1. Introduction

The evolution of hypersurfaces by their mean curvature has been studied by many authors since the appearance of Gerhard Huisken’s first paper [Hu1] on the subject in 1984. More recently, mean curvature flow of higher codimension submanifolds has also received attention. In this paper we prove a result analogous to that of [Hu1] for submanifolds of any codimension.

Let \( F_0 : \Sigma \to \mathbb{R}^{n+k} \) be a smooth immersion of a compact manifold \( \Sigma \). The mean curvature flow with initial condition \( F_0 \) is a smooth family of immersions \( F : \Sigma \times [0, T) \to \mathbb{R}^{n+k} \) satisfying

\[
\begin{aligned}
\frac{\partial}{\partial t} F(p, t) &= H(p, t), \quad p \in \Sigma, \quad t \geq 0, \\
F(\cdot, 0) &= F_0,
\end{aligned}
\]

where \( H(p, t) \) is the mean curvature vector of the submanifold \( \Sigma_t = F(\Sigma, t) \) at \( p \). We use the abbreviation “MCF” for the system (1).

High codimension MCF is the steepest descent flow for the area functional, and so arises naturally in several contexts. For example, singular sets in harmonic map heat flow move by generalized mean curvature flow [LT], Generalized solutions can be defined using minimal barriers or level sets [AS1, AS2, BN]. Huisken’s monotonicity formula [Hu3] applies in any codimension [Ha4], relating singularity formation to self-similar solutions of the flow. However rather little is known about such self-similar solutions (see [S2] for some recent results in this direction).

Much of the previous work on high codimension mean curvature flow has used assumptions on the Gauss image, focussing on graphical submanifolds [CLT, LL, W2, W4] or symplectic or Lagrangian submanifolds [SW, CL2, W1, S1, N], or making use of the fact that convex subsets of the Grassmannian are preserved [TW, W3, W5].

In this paper we work with conditions on the extrinsic curvature (second fundamental form), which have the advantage of being invariant under rigid motions. Several difficulites arise in carrying out this program: First, in high codimension the second fundamental form has a much more complicated structure than in the hypersurface case. In particular, under MCF the second fundamental form evolves according to a reaction-diffusion system in which the reaction terms are rather complicated, whereas in the hypersurface case they are quite easily understood. Thus it can be extremely difficult to determine whether the reaction terms are favourable for preserving a given curvature condition. Second, there do not seem to be any useful invariant conditions on the extrinsic curvature which define convex subsets of the space of second fundamental forms. This lack of convexity is forced by the necessity for invariance under rotation of the normal bundle. This means that the vector bundle maximum principle formulated by Hamilton in [Ha2], which states that the reaction-diffusion system will preserve an invariant convex set if the reaction terms are favourable, cannot be applied. The latter maximum principle has been extremely effective in

\[\text{Research partially supported by Discovery Grant DP0556211 of the Australian Research Council.}\]
the Ricci flow in high dimensions [BW, BS, B] where the algebraic complexity of the curvature tensor has presented similar difficulties. For arbitrary reaction-diffusion systems, the convexity condition is necessary for a maximum principle to apply. However, in our setting the Codazzi identity adds a constraint on the first derivatives of solutions which allows some non-convex sets to be preserved. A similar situation arose Huisken’s work on evolving hypersurfaces in spheres [Hu2], where a non-convex condition was preserved. Our result is as follows:

**Theorem 1.** Let \( \Sigma_0 = F_0(\Sigma^n) \) be a compact submanifold smoothly immersed in \( \mathbb{R}^{n+k} \). If \( \Sigma_0 \) has \( H \neq 0 \) everywhere and satisfies \( |h|^2 \leq c|H|^2 \), where

\[
c = \begin{cases} 
\frac{4}{3n}, & \text{if } 2 \leq n \leq 4; \\
\frac{1}{n-1}, & \text{if } n \geq 4,
\end{cases}
\]

then MCF has a unique smooth solution \( F : \Sigma \times [0,T) \) on a finite maximal time interval, and \( F_t(.) \) converges uniformly to a point \( q \in \mathbb{R}^{n+k} \) as \( t \to T \). The rescaled maps \( F_t = \frac{F_t-q}{\sqrt{2n(T-t)}} \) converge in \( C^\infty \) as \( t \to T \) to a limiting embedding \( \tilde{F}_T \) with image equal to a regular unit \( n \)-sphere in some \( (n+1) \)-dimensional subspace of \( \mathbb{R}^{n+k} \).

In dimensions greater than or equal to four Theorem 1 is optimal as the following example illustrates: Consider the submanifolds \( S^m \times S^1(1) \subset \mathbb{R}^n \times \mathbb{R}^2 \), where \( \varepsilon \) is a small positive number. The second fundamental form is given by

\[
h|_{(x,y)} = \begin{pmatrix}
\frac{1}{\varepsilon} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \frac{1}{\varepsilon}
\end{pmatrix}(x,0) + \begin{pmatrix}
0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 1
\end{pmatrix}(0,y)
\]

and so they satisfy \( |h|^2 = \frac{1}{n-1} \left( 1 + \frac{\varepsilon^{n(n-2)}}{n-1} \right) |H|^2 \). These submanifolds collapse to \( S^1 \) under MCF and do not contract to points. In dimensions two and three the size of gradient and reaction terms of equation (31) prevent the optimal result from being achieved. This is similar to the situation in [Hu2], where in dimension two the difficulty in controlling the gradient terms prevents the optimal result from being obtained. We remark that contrary to the situation in [Hu2], one cannot expect to obtain such a result with \( c = 1/(n-1) = 1 \) in the case \( n = 2 \) in arbitrary codimension, since the Veronese surface provides a counterexample: This is a surface in \( \mathbb{R}^5 \) which satisfies \( |h|^2 = \frac{2}{n} |H|^2 \), but which contracts without changing shape under mean curvature flow. However we are not aware of any such counterexamples in the case \( n = 3 \) (there are none among minimal submanifolds of spheres [CO]).

Curvature pinching conditions similar to those in our theorem have appeared previously in a number of results for special classes of submanifolds: In [O1] Okumura shows that if a submanifold of Euclidean space with parallel mean curvature vector and flat normal bundle satisfies \( |h|^2 < |H|^2/(n-1) \), then the submanifold is a sphere. The equivalent result for hypersurfaces of the sphere with \( |h|^2 < \frac{1}{n-1} |H|^2 + 2 \) (where the flat normal bundle condition is vacuous) was proved by Okumura in [O2]. Chen and Okumura [CO] later removed the assumption of flat normal bundle and so proved that if a submanifold of Euclidean space with parallel mean curvature vector satisfies \( |h|^2 < 1/(n-1)|H|^2 \), then the submanifold is a sphere (or, in the case \( n = 2 \), a minimal surface with positive intrinsic curvature in a sphere, such as the Veronese surface). A related series of results began with work of Chern, do Carmo and Kobayashi [CdCK]. They...
classified minimal submanifolds of the sphere that satisfy $|h|^2 \leq n/(2 - 1/k)$, where $k$ is the codimension. Two decades later do Carmo and Alencar [AdC] classified hypersurfaces of the sphere with constant mean curvature satisfying a certain pinching condition, and shortly afterwards Santos [Sa] extended the classification, under a suitable pinching condition, to submanifolds of the sphere with parallel mean curvature vector.

Our result is closely related to some of the above: In particular the results on minimal submanifolds of spheres relate to ours, since such submanifolds contract without change of shape under the mean curvature flow. The results for parallel mean curvature vector do not relate as directly, since such submanifolds do not behave simply under the mean curvature flow. Our theorem implies that the entire class of $n$-submanifolds satisfying the curvature pinching condition retracts onto the totally umbilic $n$-spheres, and thus onto the Grassmannian $G_{n+1,n+k}$ of $(n+1)$-dimensional subspaces of $\mathbb{R}^{n+k}$.

The broad structure of the proof of Theorem 1 is similar to that in [Hu1], which in turn, was inspired by Hamilton’s seminal paper on Ricci flow [Ha1]. We first introduce our notation and some facts from submanifold geometry of high codimension. A key aspect of this is the machinery of connections on vector bundles which we employ extensively in deriving the evolution equations for geometric quantities. In particular we introduce a novel connection on tangent and normal bundles defined over space and time, which proves very useful in deriving evolution equations and allowing simple commutation of time and space derivatives. This connection also provides a natural interpretation of the ‘Uhlenbeck trick’ introduced in [Ha2] to take into account the change in length of spatial tangent vectors under the flow. The key step in our argument is to prove that curvature pinching is preserved. This plays a role similar to Huisken’s estimate $h_{ij} - \varepsilon H g_{ij} \geq 0$ from [Hu1]. The Codazzi identity is used to derive an inequality on the derivatives of the second fundamental form analogous to that in [Hu1], in order to control the gradient terms which arise in the evolution equation. An inequality from [AMJ] also appears in the argument to control the reaction terms, which in this setting are much more complicated than in the hypersurface case. A stronger pinching estimate is deduced using a Stampacchia iteration argument following the model of [Hu1], although again the curvature terms are considerably more complicated here and the argument to control them is quite involved. The subsequent analysis to prove convergence is again parallel to that in [Hu1], with only minor differences introduced by the high codimension setting.

This result will form part of the second author’s doctoral thesis at The Australian National University. He wishes to thank Carlo Mantegazza and Giovanni Catino for many helpful discussions whilst he was a guest at the Institut Henri Poincare.

2. NOTATION AND PRELIMINARY RESULTS

To a large extent our notations are compatible with those of [Hu1]. In order to work with the normal bundle we first discuss vector bundles, including pullback bundles and sub-bundles. The machinery we develop is useful and new even in the codimension one case, as we work with the tangent and normal bundles as vector bundles over the space-time domain, and introduce natural metrics and connections on these. In particular, the connection we introduce on the ‘spatial’ tangent bundle (as a bundle over spacetime) contains more information than the Levi-Civita connections of the metrics at each time, and this proves particularly useful in computing evolution equations for geometric quantities.
2.1. Vector bundles. We denote the space of smooth sections of a vector bundle $E$ by $\Gamma(E)$. If $E$ is a vector bundle over $N$, the dual bundle $E^*$ is the bundle whose fibres are the dual spaces of the fibres of $E$. If $E_1$ and $E_2$ are vector bundles over $N$, the tensor product $E_1 \otimes E_2$ is the vector bundle whose fibres are the tensor products $(E_1)_p \otimes (E_2)_p$.

2.1.1. Metrics. A metric $g$ on a vector bundle $E$ is a section of $E^* \otimes E^*$ which is an inner product on $E_p$ for each $p$ in $N$. A metric on $E$ defines a bundle isomorphism $\#_g$ from $E$ to $E^*$, defined by
\[
(\#_g(\xi))(\eta) = g(\xi, \eta)
\]
for all $\xi, \eta \in E_p$. If $g$ is a metric on $E$, then there is a unique metric on $E^*$ (also denoted $g$) such that the identification $\#_g$ is a bundle isomorphism: For all $\xi, \eta \in E_p$,
\[
g(\#_g(\xi), \#_g(\eta)) = g(\xi, \eta).
\]
If $g_i$ is a metric on $E_i$, $i = 1, 2$, then $g = g_1 \otimes g_2 \in \Gamma((E_1^* \otimes E_1^*) \otimes (E_2^* \otimes E_2^*)) \simeq \Gamma((E_1 \otimes E_2)^* \otimes (E_1 \otimes E_2)^*)$ is the unique metric on $E_1 \otimes E_2$ such that $g(\xi_1 \otimes \eta_1, \xi_2 \otimes \eta_2) = g_1(\xi_1, \xi_2) g_2(\eta_1, \eta_2)$.

2.1.2. Connections. A connection $\nabla$ on a vector bundle $E$ over $N$ is a map $\nabla : \Gamma(TN) \times \Gamma(E) \to \Gamma(E)$ which is $C^\infty(N)$-linear in the first argument and $\mathbb{R}$-linear in the second, and satisfies
\[
\nabla_U(f \xi) = f \nabla_U \xi + (Uf) \xi
\]
for any $U \in \Gamma(TN)$, $\xi \in \Gamma(E)$, and $f \in C^\infty(N)$. Here the notation $Uf$ means the derivative of $f$ in direction $U$. Given a connection $\nabla$ on $E$, there is a unique connection on $E^*$ (also denoted $\nabla$) such that for all $\xi, \eta \in \Gamma(E)$, $\omega \in \Gamma(E^*)$, and $X \in \Gamma(TN)$,
\[
(2) \quad X(\omega(\xi)) = (\nabla_X \omega)(\xi) + \omega(\nabla_X \xi).
\]
If $\nabla^i$ is a connection on $E_i$ for $i = 1, 2$, then there is a unique connection $\nabla$ on $E_1 \otimes E_2$ such that
\[
(3) \quad \nabla_X(\xi_1 \otimes \xi_2) = (\nabla_X^1 \xi_1) \otimes \xi_2 + \xi_1 \otimes (\nabla_X^2 \xi_2)
\]
for all $X \in \Gamma(TN)$, $\xi_i \in \Gamma(E_i)$. In particular, for $S \in \Gamma(E_1^* \otimes E_2)$ (an $E_2$-valued tensor acting on $E_1$), $\nabla S \in \Gamma(TN \otimes E_1^* \otimes E_2)$ is given by
\[
(4) \quad (\nabla_X S)(\xi) = \nabla_X^{E_2} (S(\xi)) - S(\nabla_X^{E_1} \xi).
\]
A connection $\nabla$ on $E$ is compatible with a metric $g$ if for any $\xi, \eta \in \Gamma(E)$ and $X \in \Gamma(TN)$,
\[
Xg(\xi, \eta) = g(\nabla_X \xi, \eta) + g(\xi, \nabla_X \eta).
\]
If $\nabla$ is compatible with a metric $g$ on $E$, then the induced connection on $E^*$ is compatible with the induced metric on $E^*$. Similarly, if $\nabla_i$ is a connection on $E_i$ compatible with a metric $g_i$ for $i = 1, 2$, then the metric $g_1 \otimes g_2$ is compatible with the connection on $E_1 \otimes E_2$ defined above.

2.1.3. Curvature. Let $E$ be a vector bundle over $N$. If $\nabla$ is a connection on $E$, then the curvature of $\nabla$ is the section $R_\nabla \in \Gamma(T^*N \otimes T^*N \otimes E^* \otimes E)$ defined by
\[
R_\nabla(X, Y) \xi = \nabla_Y(\nabla_X \xi) - \nabla_X(\nabla_Y \xi) - \nabla_{[Y, X]} \xi.
\]
The curvature of the connection on $E^*$ given by Equation (2) is characterized by the formula
\[
0 = (R(X, Y) \omega)(\xi) + \omega(R(X, Y) \xi)
\]
for all $X, Y \in \Gamma(TN)$, $\omega \in \Gamma(E^*)$ and $\xi \in \Gamma(E)$. 
The curvature on a tensor product bundle (with connection defined by Equation (3)) can be computed in terms of the curvatures of the factors by the formula
\[ R_{XY}(\xi_1 \otimes \xi_2) = (R_{XY}(X,Y)\xi_1) \otimes \xi_2 + \xi_1 \otimes (R_{XY}(X,Y)\xi_2). \]

In particular, the curvature on \(E_1^* \otimes E_2\) (\(E_2\)-valued tensors acting on \(E_1\)) is given by
\[ (R(X,Y)S)(\xi) = R_{XY}(X,Y)(S(\xi)) - S(R_{XY}(X,Y)\xi). \]

### 2.2. Pullback bundles.
Let \(M\) and \(N\) be smooth manifolds, and let \(E\) be a vector bundle over \(N\) and \(f\) a smooth map from \(M\) to \(N\). Then \(f^* E\) is the pullback bundle of \(E\) over \(M\), which is a vector bundle with fibre \((f^*E)_x = E_{f(x)}\). If \(\xi \in \Gamma(E)\), then we denote by \(\xi_f\) the section of \(f^*E\) defined by \(\xi_f(x) = \xi(f(x))\) for each \(x \in M\) (called the restriction of \(\xi\) to \(f\)).

The pullback operation on vector bundles commutes with taking duals and tensor products, so the tensor bundles constructed from a vector bundle \(E\) pull back to give the tensor bundles of the pull-back bundle \(f^*E\). In particular, if \(g\) is a metric on \(E\), then \(g\) is a section of \(E^* \otimes E^*\), and the restriction \(g_f \in \Gamma(f^*(E^* \otimes E^*))\) defines a metric on \(f^*E\).

If \(\nabla\) is a connection on \(E\), then there is a unique connection \(f^*\nabla\) on \(f^*E\), called the pullback connection which satisfies \(f^*\nabla_u(X_f) = \nabla_{f^*u}X\) for any \(u \in TM\) and \(X \in \Gamma(E)\).

**Proposition 1.** If \(g\) is a metric on \(E\) and \(\nabla\) is a connection on \(E\) compatible with \(g\), then \(f^*\nabla\) is compatible with the restriction metric \(g_f\).

**Proof.** \(f^*\nabla\) is compatible with \(g\) if and only if \(\nabla g = 0\). We must therefore show that \(f^*\nabla g_f = 0\) if \(\nabla g = 0\). But this is immediate, since \(f^*\nabla (g_f) = \nabla_{f^*u}g = 0\).

**Proposition 2.** The curvature of the pull-back connection is the pull-back of the curvature of the original connection. Here \(R_{f^*E} \in \Gamma(T^*N \otimes T^*N \otimes E^* \otimes E)\), so that
\[ f^* (R_{XY}) = \Gamma(T^*M \otimes T^*M \otimes f^*(E^* \otimes E)) = \Gamma(T^*M \otimes T^*M \otimes (f^*E)^* \otimes f^*E). \]

**Proof.** Since curvature is torsorial, it is enough to check the formula for a basis. Choose a local frame \(\{Z_p\}_{p=1}^k\) for \(E\). Then \(\{f^*Z_p\}_f\) is a local frame for \(f^*E\). Choose local coordinates \(\{y^\alpha\}\) for \(N\) near \(f(p)\) and \(\{x^\lambda\}\) for \(M\) near \(p\), and write \(f^\alpha = y^\alpha \circ f\). Then
\[
R_{f^*E}(\partial_i, \partial_j)(f^*Z_p)_f = f^*R_{\partial_i, \partial_j}(f^*Z_p)_f = f^*\nabla_{\partial_i}(f^*\nabla_{\partial_j}(f^*Z_p)) - (i \leftrightarrow j) \\
= f^*\nabla_{\partial_j}(f^*\nabla_{\partial_i}(f^*Z_p)) - (i \leftrightarrow j) \\
= f^*\nabla_{\partial_j}(\partial_i f^\alpha \nabla_{\partial_j} Z_p) - (i \leftrightarrow j) \\
= \partial_i f^\alpha \nabla_{\partial_j, \partial_j} (\nabla_{\partial_j} Z_p) - (i \leftrightarrow j) \\
= \partial_i f^\alpha \nabla_{\partial_j, \partial_j} (\nabla_{\partial_j} Z_p) - (\alpha \leftrightarrow \beta) \\
= \partial_i f^\alpha \partial_j f^\beta R_{\partial_j, \partial_j} Z_p \\
= R_{f^*E}(\partial_i, \partial_j, \partial_j) Z_p.
\]

In the case of pulling back a tangent bundle, there is another important property:
Proposition 3. If $\nabla$ is a symmetric connection on $TN$, then the pull-back connection $f^*\nabla$ on $f^*TN$ is symmetric, in the sense that for any $U, V \in \Gamma(TM)$,

$$f^*\nabla_U(f_*V) - f^*\nabla_V(f_*U) = f_*([U, V]).$$

Proof. Choose local coordinates $x^i$ for $M$ near $p$, and $y^\alpha$ for $N$ near $f(p)$, and write $U = U^i \partial_i$ and $V = V^j \partial_j$. Then

$$f^*\nabla_U(f_*V) - f^*\nabla_V(f_*U) = U^i \partial_i (V^j f^a \partial_a) - V^j \partial_j (U^i f^a \partial_a)$$

$$= U^i \partial_i (V^j f^a \partial_a) + V^j \partial_j U^i f^a \partial_a - U^i f^a \partial_a - V^j f^a \partial_a$$

$$= (U^i \partial_i - V^j \partial_j) f^a \partial_a + U^i f^a \partial_a - V^j f^a \partial_a$$

$$= f_*([U, V]).$$

2.3. Subbundles. A subbundle $K$ of a vector bundle $E$ over $M$ is a vector bundle $K$ over $M$ with an injective vector bundle homomorphism $f_K : K \rightarrow E$ covering the identity map on $M$. We consider complementary sub-bundles $K$ and $L$, so that $E|_x = f_K(K_x) \oplus f_L(L_x)$, and denote by $\pi_K$ and $\pi_L$ the corresponding projections onto $K$ and $L$ (so $\pi_K \circ f_K = \text{Id}_K$, $\pi_L \circ f_L = \text{Id}_L$, $\pi_K \circ f_L = 0$, $\pi_L \circ f_K = 0$, and $f_K \circ \pi_K + f_L \circ \pi_L = \text{Id}_E$). If $\nabla$ is a connection on $E$, we define a connection $\tilde{\nabla}$ on $K$ and a tensor $h^K \in \Gamma(T^*M \otimes K^* \otimes L)$ (the second fundamental form of $K$) by

$$\tilde{\nabla}_u (\xi) = \pi_K (\nabla_u (f_K \xi)); \quad h^K(u, \xi) = \pi_L (\nabla_u (f_K \xi));$$

so that

$$\nabla_u (f_K \xi) = f_K \left( \tilde{\nabla}_u (\xi) + h^K(u, \xi) \right)$$

for any $u \in TM$ and $\xi \in \Gamma(K) \subset \Gamma(E)$. The curvature $R^K$ of $\tilde{\nabla}$ is related to the second fundamental form $h^K$ and the curvature of $\nabla$ via the Gauss equation:

$$R^K(u, v)\xi = \pi_K (R\nabla(u, v)(f_K \xi)) + h^K(u, h^K(v, \xi)) - h^K(v, h^K(u, \xi))$$

for all $u, v \in T_M$ and $\xi \in \Gamma(K)$. The other important identity relating the second fundamental form to the curvature is the Codazzi identity, which states:

$$\pi_L (R\nabla(v, u)(f_K \xi)) = \tilde{\nabla}_u (h^K(v, \xi)) - \tilde{\nabla}_v (h^K(u, \xi)) - h^K(u, \tilde{\nabla}_v (\xi)) + h^K(v, \tilde{\nabla}_u (\xi)) - h^K([u, v], \xi).$$

If we are supplied with an arbitrary symmetric connection on $TM$, then we can make sense of the covariant derivative $\nabla h^K$ of the second fundamental form $h^K$, and the Codazzi identity becomes

$$\nabla_u h^K(v, \xi) - \nabla_v h^K(u, \xi) = \pi_L (R\nabla(v, u)(f_K \xi)) .$$

An important case is where $K$ and $L$ are orthogonal with respect to a metric $g$ on $E$ compatible with $\nabla$. Then $\tilde{\nabla}$ is compatible with the induced metric $g^K$, and $h^K$ and $h^L$ are related by the Weingarten relation:

$$g^L(h^K(u, \xi), \eta) + g^K(\xi, h^L(u, \eta)) = 0$$

for all $\xi \in \Gamma(K)$ and $\eta \in \Gamma(L)$. 
2.4. **The tangent and normal bundles of a time-dependent immersion.** The machinery introduced above is familiar in the following setting: If \( F: M^n \to N^{n+k} \) is an immersion, then \( F_*: TM \to F^*TN \) defines the tangent sub-bundle of \( F^*TN \), and its orthogonal complement is the normal bundle \( NM = F_*(TM)^\perp \). If \( \bar{g} \) is a metric on \( TN \) with Levi-Civita connection \( \bar{\nabla} \), then the metric \( g^{TM} \) is the induced metric on \( M \), and \( \nabla^{TM} = \bar{\nabla}^{TM} \) is its Levi-Civita connection, while \( h^{TM} \in \Gamma(TM \otimes T^*M \otimes NM) \) is the second fundamental form, and \( h^{NM} \) is minus the Weingarten map. The Gauss identities (8) for \( TM \) are the usual Gauss equations for a submanifold, while those for \( NM \) are usually called the Ricci identities. The Codazzi identities for the two are equivalent to each other.

In this paper we want to apply the same machinery in a setting adapted to time-dependent immersions. If \( I \) is a real interval, then the tangent space \( T(\Sigma \times I) \) splits into a direct product \( \mathcal{H} \oplus \mathbb{R}\partial_t \), where \( \mathcal{H} = \{ u \in T(\Sigma \times I): \partial_t(u) = 0 \} \) is the ‘spatial’ tangent bundle.

We consider a smooth map \( F: \Sigma^n \times I \to N^{n+k} \) which is a time-dependent immersion, i.e. for each \( t \in I \), \( F(.,t): \Sigma \to N \) is an immersion. Then \( F^*TN \) is a vector bundle over \( \Sigma \times I \), which we can equip with the restriction metric \( \bar{g} \) and pullback connection \( F^*\bar{\nabla} \) coming from a Riemannian metric \( \bar{g} \) on \( N \) and its Levi-Civita connection \( \bar{\nabla} \). The map \( F_*: \mathcal{H} 
abla F_\ast \) over \( \Sigma \times I \), which agrees with the Levi-Civita connection \( g \) for each fixed \( t \). We denote by \( \hat{g} \) the metric induced on \( N \), given by \( \hat{g}(\xi,\eta) = \bar{g}(t\xi,t\eta) \). The construction also gives a connection \( \hat{\nabla} := \hat{\nabla} = F^*\bar{\nabla} = F^*\hat{g} \) on \( N \). We denote by \( h \in \Gamma(\mathcal{H}^* \otimes \mathcal{N}^* \otimes N) \) the restriction of \( h^{\mathcal{H}} = F^*\bar{\nabla} F_* \) to \( \mathcal{H} \) in the first argument. Proposition 3 implies that \( h \) is a symmetric bilinear form on \( \mathcal{H} \) with values in \( N \). The remaining components of \( h^{\mathcal{H}} \) are given by

\[
    h^{\mathcal{H}}(\partial_t, v) = \hat{\pi} (F^*\bar{\nabla} F_* v)
    = \hat{\pi} (F^*\bar{\nabla} F_\ast + F_* ([\partial_t, v]))
    = \hat{\nabla} v (\hat{\pi} F_\ast \partial_t) + h(v, \pi F_* \partial_t)
\]

(12)

where we used Proposition 3. Henceforward we restrict to normal variations (with \( \pi F_* \partial_t = 0 \)), since this is the situation for the mean curvature flow. We also define \( \mathcal{W} \in \Gamma(\mathcal{H}^* \otimes N \otimes \mathcal{H}) \) by \( \mathcal{W}(u,\xi) = -h^{N}(u,\xi) = -\pi (F^*\nabla_u t \xi) \) for any \( u \in \Gamma(\mathcal{H}) \) and \( \xi \in \Gamma(N) \) (we refer to this as the Weingarten map). The Weingarten relation (11) gives two identities:

\[
    \hat{g} \left( h(u, v), \xi \right) = g(v, \mathcal{W}(u,\xi)) \quad \text{(13)}
\]

\[
    g \left( h^{N}(\partial_t, \xi), v \right) = -\hat{g} \left( \hat{\nabla} v, \hat{\pi} F_* \partial_t, \xi \right) \quad \text{(14)}
\]

where the latter identity used (12). The Gauss and Codazzi identities for \( \mathcal{H} \) and \( N \) give the following identities for the second fundamental form: First, if \( u \) and \( v \) are in \( \mathcal{H} \), then the Gauss
These follow from our construction and Equation (4): For the first we have (since (24)
\[ \nabla \]
Note that here (19)
\[ \nabla \]
Finally, the Codazzi identities resolve into the tangential Codazzi identities, given by (23)
\[ \nabla \]
while if \( u = \partial_t \) and \( v \in \mathcal{H} \), then we have the identity (18)
\[ \nabla \]
Finally, the Codazzi identities resolve into the tangential Codazzi identities, given by
\[ \nabla_w h(u,v,w) - \nabla_v h(u,w) = \hat{\pi} (R(F_v,F_u)F_v,F_w) \]
for all \( u,v,w \in \Gamma(\mathcal{H}) \), and the ‘timelike’ part, where \( u = \partial_t \) and \( v,w \in \Gamma(\mathcal{H}) \):
\[ \hat{\pi} (R(F_v,F_u)\partial_t)F_w = \nabla_{\partial_t} h(v,w) - \nabla_v \nabla_w \left( \hat{\pi} F_v \partial_t \right) - h (w, W(v, \hat{\pi} F_v \partial_t)) \]
Note that here \( \nabla h \in \Gamma(T^* (\Sigma \times I) \otimes \mathcal{H}^* \otimes \mathcal{H}^* \otimes \mathbb{N}) \) is defined using the connections \( \nabla \) and \( \hat{\nabla} \) as in Equation (4), that is \( \nabla_{\partial_t} h(u,v) = \hat{\nabla}_{\partial_t} (h(u,v)) - h (\nabla_{\partial_t} u, v) - h (u, \nabla_{\partial_t} v) \).
We remark that by construction we have \( \nabla g = 0 \) and \( \hat{\nabla} \tilde{g} = 0 \). In contrast to the situation in other work on evolving hypersurfaces, we have \( \nabla_{\partial_t} g = 0 \). That is, the connections we have constructed automatically build in the so-called ‘Uhlenbeck trick’ [Ha2, Section 2].

**Proposition 4.** The tensors \( F_s \in \Gamma(\mathcal{H}^* \otimes F^* T N) \), \( t \in \Gamma(\mathcal{N}^* \otimes F^* T N) \), \( \pi \in \Gamma(F^* T N \otimes \mathcal{H}) \) and \( \hat{\pi} \in \Gamma(F^* T N \otimes \mathbb{N}) \) satisfy
\[ (\nabla_{\partial_t} F_s)(V) = \tilde{h} (U, V) \]
\[ (\nabla_{\partial_t} \tilde{t})(\xi) = -F_s W(U, \xi) \]
\[ (\nabla_{\partial_t} \tilde{\pi})(X) = \tilde{W}(U, \hat{\pi} X) \]
\[ (\nabla_{\partial_t} \hat{\pi})(X) = -h(U, \pi X) \]
for all \( U, V \in \Gamma(\mathcal{H}) \), \( \xi \in \Gamma(\mathcal{N}) \) and \( X \in \Gamma(F^* T N) \).

**Proof.** These follow from our construction and Equation (4): For the first we have (since \( F_s \) is a \( F^* T N \)-valued tensor acting on \( \mathcal{H} \))
\[ (\nabla_{\partial_t} F_s)(V) = F \nabla U (F_s V) - F_s (\nabla_U V) = F^* (\nabla_U V) + \tilde{h} (U, V) - F_s (\nabla_U V) = \tilde{h} (U, V), \]
where we used the definitions of \( h \) and \( \nabla \). The second identity is similar. For the third we have:

\[
(\nabla_U \pi)(X) = \nabla_U (\pi X) - \pi \left( F^U X \right) \\
= \nabla_U (\pi X) - \pi \left( F^U \left( F_s \pi X + \hat{\pi} X \right) \right) \\
= \nabla_U (\pi X) - \nabla_U (\pi X) + \mathcal{W}(U, \hat{\pi} X) \\
= \mathcal{W}(U, \hat{\pi} X).
\]

The fourth identity is similar to the third. \( \square \)

We illustrate the application of the above identities in the proof of Simons’ identity, which amounts to saying that the second derivatives of the second fundamental form are totally symmetric, up to corrections involving second fundamental form and the curvature of \( N \):

**Proposition 5.**

\[
\nabla_w \nabla_z h(u, v) - \nabla_v \nabla_z h(w, z) = h(v, \mathcal{W}(u, h(w, z))) - h(z, \mathcal{W}(w, h(u, v))) - h(u, \mathcal{W}(w, h(v, z))) \\
+ \nabla_w \mathcal{W}(u, h(w, z)) + \nabla_v \mathcal{W}(u, h(v, z)) - \mathcal{W}(v, \mathcal{W}(w, h(u, z))) \\
- h(u, \pi \hat{\mathcal{R}}(F_s v, F_s w) F_s z) - h(w, \pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v) \\
- h(z, \pi \hat{\mathcal{R}}(F_s u, F_s w) F_s v) - h(v, \pi \hat{\mathcal{R}}(F_s u, F_s w) F_s z) \\
+ \hat{\pi} \hat{\mathcal{R}}(\hat{\pi} h(u, v), F_s w) F_s z - \hat{\pi} \hat{\mathcal{R}}(\hat{\pi} h(w, z), F_s u) F_s v \\
+ \hat{\pi} \hat{\mathcal{R}}(\hat{\pi} h(u, z), F_s w) F_s v - \hat{\pi} \hat{\mathcal{R}}(\hat{\pi} h(w, u), F_s v) F_s z \\
+ \hat{\pi} \hat{\mathcal{W}}(\hat{\pi} F_s h(u, v), F_s w) F_s z - \hat{\pi} \hat{\mathcal{W}}(\hat{\pi} F_s h(w, z), F_s u) F_s v.
\]

**Proof.** Since the equation is tensorial, it suffices to work with \( u, v, w, z \in \Gamma(\mathcal{F}) \) for which \( \nabla u = 0 \), etc., at a given point. Computing at that point we find

\[
\nabla_u \nabla_z h(u, v) = \nabla_w \left( \nabla_u h(z, v) + \pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v \right) \\
= \nabla_u \nabla_w h(z, v) + \left( R(u, w) h \right)(v, z) + \nabla_w \left( \pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v \right) \\
= \nabla_u \nabla_w h(z, v) + \pi \hat{\mathcal{R}}(F_s v, F_s v) F_s z + \left( R(u, w) h \right)(v, z) + \nabla_w \left( \pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v \right) \\
= \nabla_u \nabla_w h(z, v) + \left( R(u, w) h \right)(v, z) + \nabla_w \left( \pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v \right) + \nabla_w \left( \pi \hat{\mathcal{R}}(F_s v, F_s v) F_s z \right)
\]

where we used the Codazzi identity in the first and third lines, and the definition of curvature in the second. Since \( h \) is a \( \mathbb{N} \)-valued tensor with arguments in \( \mathcal{F} \), the second term may be computed using the identity (5) to give

\[
\left( R(u, w) h \right)(v, z) = \hat{\mathcal{R}}(u, w)(h(v, z)) - h(\mathcal{R}(u, w)v, z) - h(v, \mathcal{R}(u, w)z).
\]

This in turn can be expanded using the Gauss identity (15a) for \( R \) and the Ricci identity (17) for \( \hat{\mathcal{R}} \). In the third term (and similarly the fourth) we apply the identity (4) to \( \hat{\pi} \) to get

\[
\nabla_w (\pi \hat{\mathcal{R}}(F_s u, F_s z) F_s v) = \nabla_w \hat{\pi} \left( \mathcal{R}(F_s u, F_s z) F_s v \right) + \hat{\pi} \left( \hat{\mathcal{F}} \nabla_w \left( \mathcal{R}(F_s u, F_s z) F_s v \right) \right).
\]
In the first term here we apply the identity (24). In the second we can expand further as follows:

\[ F_{v} \left( R(F_{u}, F_{z}) F_{v} \right) = \left( F_{v} R \right) \left( F_{u}, F_{z} \right) F_{v} + R(\nabla_{w} F_{u}, F_{z}) F_{v} + R(F_{u} F_{v}, F_{z}) (\Omega_{v} F_{u} F_{z}) \]

In the terms involving \( \nabla F_{v} \) we apply (21), and we also observe that \( F_{v} R = \tilde{\nabla}_{F_{v}} R \) by the definition of the connection \( F_{v} \). Substituting these identities gives the required result.

In subsequent computations we often work in a local orthonormal frame \( \{ e_{i} \} \) for the spatial tangent bundle \( \mathcal{N} \), and a local orthonormal frame \( \{ v_{\alpha} \} \) for the normal bundle \( \mathcal{N} \). We use greek indices for the normal bundle, and latin ones for the tangent bundle. When working in such orthonormal frames we sum over repeated indices whether raised or lowered. For example the mean curvature vector \( H \in \Gamma(\mathcal{N}) \) may be written in the various forms

\[ H = \mathbf{tr} h = g^{ij} h_{ij} = \tilde{h}_{i} = h_{ii} = g^{ij} h_{ij} \]

Similarly, we write \( |h|^{2} = g^{ik} g^{jk} h_{\alpha} h_{\beta} = h_{ij} h_{ij} \). The Weingarten relation (13) becomes

\[ \mathcal{W}(e_{i}, v_{\alpha}) = h_{ij} e_{i}, e_{j}, v_{\alpha} \]

while the Gauss equation (15a) becomes

\[ R_{ijkl} = h_{ika} h_{jlb} - h_{jka} h_{ila} + \tilde{R}_{ijkl}, \]

where we denote \( \tilde{R}_{ijkl} = \tilde{R}(F_{i} e_{j}, F_{k} e_{l}, F_{e_{e_{i}}}) \). The Ricci equations (17) give

\[ \tilde{R}_{ij} = h_{ipa} h_{jp} h_{\beta} - h_{jpa} h_{ip} h_{\beta} + \tilde{R}_{ij} \]

where \( \tilde{R}_{ij} = \tilde{R}(F_{i} e_{i}, F_{e_{j}}) \). The Codazzi identity (19) gives

\[ \nabla_{i} h_{jk} - \nabla_{j} h_{ik} = \tilde{R}_{ij} A_{V}. \]

In this notation the identity from Proposition 5 takes the following form:

\[ \nabla_{k} \nabla_{j} h_{ij} = \nabla_{i} \nabla_{j} h_{ij} + h_{ika} h_{ipa} h_{jp} - h_{ija} h_{kpa} h_{ip} + h_{jla} h_{ipa} h_{kp} - h_{jia} h_{kpa} h_{ip} + h_{ika} \tilde{R}_{ijk} h_{\beta} - h_{jka} \tilde{R}_{ij} h_{\beta} + \tilde{R}_{ik} h_{ip} h_{\beta} - \tilde{R}_{ik} h_{jp} h_{\beta} + h_{jla} \tilde{R}_{ijk} h_{\beta} + h_{jka} \tilde{R}_{ij} h_{\beta} + h_{ika} \tilde{R}_{ijk} h_{\beta} + h_{jka} \tilde{R}_{ij} h_{\beta} + \tilde{R}_{ik} h_{ip} h_{\beta} - \tilde{R}_{ik} h_{jp} h_{\beta}. \]

Particularly useful is the equation obtained by taking a trace of the above identity over \( k \) and \( l \):

\[ \Delta h_{ij} = \nabla_{i} \nabla_{j} H + H \cdot h_{ip} h_{pj} - h_{ij} h_{pq} h_{pq} + 2 h_{ij} h_{jp} h_{pq} - h_{ij} h_{jp} h_{pq} - h_{ij} h_{qp} h_{ij} + H A \tilde{R}_{ij} h_{\beta} - h_{jka} \tilde{R}_{kab} h_{\beta} + \tilde{R}_{j} h_{jp} h_{\beta} + h_{jka} \tilde{R}_{kab} h_{\beta} + \tilde{R}_{j} h_{jp} h_{\beta} - h_{jka} \tilde{R}_{kab} h_{\beta} + \tilde{R}_{j} h_{jp} h_{\beta} \]

\[ + 2 h_{ipa} \tilde{R}_{ipa} h_{\beta} + 2 h_{ipp} \tilde{R}_{ipp} h_{\beta} + \tilde{R}_{j} \tilde{R}_{j} h_{\beta} - \tilde{R}_{j} \tilde{R}_{j} h_{\beta}. \]

(25)

Here the dots represent inner products in \( \mathcal{N} \).

We finish this section with a brief comment about short time existence of MCF. It is well known that the geometric invariance of MCF introduces degeneracies into the principal symbol, so that standard parabolic theory does not immediately apply. This may be circumvented by including a tangential term corresponding to a harmonic map heat flow (the so called DeTurck trick). In [Ha3] Hamilton shows how to use the latter method to achieve short time existence to mean curvature flow of arbitrary codimension. The reader is referred there for the details.
3. Preservation of Curvature Pinching

In this section we wish to prove the following key pinching lemma:

**Lemma 1.** If a solution $F : \Sigma \times [0,T) \to \mathbb{R}^{n+k}$ of MCF satisfies $|h|^2 \leq (\frac{1}{n} + \frac{1}{3n})|H|^2$ at $t = 0$, then this remains true for all $0 \leq t < T$.

We derive evolution equations for the squared lengths of the second fundamental form and the mean curvature vector. The mean curvature flow amounts to the prescription $F, \partial_t = tH$ in the notation of the previous section. The timelike Codazzi identity (20) gives an evolution equation for second fundamental form under MCF of a submanifold in an arbitrary background space $N$:

$$\nabla \partial_t h_{ij} = \nabla \partial_t h_{ij} = \Delta h_{ij} + h_{ij}, h_{pq} h_{pq} + h_{iq}, h_{jp} h_{pj} + h_{jq}, h_{ip} h_{pi} - 2h_{ip}, h_{jq} h_{pq} + 2\bar{R}_{ijpq} h_{pq} - R_{ijkp} h_{pi} - R_{ikjp} h_{pj} + h_{ij}, R_{kikp} v_{\beta} + 2h_{jp}, R_{ipaq} v_{\beta} - 2h_{ip}, R_{jpaq} v_{\beta} + \bar{\nabla} h_{i}, R_{kikp} v_{\beta} - \bar{\nabla} h_{k}, R_{jkip} v_{\beta}.$$

This converts to a reaction-diffusion equation using the identity (25):

$$\nabla \partial_t h_{ij} = \nabla \partial_t h_{ij} = \Delta h_{ij} + h_{ij}, h_{pq} h_{pq} + h_{iq}, h_{jp} h_{pj} + h_{jq}, h_{ip} h_{pi} - 2h_{ip}, h_{jq} h_{pq}.$$

Taking the trace with respect to $g$ we obtain an evolution equation for the mean curvature vector:

$$\nabla \partial_t H = \Delta H + H, h_{pq} h_{pq}.$$

The evolution equations for $|h|^2$ and $|H|^2$ follow from equations (27) and (28):

$$\frac{\partial}{\partial t} |h|^2 = \Delta |h|^2 - 2|\nabla h|^2 + 2 \sum_{i,j} \left( \sum_{\alpha, \beta} h_{ij, \alpha h_{i j, \beta}} \right)^2 + 2 \sum_{i,j, \alpha, \beta} \left( \sum_{p} h_{i p a h_{j p, \beta} - h_{j p a h_{i p, \beta}} \right)^2;$$

$$\frac{\partial}{\partial t} |H|^2 = \Delta |H|^2 - 2|\nabla H|^2 + 2 \sum_{i,j} \left( \sum_{\alpha} H_{\alpha h_{i j, \alpha}} \right)^2.$$
In order to estimate the reaction terms of (31) it is convenient to work with the traceless part of second fundamental form \( \hat{h} = h - \frac{1}{n}gH \). The lengths of \( h \) and \( \hat{h} \) are related by \( |\hat{h}|^2 = |h|^2 - \frac{1}{n}|H|^2 \).

We will often work with a local orthonormal frame \( \{v_\alpha, 1 \leq \alpha \leq k\} \) for \( N \), such that \( v_1 = H/|H| \). With this choice of frame the traceless second fundamental form takes the form

\[
\begin{align*}
\hat{h}_1 &= h_1 - \frac{|H|}{n}Id \\
\hat{h}_\alpha &= h_\alpha & \alpha > 1,
\end{align*}
\]

and

\[
\begin{align*}
\text{tr} \hat{h}_1 &= |H| \\
\text{tr} \hat{h}_\alpha &= 0 & \alpha > 1.
\end{align*}
\]

At a point we may choose a basis for the tangent bundle such that \( h_1 \) is diagonal. We denote by \( \lambda_\iota \) and \( \hat{\lambda}_\iota \) the diagonal entries of \( h_1 \) and \( \hat{h}_1 \) respectively. We denote the norm of the \( (\alpha \neq 1) \)-directions of the second fundamental form by \( |\hat{h}_-|^2 \), i.e. \( |\hat{h}|^2 = |h_1|^2 + |\hat{h}_-|^2 \). One final piece of notation we adopt from [CdCK]: For a matrix \( A = (a_{ij}) \), we denote

\[
N(A) = \text{tr} A \cdot A^t = \sum_{ij}(a_{ij})^2.
\]

In particular, we have

\[
\sum_{\alpha, \beta} N(\hat{h}_\alpha \hat{h}_\beta - \hat{h}_\beta \hat{h}_\alpha) = |\hat{R}|^2.
\]

We begin by estimating the gradient terms: We certainly have \( |\nabla h|^2 - \frac{1}{n}|\nabla H|^2 = |\nabla \hat{h}|^2 \geq 0 \), but we need a stronger inequality:

**Proposition 6.** We have the estimates

\[
\begin{align*}
|\nabla h|^2 &\geq \frac{3}{n+2}|\nabla H|^2 \\
|\nabla h|^2 - \frac{1}{n}|\nabla H|^2 &\geq \frac{2(n-1)}{3n}|\nabla h|^2.
\end{align*}
\]

**Proof.** In exactly the same way as [Hu1] and [Ha1], we decompose the tensor \( \nabla h \) into orthogonal components \( \nabla h_{jk} = E_{ijk} + f_{ijk} \), where

\[
E_{ijk} = \frac{1}{n+2}(g_{ij}\nabla_k H + g_{ik}\nabla_j H + g_{jk}\nabla_i H).
\]

Then \( |\nabla h|^2 \geq |E|^2 = \frac{3}{n+2}|\nabla H|^2 \). The second estimate follows easily from the first. \( \square \)

To estimate the reaction terms, we work with the bases described above and separate the \( \alpha = 1 \) components from the others. The reaction terms of (31) become

\[
\sum_{\alpha, \beta} \left( \sum_{i,j} h_{ij\alpha}h_{ij\beta} \right)^2 = |\hat{h}_1|^4 + \frac{2}{n} |\hat{h}_1|^2 |H|^2 + \frac{1}{n^2} |H|^4 + 2 \sum_{\alpha > 1} \left( \sum_{i,j} \hat{h}_{ij1} \hat{h}_{ij\alpha} \right)^2 + \sum_{\alpha, \beta > 1} \left( \sum_{i,j} \hat{h}_{ij\alpha} \hat{h}_{ij\beta} \right)^2;
\]

\[
|\hat{R}|^2 = 2 \sum_{\alpha > 1} N(h_1 \hat{h}_\alpha - \hat{h}_\alpha h_1) + \sum_{\alpha, \beta > 1} N(\hat{h}_\alpha \hat{h}_\beta - \hat{h}_\beta \hat{h}_\alpha);
\]

\[
\sum_{i,j} \left( \sum_{\alpha} H_{\alpha} h_{ij\alpha} \right)^2 = |\hat{h}_1|^2 |H|^2 + \frac{1}{n} |H|^4.
\]
Writing out all the reaction terms we now have

\[ 2 \sum_{\alpha, \beta} \left( \sum_{i,j} h_{ij\alpha} h_{ij\beta} \right)^2 + 2 |\dot{R}|^2 - 2c \sum_{i,j} \left( \sum_{\alpha} H_\alpha h_{ij\alpha} \right)^2 \]

\[ = 2|h_1|^4 - 2(c - \frac{2}{n})|\dot{h}_1|^2 |H|^2 - \frac{2}{n} (c - \frac{1}{n}) |H|^4 \]

\[ + 4 \sum_{\alpha > 1} \left( \sum_{i,j} \dot{h}_{ij\alpha} h_{ij\alpha} \right)^2 + 4 \sum_{\alpha > 1} N(h_1 h_\alpha - \dot{h}_\alpha h_1) \]

\[ + 2 \sum_{\alpha, \beta > 1} \left( \sum_{i,j} \dot{h}_{ij\alpha} h_{ij\beta} \right)^2 + 2 \sum_{\alpha, \beta > 1} N(h_\alpha h_\beta - \dot{h}_\beta h_\alpha) \]

\[ \leq 2|h_1|^4 - 2|\dot{h}_1|^2 (|\dot{h}_1|^2 + |\dot{h}_-|^2) - \frac{2}{n(c - 1/n)} |\dot{h}_-|^2 (|\dot{h}_1|^2 + |\dot{h}_-|^2) \]

\[ + 4 \sum_{\alpha > 1} \left( \sum_{i,j} \dot{h}_{ij\alpha} h_{ij\alpha} \right)^2 + 4 \sum_{\alpha > 1} N(h_1 h_\alpha - \dot{h}_\alpha h_1) \]

\[ + 2 \sum_{\alpha, \beta > 1} \left( \sum_{i,j} \dot{h}_{ij\alpha} h_{ij\beta} \right)^2 + 2 \sum_{\alpha, \beta > 1} N(h_\alpha h_\beta - \dot{h}_\beta h_\alpha) \]

We need to control the last two lines of the above inequality. In the second last line, we proceed by expanding the terms and using the fact that $\dot{h}_1$ is diagonal:

\[ \sum_{\alpha > 1} \left( \sum_{i,j} \dot{h}_{ij\alpha} h_{ij\alpha} \right)^2 = \sum_{\alpha > 1} \left( \sum_{i} \dot{h}_{i\alpha} h_{i\alpha} \right)^2 \]

\[ \leq \left( \sum_{i} \dot{h}_{i\alpha} \right)^2 \left( \sum_{\alpha > 1} (h_{i\alpha})^2 \right) \]

\[ = |\dot{h}_1|^2 \sum_{\alpha > 1} (h_{i\alpha})^2. \]

Furthermore,

\[ \sum_{\alpha > 1} N(h_1 h_\alpha - \dot{h}_\alpha h_1) = \sum_{i \neq j} \left( \lambda_i - \lambda_j \right)^2 (h_{ij\alpha})^2 \]

\[ = \sum_{i \neq j} \left( \lambda_i - \lambda_j \right)^2 (\dot{h}_{ij\alpha})^2 \]

\[ \leq \sum_{i \neq j} 2(\lambda_i^2 + \lambda_j^2) (\dot{h}_{ij\alpha})^2 \]

\[ \leq 2|h_1|^2 \sum_{i \neq j} (h_{ij\alpha})^2 \]

\[ = 2|h_1|^2 (|\dot{h}_-|^2 - \sum_{\alpha > 1} (h_{i\alpha})^2), \]
and so
\[
\sum_{a>1} \left( \sum_{i,j} h_{ij} h_{ij\alpha} \right)^2 + \sum_{a>1} N(h_{\alpha} h_{\alpha} - \hat{h}_{\alpha} h_{\alpha}) \leq 2|\hat{h}_1|^2|h_\perp|^2 - |\hat{h}_1|^2 \sum_{a>1} (h_{i\alpha})^2 \\
\leq 2|\hat{h}_1|^2|h_\perp|^2.
\]

To estimate the last line, we use an inequality first derived in [CdCK] and later improved to the version we use in [AMJ]. In our notation we have

**Lemma 2.**
\[
\sum_{a,b>1} \left( \sum_{i,j} h_{ij\alpha} h_{ij\beta} \right)^2 + \sum_{a,b>1} N(h_{\alpha} h_{\beta} - \hat{h}_{\alpha} \hat{h}_{\beta}) \leq \frac{3}{2}|\hat{h}_\perp|^4.
\]

**Proof of Theorem 3.** Using the above inequalities we estimate the reaction terms by
\[
2 \sum_{a,b>1} \left( \sum_{i,j} h_{ij\alpha} h_{ij\beta} \right)^2 + 2|\hat{h}_1|^2 - 2c \sum_{i,j} \left( \sum_{\alpha} H_{\alpha} h_{ij\alpha} \right)^2 \\
\leq 2|\hat{h}_1|^4 - 2|\hat{h}_1|^2 |\hat{h}_\perp|^2 - \frac{2}{n(c-1/n)} |\hat{h}_\perp|^2 (|\hat{h}_1|^2 + |\hat{h}_\perp|^2) \\
+ 8|\hat{h}_1|^2|h_\perp|^2 + 3|h_\perp|^4.
\]

Equating coefficients, we find that the $|\hat{h}_1|^4$ terms cancel exactly, the $|\hat{h}_1|^2|h_\perp|^2$ terms are nonpositive for $c \leq \frac{1}{n} + \frac{1}{3n}$ and the $|\hat{h}_\perp|^4$ terms nonpositive for $c \leq \frac{1}{n} + \frac{2}{3n}$. The gradient terms are nonpositive for $c \leq \frac{1}{n+2}$, so both reaction and gradient terms are nonpositive for $c \leq \frac{1}{n} + \frac{1}{3n}$. The maximum principle implies that $|h|^2 \leq c|H|^2$ is preserved by MCF for $c \leq \frac{1}{n} + \frac{1}{3n}$. □

4. **Higher derivative estimates and long time existence**

Here we consider the long time behaviour of MCF and establish the existence of a solution on a finite maximal time interval determined by the blowup of the second fundamental form.

**Theorem 2.** Under the assumptions of Theorem 1, MCF has a unique solution on a finite maximal time interval $0 \leq t < T < \infty$. Moreover, $\max_t |h|^2 \to \infty$ as $t \to T$.

**Lemma 3.** The maximal time of existence $T$ is finite.

**Proof.** Define $Q = |H|^2 - a|h|^2 - b(t)$, where $a = \frac{3n}{n+2}$. The assumptions of Theorem 1 guarantee that for $n \geq 4$ we can choose $t_0 = 0$ and $b(0) = b_0 > 0$ such that $Q \geq 0$ for $t = t_0$. For $n \leq 4$, the strong maximum principle implies $|h|^2 < \frac{4}{3n}|H|^2$ for any $t > 0$, so we can find a small $t_0 > 0$ and $b(t_0) = b_0 > 0$ such that $Q \geq 0$ for $t = t_0$ (the application of the strong maximum principle is detailed at the end of Section 5). We have
\[
\frac{\partial}{\partial t} Q = \Delta Q - 2(|\nabla H|^2 - a|\nabla h|^2) + 2R_2 - 2aR_1 - b'(t).
\]
Estimating the reaction terms as before we obtain
\[
2R_2 - 2aR_1 - b'(t) = 2\sum_{i,j} \left( \sum_a H_a h_{i,j,a} \right)^2 - 2a \sum_{\alpha, \beta} \left( \sum_{i,j} h_{i,j,a} h_{i,j,\beta} \right)^2 - 2a |\hat{R}|^2 - b'(t)
\]
\[
\geq 2|\hat{h}_1|^2 (a|\hat{h}_1|^2 + a|\hat{h}_2|^2 + b) + \frac{2}{n(1-a/n)} (a|\hat{h}_2|^2 + b)(a|\hat{h}_1|^2 + a|\hat{h}_2|^2 + b)
\]
\[
- 2a |\hat{h}_1|^4 - 8a |\hat{h}_1|^2 |\hat{h}_2|^2 - 3a |\hat{h}_2|^4 - b'(t).
\]
Equating coefficients, we find \( Q \geq 0 \) is preserved if \( \frac{db}{dt} \leq \frac{8b^2}{n} \). Since \( b(t_0) = b_0 \), we can take
\[
b(t) = \frac{nb_0}{n - 8b_0(t - t_0)}.
\]
This is unbounded as \( t \to t_0 + \frac{n}{8b_0} \), so we must have \( T \leq t_0 + \frac{n}{8b_0} \). \( \square \)

We next want to prove interior-in-time higher derivative estimates for the second fundamental form. We use Hamilton’s * notation: For tensors \( S \) and \( T \) (that is, sections of bundles constructed from \( \mathcal{H} \) and \( \mathcal{N} \) by taking duals and tensor products) the product \( S \ast T \) denotes any linear combination of contractions of \( S \) with \( T \).

**Proposition 7.** The evolution of the \( m \)-th covariant derivative of \( h \) is of the form
\[
\nabla_i \nabla^m h = \Delta \nabla^m h + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h.
\]

**Proof.** We argue by induction on \( m \). The case \( m = 0 \) is given by the evolution equation for the second fundamental form. Now suppose that the result holds up to \( m - 1 \). Differentiating the \( m \)-th covariant derivative of \( h \) in time and using the timelike Gauss and Ricci equations to interchange derivatives we find
\[
\nabla_i \nabla^m h = \nabla \nabla_i \nabla^{m-1} h + \nabla^{m-1} h \ast \nabla^m h \ast \nabla h
\]
\[
= \nabla (\Delta \nabla^{m-1} h + \sum_{i+j+k=m-1} \nabla^p h \ast \nabla^q h \ast \nabla^r h) + \nabla^{m-1} h \ast \nabla^m h \ast \nabla^k h
\]
\[
= \nabla \Delta \nabla^{m-1} h + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h.
\]
The formula for commuting the Laplacian and gradient of a normal-valued tensor is given by:
\[
\Delta V_{\ell} T = V_{\ell} \Delta T + \nabla_m (R(\partial_\ell, \partial_m)T) + \left( (R(\partial_\ell, \partial_m)(\nabla T)) (\partial_m) \right).
\]
Since \( T \) and \( \nabla T \) are \( N \)-valued tensors acting on \( \mathcal{H} \), equation (5) gives expressions for \( R(\partial_\ell, \partial_m)T \) as \( R \ast T + \hat{R} \ast T \), and similarly \( R(\partial_\ell, \partial_m)\nabla T = R \ast \nabla T + \hat{R} \ast \nabla T \), where \( R \) and \( \hat{R} \) are the curvature tensors on \( \mathcal{H} \) and \( \mathcal{N} \), which are both of the form \( h \ast h \). The terms arising in commuting the gradient and Laplacian of \( \nabla^{m-1} h \) are of the form \( \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h \), so we obtain
\[
\nabla_i \nabla^m h = \Delta \nabla^m h + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h
\]
as required. \( \square \)
Proposition 8. The evolution of $|\nabla^m h|^2$ is of the form

$$\frac{\partial}{\partial t} |\nabla^m h|^2 = \Delta |\nabla^m h|^2 - 2|\nabla^{m+1} h|^2 + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h \ast \nabla^m h.$$

Proof. Denoting by angle brackets the inner product on $\otimes^{m+2} T^\ast \otimes N$, which is compatible with the connection on the same bundle, we have

$$\frac{\partial}{\partial t} |\nabla^m h|^2 = \frac{\partial}{\partial t} \langle \nabla^m p h, \nabla^m p h \rangle$$

$$= 2\langle \nabla^m p h, \nabla^m p h \rangle$$

$$= 2\langle \nabla^m p h, \nabla^m p h + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h \rangle$$

$$= \Delta |\nabla^m h|^2 - 2|\nabla^{m+1} h|^2 + \sum_{i+j+k=m} \nabla^i h \ast \nabla^j h \ast \nabla^k h \ast \nabla^m h$$

as required. $\Box$

Lemma 4. Suppose that mean curvature flow of a given submanifold $\Sigma_0$ has a solution on a time interval $t \in [0, \tau]$. If $|h|^2 \leq K$ for all $t \in [0, \tau]$, then $|\nabla^m h|^2 \leq C_m (1 + 1/t^m)$ for all $t \in (0, \tau]$, where $C_m$ is a constant that depends on $m$, $n$ and $K$.

The strength of this estimate is that assuming only a bound on the second fundamental form (and no information about it’s derivatives) we can bound all higher derivatives. The fact that these estimates blow up as $t$ approaches zero poses no difficulty, since the short-time existence result bounds all derivatives of $h$ for a short time. While not crucial here, the interior-in-time estimates are useful in the singularity analysis of Section 7.

Proof. The proof is by induction on $m$. We first prove the Lemma for $m = 1$. We consider the quantity $G = t|\nabla h|^2 + |h|^2$, which has a bound at $t = 0$ depending only on curvature. The strategy is now to use the good term from the evolution of $|h|^2$ to control the bad term in the evolution of $|\nabla h|^2$: Differentiating $G$ we get

$$\frac{\partial G}{\partial t} = |\nabla h|^2 + t(\Delta |\nabla h|^2 - 2|\nabla^2 h|^2 + h \ast h \ast \nabla h \ast \nabla h)$$

$$+ (\Delta |h|^2 - 2|\nabla h|^2 + h \ast h \ast h \ast h)$$

$$\leq \Delta G + (c_1 t|h|^2 - 1)|\nabla h|^2 + c_2 |h|^4.$$

for $t \leq 1/(c_1 K)$ we can estimate

$$\frac{\partial}{\partial t} G \leq \Delta G + c_2 K^2,$$

and the maximum principle implies $\max_{\Sigma} G \leq K + c_2 K^2 t$. Then $|\nabla h|^2 \leq G/t \leq K/t + c_2 K^2$ for $t \in (0, 1/(c_1 K)]$. If $t > 1/(c_1 K)$ we apply the same argument on the interval $[t - 1/(c_1 K), t]$, yielding $|\nabla h|^2(t) \leq (c_1 + c_2)K^2$. This completes the proof for $m = 1$. Now suppose the estimate
The difference \( T \) time-independent metric on \( H \) bundle induced by the connections \( \tilde{\nabla} \) constructed from \( H \).

Proof of Theorem 2. Fix a smooth metric \( \tilde{g} \) on \( \Sigma \) with Levi-Civita connection \( \tilde{\nabla} \). \( \tilde{g} \) extends to a time-independent metric on \( \mathcal{H} \), and \( \tilde{\nabla} \) extends to \( \mathcal{H} \) by taking \( \tilde{\nabla}_{\partial_t} u = 0 \) whenever \([\partial_t, u] = 0 \).

The difference \( T = \nabla - \tilde{\nabla} \) restricts to a section of \( \mathcal{H}^* \otimes \mathcal{H}^* \otimes \mathcal{H} \). If \( S \) is a section of a bundle constructed from \( \mathcal{H}, \mathcal{N} \) and \( F^* TN \), \( \nabla S \) denotes the derivative of \( S \) with the connection on this bundle induced by the connections \( \tilde{\nabla} \) on \( \mathcal{H} \), \( \tilde{\nabla} \) on \( \mathcal{N} \), and \( \nabla \) on \( F^* TN \), so that \( \nabla S - \nabla S = S \cdot T \).

To prove Theorem 2 we assume that \( |h| \) remains bounded on the interval \([0, T] \), and derive a contradiction. This suffices to prove the Theorem, since if \(|h| \) is bounded on any subsequence of times approaching \( T \), then Equation (29) implies that \(|h| \) is bounded on \( \Sigma \times [0, T] \). Under this assumption the boundedness of \( \tilde{\nabla}_t g = -2H \cdot h \) implies that the metric \( g \) remains comparable to \( \tilde{g} \): we have for any non-zero vector \( v \in T \Sigma \)

\[
\left| \frac{\partial}{\partial t} \left( \frac{g(v,v)}{\tilde{g}(v,v)} \right) \right| = \left| \nabla_t g(v,v) \frac{g(v,v)}{\tilde{g}(v,v)} \right| \leq 2|H||h| \frac{g(v,v)}{\tilde{g}(v,v)},
\]

so that the ratio of lengths is controlled above and below by exponential functions of time, and hence since the time interval is bounded, there exists a positive constant \( c_0 \) such that

\[
(33) \quad \frac{1}{c_9} \tilde{g} \leq g \leq c_9 \tilde{g}.
\]
Next we observe that covariant derivatives of all orders of $F$ with respect to $\tilde{\nabla}$ can be expressed in terms of $\tilde{h}$ and $T$ and their derivatives: We prove by induction that

$$\tilde{\nabla}^k F = F_* \tilde{\nabla}^{k-2} T + F_* \left( \sum_{i_0 + 2i_1 + \cdots + (k-2)i_{k-3} = k-1} T^{i_0} \ast (\tilde{\nabla} T)^{i_1} \ast \cdots \ast (\tilde{\nabla}^{k-3} T)^{i_{k-3}} \right) + (t + F_* \ast \sum_{j=1}^{k-1} \left( \sum_{\Sigma(n+1)_{(j-1)-j} = n} \prod_{m=0}^{k-2-j} (\tilde{\nabla}^m T)^{j_m} \right) \ast \left( \sum_{\Sigma(m+1)_{(j)-j} = m} \prod_{m=0}^{j-1} (\nabla^m h)^{p_m} \right).$$

This is true for $k = 2$, since

$$\tilde{\nabla}^2_{u,v} F = F_* \tilde{\nabla} u (F_* v) - F_* (\tilde{\nabla} u v) = F_* (\tilde{\nabla} u v - \tilde{\nabla} u v) + t h_{u,v} = F_* T_{u,v} + t h_{u,v}.$$  

To deduce the result for higher $k$ by induction, we note that Equation (35) implies a formula for the derivative of $F_*:\nabla F_*(V) = F_* T(,V) + t h(,V) = F_* \nabla V + t h * V$,

while Equation (22) gives

$$\tilde{\nabla} F_*(\xi) = -F_* \tilde{\nabla} (. \xi) = F_* h * \xi.$$  

The result for $k + 1$ now follows by differentiating the expression (34), and writing $\tilde{\nabla} (\nabla^m h) = \nabla^{m+1} h + \nabla^m h * T$. It follows that if $|\tilde{\nabla}^j F|_{g}$ is bounded for $j = 1, \ldots, k - 1$, then

$$|\tilde{\nabla}^{k-2} T|_{g} \leq C \left( 1 + |\tilde{\nabla}^k F|_{g} \right).$$

The above observations allow us to prove $C^k$ convergence of $F$ as $t \to T$ for every $k$: We have $\tilde{\nabla} F = t H$, so the boundedness of $H$ implies that $F$ remains bounded and converges uniformly as $t \to T$. Differentiating as above, we find by induction that

$$\tilde{\nabla}^k \tilde{\nabla} F = (F_* + 1) * \sum_{\Sigma(n+1)_{(j-1)-j} = n} \prod_{m=0}^{k-2-j} (\tilde{\nabla}^m T)^{j_m} \ast \left( \sum_{\Sigma(m+1)_{(j-2)-j} = m} \prod_{m=0}^{j-1} (\nabla^m h)^{p_m} \right).$$

Suppose we have established a bound on $|\tilde{\nabla}^j F|_{g}$ for $j \leq k - 1$. Then using the estimate (36), the bounds on $|\nabla^m h|_{g}$ from Lemma 4, and the comparability of $g$ and $\tilde{g}$ from (33) we can estimate

$$|\tilde{\nabla} F|_{g} \leq C \left( 1 + |\tilde{\nabla}^{k-2} T|_{g} \right) \leq C \left( 1 + |\tilde{\nabla}^k F|_{g} \right),$$

so that $|\tilde{\nabla} F|_{g}$ remains bounded, and $\tilde{\nabla} F$ converges uniformly as $t \to T$. This completes the induction, proving that $F(,.t)$ converges in $C^m$ to a limit $F(,.T)$ which is an immersion.

Finally, applying the short time existence result with initial data $F(,.T)$, we deduce that the solution can be continued to a larger time interval, contradicting the maximality of $T$. This completes the proof of Theorem 2. \hfill \Box

5. A PINCHING ESTIMATE FOR THE TRACELESS SECOND FUNDAMENTAL FORM

In this section we prove a pinching estimate for the traceless second fundamental form. This is the key estimate that will imply that the submanifold is evolving to a “round” point.
Theorem 3. There exist constants $C_0 < \infty$ and $\delta > 0$ both depending only on $\Sigma_0$ such that for all time $t \in [0, T)$ we have the estimate

$$|\dot{h}|^2 \leq C_0|H|^{2-\delta}.$$  

We wish to bound the function $f_\sigma = (|h|^2 - |H|^2/n)/|H|^{2(1-\sigma)}$ for sufficiently small $\sigma$. As in the hypersurface case, a distinguishing feature of mean curvature flow when compared to Ricci flow is that this result cannot be proved by a maximum principle argument alone. Some-what more technical integral estimates and a Stampacchia iteration procedure are required. We proceed by first deriving an evolution equation for $f_\sigma$.

Proposition 9. For any $\sigma \in (0, 1)$ we have the evolution equation

$$\frac{\partial}{\partial t} f_\sigma \leq \Delta f_\sigma + \frac{4(1-\sigma)}{|H|^2} \langle \nabla |H|, \nabla f_\sigma \rangle - \frac{2\epsilon_\nu}{|H|^{2(1-\sigma)}} |\nabla H|^2 + 2\sigma |h|^2 f_\sigma. $$

Proof. Differentiating $f_\sigma$ in time and substituting in the evolutions equations for the lengths of the second fundamental form and mean curvature squared we get

$$\frac{\partial}{\partial t} f_\sigma = \frac{\Delta |h|^2 - 2\nabla |H|^2 + 2R_1}{(|H|^2)^{1-\sigma}} - \frac{1}{n} \frac{(\Delta |H|^2 - 2|\nabla H|^2 + 2R_2)}{(|H|^2)^{1-\sigma}} - \frac{(1-\sigma)(|h|^2 - 1/n|H|^2)}{(|H|^2)^{2-\sigma}} (\Delta |H|^2 - 2|\nabla H|^2 + 2R_2).$$

Substituting in the Laplacian of $f_\sigma$:

$$\Delta f_\sigma = \frac{\Delta |h|^2 - 1/n|H|^2}{(|H|^2)^{1-\sigma}} - \frac{2(1-\sigma)}{(|H|^2)^{2-\sigma}} \langle \nabla |h|^2 - 1/n|H|^2, \nabla |H|^2 \rangle - \frac{(1-\sigma)(|h|^2 - 1/n|H|^2)}{(|H|^2)^{2-\sigma}} \Delta |H|^2 + \frac{(2-\sigma)(1-\sigma)(|h|^2 - 1/n|H|^2)}{(|H|^2)^{1-\sigma}} |\nabla |H|^2|^2$$

and using the following identity:

$$- \frac{2(1-\sigma)}{(|H|^2)^{2-\sigma}} \langle \nabla |h|^2 - 1/n|H|^2, \nabla |H|^2 \rangle = \frac{2(1-\sigma)}{|H|^2} \langle \nabla |h|^2, \nabla f_\sigma \rangle - \frac{8(1-\sigma)^2}{(|H|^2)^{2-\sigma}} f_\sigma |H|^2 |\nabla |H|^2|^2,$$

equation (39) can be manipulated into the form

$$\frac{\partial}{\partial t} f_\sigma = \Delta f_\sigma + \frac{2(1-\sigma)}{|H|^2} \langle \nabla |h|^2, \nabla f_\sigma \rangle - \frac{2}{(|H|^2)^{1-\sigma}} \left( |\nabla h|^2 - \frac{|h|^2}{|H|^2} |\nabla H|^2 \right) + \frac{2\sigma R_2 f_\sigma}{|H|^2} - 4\sigma f_\sigma |H|^2 |\nabla |H|^2|^2 + \frac{2}{(|H|^2)^{1-\sigma}} |R_1 - |h|^2/H|^2 R_2|.$$  

We discard the terms on the last line as these are nonpositive under our pinching assumption. The gradient terms on the first line may be estimated as follows:

$$- \frac{2}{(|H|^2)^{1-\sigma}} \left( |\nabla h|^2 - \frac{|h|^2}{|H|^2} |\nabla H|^2 \right) \leq - \frac{2}{(|H|^2)^{1-\sigma}} \left( \frac{3}{n+2} - c \right) |\nabla H|^2,$$

and also $R_2 \leq |h|^2 |H|^2$. Importantly, observe that if $c \leq 4/(3n)$, then $\epsilon_\nu := 3/(n+2) - c$ is strictly positive.  \(\square\)
The reaction term $2\sigma |h|^2 f_\alpha$ in this evolution equation is positive and hence we cannot apply the maximum principle. As in the hypersurface case, we exploit the negative term involving the gradient of the mean curvature by integrating a suitable form of Simons' identity: Contracting equation (25) with the second fundamental form we obtain the following:

$$
\frac{1}{2} \Delta |h|^2 = \hat{h}_{ij} \cdot \nabla \nabla_j H + (|\nabla h|^2 - \frac{1}{n} |\nabla H|^2) + \frac{1}{n} |H| \Delta |H| + Z,
$$

where

$$
Z = - \sum_{\alpha, \beta} \left( \sum_{i,j} h_{ij\alpha} h_{ij\beta} \right)^2 - |\hat{R}|^2 + \sum_{i,j,p} \sum_{\alpha, \beta} h_{i\alpha} h_{j\beta} h_{p\alpha} h_{p\beta}.
$$

**Lemma 5.** For a submanifold $\Sigma^n$ of $\mathbb{R}^{n+k}$, if the second fundamental form of $\Sigma^n$ satisfies $|h|^2 \leq c|H|^2$, where

$$
c < \begin{cases} 
\frac{3}{16}, & \text{if } n = 2; \\
\frac{181}{384}, & \text{if } n = 3; \\
\frac{1}{n-1}, & \text{if } n \geq 4,
\end{cases}
$$

then there exists a strictly positive constant $\varepsilon$ depending only on $\Sigma_0$ such that $Z \geq \varepsilon |\hat{h}|^2 |H|^2$.

In dimension two and three our pinching condition is preserved for $n \leq 2/3$ and $n \leq 4/9$ respectively. As $2/3 < 3/4$ and $4/9 < 181/384$ we do not need to assume any pinching beyond that of the pinching lemma to prove Theorem 3. In dimensions greater than or equal to four, we will soon see that we need the stronger condition that $c < 1/(n-1)$.

**Proof of Lemma 5.** Working with the local orthonormal frames of Section 3 we expand $Z$ to get

$$
Z = -|\hat{h}|^4 + \frac{1}{n} |\hat{h}|^2 |H|^2 + \frac{1}{n} |\hat{h}^-|^2 |H|^2 - 2 \sum_{\alpha, \beta} \left( \sum_{i,j} \hat{h}_{ij\alpha} \hat{h}_{ij\beta} \right)^2 - 2 \sum_{\alpha, \beta} N(\hat{h}_{\alpha} \hat{h}_{\alpha} - \hat{h}_{\beta} \hat{h}_{\beta})
$$

$$
+ \sum_{\alpha, \beta} |H| \lambda_{\alpha}^2 + \sum_{\alpha, \beta} |H| \lambda_{\beta}^2 + \sum_{\alpha, \beta} |H| \lambda_{\alpha} \lambda_{\beta}.
$$

For dimension $n = 2$ all the above terms can be computed explicitly to give $Z \geq c_1 |\hat{h}|^4$ for $c < 3/4$, where $c_1 > 0$ depends only on $\Sigma_0$. The Lemma then follows using the same argument we use below for the cases $n \geq 3$. For the cases $n \geq 3$ we cannot easily calculate the terms of $Z$ explicitly and we proceed by estimating the various terms. We can estimate the first summation term on line one and the two terms on line two as before, that is

$$
-2 \sum_{\alpha, \beta} \left( \sum_{i,j} \hat{h}_{ij\alpha} \hat{h}_{ij\beta} \right)^2 \geq -2 |\hat{h}|^2 \sum_{\alpha, \beta} (\hat{h}_{\alpha} \hat{h}_{\beta})^2
$$

$$
- \sum_{\alpha, \beta} \left( \sum_{i,j} \hat{h}_{ij\alpha} \hat{h}_{ij\beta} \right)^2 - \sum_{\alpha, \beta} N(\hat{h}_{\alpha} \hat{h}_{\beta} - \hat{h}_{\beta} \hat{h}_{\alpha}) \geq \frac{3}{2} |\hat{h}^-|^4,
$$

however we need to work somewhat harder with the remaining summation terms.
Proposition 10. For any $\eta \geq 8$ we have the following estimate

$$-2 \sum_{\alpha > 1} N(h_1 \hbar_{\alpha} - \bar{h}_{\alpha} h_1) + \sum_{\alpha \neq j} |H| \hat{\lambda}_i(h_{ij\alpha})^2 \geq -\frac{\eta}{2} |h_1|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2) - \frac{1}{4\eta} |H|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2).$$

Proof. We estimate

$$-2 \sum_{\alpha > 1} N(h_1 \hbar_{\alpha} - \bar{h}_{\alpha} h_1) + \sum_{\alpha \neq j} |H| \hat{\lambda}_i(h_{ij\alpha})^2$$

$$= -\sum_{\alpha \neq j} \left\{ 2(\hat{\lambda}_i - \hat{\lambda}_j)^2 - \frac{|H|}{2} (\hat{\lambda}_i + \hat{\lambda}_j) \right\} (\hbar_{ij\alpha})^2$$

$$\geq -\sum_{\alpha \neq j} \left\{ (2 + \eta/4) \hat{\lambda}_i^2 \hat{\lambda}_j^2 + (\eta/2 - 4) \hat{\lambda}_i \hat{\lambda}_j \right\} (\hbar_{ij\alpha})^2 - \frac{1}{4\eta} |H|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2)$$

$$= -\sum_{\alpha \neq j} \left\{ (2 + \eta/4) \hat{\lambda}_i^2 \hat{\lambda}_j^2 + (\eta/2 - 2) \hat{\lambda}_i \hat{\lambda}_j \right\} (\hbar_{ij\alpha})^2 - \frac{1}{4\eta} |H|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2)$$

$$= -\frac{\eta}{2} \sum_{\alpha \neq j} (\hat{\lambda}_i^2 + \hat{\lambda}_j^2) (\hbar_{ij\alpha})^2 - \frac{1}{4\eta} |H|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2)$$

$$\geq -\frac{\eta}{2} |h_1|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2) - \frac{1}{4\eta} |H|^2 (|\hbar_{\cdot \cdot}|^2 - \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2)$$

as required. \qed

To estimate the remaining two terms we take the following two inequalities from [AdC] and [Sa]

$$\sum_{\alpha} |H| \hat{\lambda}_i \hat{\lambda}_i \hat{\lambda}_i \geq -\frac{n - 2}{\sqrt{n(n - 1)}} |H| |h_1|^3$$

$$\sum_{\alpha} |H| \hat{\lambda}_i (\hbar_{i\alpha})^2 \geq -\frac{n - 2}{\sqrt{n(n - 1)}} |H| |h_1| \sum_{\alpha \neq i} (\hat{h}_{i\alpha})^2,$$
and further estimate using Peter-Paul to get
\[
\sum_{\alpha > 1} |H| \hat{\lambda}_\alpha^3 \geq -\frac{\mu}{2} |\hat{h}_1|^4 - \frac{1}{2\mu \xi} \left( \frac{n-2}{2} \right)^2 |\hat{h}_1|^2 |H|^2
\]
\[
\sum_{\alpha > 1} |H| \hat{\lambda}_\alpha (\hat{h}_{i\alpha})^2 \geq -\rho |\hat{h}_1|^3 \sum_{i > 1} (\hat{h}_{i\alpha})^2 - \frac{1}{4\rho \eta} \left( \frac{n-2}{2} \right)^2 |H|^2 \sum_{i > 1} (\hat{h}_{i\alpha})^2.
\]
Putting everything together we obtain
\[
Z \geq -|\hat{h}_1|^4 + \frac{1}{n} |\hat{h}_1|^2 |H|^2 + \frac{1}{n} |\hat{h}_-|^2 |H|^2 - 2|\hat{h}_1|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2 - \frac{3}{2} |\hat{h}_-|^4
\]
\[-\frac{\eta}{2} |\hat{h}_1|^2 |\hat{h}_-|^2 + \frac{\eta}{2} |\hat{h}_1|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2 - \frac{1}{4\eta} |H|^2 |\hat{h}_-|^2 + \frac{1}{4\eta} |H|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2
\[-\frac{\mu}{2} |\hat{h}_1|^4 - \frac{1}{2\xi} \left( \frac{n-2}{2} \right)^2 |\hat{h}_1|^2 |H|^2 - \rho |\hat{h}_1|^3 \sum_{i > 1} (\hat{h}_{i\alpha})^2 - \frac{1}{4\rho \eta} \left( \frac{n-2}{2} \right)^2 |H|^2 \sum_{i > 1} (\hat{h}_{i\alpha})^2.
\]
We now need to choose the optimal values of the constants \(\eta, \mu, \rho\). First we choose \(\mu\) to be equal to \(n - 2\). This is valid for all \(n \geq 3\). We next choose \(\rho\) to cancel the \(|\hat{h}_1|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2\) terms, that is we want
\[
\left( \frac{\eta}{2} - 2 - \rho \right) |\hat{h}_1|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2 = 0
\]
and so we choose
\[
\rho = \frac{\eta}{2} - 2 = \frac{1}{2} (\eta - 4).
\]
This is always valid since \(\eta \geq 8\). The only mildly troublesome term that remains is
\[
\left( \frac{1}{4\eta} - \frac{1}{2(\eta - 4) n(n-1)} \right) |H|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2.
\]
The optimal choice for \(\eta\) is \(n + 2\), however we still also require that \(\eta \geq 8\). Consequently in dimensions three, four and five we choose \(\eta = 8\) and for all higher dimensions choose \(\eta = n + 2\).
In dimension three our choice of \(\eta\) makes the above term positive and we discard it in this case.
In dimensions four and higher the term is negative and we estimate
\[
-\left( \frac{1}{2(\eta - 4) n(n-1)} - \frac{1}{4\eta} \right) |H|^2 \sum_{\alpha > 1} (\hat{h}_{i\alpha})^2 \geq - \left( \frac{1}{2(\eta - 4) n(n-1)} - \frac{1}{4\eta} \right) |H|^2 |\hat{h}_-|^2.
\]
After substituting in our choices for \(\rho, \mu, \beta\) we have, in dimension three
\[
Z = -|\hat{h}_1|^4 + \frac{1}{n} |\hat{h}_1|^2 |H|^2 + \frac{1}{n} |\hat{h}_-|^2 |H|^2 - 3 |\hat{h}_-|^4 - 4 |\hat{h}_1|^2 |\hat{h}_-|^2
\]
\[-\frac{1}{32} |H|^2 |\hat{h}_-|^2 - \frac{n-2}{2} |\hat{h}_1|^4 - \frac{n-2}{2n(n-1)} |\hat{h}_1|^2 |H|^2,
\]
in dimensions four and five

\[
Z \geq -|\hat{h}_1|^4 + \frac{1}{n}|\hat{h}_1|^2|H|^2 + \frac{1}{n}|\hat{h}_-|^2|H|^2 - \frac{3}{2} |\hat{h}_-|^4 - 4|\hat{h}_1|^2|\hat{h}_-|^2
- \frac{n-2}{2}|\hat{h}_1|^4 - \frac{n-2}{2n(n-1)}|\hat{h}_1|^2|H|^2 - \frac{(n-2)^2}{8n(n-1)}|H|^2|\hat{h}_-|^2, 
\]

and in dimensions six and higher

\[
Z \geq -|\hat{h}_1|^4 + \frac{1}{n}|\hat{h}_1|^2|H|^2 + \frac{1}{n}|\hat{h}_-|^2|H|^2 - \frac{3}{2} |\hat{h}_-|^4 - \frac{n+2}{2} |\hat{h}_1|^2|\hat{h}_-|^2 
- \frac{n/2}{2n(n-1)}|\hat{h}_1|^4 - \frac{n-2}{2n(n-1)}|\hat{h}_1|^2|H|^2 - \frac{n-2}{2n(n-1)}|H|^2|\hat{h}_-|^2. 
\]

We now group like terms, estimate $|H|^2$ from below by $|\hat{h}_1|^2/(c - 1/n)$ and calculate the maximum value of $c$ permissible in each case such that the coefficients are all strictly positive. For $n = 3$ we have

\[
Z \geq \left( -1 + \frac{1}{n(c - 1/n)} \right) - \frac{n-2}{2} - \frac{n-2}{2n(n-1)(c - 1/n)} |\hat{h}_1|^4 
+ \left( \frac{2}{n(c - 1/n)^2} - 4 \frac{1}{32(c - 1/n)^2} - \frac{n-2}{2n(n-1)(c - 1/n)} \right) |\hat{h}_1|^2|\hat{h}_-|^2 
+ \left( \frac{1}{n(c - 1/n)} - \frac{3}{2} - \frac{1}{32(c - 1/n)^2} \right) |\hat{h}_-|^4. 
\]

The $|\hat{h}_1|^4$ terms are strictly positive for $c < 1/(n-1) = 1/2$, the $|\hat{h}_1|^2|\hat{h}_-|^2$ terms for $c < 181/384$ and the $|\hat{h}_-|^4$ terms for $c < 77/144$. Note that in this case the smallest term is the mixed term; this too is the case when $n = 2$. The higher dimensional cases follow similarly however the smallest term in all these cases is now the $|\hat{h}_1|^4$ term. Furthermore, we find this term is always identically zero for all dimensions $n \geq 4$ when $c = 1/(n-1)$. We have now shown for the values of $c$ stated in the Proposition and strictly positive constants $c_2$, $c_3$, and $c_4$ depending on $\Sigma_0$ that

\[
Z \geq c_2|\hat{h}_1|^4 + c_3|\hat{h}_1|^2|\hat{h}_-|^2 + c_4|\hat{h}_-|^4 
\geq c_5|\hat{h}|^4, 
\]

(41)

where $c_5 = \min\{c_2, c_3/2, c_4\}$. To prove the desired estimate we note that by using Peter-Paul on various terms of $Z$ we can estimate

\[
Z \geq c_6|\hat{h}|^2|H|^2 - c_7|\hat{h}|^4.
\]

Combining this with (41) gives for any $a \in [0, 1]$ that

\[
Z \geq a(c_6|\hat{h}|^2|H|^2 - c_7|\hat{h}|^4) + (1 - a)c_5|\hat{h}|^4.
\]

Choosing $a = c_5/(c_5 + c_7)$ gives

\[
Z \geq \frac{c_5c_6}{c_5 + c_7}|\hat{h}|^2|H|^2
\]

and the Lemma is complete by setting $\varepsilon = c_5c_6/(c_5 + c_7).$ \qed
Proposition 11. For any $p \geq 2$ and $\eta > 0$ we have the estimate

$$\int_{\Sigma} f_{\sigma}^p |H|^2 d\mu_g \leq \frac{(p\eta + 4)}{e} \int_{\Sigma} f_{\sigma}^{p-1} |\nabla H|^2 d\mu_g + \frac{p-1}{e\eta} \int_{\Sigma} f_{\sigma}^{p-2} |\nabla f_{\sigma}|^2 d\mu_g.$$ 

Proof. Using the contracted form of Simons’ identity we can express the Laplacian of $f_{\sigma}$ as

$$\Delta f_{\sigma} = \frac{2}{|H|^{2(1-\sigma)}} \langle \dot{h}_{ij}, \nabla f_{\sigma} \rangle + \frac{2}{|H|^{2(1-\sigma)}} Z - \frac{4(1-\sigma)}{|H|} \langle \nabla |H|, \nabla f_{\sigma} \rangle - \frac{2(1-\sigma)}{|H|} f_{\sigma} \Delta |H|$$

$$+ \frac{2}{|H|^{2(1-\sigma)}} \left( |\nabla h|^2 - \frac{1}{n} |\nabla H|^2 \right) + \frac{4\sigma(1-\sigma)}{|H|^2} |\nabla H|^2.$$ 

The two terms on the last line are non-negative and we discard them. We multiply the remaining terms by $f_{\sigma}^{p-1}$ and integrate over $\Sigma$. On the left, and in the last term on line one we use Green’s first identity, and in integrating the first term on the right we use the Divergence Theorem and the Codazzi equation. The term arising from integrating on the left is non-negative and we discard it. Two other terms arising from the integration combine, ultimately giving

$$2 \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} Z d\mu_g \leq 2(p-1) \int_{\Sigma} \frac{f_{\sigma}^{p-2}}{|H|^{2(1-\sigma)}} \langle \nabla f_{\sigma} \cdot \dot{h}_{ij}, \nabla f_{\sigma} \rangle d\mu_g$$

$$- 4(1-\sigma) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)+1}} \langle \nabla |H|, \dot{h}_{ij}, \nabla f_{\sigma} \rangle d\mu_g + \frac{2(n-1)}{n} \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g$$

$$- 2(1-\sigma)(p-2) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} \langle \nabla |H|, \nabla f_{\sigma} \rangle d\mu_g + 2(1-\sigma) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla f_{\sigma}|^2 d\mu_g.$$ 

The terms with an inner product do not have a sign. Using the Cauchy-Schwarz and Young’s inequality, the inequalities $f_{\sigma} \leq c |H|^{2\sigma}$, $|\nabla H|^2 \leq |\nabla H|^2$, $1-\sigma \leq 1$ and $|\dot{h}| = f_{\sigma}|H|^{2(1-\sigma)}$ we estimate each term as follows:

$$\begin{cases} 
2(p-1) \int_{\Sigma} \frac{f_{\sigma}^{p-2}}{|H|^{2(1-\sigma)}} \langle \nabla f_{\sigma} \cdot \dot{h}_{ij}, \nabla f_{\sigma} \rangle d\mu_g \\
\leq \frac{p-1}{\eta} \int_{\Sigma} \frac{f_{\sigma}^{p-2}}{|H|^{2(1-\sigma)}} |\nabla f_{\sigma}|^2 d\mu_g + (p-1) \eta \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g \\
- 4(1-\sigma) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)+1}} \langle \nabla |H|, \dot{h}_{ij}, \nabla f_{\sigma} \rangle d\mu_g \leq 4c \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g \\
- 2(1-\sigma)(p-2) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} \langle \nabla |H|, \nabla f_{\sigma} \rangle d\mu_g \\
\leq \frac{p-2}{\mu} \int_{\Sigma} \frac{f_{\sigma}^{p-2}}{|H|^{2(1-\sigma)}} |\nabla f_{\sigma}|^2 d\mu_g + (p-2) \mu c \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g \\
(2-\sigma) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla f_{\sigma}|^2 d\mu_g \leq 2c \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g.
\end{cases}$$ 

Putting all the estimates together we obtain

$$2 \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} Z d\mu_g \leq \left( 6c + \frac{2(n-1)}{n} + (p-1) \eta + (p-2) \mu c \right) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 d\mu_g$$

$$+ \left( \frac{p-1}{\eta} + \frac{p-2}{\mu} \right) \int_{\Sigma} \frac{f_{\sigma}^{p-2}}{|H|^{2(1-\sigma)}} |\nabla f_{\sigma}|^2 d\mu_g.$$ 

Our use for this inequality will be to show that sufficiently high $L^p$ norms of $f_{\sigma}$ are bounded. We are not interested in finding optimal values of $p$ and consequently we are going to be a little
rough with the final estimates in order to put the Lemma into a convenient form. Setting $\mu = \eta$, and using $c \leq 1$, $p - 2 \leq p - 1 \leq p$ and Lemma 5 we get

$$2\varepsilon \int_{\Sigma} f_\sigma^p |H|^2 \leq (2p\eta + 8) \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|^{2-\sigma}} |\nabla H|^2 \, d\mu_\sigma + \frac{2(p - 1)}{\eta} \int_{\Sigma} f_\sigma^{p-2} |\nabla f_\sigma|^2 \, d\mu_\sigma.$$ 

Dividing through by $2\varepsilon$ completes the Lemma. $\square$

**Proposition 12.** For any $p \geq \max\{2, 8\varepsilon/(\varepsilon \sigma + 1)\}$ we have the estimate

$$\frac{d}{dt} \int_{\Sigma} f_\sigma^p \, d\mu_\sigma \leq -p(p - 1) \int_{\Sigma} \frac{f_\sigma^{p-2}}{|H|} |\nabla f_\sigma|^2 \, d\mu_\sigma - p\varepsilon \int_{\Sigma} f_\sigma^{p-1} \frac{|\nabla H|^2}{|H|} \, d\mu_\sigma + p\sigma c \int_{\Sigma} |H|^2 f_\sigma^p \, d\mu_\sigma.$$ 

**Proof.** Differentiating under the integral sign and substituting in the evolution equations for $f_\sigma$ and the measure $d\mu_\sigma$ gives

$$= \int_{\Sigma} (p f_\sigma^{p-1} \frac{\partial f_\sigma}{\partial t} - |H|^2 f_\sigma^p) \, d\mu_\sigma$$

$$\leq \int_{\Sigma} p f_\sigma^{p-1} \frac{\partial f_\sigma}{\partial t} \, d\mu_\sigma$$

$$\leq -p(p - 1) \int_{\Sigma} \frac{f_\sigma^{p-2}}{|H|} |\nabla f_\sigma|^2 \, d\mu_\sigma + 4(1 - \sigma) p \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|} |\nabla f_\sigma| |\nabla H| \, d\mu_\sigma$$

$$- 2p\varepsilon \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 \, d\mu_\sigma + p\sigma c \int_{\Sigma} |H|^2 f_\sigma^p \, d\mu_\sigma.$$ 

We estimate the second integral by

$$4(1 - \sigma) p \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|} |\nabla f_\sigma| \, d\mu_\sigma \leq \frac{2p}{\rho} \int_{\Sigma} f_\sigma^{p-2} |\nabla f_\sigma|^2 \, d\mu_\sigma + 2p c \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 \, d\mu_\sigma,$$

and then substituting this estimate back into (42) gives

$$\frac{d}{dt} \int_{\Sigma} f_\sigma^p \, d\mu_\sigma \leq \left( -p(p - 1) + \frac{2p}{\rho} \right) \int_{\Sigma} f_\sigma^{p-2} |\nabla f_\sigma|^2 \, d\mu_\sigma - 2p\varepsilon |\nabla f_\sigma|^2 \, d\mu_\sigma$$

$$+ p\sigma c \int_{\Sigma} |H|^2 f_\sigma^p \, d\mu_\sigma$$

$$= -p(p - 1) \left( 1 - \frac{2}{\rho(p - 1)} \right) \int_{\Sigma} f_\sigma^{p-2} |\nabla f_\sigma|^2 \, d\mu_\sigma$$

$$- 2p\varepsilon \left( 1 - \frac{pc}{\varepsilon \sigma} \right) \int_{\Sigma} \frac{f_\sigma^{p-1}}{|H|^{2(1-\sigma)}} |\nabla H|^2 \, d\mu_\sigma + p\sigma c \int_{\Sigma} |H|^2 f_\sigma^p \, d\mu_\sigma.$$ 

We now want to choose $\rho$ so that $1 - 2/(\rho(p - 1)) \geq 1/2$ and $p$ so that $1 - pc/\varepsilon \sigma \geq 1/2$. Choosing $\rho = 4/(p - 1)$ and $p \geq \max\{2, 8\varepsilon/(\varepsilon \sigma + 1)\}$ gives the result. $\square$

**Lemma 6.** There exist constants $c_8$ and $c_9$ depending only on $\Sigma_0$ such that if $p \geq c_8$ and $\sigma \leq c_9/\sqrt{p}$, then for all time $t \in [0, T)$ we have the estimate

$$\frac{d}{dt} \int_{\Sigma} f_\sigma^p \, d\mu_\sigma \leq 0.$$
Proof. Combining Propositions 11 and 12 we get
\[
\frac{d}{dt} \int_{\Sigma} f_{\sigma}^p \, d\mu_k \leq -p(p-1) \left( \frac{1 - \frac{\sigma c}{\epsilon \eta}}{2} \right) \int_{\Sigma} f_{\sigma}^{p-2} |\nabla f_{\sigma}|^2 \, d\mu_k \\
- \left( p\epsilon \eta - \frac{p\sigma c(p\eta + 4)}{\epsilon} \right) \int_{\Sigma} \frac{f_{\sigma}^{p-1}}{|H|^{(2(1-\sigma))}} \, |\nabla H|^2 \, d\mu_k.
\]
Recall that we are already assuming that \( p \geq \max \{2, 8\epsilon/(\epsilon \eta + 1)\} \). Now suppose that
\[
\sigma \leq \frac{\epsilon}{4c} \sqrt{\frac{\epsilon \eta}{p}}.
\]
Setting \( \eta = 4c \sigma / \epsilon \) we obtain
\[
\left\{ \begin{array}{l}
\frac{\sigma c}{\epsilon \eta} = \frac{1}{4} < \frac{1}{2} \\
\frac{p\sigma c(p\eta + 4)}{\epsilon} \leq \frac{1}{4} \sqrt{p\epsilon \eta} (\sqrt{p\epsilon \eta} + 4) \leq \frac{p\epsilon \eta}{2} < p\epsilon \eta.
\end{array} \right.
\]
For the second last inequality to hold we must assume \( p \geq 16/\epsilon \eta \). We conclude that
\[
\frac{d}{dt} \int_{\Sigma} f_{\sigma}^p \, d\mu_k \leq 0.
\]
The Lemma holds with \( c_8 = \max \{2, 8\epsilon/(\epsilon \eta + 1), 16/\epsilon \eta\} \) and \( c_9 = \epsilon \sqrt{\epsilon \eta}/(4c) \). \qed

Lemma 6 shows that for \( \sigma \) small enough, high \( L^p \) norms of \( f_{\sigma} \) are bounded. We proceed as in [Hu1] to derive a bound on the supremum of \( f_{\sigma} \) by Stampacchia iteration. The argument follows line for line the argument in [Hu1] and the reader is referred there for the details (see also [HS]).

We have now arrived at the values of the constants appearing in the statement of the Main Theorem, except that we have not yet included the boundary case \( c = 1/(n-1) \) for \( n \geq 4 \). Although the pinching lemma shows that \(|h|^2 \geq c|H|^2\) is preserved along the flow for \( c \leq 4/(3n) \), for \( n \geq 4 \) we have needed to make the stronger assumption that \( c < 1/(n-1) \) in order for Theorem 3 to hold. Let us now show that we can include the case \( c = 1/(n-1) \) in dimensions greater than or equal to four. Suppose that \( \Sigma_0 \) satisfies \(|h|^2 \leq 1/(n-1)|H|^2\), and furthermore suppose that the pinching does not strictly improve, that is, at some future time suppose that \( c = 1/(n-1) \) at some point of \( \Sigma_0 \). From the proof of the pinching lemma we have

\[
\frac{\partial}{\partial t} (|h|^2 - \frac{1}{n-1}|H|^2) \leq \Delta(|h|^2 - \frac{1}{n-1}|H|^2) - c_{10} |\nabla h|^2 - c_{11} |\hat{h}_\perp|^2 - c_{12} |\hat{h}_\perp|^{4},
\]
where importantly \( c_{10} \) and \( c_{12} \) are positive for all \( n \geq 4 \). The strong maximum principle implies \( |h_\perp| = 0 \) and \( |\nabla h| = 0 \). The condition on \( |\hat{h}_\perp| \) implies that \( \Sigma_0 \) must lie in a \((n+1)\)-subspace of \( \mathbb{R}^{n+k} \), and then this combined with the fact that \( |\nabla h| = 0 \) means that \( \Sigma_0 \) must be a product \( S^p \times \mathbb{R}^{n-p} \rightarrow \mathbb{R}^{n+k} \) (see [L], Theorem 4). These submanifolds are of course not compact. We have now shown that if the (compact) submanifold \( \Sigma_0 \) satisfies \(|h| \leq 1/(n-1)|H|^2\), then it immediately becomes strictly \( 1/(n-1) \)-pinched, and so the results of this section still apply.

6. A GRADIENT ESTIMATE FOR THE MEAN CURVATURE

In this section we derive a gradient estimate for the mean curvature. This will be used in the following section to compare the mean curvature of the submanifold at different points.
Theorem 4. For each $\eta > 0$ there exists a constant $C(\eta)$ depending only on $\Sigma_0$ such that
\[ |\nabla H|^2 \leq \eta |H|^{4-\frac{\delta}{2}} + C(\eta) \quad \text{on } \Sigma \times [0, T). \]

We begin by deriving a number of evolution equations.

Proposition 13. There exists a constant $c_1$ depending only on $\Sigma_0$ such that
\[ \frac{\partial}{\partial t} |\nabla h|^2 \leq \Delta |\nabla h|^2 - 2 |\nabla^2 H|^2 + c_1 |H|^2 |\nabla h|^2. \]

Proof. Differentiating the length of the gradient squared in time gives
\[
\frac{\partial}{\partial t} |\nabla H|^2 = \frac{\partial}{\partial t} \langle \nabla H, \nabla H \rangle \\
= 2 \langle \nabla^2 H, \nabla H \rangle \\
= 2 \tilde{g}(\tilde{V}_k \nabla H + \tilde{R}(\tilde{\partial}_k, \tilde{\partial}_i) H, \nabla_k H) \\
(43)
= 2 \tilde{g}(\tilde{V}_k (\tilde{\Delta} H + H \cdot h_i, h_j), \nabla_k H) + 2 \tilde{g} (\tilde{H}(\tilde{\partial}_k, \tilde{\partial}_i) H, \nabla_k H).
\]

To manipulate this into the desired form we use the following formulae for the Laplacian of the gradient squared and the commutator of the Laplacian and gradient of the mean curvature:
\[ \Delta |\nabla H|^2 = 2 \tilde{g}(\Delta \nabla_k H, \nabla_k H) + 2 |\nabla^2 H|^2; \]
\[ \Delta \nabla_k H = \nabla_k \Delta H + \tilde{V}_p (\tilde{H}(\tilde{\partial}_k, \tilde{\partial}_p) H) + \tilde{R} (\tilde{\partial}_k, \tilde{\partial}_p) \tilde{V}_p H + \tilde{R}c_{pk} \tilde{V}_p H. \]

Substituting these into (43) and observing that the Gauss equation (15a) and the Ricci equation (17) are of the form $R = h * h$ and $\tilde{R} = h * h$, and that the timelike Ricci equation (18) is of the form $\tilde{R}(., \partial_i) = h * \nabla h * \nabla h$, we find
\[ \frac{\partial}{\partial t} |\nabla H|^2 = \Delta |\nabla H|^2 - 2 |\nabla^2 H|^2 + h * h * \nabla h * \nabla h. \]

The proposition then follows by Cauchy-Schwarz and the inequality $|h|^2 \leq c |H|^2$. 

Proposition 14. For any $N > 0$ we have the evolution equation
\[
\frac{\partial}{\partial t} \left( (N_1 + N_2 |H|^2) |\dot{h}|^2 \right) \leq \Delta \left( (N_1 + N_2 |H|^2) |\dot{h}|^2 \right) - \frac{4(n-1)}{3n} (N_2 - 1) |H|^2 |\nabla h|^2 \\
- \frac{4(n-1)}{3n} (N_1 - c_2 N_2) |\nabla h|^2 + c_1 (N_1, N_2) \langle H^{6-\delta} + c_2 (N_1, N_2), \nabla h \rangle,
\]
where all constants depend only on $\Sigma_0$, and possibly $N_1$ and $N_2$ as well.

Proof. From the evolution equations for $|h|^2$ and $|H|^2$ we derive
\[
\frac{\partial}{\partial t} \left( (N_1 + N_2 |H|^2) |\dot{h}|^2 \right) = \Delta \left( (N_1 + N_2 |H|^2) |\dot{h}|^2 \right) - 2N_2 \langle \nabla_i |H|^2, \nabla_i |\dot{h}|^2 \rangle - 2N_2 |\dot{h}|^2 |H|^2 + 2N_2 R_2 |h|^2 \\
- 2(\Delta N_1 + N_2 |H|^2) (|\nabla h|^2 - 1/n |\nabla H|^2) + 2(N_1 + N_2 |H|^2) (R_1 - 1/n R_2).
\]

We now estimate the various terms. Using Theorem 3 and Young’s inequality we estimate
\[
2N_2 \langle \nabla_i |H|^2, \nabla_i |\dot{h}|^2 \rangle \leq 8N_2 \sqrt{c_0} |H|^2 |\nabla h|^2 \\
\leq \frac{4(n-1)}{3n} |H|^2 |\nabla h|^2 + \frac{4(n-1)}{3n} c_2 N_2 |\nabla h|^2,
\]
and using Theorem 3, Young’s inequality and the inequalities $R_2 \leq c_3|H|^4$ and $R_1 - 1/nR_2 \leq c_4|h|^2|H|^2$, we estimate

$$2N_2R_2|h|^2 + 2(N_1 + N_2|H|^2)(R_1 - 1/nR_2) \leq c_1(N_1,N_2)|H|^{6-\delta} + c_2(N_1,N_2).$$

Proposition 15. We have the evolution equation

\begin{equation}
\frac{\partial}{\partial t}|H|^{4-\frac{\delta}{2}} = \Delta |H|^{4-\frac{\delta}{2}} - 4(2 - \delta/4)(1 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 \\
- 2(2 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 + 2(2 - \delta/2)|H|^{2-\frac{\delta}{2}}R_2.
\end{equation}

Proof. The proof follows directly from the evolution equation for $|H|^2$. \hfill \Box

We are now ready to prove the gradient estimate.

Proof of Theorem 4. Consider the quantity $f = |\nabla H|^2 + (N_1 + N_2|H|^2)|h|^2$. From the evolution equations derived above we see $f$ satisfies the evolution equation

$$\frac{\partial}{\partial t} f \leq \Delta f + c_1|H|^2|\nabla h|^2 - \frac{4(n-1)}{3n}(N_2 - 1)|H|^2|\nabla h|^2 - \frac{4(n-1)}{3n}(N_1 - c_2N_2)|\nabla h|^2 + c_1(N_1,N_2)|H|^{6-\delta} + c_2(N_1,N_2).$$

We choose $N_2$ large enough to absorb the positive $|H|^2|\nabla h|^2$ term, leaving

$$\frac{\partial}{\partial t} f \leq \Delta f - csN_2|H|^2|\nabla h|^2 - \frac{4(n-1)}{3n}(N_1 - c_2N_2)|\nabla h|^2 + c_1(N_1,N_2)|H|^{6-\delta} + c_2(N_1,N_2).$$

Now consider $f - \eta |H|^{4-\frac{\delta}{2}}$, where $\eta > 0$. From the above evolution equations we get

$$\frac{\partial}{\partial t}(f - \eta |H|^{4-\frac{\delta}{2}}) \leq \Delta(f - \eta |H|^{4-\frac{\delta}{2}}) + 4\eta(2 - \delta/4)(1 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 \\
+ 2\eta(2 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 - 2\eta(2 - \delta/2)|H|^{2-\frac{\delta}{2}}R_2 \\
- csN_2|H|^2|\nabla h|^2 - \frac{4(n-1)}{3n}(N_1 - c_2N_2)|\nabla h|^2 + c_1(N_1,N_2)|H|^{6-\delta} + c_2(N_1,N_2).$$

Working first with the $|H|^2|\nabla h|^2$ terms, using $|\nabla |H||^2 \leq |\nabla H|^2$, $|\nabla H|^2 \leq n|\nabla h|^2$ and Young’s inequality we estimate

$$4\eta(2 - \delta/4)(1 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 + 2\eta(2 - \delta/4)|H|^{2-\frac{\delta}{2}}|\nabla |H||^2 \leq |H|^2|\nabla h|^2 + c_1(\eta)|\nabla h|^2,$$

and then by choosing $N_2$ larger again we can absorb this positive gradient term too. Working now with the pure reaction terms, using $R_2 \geq 1/n|H|^4$ and Young’s inequality again we estimate

$$-2(2 - \delta/4)|H|^{2-\frac{\delta}{2}}R_2 + c_1(N_1,N_2)|H|^{6-\delta} \leq c_6.$$

Lastly, we choose $N_1$ large enough to absorb the positive $|\nabla h|^2$ terms, and putting everything together we obtain

$$\frac{\partial}{\partial t}(f - \eta |H|^{4-\frac{\delta}{2}}) \leq \Delta(f - \eta |H|^{4-\frac{\delta}{2}}) + c_2(\eta).$$
The maximum principle implies that \( f - \eta |H|^{n-\frac{2}{2}} \leq c_2(\eta)T + c_3(\eta) \) and then from the definition of \( f \) we conclude that
\[
|\nabla H|^2 \leq \eta |H|^{n-\frac{2}{2}} + c_2(\eta)T + c_3(\eta).
\]

\[\square\]

7. Contraction to a Point and Convergence

In Section 4 we established that MCF has a unique solution on a finite maximal time interval \( 0 \leq t < T \) determined by the blowup of the second fundamental form. With the results of the previous two sections in place, we can now show that the diameter of the submanifold approaches zero as \( t \to T \), or put another away, the submanifold is shrinking to a point. This combined with Theorem 2 then completes the first part of the Main Theorem.

**Theorem 5.** Suppose that mean curvature flow of a given submanifold \( \Sigma_0 \) exists on a maximal time interval \( 0 \leq t < T \). Then \( \text{diam}\Sigma \to 0 \) as \( t \to T \).

The proof is of course motivated by Hamilton’s idea in [Ha1] to use Myer’s Theorem, however here our pinching condition gives a strictly positive lower bound on the sectional curvature of \( \Sigma \), and we can use Bonnet’s Theorem instead. The reader is referred to [P, page 170] for a proof of Bonnet’s Theorem:

**Theorem 6** (Bonnet, Hopf-Rinow, Myers). Let \( M \) be a complete Riemannian manifold and suppose that \( x \in M \) such that the sectional curvature satisfies \( K \geq K_{\min} > 0 \) along all geodesics of length \( \pi / \sqrt{K_{\min}} \) from \( x \). Then \( M \) is compact and \( \text{diam}M \leq \pi / \sqrt{K_{\min}} \).

We will also need the following result due to Bang-Yen Chen:

**Proposition 16.** For \( n \geq 2 \), if \( \Sigma^n \) is a submanifold of \( \mathbb{R}^{n+k} \), then at each point \( p \in \Sigma^n \) the smallest sectional curvature \( K_{\min} \) satisfies
\[
K_{\min}(p) \geq \frac{1}{2} \left( \frac{1}{n-1} |H(p)|^2 - |h(p)|^2 \right).
\]

**Proof.** The proof is a consequence of the Gauss equations and a custom-made inequality and can be found in [C, Lemma 3.2]

Combining this with our pinching assumption we see
\[
K_{\min}(p) \geq \frac{1}{2} \left( \frac{1}{n-1} - \varepsilon \right) |H(p)|^2 := \varepsilon^2 |H(p)|^2 > 0.
\]

**Lemma 7.** The ratio \( |H|_{\max}/|H|_{\min} \to 1 \) as \( t \to T \).

**Proof.** From Theorem 4 we know that for any \( \eta > 0 \) there exists a constant \( C(\eta) \) such that \( |\nabla H| \leq \eta |H|^2 + C(\eta) \) on \( 0 \leq t < T \). Since \( |H|_{\max} \to \infty \) as \( t \to T \), there exists a \( \tau(\eta) \) such that \( C(\eta/2) \leq 1/2n |H|_{\max}^2 \) for all \( \tau \leq t < T \), and so \( |\nabla H| \leq \eta |H|_{\max}^2 \) for all \( t \geq \tau \). For any \( \sigma \in (0, 1) \) we choose \( \eta = \frac{\sigma(1-\varepsilon)k}{\pi} \). Let \( t \in [\tau(\eta), T) \), and let \( x \) be a point with \( |H(x)| = |H|_{\max} \). Then along any geodesic of length \( \frac{\pi}{\varepsilon \sigma |H|_{\max}^2} \) from \( x \), we have \( |H| \geq |H|_{\max} - \frac{\pi}{\varepsilon \sigma |H|_{\max}^2} \eta |H|_{\max}^2 = \sigma |H|_{\max} \), and consequently the sectional curvatures satisfy \( K \geq \varepsilon^2 \sigma^2 |H|_{\max} \). The Bonnet Theorem applies to prove that \( \text{diam}M \leq \frac{\pi}{\varepsilon \sigma |H|_{\max}} \), so that \( |H|_{\min} \geq \sigma |H|_{\max} \) for \( t \in [\tau(\eta), T) \).

\[\square\]
Theorem 5 is now also proved in the last line of the proof above, so the first part of Theorem 1 is complete.

We now have all the necessary results in place to proceed as in sections 9 and 10 of [Hu1] (see also Section 17 of [Ha1]) to obtain smooth convergence of the rescaled maps to a sphere. The reader is referred to these sources for the details.

REFERENCES


[BW] Christoph Böhm and Burkhard Wilking, Manifolds with positive curvature operators are space forms, Ann. of Math. (2) 167 (2008), no. 3, 1079–1097. MR2415394


MEAN CURVATURE FLOW OF PINCHED SUBMANIFOLDS TO SPHERES


[LL] Jiayu Li and Ye Li, Mean curvature flow of graphs in $\Sigma_1 \times \Sigma_2$, J. Partial Differential Equations 16 (2003), no. 3, 255–265. MR1995747 (2004e:53102)


CENTRE FOR MATHEMATICS AND ITS APPLICATIONS, AUSTRALIAN NATIONAL UNIVERSITY, ACT 0200 AUSTRALIA

E-mail address: Ben.Andrews@maths.anu.edu.au

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, AUSTRALIAN NATIONAL UNIVERSITY, ACT 0200 AUSTRALIA

E-mail address: Charles.Baker@maths.anu.edu.au