hui84, Lemma 5.4], in that we combine Simons’ identity with an integration by parts argument. We also incorporate an idea from [BH17, Proposition 3.1], where the authors symmetrise and then take the square of Simons’ identity to fully exploit the structure of the cubic zeroth-order terms.

Simons’ identity for high codimension submanifolds states that
\[
\nabla_k \nabla_l A_{ij\alpha} = \nabla_i \nabla_j A_{k\alpha} + A_{kl\mu} A_{ip\beta} A_{jp\alpha} - A_{ij\beta} A_{kp\beta} A_{ip\alpha}
\]
\[
+ A_{ji\beta} A_{ip\beta} A_{kp\alpha} + A_{jk\beta} A_{ip\beta} A_{lp\alpha} - A_{il\beta} A_{kp\beta} A_{jp\alpha} - A_{jl\beta} A_{kp\beta} A_{ip\alpha}.
\]

We symmetrise to get
\[
\nabla_k \nabla_l A_{ij\alpha} = \nabla_i \nabla_j A_{k\alpha} + \nabla_l \nabla_k A_{j\alpha} + E_{kl\alpha}.
\]
where
\[
E_{kl\alpha j} = A_{kl\beta} A_{ip\beta} A_{jpa} + A_{lk\beta} A_{jp\beta} A_{ipa}
- A_{ij\beta} A_{kp\beta} A_{lpa} - A_{ji\beta} A_{lp\beta} A_{kpa}
+ A_{jl\beta} A_{ip\beta} A_{kpa} + A_{ik\beta} A_{jp\beta} A_{ipa}
+ A_{jk\beta} A_{ip\beta} A_{lpa} + A_{il\beta} A_{jp\beta} A_{kpa}
- A_{il\beta} A_{kp\beta} A_{jpa} - A_{jk\beta} A_{ip\beta} A_{jpa}.
\]

Using the relation
\[
R^\perp_{i\rho j\alpha \beta} = A_{ip\alpha} A_{j\rho \beta} - A_{i\rho \beta} A_{j\rho \alpha},
\]
we can rewrite the components of \( E \) as
\[
E_{kl\alpha j} = A_{kl\beta} (A_{ip\beta} A_{jpa} + A_{jp\beta} A_{ipa}) - A_{ij\beta} (A_{kp\beta} A_{lpa} + A_{lp\beta} A_{kpa})
+ A_{jl\beta} (A_{ip\beta} A_{kpa}) + A_{ik\beta} (A_{jp\beta} A_{ipa})
+ A_{jk\beta} (A_{lpa} A_{ip\beta}) - A_{il\beta} A_{jp\beta} A_{kpa}
= A_{kl\beta} (A_{ip\beta} A_{jpa} + A_{jp\beta} A_{ipa}) - A_{ij\beta} (A_{ kp\beta} A_{lpa} + A_{lp\beta} A_{kpa})
+ A_{jl\beta} R^\perp_{i\rho j\alpha \beta} + A_{ik\beta} R^\perp_{i\rho j\alpha \beta} + A_{jk\beta} R^\perp_{i\rho j\alpha \beta} + A_{il\beta} R^\perp_{i\rho j\alpha \beta}
= 2 A_{ kl\beta} A_{ ip\beta} A_{ jpa} - 2 A_{ ij\beta} A_{ kp\beta} A_{ lpa} + A_{ kl\beta} R^\perp_{ i\rho j\alpha \beta} - A_{ ij\beta} R^\perp_{ i\rho j\alpha \beta}
+ A_{ jl\beta} R^\perp_{ i\rho j\alpha \beta} + A_{ ik\beta} R^\perp_{ i\rho j\alpha \beta} + A_{ il\beta} R^\perp_{ i\rho j\alpha \beta}.
\]

**Lemma 3.1.** There is a positive constant \( C = C(n) \) such that
\[
|E|^2 \geq 8|\mathbf{h}|^2 \text{tr}(\mathbf{h}^4) - 8 \text{tr}(\mathbf{h}^3)^2 - C|\mathbf{A}|^5|\hat{\mathbf{A}}|.
\]

**Proof.** Let us decompose \( E \) as
\[
E_{kl\alpha j} = U_{kl\alpha j} + V_{kl\alpha j}
\]
where
\[
U_{kl\alpha j} := 2A_{kl\beta} A_{ip\beta} A_{jpa} - 2A_{ij\beta} A_{kp\beta} A_{lpa},
V_{kl\alpha j} := A_{kl\beta} R^\perp_{i\rho j\alpha \beta} - A_{ij\beta} R^\perp_{i\rho j\alpha \beta} + A_{jl\beta} R^\perp_{i\rho j\alpha \beta}
+ A_{ik\beta} R^\perp_{i\rho j\alpha \beta} + A_{jk\beta} R^\perp_{i\rho j\alpha \beta} + A_{il\beta} R^\perp_{i\rho j\alpha \beta}.
\]

There then holds
\[
|E|^2 = |U|^2 + 2\langle U, V \rangle + |V|^2.
\]

Breaking \( U \) into components parallel and orthogonal to the mean curvature vector we obtain
\[
U_{kl\alpha j} = 2(A_{kl\beta} A_{ip\beta} A_{jpa} - 2A_{ij\beta} A_{kp\beta} A_{lpa})
= 2h_{kl} h_{ip} A_{jp} - 2h_{ij} h_{kp} A_{lp} + 2(A_{kl}, \hat{A}_{ip}) A_{jpa} - 2(A_{ij}, \hat{A}_{lp}) A_{jpa}
= 2h_{kl} h_{ip} A_{jp} - 2h_{ij} h_{kp} A_{lp} + 2h_{kl} h_{ip} \hat{A}_{jp} - 2h_{ij} h_{kp} \hat{A}_{lp}
+ 2(\hat{A}_{kl}, \hat{A}_{lp}) A_{jpa} - 2(A_{ij}, \hat{A}_{lp}) A_{jpa},
\]

hence
\[
|U|^2 \geq 4|h_{kl} h_{ip} A_{jp} - h_{ij} h_{kp} h_{lp}|^2 - C(n)|\hat{\mathbf{A}}| |\mathbf{A}|^5 + 4\langle \hat{A}_{kl}, \hat{A}_{lp} \rangle A_{jpa} - \langle A_{ij}, \hat{A}_{lp} \rangle A_{jpa}^2
\geq 8|h|^2 \text{tr}(\mathbf{h}^4) - 8 \text{tr}(\mathbf{h}^3)^2 - C(n)|\hat{\mathbf{A}}| |\mathbf{A}|^5.
\]

Substituting this back in we arrive at
\[
|E|^2 \geq 8|h|^2 \text{tr}(\mathbf{h}^4) - 8 \text{tr}(\mathbf{h}^3)^2 - 2|U||V| - C(n)|\hat{\mathbf{A}}| |\mathbf{A}|^5.
\]
There is a purely dimensional constant $C$ such that
\[ |V| \leq C|A||R^\perp|, \]
and we have
\[ R^\perp_{ij} = h_{ip} \hat{A}_{jp} - \hat{A}_{ip} h_{jp} \]
and
\[ R^\perp_{ij\alpha\beta} = \hat{A}_{i\alpha\beta \gamma} A_{j\gamma} - A_{i\alpha\beta \gamma} \hat{A}_{j\gamma}, \quad \alpha \geq 2, \]
so for a larger constant $C$ there holds
\[ |V| \leq C|A|^2|\hat{A}|. \]

Since $|U| \leq C|A|^3$ we have
\[ |E|^2 \geq 8|h|^2 \text{tr}(h^4) - 8 \text{tr}(h^3)^2 - C|A|^5|\hat{A}|. \]
\[ \square \]

We are now ready to prove the Poincaré inequality. The proof does not actually use the fact that $M_t$ moves by mean curvature, so this result can be viewed as a general statement about high codimensional submanifolds.

**Proposition 3.2.** Fix $t \in [0, T)$ and let $u : M \to \mathbb{R}$ be a nonnegative Lipschitz function which is supported in $\text{supp}(f(\cdot, t))$. Then there is a positive constant $C = C(n, \varepsilon, \varepsilon, \Lambda)$ such that
\[ \int_M |h|^2 u^2 d\mu_t \leq C \int_M u^2 \frac{\nabla A}{|A|^2} d\mu_t + C \int_M u|\nabla u| \frac{\nabla A}{|A|} d\mu_t + C \int_M |A|\hat{A}|u^2 d\mu_t. \]

**Proof.** For a symmetric matrix $B$ with eigenvalues $\mu_i$ there holds
\[ |B|^2 \text{tr}(B^4) - \text{tr}(B^3)^2 = \frac{1}{2} \sum_{i,j} (\mu_i^2 \mu_j^4 - \mu_i^3 \mu_j^3) + \frac{1}{2} \sum_{i,j} (\mu_i^2 \mu_j^3 - \mu_i^3 \mu_j^2) \]
\[ = \frac{1}{2} \sum_{i,j} \mu_i^2 \mu_j^2 (\mu_i^2 + \mu_j^2 - 2\mu_i \mu_j) \]
\[ = \frac{1}{2} \sum_{i,j} \mu_i^2 \mu_j^2 (\mu_i - \mu_j)^2. \]

Observe that the right-hand side vanishes if and only if $B$ is the second fundamental form of a codimension-one cylinder. Let us define
\[ B(x, t) = h(x, t) - \Lambda u(x, t)g(x, t), \]
which has as eigenvalues $\mu_i = \lambda_i - \Lambda \varepsilon$. In particular, the computation above shows that
\[ |B|^2 \text{tr}(B^4) - \text{tr}(B^3)^2 \geq \mu_n^2 \mu_1^2 (\mu_n - \mu_1)^2 \]
\[ = \lambda_n^2 \mu_n^2 (\mu_n - \mu_1)^2 - 2\Lambda \mu_n \mu_1^2 (\mu_n - \mu_1)^2 v + 2\Lambda^2 \mu_1^2 (\mu_n - \mu_1)^2 v^2 \]
\[ \geq \frac{1}{C} |h|^2 \mu_n^2 (\mu_n - \mu_1)^2 - C|A|^5|\hat{A}| \]
where $C = C(n, \Lambda)$. At any point where $f(x, t) > 0$ we have
\[ \lambda_1(x, t) < -\varepsilon w(x, t) + \Lambda u(x, t), \]
which is to say that $\mu_1(x, t) \leq -\varepsilon w(x, t)$. Furthermore, since
\[ 0 < H_1(x, t) = \lambda_1(x, t) + \cdots + \lambda_n(x, t) \leq \lambda_1(x, t) + (n - 1)\lambda_n(x, t) \]
there holds
\[ \mu_n(x, t) - \mu_1(x, t) = \lambda_n(x, t) - \lambda_1(x, t) \]
\[ \geq - \left( 1 + \frac{1}{n-1} \right) \lambda_1(x, t) \]
\[ \geq \frac{n}{n-1} \varepsilon w(x, t) - \frac{n}{n-1} \Lambda v(x, t). \]

If the right-hand side is nonnegative then we can square both sides to get an estimate of the form
\[ (\mu_n(x, t) - \mu_1(x, t))^2 \geq \frac{\varepsilon^2}{C} w(x, t)^2 - C \varepsilon \Lambda ||A|| \hat{A}^2 \]
where \( C = C(n) \). On the other hand if
\[ \frac{n}{n-1} \varepsilon w(x, t) - \frac{n}{n-1} \Lambda v(x, t) < 0 \]
then trivially there holds
\[ (\mu_n(x, t) - \mu_1(x, t))^2 \geq 0 \geq \frac{n^2}{(n-1)^2} \varepsilon^2 w(x, t)^2 - \frac{n^2}{(n-1)^2} \Lambda^2 v(x, t)^2, \]
so in either case we can bound
\[ (\mu_n(x, t) - \mu_1(x, t))^2 \geq \frac{\varepsilon^2}{C} w(x, t)^2 - C ||A|| \hat{A} \]
with \( C = C(n, \varepsilon, \Lambda) \).

Putting these estimates together we find on the support of \( f \) there holds
\[ |B|^2 \text{tr}(B^4) - \text{tr}(B^3)^2 \geq \mu_n^2 \mu_1^2 (\mu_n - \mu_1)^2 \]
\[ \geq \frac{\varepsilon^4}{C} |h|^2 w^4 - C ||A||^5 \hat{A}, \]
and since
\[ |B|^2 \text{tr}(B^4) - \text{tr}(B^3)^2 \leq |h|^2 \text{tr}(h^4) - \text{tr}(h^3)^2 + C ||A||^5 \hat{A} \]
we finally get
\[ |h|^2 \text{tr}(h^4) - \text{tr}(h^3)^2 \geq C^{-1} |h|^2 ||A||^4 - C ||A||^5 \hat{A} \]
where \( C = C(n, \varepsilon_0, \varepsilon, \Lambda) \).

Combining this inequality with the result of the last lemma we find on the support of \( f \) there holds
\[ |h|^2 ||A||^4 \leq C (|h|^2 \text{tr}(h^4) - \text{tr}(h^3)^2) + C ||A||^5 \hat{A} \]
\[ \leq C||E||^2 + C ||A||^5 \hat{A}. \]

Let \( u \) be a nonnegative Lipschitz function supported in \( \text{supp}(f) \). Then we can multiply this inequality by \( ||A||^{-4} u^2 \) and integrate over \( M \) to get
\[ \int_M |h|^2 u^2 d\mu \]
\[ \leq C \int_M ||A||^{-4} u^2 |E|^2 + ||A|| \hat{A} u^2 d\mu \]
\[ = C \int_M ||A||^{-4} u^2 E_{\kappa \lambda \nu} (\nabla_k \nabla_l A_{\lambda \nu} + \nabla_l \nabla_k A_{\kappa \lambda} - \nabla_i \nabla_j A_{\kappa \nu} - \nabla_j \nabla_i A_{\kappa \lambda}) d\mu \]
\[ + C \int_M ||A|| \hat{A} u^2 d\mu. \]
We are going to estimate each of the four Hessian terms on the right. Since each of these is handled in the same way, we only give the argument for the first one. Defining
\[ T_k := |A|^{-4}u^2 E_{ijkl} \nabla_k A_{ij}, \]
we can write
\[ |A|^{-4}u^2 E_{ijkl} \nabla_k \nabla_l A_{ij} a = \nabla_k T_k + 4|A|^{-5}u^2 E_{ijkl} \nabla_l A_{ij} \nabla_k |A| \]
\[ - 2|A|^{-4}u E_{ijkl} \nabla_l A_{ij} \nabla_k u - |A|^{-4}u^2 \nabla_k E_{ijkl} \nabla_l A_{ij}. \]
The divergence term vanishes upon integration, and there is a purely dimensional constant \( C \) such that
\[ |E| \leq C|A|^3, \quad |\nabla E| \leq C|A|^2 |\nabla A|, \quad |\nabla|A|| \leq C|\nabla A|, \]
so making \( C \) a bit larger, we have
\[ \int_M |A|^{-4}u^2 E_{ijkl} \nabla_k A_{ij} d\mu_t \leq C \int_M u^2 |A|^{-5}|A|^3 |\nabla A|^2 d\mu_t \]
\[ + C \int_M u|\nabla u||A|^{-4}|A|^3 |\nabla A| d\mu_t \]
\[ + C \int_M u^2 |A|^{-4}|A|^2 |\nabla A|^2 d\mu_t. \]
Estimating the remaining Hessian terms in the same way and substituting back in we arrive at
\[ \int_M |h|^2 u^2 d\mu_t \leq C \int_M u^2 |\nabla A|^2 |A|^2 d\mu_t + C \int_M u|\nabla u||A| d\mu_t + C \int_M |A||\hat{A}| |u|^2 d\mu_t. \]

\[ \square \]

4. Stampacchia iteration

In this section we establish the convexity estimate by proving an a priori supremum estimate for the function
\[ f_\sigma := \frac{f}{|H|^{1-\sigma}} \]
where \( \sigma \in (0,1) \) is chosen small depending on \( n \) and \( M_0 \). Recall from Lemma 2.10 that at each point in \( Q \) there holds
\[ (\partial_t - \Delta) f \leq |h|^2 f + C(1 + \Lambda)|A|^2 |\hat{A}| - (f + \varepsilon \nu_1)|\nabla \nu_1|^2 \]
\[ - \left( \varepsilon \delta_0 - \frac{C}{H_1} \right) |\nabla A|^2 \frac{|A|}{H_1} - \left( 2\Lambda - \frac{C}{\varepsilon \delta_0} \right) |\nabla \hat{A}|^2 \frac{|\hat{A}|}{H_1}, \]
where \( C = C(n) \). Let us fix
\[ \Lambda = \frac{C}{2\varepsilon \delta_0} \]
so that the last term vanishes. Then using
\[ (\partial_t - \Delta) |H|^{1-\sigma} = (1 - \sigma)|h|^2 H_1^{1-\sigma} - (1 - \sigma)H_1^{1-\sigma} |\nabla \nu_1|^2 + \sigma(1 - \sigma)H_1^{-\sigma-1} |\nabla H_1|^2 \]
we compute that
\[
(\partial_t - \Delta) f_\sigma \leq \sigma |h|^2 f_\sigma + C(1 + \Lambda)|A|^2 - (\frac{\sigma f_\sigma + \varepsilon \frac{w}{H_{2n+1}}} {H_{2n+1}}) |\nabla \nu_1|^2 \\
- \left( \frac{\varepsilon \delta_0}{2} - CA \frac{|A|}{H_{1}} \right) H^\sigma |\nabla A|^2 - \sigma (1 - \sigma) f_\sigma \frac{|\nabla H|^2}{H_{2}} \\
+ 2(1 - \sigma) \left( \nabla f_\sigma, \frac{\nabla H_1}{H_{1}} \right).
\]

Hence at points in \( Q \cap \text{supp}(f_\sigma) \) we have
\[
(\partial_t - \Delta) f_\sigma \leq \sigma |h|^2 f_\sigma + C(1 + \Lambda)|A|^2 - (\frac{\sigma f_\sigma + \varepsilon \frac{w}{H_{2n+1}}} {H_{2n+1}}) |\nabla \nu_1|^2 \\
- \left( \frac{\varepsilon \delta_0}{2} - CA \frac{|A|}{H_{1}} \right) H^\sigma |\nabla A|^2 - \sigma (1 - \sigma) f_\sigma \frac{|\nabla H|^2}{H_{2}} \\
+ 2(1 - \sigma) \left( \nabla f_\sigma, \frac{\nabla H_1}{H_{1}} \right).
\]

where \( C = C(n) \).

All of the computations until now were for a quadratically pinched solution with \( c \leq \frac{4}{3n} - \varepsilon_0 \).

From here on we assume \( n \geq 5 \) and the more restrictive condition \( c \leq c_n - \varepsilon_0 \) where
\[
c_n := \begin{cases} 
\frac{3(n+1)}{2n(n+2)} & n = 5, 6, 7, \\
\frac{4}{3n} & n \geq 8.
\end{cases}
\]

This is the range of pinching constants for which Naff’s codimension estimate is valid.

**Theorem 4.1** ([Naf19a]). Let \( F : M \times [0, T) \to \mathbb{R}^{n+k}, n \geq 5, \) be a quadratically pinched mean curvature flow with \( c \leq c_n - \varepsilon_0 \). Then there is a constant \( \eta = \eta(n, \varepsilon_0) \) in \((0, 1)\) such that
\[
\max_{M_t} \frac{|\hat{A}|^2}{|H|^{2-2\eta}} \leq \max_{M_0} \frac{|\hat{A}|^2}{|H|^{2-2\eta}}
\]
for each \( t \in [0, T) \).

Hence if we set
\[
L := \max_{M_0} |H|
\]
then the inequality
\[
\frac{|\hat{A}|}{|H|} \leq C(n)L^\eta |H|^{-\eta}
\]
holds on \( M_t \) for every \( t \in [0, T) \). Inserting this estimate into (4.1) we find
\[
(\partial_t - \Delta) f_\sigma \leq \sigma |h|^2 f_\sigma + C(1 + \Lambda)L^\eta |A|^2 - (\frac{\varepsilon \delta_0}{2} - CA \frac{|A|}{H_{1}} \right) H^\sigma \frac{|\nabla A|^2}{H_{2}} \\
+ 2(1 - \sigma) \left( \nabla f_\sigma, \frac{\nabla H_1}{H_{1}} \right).
\]

on \( Q \cap \text{supp}(f_\sigma) \), where \( C = C(n) \).
4.1. $L^p$-estimates. For each $k > 0$ let us define

$$f_{\sigma,k}(x,t) := \max\{f_{\sigma}(x,t) - k, 0\}.$$ 

Using the Poincaré inequality we now establish an $L^p$-estimate for $f_{\sigma,k}$. In the codimension one case similar estimates have appeared in [Hui84] and [HS99].

**Proposition 4.2.** There are positive constants $p_0$ and $\ell_0$ depending on $n$, $\varepsilon_0$, $\eta$, $\varepsilon$ and $\Lambda$, and a positive constant $k_0 = k_0(n, \varepsilon_0, \eta, \varepsilon, \Lambda, L)$, with the following property. For every

$$p \geq p_0, \quad \sigma \leq \ell_0 p^{-\frac{1}{2}}, \quad k \geq k_0,$$

we have

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

where $C = C(n, \varepsilon_0, \eta, \varepsilon, \Lambda, L, \mu_0(M), T, k, \sigma, p)$.

**Proof.** Suppose for now that $p_0 \geq 4$ and $\ell_0 \leq \eta$. Then the condition $\sigma \leq \ell_0 p^{-\frac{1}{2}}$ ensures that $\sigma \leq \eta/2$. On $\text{supp}(f_{\sigma,k})$ we have

$$k < \frac{f}{H_1} H_1^q \leq C_0(n, \Lambda) H_1^q,$$

so if we take $k_0 \geq C_0$ and impose $k \geq k_0$ then on $\text{supp}(f_{\sigma,k})$ there holds

$$H_1 \geq (k/C_0)^{\frac{1}{p'}} \geq \max\{k/C_0, 1\}.$$

Substituting this into (4.2) we find

$$(\partial_t - \Delta) f_{\sigma} \leq \sigma |h|^2 f_{\sigma} + C(1 + \Lambda) L^n |A|^2 H_1^{\frac{\sigma}{2}} - \left( \frac{\varepsilon \delta_0}{2} - C_1 k^{-\eta} \right) H_1^\sigma |\nabla A|^2 H_1^2$$

$$+ 2(1 - \sigma) \left\langle \nabla f_{\sigma}, \frac{\nabla H_1}{H_1} \right\rangle$$

on $Q \cap \text{supp}(f_{\sigma,k})$, where $C = C(n)$ and $C_1 = C_1(n, \eta, \Lambda, L)$. Choosing $k_0$ a bit larger so that

$$k_0 \geq \max \left\{ 1, C_0, \left( \frac{4C_1}{\varepsilon \delta_0} \right)^{\frac{1}{\eta}} \right\}$$

and using $f/H_1 \leq C_0$, we find on $Q \cap \text{supp}(f_{\sigma,k})$,

$$(\partial_t - \Delta) f_{\sigma} \leq \sigma |h|^2 f_{\sigma} + C(1 + \Lambda) L^n |A|^2 H_1^{\frac{\sigma}{2}} - \frac{\varepsilon \delta_0}{4C_0} f_{\sigma} |\nabla A|^2 H_1^2$$

$$+ 2(1 - \sigma) \left\langle \nabla f_{\sigma}, \frac{\nabla H_1}{H_1} \right\rangle.$$

By Young’s inequality we have

$$2(1 - \sigma) \left\langle \nabla f_{\sigma}, \frac{\nabla H_1}{H_1} \right\rangle \leq C_2 \frac{|\nabla f_{\sigma}|^2}{f_{\sigma}} + \frac{\varepsilon \delta_0}{8C_0} f_{\sigma} |\nabla A|^2 H_1^2$$

on $\text{supp}(f_{\sigma})$, where $C_2 = C_2(n, \varepsilon_0, \varepsilon, C_0)$. Hence on $Q \cap \text{supp}(f_{\sigma,k})$,

$$(\partial_t - \Delta) f_{\sigma} \leq \sigma |h|^2 f_{\sigma} + C(1 + \Lambda) L^n |A|^2 H_1^{\frac{\sigma}{2}} - \frac{\varepsilon \delta_0}{8C_0} f_{\sigma} H_1^2 |\nabla A|^2 H_1^2 + C_2 \frac{|\nabla f_{\sigma}|^2}{f_{\sigma}}.$$

Applying the pinching we can bound

$$C(1 + \Lambda) L^n |A|^2 H_1^{\frac{\sigma}{2}} \leq C_3(n, \eta, \Lambda)L |h|^2 H_1^{-\eta/2},$$
and by Young’s inequality
\[ H_1^{-\eta/2} \leq \frac{4 - \eta}{4} s^{4/(4 - \eta)} + \frac{\eta}{s^{4/\eta}} \frac{1}{H_1} \leq s^{4/(4 - \eta)} + \frac{1}{s^{4/\eta}} \frac{1}{H_1} \]
for every positive \( s \). Setting \( s = \sigma^{(4 - \eta)/4} \) gives
\[ H_1^{-\eta/2} \leq \sigma + \frac{1}{\sigma^{(4 - \eta)/\eta}} \frac{1}{H_1} \leq \sigma + \sigma^{-4/\eta} H_1^{-2}, \]
so using the pinching we get
\[ C(1 + \Lambda)L^\eta |A|^2 H_1^{-\hat{\Phi}} \leq C_3 \sigma |h|^2 + C_4 \sigma^{-4/\eta} \]
for some \( C_4 = C_4(n, \eta, \Lambda, L) \). Substituting back in, we have
\[ (\partial_t - \Delta)f_\sigma \leq |h|^2 f_\sigma + C_3 \sigma |h|^2 - c_0 f_\sigma \frac{\nabla A|^2}{H_1} + C_2 \frac{\nabla f_\sigma|^2}{f_\sigma} + C_4 \sigma^{-4/\eta} \]
on \( Q \cap \text{supp}(f_\sigma, k) \), where
\[ c_0 := \frac{\epsilon \delta_0}{SC_0}. \]
If \( \varphi \) is any nonnegative Lipschitz function supported in \( \text{supp}(f_\sigma, k) \), then on almost every timeslice we can multiply the last inequality by \( \varphi \) and integrate to get
\[ \int_M \partial_t f_\sigma \cdot \varphi \, d\mu_t \leq \int_M \Delta f_\sigma \cdot \varphi \, d\mu_t + \sigma \int_M |h|^2 f_\sigma \varphi \, d\mu_t + C_3 \sigma \int_M |h|^2 \varphi \, d\mu_t - c_0 \int_M f_\sigma \varphi \frac{\nabla A|^2}{H_1^2} \, d\mu_t + C_2 \int_M \frac{\nabla f_\sigma|^2}{f_\sigma} \, d\mu_t + C_4 \sigma^{-4/\eta} \int_M \varphi \, d\mu_t. \]
Since \( f_\sigma \) is a locally semiconvex function we can use Lemma 2.6 to integrate by parts, and so obtain
\[ \int_M \partial_t f_\sigma \cdot \varphi \, d\mu_t \leq -\int_M \langle \nabla f_\sigma, \nabla \varphi \rangle \, d\mu_t + \sigma \int_M |h|^2 f_\sigma \varphi \, d\mu_t + C_3 \sigma \int_M |h|^2 \varphi \, d\mu_t - c_0 \int_M f_\sigma \varphi \frac{\nabla A|^2}{H_1^2} \, d\mu_t + C_2 \int_M \varphi \frac{\nabla f_\sigma|^2}{f_\sigma} \, d\mu_t + C_4 \sigma^{-4/\eta} \int_M \varphi \, d\mu_t. \]
We set \( \varphi = pf_\sigma^{-1} \) in this inequality and use
\[ \frac{d}{dt} \int_M f_\sigma^{p} \, d\mu_t = p \int_M \partial_t f_\sigma \cdot f_\sigma^{p-1} \, d\mu_t - \int_M |h|^2 f_\sigma^{p} \, d\mu_t \]
to estimate
\[ \frac{d}{dt} \int_M f_\sigma^{p} \, d\mu_t \leq -p(p - 1) \int_M f_\sigma^{p-2} \nabla f_\sigma|^2 \, d\mu_t + \sigma p \int_M |h|^2 f_\sigma f_\sigma^{p-1} \, d\mu_t + C_3 \sigma p \int_M |h|^2 f_\sigma^{p-1} \, d\mu_t - c_0 p \int_M f_\sigma f_\sigma^{p-1} \frac{\nabla A|^2}{H_1^2} \, d\mu_t + C_2 p \int_M f_\sigma^{p-1} \frac{\nabla f_\sigma|^2}{f_\sigma} \, d\mu_t + C_4 \sigma^{-4/\eta} \int_M f_\sigma^{p-1} \, d\mu_t \]
for almost every \( t \in [0, T) \). Using that \( f_{\sigma,k} = f_{\sigma} - k \) on \( \text{supp}(f_{\sigma,k}) \) and rearranging slightly, this gives

\[
\frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq -(p(p-1) - C_2 p) \int_M f_{\sigma,k}^{p-2} |\nabla f_{\sigma,k}|^2 \, d\mu_t - c_0 p \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t \\
+ \sigma p \int_M |h|^2 f_{\sigma,k}^p \, d\mu_t + (C_3 + k) \sigma p \int_M |h|^2 f_{\sigma,k}^{p-1} \, d\mu_t \\
+ C_4 \sigma^{-4/\eta} p \int_M f_{\sigma,k}^{p-1} \, d\mu_t.
\]

Using Young’s inequality we estimate

\[
(C_3 + k) \sigma p \int_M |h|^2 f_{\sigma,k}^{p-1} \, d\mu_t \leq \sigma (p-1) \int_M |h|^2 f_{\sigma,k}^p \, d\mu_t + (C_3 + k) \sigma p \int_M |h|^2 \, d\mu_t
\]

and

\[
C_4 \sigma^{-4/\eta} p \int_M f_{\sigma,k}^{p-1} \, d\mu_t \leq C_4 \sigma^{-4/\eta} (p-1) \int_M f_{\sigma,k}^p \, d\mu_t + C_4 \sigma^{-4/\eta} \mu_t(M).
\]

Inserting these inequalities we arrive at

\[
\frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq -(p(p-1) - C_2 p) \int_M f_{\sigma,k}^{p-2} |\nabla f_{\sigma,k}|^2 \, d\mu_t - c_0 p \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t \\
+ 2 \sigma p \int_M |h|^2 f_{\sigma,k}^p \, d\mu_t + (C_3 + k) \sigma p \int_M |h|^2 \, d\mu_t \\
+ C_4 \sigma^{-4/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t + C_4 \sigma^{-4/\eta} \mu_t(M).
\]

Since \( f_{\sigma,k} \) is supported in \( \text{supp}(f) \), we can apply the Poincaré inequality with \( u = f_{\sigma,k}^p \) to obtain

\[
\int_M |h|^2 f_{\sigma,k}^p \, d\mu_t \leq C_5 \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t + C_5 p \int_M f_{\sigma,k}^{p-1} |\nabla f_{\sigma,k}| \frac{|\nabla A|}{|H|} \, d\mu_t \\
+ C_5 \int_M |A| |\tilde{A}| f_{\sigma,k}^p \, d\mu_t,
\]

where the constant \( C_5 \) depends on \( n, \varepsilon_0, \varepsilon \) and \( \Lambda \). Applying Young’s inequality we obtain

\[
\int_M |h|^2 f_{\sigma,k}^p \, d\mu_t \leq C_5 (1 + p^{1/2}) \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t + C_5 p^{3/2} \int_M f_{\sigma,k}^{p-2} |\nabla f_{\sigma,k}|^2 \, d\mu_t \\
+ C_5 \int_M |A| |\tilde{A}| f_{\sigma,k}^p \, d\mu_t.
\]

Inserting the codimension estimate and quadratic pinching we get

\[
C_5 \int_M |A| |\tilde{A}| f_{\sigma,k}^p \, d\mu_t \leq C_0(n, L, C_5) \int_M |h|^2 |H|^{-\eta} f_{\sigma,k}^p \, d\mu_t,
\]

and we know that \( |H| \geq k/C_0 \) on \( \text{supp}(f_{\sigma,k}) \), so if we take

\[
k_0 \geq \max \left\{ 1, C_0, \left( \frac{4C_1}{\varepsilon \delta_0} \right)^{1/\eta}, C_0(2C_6)^{1/\eta} \right\}
\]

then

\[
C_5 \int_M |A| |\tilde{A}| f_{\sigma,k}^p \, d\mu_t \leq \frac{1}{2} \int_M |h|^2 f_{\sigma,k}^p \, d\mu_t.
\]
In this case
\[ \frac{1}{2} \int_M |h|^2 f_{\sigma,k}^p \, d\mu_t \leq C_5 (1 + p^2) \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t + C_5 p^2 \int_M f_{\sigma,k}^{p-2} |\nabla f_\sigma|^2 \, d\mu_t. \]

Multiplying this inequality through by \(4\sigma p\) and substituting back into (4.3) gives
\[ \frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq -(p(p - 1) - C_2 p - 4C_5 \sigma p \frac{2}{p}) \int_M f_{\sigma,k}^{p-2} |\nabla f_\sigma|^2 \, d\mu_t \]
\[ - (c_0 p - 4C_5 \sigma p - 4C_5 \sigma p \frac{2}{p}) \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t \]
\[ + (C_3 + k)p \sigma \int_M |h|^2 \, d\mu_t + C_4 \sigma^{-1/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t \]
\[ + C_4 \sigma^{-4/\eta} \mu_t(M). \]

Now we insert the assumption \(\sigma \leq \ell_0 p^{-\frac{2}{3}}\) and thus obtain
\[ \frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq -(p(p - 1) - C_2 p - 4C_5 \ell_0 p^2) \int_M f_{\sigma,k}^{p-2} |\nabla f_\sigma|^2 \, d\mu_t \]
\[ - (c_0 p - 4C_5 \ell_0 p^2 - 4C_5 \ell_0 p) \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t \]
\[ + (C_3 + k)p \sigma \int_M |h|^2 \, d\mu_t + C_4 \sigma^{-1/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t \]
\[ + C_4 \sigma^{-4/\eta} \mu_t(M). \]

Decreasing \(\ell_0\) so that
\[ \ell_0 \leq \min \left\{ \eta, \frac{c_0}{8C_5} \frac{1}{8C_5} \right\} \]
now gives
\[ \frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq -(p^2/2 - p - C_2 p) \int_M f_{\sigma,k}^{p-2} |\nabla f_\sigma|^2 \, d\mu_t \]
\[ - (c_0 p / 2 - 2C_5 \ell_0 p^2) \int_M f_{\sigma,k}^p \frac{|\nabla A|^2}{|H|^2} \, d\mu_t \]
\[ + (C_3 + k)p \sigma \int_M |h|^2 \, d\mu_t + C_4 \sigma^{-1/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t \]
\[ + C_4 \sigma^{-4/\eta} \mu_t(M). \]

We can now take \(p_0\) large depending only on \(c_0\) and \(C_5\) to ensure that for \(p \geq p_0\) the inequality
\[ \frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq (C_3 + k)p \sigma \int_M |h|^2 \, d\mu_t + C_4 \sigma^{-1/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t \]
\[ + C_4 \sigma^{-4/\eta} \mu_t(M). \]
holds for almost every \(t \in [0, T)\).
Taking \(k_0\) a bit larger depending on \(n\) and \(C_3\), using \(k \geq k_0\) we can bound
\[ \frac{d}{dt} \int_M f_{\sigma,k}^p \, d\mu_t \leq 2p k^p \sigma \int_M |H|^2 \, d\mu_t + C_4 \sigma^{-1/\eta} p \int_M f_{\sigma,k}^p \, d\mu_t \]
\[ + C_4 \sigma^{-4/\eta} \mu_t(M). \]

Since
\[ \frac{d}{dt} \int_M 2p k^p \sigma \, d\mu_t = -2p k^p \sigma \int_M |H|^2 \, d\mu_t \]
In particular, this implies
\[
\frac{d}{dt} \int_M f_{\sigma,k}^p + 2^p k^p \sigma \, d\mu_t \leq C_4 \sigma^{-4} \gamma p \int_M f_{\sigma,k}^p \, d\mu_t + C_4 \sigma^{-4} \gamma \mu_t(M)
\]
\[
= C_4 \sigma^{-4} \gamma p \int_M f_{\sigma,k}^p + k^{-1} \, d\mu_t.
\]
Hence the function
\[
\phi(t) := \int_M f_{\sigma,k}^p + 2^p k^p \sigma + k^{-1} \, d\mu_t
\]
satisfies
\[
\phi'(t) \leq C_4 \sigma^{-4} \gamma p \phi(t)
\]
for almost every \( t \in [0, T) \). Since \( \phi \) is Lipschitz continuous in time it follows that
\[
\phi(t) \leq \phi(0) \exp(C_4 \sigma^{-4} \gamma pt).
\]
In particular, \( \phi \) can be bounded from above in terms of its value at the initial time, and the constants \( C_4, \eta, \sigma, p \) and \( T \). Recall that \( C_4 \) depends only on \( n, \eta, \Lambda \) and \( L \). Also,
\[
f_{\sigma,k}^p \leq C_0 |H|^{\sigma^p},
\]
so \( \phi(0) \) can be bounded purely in terms of \( n, \Lambda, \sigma, p, L \) and \( \mu_0(M) \). This completes the proof. \( \square \)

4.2. The supremum estimate. Combining the \( L^p \)-estimates just established with the Michael-Simon Sobolev inequality \([MS73]\) we obtain the following iteration inequality. The proof is very similar to that of Theorem 5.1 in \([Hui84]\), so we omit the details.

**Proposition 4.3.** There are positive constants \( p_1 \geq p_0 \) and \( \ell_1 \leq \ell_0 \) depending on \( n, \varepsilon_0, \eta, \varepsilon \) and \( \Lambda \), and a positive constant \( k_1 \geq k_0 \) depending on \( n, \varepsilon_0, \eta, \varepsilon \) and \( \Lambda \), with the following property. Suppose \( p \geq p_1 \) and \( \sigma \leq \ell_1 p^{\frac{1}{\gamma}} \) and set
\[
A(k) := \int_0^T \int_{\text{supp}(f_{\sigma,k}^p)} d\mu_t dt.
\]
Then for every \( h > k \geq k_1 \) we have
\[
A(h) \leq \frac{C}{(h - k)^{\sigma^p}} A(k)^{\gamma},
\]
where \( \gamma > 1 \) depends on \( n \) and \( C = C(n, \varepsilon_0, \eta, \varepsilon, \Lambda, L, \mu_0(M), T, \sigma, p) \).

Appealing to Stampacchia’s lemma (see for example Lemma B.1 in \([KS80]\)) we obtain:

**Corollary 4.4.** There is a constant \( k_2 = k_2(n, \varepsilon_0, \eta, \varepsilon, \Lambda, L, \mu_0(M), T) \) such that
\[
f_{\sigma_0,k_2} = 0
\]
on \( M \times [0, T) \), where \( \sigma_0 := \ell_1 p_1^{-\frac{1}{\gamma}} \) depends only on \( n, \varepsilon_0, \eta, \varepsilon \) and \( \Lambda \).

Recall that \( \eta \) depends only on \( n \) and \( \varepsilon_0 \), and we chose \( \Lambda \) depending only on \( n, \varepsilon_0 \) and \( \varepsilon \). Therefore, by the corollary we have an estimate of the form
\[
\frac{\lambda_1 + \varepsilon w}{|H|} \geq -C |H|^{-\sigma_0} - \Lambda \frac{|A|^2}{|H|^2}
\]
on \( M \times [0, T) \), where \( C = C(n, \varepsilon_0, \varepsilon, L, \mu_0(M), T) \). Appealing to the codimension estimate of Theorem 4.1, we finally obtain
\[
(4.4) \quad \frac{\lambda_1 + \varepsilon w}{|H|} \geq -C |H|^{-\sigma_0} - C |H|^{-2\eta},
\]
where $C$ has the same dependencies as before. From here, since $\varepsilon$ can be made arbitrarily small, an application of Young’s inequality to the two lower-order terms on the right-hand side gives the convexity estimate of Theorem 1.1 (note that $T$ can be bounded in terms of $M_0$ by applying the maximum principle to the evolution equation of $W$).

**Remark 4.5.** There is another way to prove Theorem 1.1 using compactness and the strong maximum principle, which we now sketch. Let $F : M \times [0, T) \to \mathbb{R}^{n+k}$ be a quadratically pinched mean curvature flow with $n \geq 5$ and $c \leq c_n$, fix a constant $\varepsilon > 0$, and set

$$f_\varepsilon := \frac{\lambda_1 + \varepsilon w}{|H|}.$$

For each $j \in \mathbb{N}$ set

$$\delta_j := \inf \{ f_\varepsilon(x, t) : |H|(x, t) \geq j \}.$$

Then the $\delta_j$ form a bounded nondecreasing sequence and therefore converge to some $\delta$. Choose a sequence $(x_j, t_j) \in P_j$ such that $f_\varepsilon(x_j, t_j) \to \delta$ and form a sequence of solutions by shifting $F(x_j, t_j)$ to the spacetime origin in $\mathbb{R}^{n+k} \times \mathbb{R}$ and parabolically rescaling by $|H|(x_j, t_j)$. Then the gradient estimates established in Section 3 of [Ngu18] ensure that this sequence converges smoothly in a small spacetime neighbourhood about the origin. The limit lies in $\mathbb{R}^{n+1}$ by Naff’s codimension estimate, and the quantity $f_\varepsilon$ attains its minimum $\delta$ at the spacetime origin. The strong maximum principle applied to the evolution equation for $f_\varepsilon$ shows that $|\nabla h| \equiv 0$ on the limiting solution, which is consequently either a piece of shrinking sphere or cylinder. In either case we have $\lambda_1 \geq 0$, so $\delta > 0$. In other words, on the original solution, there exists a threshold $C_\varepsilon$ depending on $\varepsilon$ and $M_0$ such that whenever $|H|(x, t) \geq C_\varepsilon$ there holds

$$\frac{\lambda_1(x, t)}{|H|(x, t)} \geq -\varepsilon \frac{w(x, t)}{|H|(x, t)} \geq -C(n) \varepsilon.$$

Note that this argument does not yield a quantitative blow-up rate for the negative part of the second fundamental form, in contrast to the estimate (4.4) proven above. Moreover, the gradient estimates in [Ngu18] are difficult to establish and will not be available in other situations where the Stampacchia iteration goes through. Indeed, an estimate showing asymptotic positivity of curvature is often needed to establish gradient estimates; this is the case for the fully nonlinear flow studied in [BH17] and for three-dimensional Ricci flow [Per02].

## 5. Singularity Formation

In the study of parabolic evolution equations it is natural to distinguish between singularities which form at different rates. For a solution of mean curvature flow $F : M \times [0, T) \to \mathbb{R}^{n+k}$ where $T$ is the maximal time we say that a type I singularity forms as $t \to T$ if there is a positive constant $C$ such that

$$\max_{M_t} |A|^2 \leq \frac{C}{T - t}.$$

Note that this is the blow-up rate for solutions which shrink homothetically (such as shrinking spheres and cylinders). If on the other hand

$$\limsup_{t \to T} \left( (T - t) \max_{M_t} |A|^2 \right) = \infty$$

then the singularity forming at time $T$ is said to be of type II.

For certain type I singularities, Baker used Huisken’s monotonicity formula to show that appropriate rescalings about the singularity converge to a homothetically shrinking solution [Bak11]. Moreover, Baker could show that the only such solutions satisfying the quadratic pinching condition are shrinking spheres and (generalised) cylinders. The analogous result for mean-convex solutions of codimension one was proven earlier by Huisken [Hui90]. In [HS99] Huisken and Sinestrari used
their convexity estimate to show that at a type II singularity, appropriate rescalings about the maximum of the curvature converge to a convex translating solution. In this section we use our convexity estimate to generalise their result to higher codimensions.

Fix a smooth mean curvature flow $F : M \times [0, T) \to \mathbb{R}^{n+k}$ of dimension $n \geq 5$ which is quadratically pinched with $c < c_0$, and suppose a type II singularity is forming as $t \to T$. Consider a sequence of times $\tilde{t}_j \to T$ and let $(x_j, t_j)$ be such that

$$(\tilde{t}_j - t_j)|H|^2(x_j, t_j) := \max_{M \times [0, t_j]} (\tilde{t}_j - t)|H|^2(x, t).$$

Then we have

$$|H|^2(x_j, t_j) = \max_M |H|^2(x, t_j)$$

By the type II assumption, for each $K > 0$ there is a point $(y, \tau) \in M \times [0, T)$ such that

$$(T - \tau)|H|^2(y, \tau) \geq K.$$ 

If $j$ is large enough so that $\tilde{t}_j > \tau$ then we have

$$(\tilde{t}_j - t_j)|H|^2(x_j, t_j) = (\tilde{t}_j - \tau)|H|^2(y, \tau) \geq K - (T - \tilde{t}_j)|H|^2(y, \tau).$$

Hence if $j$ is sufficiently large there holds

$$(\tilde{t}_j - t_j)|H|^2(x_j, t_j) \geq K/2,$$

and since $K$ can be made arbitrarily large this shows that

$$(\tilde{t}_j - t_j)|H|^2(x_j, t_j) \to \infty.$$ 

It follows that $t_j \to T$.

Let $L_j^2 := |H|^2(x_j, t_j)$ and consider the sequence of rescaled solutions defined by

$$F_j(x, t) := L_j(F(x, L_j^{-2}t + t_j) - F(x_j, t_j)),$$

which satisfy the conditions

$$F_j(0, 0) = 0, \quad |H_j|^2(0, 0) = 1,$$

where $H_j$ is the mean curvature vector of $F_j$. More generally, for $t \leq L_j^2(\tilde{t}_j - t_j)$ there holds

$$|H_j|^2(x, t) = L_j^{-2}|H|^2(x, L_j^{-2}t + t_j)$$

$$= L_j^{-2}(\tilde{t}_j - L_j^{-2}t - t_j)|H|^2(x, L_j^{-2}t + t_j)$$

$$\leq L_j^{-2}(\tilde{t}_j - t_j)|H|^2(x_j, t_j)$$

$$= \frac{\tilde{t}_j - t_j}{\tilde{t}_j - t_j - L_j^{-2}t}.$$ 

Therefore, for times $t \leq \delta L_j^2(\tilde{t}_j - t_j)$ with $\delta < 1$ we have

$$\max_M |H_j|^2(\cdot, t) \leq \frac{1}{1 - \delta}.$$ 

Passing to a subsequence in $j$, we can guarantee that there is a sequence $\tau_j \to \infty$ such that

$$\max_M |H_j|^2(\cdot, t) \leq 1 + \frac{1}{j}, \quad \forall t \in [-\tau_j, \tau_j].$$ 

It is well known that for a compact solution of mean curvature flow, a global upper bound for $|A|$ implies bounds on all of the higher derivatives of $A$. This follows from the estimates in [EH91] in the codimension-one case, and similar arguments work in higher codimensions (the details can be
found in Section 4.3 of [Bak11]). Standard compactness theorems therefore imply that there is a smooth solution
\[ \tilde{F} : \tilde{M} \times (-\infty, \infty) \to \mathbb{R}^{n+k} \]
such that the sequence \( F_j \) subconverges smoothly to \( \tilde{F} \) in the following local sense. There is a sequence of nested open sets \( U_l \subset \tilde{M} \) such that
\[ \tilde{M} = \bigcup_{l \in \mathbb{N}} U_l \]
and local diffeomorphisms \( \varphi_l : U_l \to M \) such that the sequence
\[ (x, t) \mapsto F_j(\varphi_l(x), t), \quad (x, t) \in U_l \times [-l, l] \]
converges smoothly to
\[ (x, t) \mapsto \tilde{F}(x, t), \quad (x, t) \in U_l \times [-l, l] \]
as \( j \to \infty \) for every \( l \in \mathbb{N} \). This follows for example from Hamilton’s compactness theorem [Ham95a], as is illustrated in Section 6.1 of [Bak11].

**Theorem 5.1.** The smooth limiting solution \( \tilde{F} : \tilde{M} \times (-\infty, \infty) \to \mathbb{R}^{n+k} \) obtained by the above rescaling procedure lies in an \((n+1)\)-dimensional affine subspace and is either: a strictly convex translating solution; or the isometric product of \( \mathbb{R}^m \) with a strictly convex translating solution of dimension \( n-m \).

**Proof.** We denote the second fundamental form of \( \tilde{F} \) by \( \tilde{A} \), and use this convention for other curvature quantities as well. We know that the mean curvature vector of \( \tilde{F} \) satisfies \( |\tilde{H}(0,0)| = 1 \), but a priori, there could be a point on the limiting solution where \( |\tilde{H}| = 0 \). However the evolution equation
\[ (\partial_t - \Delta)\tilde{W} \geq 2|h|^2\tilde{W} \geq \frac{3}{2} \tilde{W}^3 \]
is still valid, so by the strong maximum principle we either have \( \tilde{W} > 0 \) or \( \tilde{W} \equiv 0 \). In the latter situation \( |\tilde{H}|^2(0,0) = 0 \), which is a contradiction, so appealing to the pinching we conclude that
\[ \frac{4}{3n}|\tilde{H}|^2 > \tilde{W} > 0 \]
on \( \tilde{M} \times (-\infty, \infty) \). Hence Naff’s codimension estimate implies \( h \equiv \tilde{A} \), and consequently, the image of \( \tilde{F} \) lies in an \((n+1)\)-dimensional subspace of \( \mathbb{R}^{n+1} \). By the convexity estimate, \( h \geq 0 \) on \( \tilde{M} \times (-\infty, \infty) \).

To recap, the blow-up limit \( \tilde{F} \) is codimension one, has nonnegative second fundamental form, its scalar mean curvature \( |\tilde{H}| \) is globally bounded from above by one, and this global upper bound is attained at the spacetime origin. This is exactly the situation considered in Section 4 of [HS99]. Applying Hamilton’s strong maximum principle for tensors to the evolution of the second fundamental form,
\[ (\partial_t - \Delta)\tilde{h}^{ij} = |\tilde{h}|^2\tilde{h}^{ij}, \]
we conclude that the solution \( M_t := \tilde{F}(\tilde{M}, t) \) splits as an isometric product \( \mathbb{R}^m \times N_t \), where \( N_t \) is a strictly convex solution of dimension \( n - m \) which exists for all \( t \in (-\infty, \infty) \). Since the spacetime maximum of the mean curvature of \( N_t \) is attained at the spacetime origin, the rigidity case of Hamilton’s Harnack inequality [Ham95b] implies the family \( N_t \) moves by translation. \( \square \)

**Remark 5.2.** By the gradient estimate in [Ngu18], the limiting flow \( \tilde{F} \) cannot be the product of \( \mathbb{R}^{n-1} \) with a grim reaper. On the other hand, the grim reaper is the only strictly convex translator in \( \mathbb{R}^2 \), so we conclude that \( m \leq n - 2 \).
Remark 5.3. If $N_t$ is uniformly two-convex in the sense that the smallest two principal curvatures satisfy
$$\lambda_1 + \lambda_2 \geq \alpha H$$
globally for some $\alpha > 0$, then by the gradient estimate in [Ngu18] and work of Bourni-Langford [BL16], $N_t$ must be rotationally symmetric and hence a bowl soliton of dimension $m$. See also the paper [Naf19b].

REFERENCES

[AB10] Ben Andrews and Charles Baker. Mean curvature flow of pinched submanifolds to spheres. *J. Differential Geom.*, 85(3):357–395, 2010.

[Bak11] Charles Baker. The mean curvature flow of submanifolds of high codimension. *arXiv preprint arXiv:1104.4409*, 2011.

[BH17] Simon Brendle and Gerhard Huisken. A fully nonlinear flow for two-convex hypersurfaces in Riemannian manifolds. *Invent. Math.*, 210(2):559–613, 2017.

[BL16] Theodora Bourni and Mat Langford. Type-II singularities of two-convex immersed mean curvature flow. *Geometric Flows*, 2(1), 2016.

[Bre15] Simon Brendle. A sharp bound for the inscribed radius under mean curvature flow. *Invent. Math.*, 202(1):217–237, 2015.

[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.

[Ham95a] Richard S Hamilton. A compactness property for solutions of the Ricci flow. *American journal of mathematics*, 117(3):545–572, 1995.

[Ham95b] Richard S Hamilton. Harnack estimate for the mean curvature flow. *Journal of Differential Geometry*, 41(1):215–226, 1995.

[HS99] Gerhard Huisken and Carlo Sinestrari. Convexity estimates for mean curvature flow and singularities of mean convex surfaces. *Acta mathematica*, 183(1):45–70, 1999.

[HS09] Gerhard Huisken and Carlo Sinestrari. Mean curvature flow with surgeries of two-convex hypersurfaces. *Invent. Math.*, 175(1):137–221, 2009.

[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. *Journal of Differential Geometry*, 31(1):285–299, 1990.

[KS80] David Kinderlehrer and Guido Stampacchia. An introduction to variational inequalities and their applications, volume 31. Siam, 1980.

[Lan17] Mat Langford. A general pinching principle for mean curvature flow and applications. *Calc. Var. Partial Differential Equations*, 56(4):Paper No. 107, 31, 2017.

[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.

[Naf19a] Keaton Naff. Codimension estimates in mean curvature flow. arXiv:1906.08184 [math.DG], 2019.

[Naf19b] Keaton Naff. Singularity models of pinched solutions of mean curvature flow in higher codimension. arXiv:1910.03968 [math.DG], 2019.

[Ngu18] Huy The Nguyen. Cylindrical estimate for high codimension mean curvature flow. arXiv:1805.11808 [math.DG], 2018.

[Ngu20] Huy The Nguyen. High codimension mean curvature flow with surgery. arXiv:2004.07163 [math.DG], 2020.

[Per02] Grisha Perelman. The entropy formula for the riemann flow and its geometric applications. *arXiv preprint math/0211159*, 2002.

[Whi03] Brian White. The nature of singularities in mean curvature flow of mean-convex sets. *J. Amer. Math. Soc.*, 16(1):123–138, 2003.

EBERHARD KARLS UNIVERSITÄT TÜBINGEN, AUF DER MORGENSTELLE 10, 72076 TÜBINGEN, GERMANY

E-mail address: stephen.lynch@math.uni-tuebingen.de

QUEEN MARY UNIVERSITY OF LONDON, MILE END ROAD, LONDON E1 4NS, UNITED KINGDOM

E-mail address: h.nguyen@qmul.ac.uk