QUADRATICALLY PINCHED SUBMANIFOLDS OF THE SPHERE VIA MEAN CURVATURE FLOW WITH SURGERY

MAT LANGFORD, STEPHEN LYNCH, AND HUY THE NGUYEN

Abstract. We study mean curvature flow of n-dimensional submanifolds of $S_K^{n+\ell}$, the round $(n+\ell)$-sphere of sectional curvature $K>0$, under the quadratic curvature pinching condition $|\mathbf{A}|^2 < \frac{1}{n-2}|\mathbf{H}|^2 + 4K$ when $n \geq 8$, $|\mathbf{A}|^2 < \frac{4}{3n}|\mathbf{H}|^2 + \frac{n}{2}K$ when $n=7$, and $|\mathbf{A}|^2 < \frac{3(n+1)}{2n(n+2)}|\mathbf{H}|^2 + \frac{2n(n-1)}{3(n+1)}K$ when $n=5$ or 6. This condition is related to a theorem of Li and Li [Arch. Math., 58:582–594, 1992] which states that the only n-dimensional minimal submanifolds of $S_K^{n+\ell}$ satisfying $|\mathbf{A}|^2 < \frac{2n}{3}K$ are the totally geodesic n-spheres. We prove the existence of a suitable mean curvature flow with surgeries starting from initial data satisfying the pinching condition. As a result, we conclude that any smoothly, properly immersed submanifold of S_K^{n+1} satisfying the pinching condition is diffeomorphic either to the sphere S^n or to the connected sum of a finite number of handles $S^1 \times S^{n-1}$. The results are sharp when $n \geq 8$ due to hypersurface counterexamples.

Contents

1. Introduction

A famous theorem of Simons [25] states that any minimal hypersurface of the round sphere S^{n+1} with squared second fundamental form $|\mathbf{A}|^2$ less than n is necessarily a hyper-equator. Simons' methods have been generalized in various directions [1,2,8,10,11,23,24], in particular to higher codimension minimal immersions. In this setting, the algebraic structure of Simons' equation becomes much more complicated, primarily due to the possibility of a non-trivially curved normal bundle. Nonetheless, Li–Li [2], building on work of Chern–do Carmo–Kobayashi [11], were able to show that a minimal immersion in S^{n+1} satisfying $|\mathbf{A}|^2 < \frac{2n}{3}$ is necessarily totally geodesic.

2000 *Mathematics Subject Classification.* Primary 53C44.

Such results can be improved upon using geometric flows [3, 6, 14, 16]. Building on work of Andrews–Baker [4] and Baker [6], Baker–Nguyen [7] showed that, under mean curvature flow, n-dimensional submanifolds of the sphere $S_K^{n+\ell}$ of sectional curvature[1] K satisfying the quadratic curvature pinching condition

$$|\mathbf{A}|^2 < \begin{cases} \frac{1}{n-1}|\mathbf{H}|^2 + 2K & \text{if } n \geq 4\,, \\ \frac{4}{3n}|\mathbf{H}|^2 + \frac{n}{2}K & \text{if } n = 2, 3\,, \end{cases} \tag{1.1}$$

where $\mathbf{H}$ is the mean curvature vector, converge (preserving the inequality) either to a "round" point in finite time or to a totally geodesic subsphere in infinite time. In case $n \geq 4$, this behaviour is sharp in the sense that, for each $\varepsilon > 0$, the Clifford embedding

$$\mathcal{M}_\varepsilon^{1,n-1} \doteqdot \left\{(x,y) \in \mathbb{R}^2 \times \mathbb{R}^n : |x|^2 = \frac{1}{1+\varepsilon^2} \text{ and } |y|^2 = \frac{\varepsilon^2}{1+\varepsilon^2}\right\}$$

of $S^1 \times S^{n-1}$ into S^{n+1} satisfies $|\mathbf{A}|^2 - \frac{1}{n-1}|\mathbf{H}|^2 - 2 = \frac{n-2}{n-1}\varepsilon^2$.

Baker–Nguyen [7] (see also [8]) refined this result in the context of surfaces in S^4 by including the curvature of the normal bundle in the pinching condition. Their pinching condition is less restrictive than (1.1) (with $n = \ell = 2$), and is almost sharp in that the Veronese surface, a minimal embedding of the projective plane into S^4, lies close to its boundary (they conjecture that the Veronese surface represents the sharp condition).

We will develop these results further by allowing a weaker curvature pinching condition. Namely, we study, for $n \geq 5$ and[2] $\ell \geq 2$, n-dimensional submanifolds of $S_K^{n+\ell}$ satisfying the quadratic pinching condition

$$|\mathbf{A}|^2 < \begin{cases} \frac{1}{n-2}|\mathbf{H}|^2 + 4K & \text{if } n \geq 8\,, \\ \frac{4}{3n}|\mathbf{H}|^2 + \frac{n}{2}K & \text{if } n = 7 \\ \frac{3(n+1)}{2n(n+2)}|\mathbf{H}|^2 + \frac{2n(n-1)}{3(n+1)}K & \text{if } n = 5, 6\,. \end{cases} \tag{1.2}$$

By constructing an appropriate mean curvature flow-with-surgeries, we are able to prove the following topological classification of submanifolds satisfying (1.2).

Theorem 1.1. *Every properly isometrically immersed n-dimensional submanifold $X : M \to S_K^{n+\ell}$ of $S_K^{n+\ell}$, $n \geq 5$, satisfying (1.2) is diffeomorphic either to S^n or to a connected sum of finitely many copies of $S^1 \times S^{n-1}$.*

This theorem is sharp when $n \geq 8$ in the sense that, for each $\varepsilon > 0$, the Clifford embedding

$$\mathcal{M}_\varepsilon^{1,n-2} \doteqdot \left\{(x,y) \in \mathbb{R}^3 \times \mathbb{R}^{n-1} : |x|^2 = \frac{1}{1+\varepsilon^2} \text{ and } |y|^2 = \frac{\varepsilon^2}{1+\varepsilon^2}\right\}$$

of $S^2 \times S^{n-2}$ into S^{n+1} satisfies $|\mathbf{A}|^2 - \frac{1}{n-2}|\mathbf{H}|^2 - 4 = 2\frac{n-4}{n-2}\varepsilon^2$.

[1] We find it convenient to work without normalizing the curvature K, as it serves as a natural scale parameter.

[2] Our arguments also apply in the codimension one case, $\ell = 1$; however in this case the results obtained are weaker than known results [16].

The first step in proving Theorem 1.1 is to prove that the pinching condition is preserved. This is achieved by a fairly straightforward application of the maximum principle (cf. [7, Lemma 3.1]). We then show that the conormal component of the second fundamental form is small compared to the mean curvature when the latter is large. Such a “codimension estimate” was first established by Naff [21] for mean curvature flow in Euclidean spaces. His argument relies on the maximum principle but requires very careful accounting of first order terms. Our proof is inspired by Naff’s but requires some effort to overcome the bad ambient curvature terms as well as the possible presence of points where the mean curvature is zero (at which the conormal subspace is not even defined). The codimension estimate is also a crucial ingredient in our “cylindrical estimate”, in that it is needed to establish a suitable “Poincaré-type inequality” (cf. [18, Proposition 3.2]), which is the key ingredient in a Huisken–Stampacchia iteration argument. The cylindrical estimate allows us to obtain pointwise estimates for the gradient and Hessian of the second fundamental form using the maximum principle. These estimates imply that the flow becomes either uniformly convex or quantitatively cylindrical with respect to a codimension one subspace in regions of high curvature. Once they are in place, we are able to apply the surgery apparatus developed by Huisken–Sinestrari [15], as extended by Nguyen [22] to the high codimension setting.

Acknowledgements. M. Langford was supported by the Australian Research Council grant DE200101834. H. T. Nguyen was supported by the EPSRC grant EP/S012907/1.

2. Preliminaries

We begin by recalling the fundamental machinery regarding submanifolds of the sphere, and their evolution by mean curvature.

2.1. **Spaceforms.** Let $N_K = (N, g)$ be a complete Riemannian manifold of constant sectional curvature $K \in \mathbb{R}$, equipped with its Levi–Civita connection D. Recall that the curvature tensor Rm of N is given by

$$\mathrm{Rm}(u,v)w = K(g(u,w)v - g(v,w)u),$$

where we use the convention

$$\mathrm{Rm}^{\nabla}(U,V)W \doteqdot \nabla_{[U,V]}W - \nabla_U(\nabla_V W) + \nabla_V(\nabla_U W)$$

for the curvature operator Rm^{∇} of a connection ∇.

2.2. **Submanifolds.** We follow [4]. Consider an immersed submanifold $X : M^n \to N_K^{n+\ell}$ of a spaceform $N_K^{n+\ell}$. The *normal bundle* NM of X is determined by the orthogonal decomposition

$$X^*TN = dX(TM) \oplus^{\perp} NM,$$

where X^*TN is the pullback of TN and $dX : TM \to TN$ is the derivative of X. The pullback X^*g of g induces positive definite bilinear forms $g^{\top}$ on TM (the *first fundamental form*) and $g^{\perp}$ on NM, respectively. The pullback connection ${}^X D$ on X^*TN induces connections $\nabla^\top$ on TM and $\nabla^\perp$ on NM via

$$\nabla^\top_u V \doteqdot ({}^X\!D_u[dX(V)])^\top \;\text{ and }\; \nabla^\perp_u N \doteqdot ({}^X\!D_u N)^\perp,$$

where $\cdot^\top : X^*TN \to TM$ and $\cdot^\perp : X^*TN \to NM$ denote the tangential and normal (orthogonal) projections, respectively. The connection $\nabla^\top$ induced on TM is the Levi-Civita connection of $g^\top$ and $\nabla^\perp$ is compatible with $g^\perp$.

The *second fundamental form* $\mathbf{A} \in \Gamma(T^*M \otimes T^*M \otimes NM)$ and the *Weingarten tensor* $\mathbf{L} \in \Gamma(T^*M \otimes NM \otimes TM)$ are defined by the *Gauss–Weingarten equations*

$$ {}^X\!D_u[dX(V)] = dX(\nabla^\top_u V) + \mathbf{A}(u,v) \tag{2.1}$$

and

$$ {}^X\!D_v N = \nabla^\perp_v N - dX(\mathbf{L}(v,\nu))\,, \tag{2.2}$$

respectively, so that

$$g^\perp(\mathbf{A}(u,v),\nu) = g^\top(\mathbf{L}(u,\nu),v)\,.$$

These equations hold at any point $p \in M$ and for any tangent vectors $u, v \in T_pM$, normal vectors $\nu \in N_pM$, and extensions $V \in \Gamma(TM)$ of v and $N \in \Gamma(NM)$ of ν.

We will continue to use these conventions below; that is, given $p \in M$, the lower case latin letters u, v, w, z will denote tangent vectors at p, the lower case greek letters μ, ν will denote normal vectors at p, and corresponding upper case letters will denote corresponding extension fields, all of which may be arbitrarily chosen.

Observe that

$$(\nabla_u dX)(v) = \mathbf{A}(u,v)\,, \tag{2.3}$$

where ∇ denotes the connection induced on the bundle $T^*M \otimes X^*TN$ by $\nabla^\top$ and ${}^X\!D$. The covariant derivatives of the tangential and normal projections $\cdot^\top$ and $\cdot^\perp$ are given by

$$(\nabla_u(\cdot^\top))(\phi) = \mathbf{L}(u,\phi^\perp) \tag{2.4}$$

and

$$(\nabla_u(\cdot^\perp))(\phi) = -\mathbf{A}(u,\phi^\top), \tag{2.5}$$

for any $\phi \in X^*TN$, where ∇ denotes the connection induced on the relevant bundle by the connections ${}^X\!D$, $\nabla^\top$ and $\nabla^\perp$.

Differentiating (2.1) and decomposing the result into tangential and normal components yields the *Gauss equation*,

$$K(g^\top(u,w)v - g^\top(v,w)u) = \mathrm{Rm}^\top(u,v)w + \mathbf{L}(v,\mathbf{A}(u,w)) - \mathbf{L}(u,\mathbf{A}(v,w)), \tag{2.6}$$

or, equivalently,

$$\begin{aligned} &K(g^\top(u,w)g^\top(v,z) - g^\top(u,z)g^\top(v,w)) \\ &\quad = \mathrm{Rm}^\top(u,v,w,z) + g^\perp\big(\mathbf{A}(u,z),\mathbf{A}(v,w)\big) - g^\perp\big(\mathbf{A}(u,w),\mathbf{A}(v,z)\big), \end{aligned} \tag{2.7}$$

and the *Codazzi–Mainardi equation*,

$$0 = \nabla_v \mathbf{A}(u,w) - \nabla_u \mathbf{A}(v,w)\,, \tag{2.8}$$

where, for any $\phi \in \Gamma(TN)$,

$$ {}^X\mathrm{Rm}(u,v)X^*\phi = \mathrm{Rm}(dX(u), dX(v))\phi $$

defines the curvature tensor ${}^X\mathrm{Rm}$ of ${}^XD$, $\mathrm{Rm}^\top$ denotes the curvature tensor of $\nabla^\top$, and the covariant derivative of $\mathbf{A}$ is defined in the canonical way:

$$ \nabla_u\mathbf{A}(v,w) \doteqdot \nabla^\perp_u(\mathbf{A}(V,W)) - \mathbf{A}(\nabla^\top_u V, w) - \mathbf{A}(v, \nabla^\top_u W). $$

The *mean curvature*, $\mathbf{H}$, is the trace of the second fundamental form,

$$ \mathbf{H} \doteqdot \mathrm{tr}_{g^\top}\mathbf{A}. $$

The trace-free part is denoted by

$$ \mathring{\mathbf{A}} \doteqdot \mathbf{A} - \frac{1}{n}\mathbf{H}\otimes g^\top. $$

Decomposing $\nabla\mathbf{A}$ into its trace and trace-free parts yields the *Kato inequality*,

$$ |\nabla\mathbf{A}|^2 \geq \frac{3}{n+2}|\nabla^\perp\mathbf{H}|^2. \tag{2.9} $$

Differentiating (2.2) and decomposing the result into normal and tangential components yields the *Ricci equation*,

$$ 0 = \mathrm{Rm}^\perp(u,v)\nu + \mathbf{A}(v, \mathbf{L}(u,\nu)) - \mathbf{A}(u, \mathbf{L}(v,\nu)), \tag{2.10} $$

and the Codazzi–Mainardi equation, respectively, where $\mathrm{Rm}^\perp$ denotes the curvature tensor of $\nabla^\perp$. Contracting the Ricci equation yields the (*contracted*) *Ricci equation*,

$$ 0 = \mathrm{Rm}^\perp(u,v,\nu,\mu) + g^\top(\mathbf{L}(u,\mu), \mathbf{L}(v,\nu)) - g^\top(\mathbf{L}(v,\mu), \mathbf{L}(u,\nu)). $$

Given a pair of tensors $S, T \in \Gamma(NM\otimes T^*M\otimes T^*M)$, we define a new tensor $S\wedge T \in \Gamma(NM\otimes NM\otimes T^*M\otimes T^*M)$ by

$$ (S\wedge T)(u,v) \doteqdot \mathrm{tr}_{g^\top}\big(S(u,\cdot)\otimes T(v,\cdot) - S(v,\cdot)\otimes T(u,\cdot)\big). \tag{2.11} $$

With this notation, the Ricci equation becomes (after identifying $\mathrm{Rm}^\perp$ with a section of $T^*M\otimes T^*M\otimes NM\otimes NM$)

$$ 0 = \mathrm{Rm}^\perp + \mathbf{A}\wedge\mathbf{A}. \tag{2.12} $$

Since

$$ \mathbf{A}\wedge\mathbf{A} = \mathring{\mathbf{A}}\wedge\mathring{\mathbf{A}}, $$

we may also write the Ricci equation as

$$ 0 = \mathrm{Rm}^\perp + \mathring{\mathbf{A}}\wedge\mathring{\mathbf{A}}. \tag{2.13} $$

Combining all of the above identities yields *Simons' equation*,

$$\begin{aligned}\nabla_u\nabla_v\mathbf{A}(w,z)-\nabla_w\nabla_z\mathbf{A}(u,v)=\;&\mathbf{A}\big(u,\mathbf{L}\big(w,\mathbf{A}(v,z)\big)\big)-\mathbf{A}\big(w,\mathbf{L}\big(u,\mathbf{A}(v,z)\big)\big)\\&+\mathbf{A}\big(\mathbf{L}\big(w,\mathbf{A}(u,v)\big),z\big)-\mathbf{A}\big(\mathbf{L}\big(u,\mathbf{A}(w,v)\big),z\big)\\&+\mathbf{A}\big(v,\mathbf{L}\big(w,\mathbf{A}(u,z)\big)\big)-\mathbf{A}\big(v,\mathbf{L}\big(u,\mathbf{A}(w,z)\big)\big)\\&+K\Big(g^\top(u,v)\mathbf{A}\big(z,w\big)-g^\top(w,z)\mathbf{A}\big(v,u\big)\\&\quad+g^\top(u,z)\mathbf{A}\big(v,w\big)-g^\top(w,v)\mathbf{A}\big(z,u\big)\Big).\end{aligned}\tag{2.14}$$

Taking the trace of (2.14) yields

$$\begin{aligned}\nabla^\perp_u\nabla^\perp_v\mathbf{H}-\Delta\mathbf{A}(u,v)=\;&\mathrm{tr}_{g^\top}\Big(\mathbf{A}\big(u,\mathbf{L}\big(\cdot,\mathbf{A}(v,\cdot)\big)\big)+\mathbf{A}\big(v,\mathbf{L}\big(\cdot,\mathbf{A}(u,\cdot)\big)\big)\\&-2\mathbf{A}\big(\cdot,\mathbf{L}\big(u,\mathbf{A}(v,\cdot)\big)\big)+\mathbf{A}\big(\cdot,\mathbf{L}\big(\cdot,\mathbf{A}(u,v)\big)\big)\Big)\\&-\mathbf{A}\big(v,\mathbf{L}\big(u,\mathbf{H}\big)\big)+K\big(g^\top(u,v)\mathbf{H}-n\mathbf{A}\big(u,v\big)\big).\end{aligned}\tag{2.15}$$

2.3. **Mean convex submanifolds.** If $H \doteqdot |\mathbf{H}| > 0$ on M, then the mean curvature vector defines a canonical normal vector field $\mathbf{N}\in\Gamma(NM)$, called the *principal normal*, via

$$\mathbf{N}\doteqdot\frac{\mathbf{H}}{|\mathbf{H}|}.$$

So the normal bundle splits globally as an orthogonal sum

$$NM=\mathbf{N}M\oplus^\perp\widehat{NM},$$

where the *principal normal bundle* $\mathbf{N}M\doteqdot\mathbb{R}\mathbf{N}$ is the span of $\mathbf{N}$ in NM, and the *conormal bundle* $\widehat{NM}$ is its orthogonal compliment in NM.

The form $g^\perp$ induces positive definite forms $g^{\mathbf{N}}$ and $\hat{g}$ on $\mathbf{N}M$ and $\widehat{NM}$, respectively, and the connection $\nabla^\perp$ induces connections $\nabla^{\mathbf{N}}$ and $\hat\nabla$ on $\mathbf{N}M$ and $\widehat{NM}$, respectively, in the usual way.

Observe that $\nabla^\perp_u\mathbf{N}\in\widehat{NM}_p$ for any $u\in T_pM$, since $g^\perp(\mathbf{N},\mathbf{N})\equiv 1$. Define the *torsion* tensors $\mathrm{T}\in\Gamma\big(T^*M\otimes\widehat{NM}^*\big)$ and $\hat{\mathbf{T}}\in\Gamma(T^*M\otimes\widehat{NM})$ by

$$\mathrm{T}(u,\mu)\mathbf{N}\doteqdot(\nabla^\perp_u M)^{\mathbf{N}}\;\text{ and }\;\hat{\mathbf{T}}(u)\doteqdot(\nabla^\perp_u\mathbf{N})^{\hat\perp}=\nabla^\perp_u\mathbf{N},$$

where $\cdot^{\mathbf{N}}$ and $\cdot^{\hat\perp}$ denote the projections onto $\mathbf{N}M$ and $\widehat{NM}$, respectively. Observe that

$$\mathrm{T}(u,\mu)+\hat{g}(\hat{\mathbf{T}}(u),\mu)=0.$$

We shall denote the principal normal component of $\mathbf{A}$ by h and the conormal projection by $\hat{\mathbf{A}}$; that is,

$$h\doteqdot\langle\mathbf{A}(u,v),\mathbf{N}\rangle\quad\text{and}\quad\hat{\mathbf{A}}(u,v)\doteqdot\mathbf{A}(u,v)-\langle\mathbf{A}(u,v),\mathbf{N}\rangle\,\mathbf{N},$$

so that

$$\mathbf{A}=\mathbf{N}\otimes h+\hat{\mathbf{A}}.\tag{2.16}$$

With the notation of (2.11), the Ricci equation (2.10) then becomes

$$\begin{aligned} 0 =&\ \mathrm{Rm}^\perp + (\mathbf{N}\otimes \mathring{h} + \hat{\mathbf{A}}) \wedge (\mathbf{N}\otimes \mathring{h} + \hat{\mathbf{A}}) \\ =&\ \mathrm{Rm}^\perp + (\mathbf{N}\otimes \mathring{h}) \wedge \hat{\mathbf{A}} + \hat{\mathbf{A}} \wedge (\mathbf{N}\otimes \mathring{h}) + \hat{\mathbf{A}}\wedge\hat{\mathbf{A}}\,, \end{aligned} \tag{2.17}$$

where $\mathring{h} \doteqdot h - \frac{1}{n}Hg^\top$ is the trace-free part of h. In particular,

$$|\mathrm{Rm}^\perp|^2 = 2|(\mathbf{N}\otimes\mathring{h})\wedge\hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}}\wedge\hat{\mathbf{A}}|^2. \tag{2.18}$$

Differentiating (2.16) yields

$$\begin{aligned} \nabla_u\mathbf{A} &= \nabla_u^\perp\mathbf{N}\otimes h + \mathbf{N}\otimes\nabla_u^\top h + \nabla_u\hat{\mathbf{A}} \\ &= (\nabla_u\hat{\mathbf{A}})^{\mathbf{N}} + \mathbf{N}\otimes\nabla^\top_u h + (\nabla_u\hat{\mathbf{A}})^{\hat\perp} + \hat{\mathbf{T}}(u)\otimes h\,. \end{aligned}$$

Since

$$\begin{aligned} (\nabla_u\hat{\mathbf{A}})^{\mathbf{N}}(v,w) &\doteqdot g^\perp(\nabla_u\hat{\mathbf{A}}(v,w),\mathbf{N})\mathbf{N} \\ &= \mathrm{T}(u,\hat{\mathbf{A}}(v,w))\mathbf{N}, \end{aligned}$$

and

$$\begin{aligned} (\nabla_u\hat{\mathbf{A}})^{\hat\perp}(v,w) &\doteqdot \nabla_u\hat{\mathbf{A}}(v,w) - g^\perp(\nabla_u\hat{\mathbf{A}}(v,w),\mathbf{N})\mathbf{N} \\ &= \hat\nabla_u\hat{\mathbf{A}}(v,w)\,, \end{aligned}$$

we obtain

$$\nabla_u\mathbf{A} = \mathbf{N}\otimes\big(\mathrm{T}(u,\hat{\mathbf{A}}) + \nabla_u^\top h\big) + \hat\nabla_u\hat{\mathbf{A}} + \hat{\mathbf{T}}(u)\otimes h. \tag{2.19}$$

By the Codazzi–Mainardi equation, the tensors

$$\mathrm{T}(\cdot,\hat{\mathbf{A}}) + \nabla^\top h \ \text{ and } \ \hat\nabla\hat{\mathbf{A}} + \hat{\mathbf{T}}\otimes h$$

are totally symmetric in their TM components. Decomposing them into their trace and trace-free components yields the Kato inequalities [21]

$$|\mathrm{T}(\cdot,\hat{\mathbf{A}}) + \nabla^\top\mathring{h}|^2 \geq \frac{2(n-1)}{n(n+2)}|\nabla H|^2 \tag{2.20}$$

and

$$|\hat\nabla\hat{\mathbf{A}} + \hat{\mathbf{T}}\otimes\mathring{h}|^2 \geq \frac{2(n-1)}{n(n+2)}H^2|\hat{\mathbf{T}}|^2. \tag{2.21}$$

We can play the same game with $\nabla^2\mathbf{A}$: differentiating (2.19) yields

$$\begin{aligned} \nabla_u\nabla_v\mathbf{A} =&\ \hat{\mathbf{T}}(u)\otimes\big(\mathrm{T}(v,\hat{\mathbf{A}}) + \nabla_v^\top h\big) + \mathbf{N}\otimes\big(\nabla_u\mathrm{T}(v,\hat{\mathbf{A}}) + \mathrm{T}(v,\hat\nabla_u\hat{\mathbf{A}}) + \nabla_u^\top\nabla_v^\top h\big) \\ &+\mathbf{N}\otimes\mathrm{T}(u,\hat\nabla_v\hat{\mathbf{A}}) + \hat\nabla_u\hat\nabla_v\hat{\mathbf{A}} + \big(\mathbf{N}\otimes\mathrm{T}(u,\hat{\mathbf{T}}(v)) + \nabla_u\hat{\mathbf{T}}(v)\big)\otimes h + \hat{\mathbf{T}}(v)\otimes\nabla_u^\top h, \end{aligned}$$

where

$$\nabla_u\mathrm{T}(v,\mu) \doteqdot u(\mathrm{T}(V,M)) - \mathrm{T}(\nabla_u^\top V,\mu) - \mathrm{T}(v,\hat\nabla_u M)$$

and

$$\nabla_u \hat{\mathbf{T}}(v) \doteqdot \hat{\nabla}_u(\hat{\mathbf{T}}(V)) - \hat{\mathbf{T}}(\nabla^\top_u V).$$

Taking the conormal projection yields

$$(\nabla_u \nabla_v \mathbf{A})^{\hat{\perp}} = \hat{\mathbf{T}}(u) \otimes \big(\mathrm{T}(v, \hat{\mathbf{A}}) + \nabla^\top_v h\big) + \hat{\nabla}_u \hat{\nabla}_v \hat{\mathbf{A}} + \nabla_u \hat{\mathbf{T}}(v) \otimes h + \hat{\mathbf{T}}(v) \otimes \nabla^\top_u h,$$

and hence

$$(\Delta \mathbf{A})^{\hat{\perp}} = \operatorname{tr}_{g^\top} \big(\hat{\mathbf{T}}(\cdot) \otimes \mathrm{T}(\cdot, \hat{\mathbf{A}}) + 2\hat{\mathbf{T}}(\cdot) \otimes \nabla^\top_\cdot h\big) + \operatorname{div} \hat{\mathbf{T}} \otimes h + \hat{\Delta} \hat{\mathbf{A}}. \tag{2.22}$$

2.4. **Mean curvature flow.** Now consider a family $X : M^n \times I \to N^{n+\ell}_K$ of immersed submanifolds $X(\cdot, t) : M^n \to N^{n+\ell}_K$ of $N^{n+\ell}_K$ which evolve by mean curvature flow. That is,

$$\partial_t X = \mathbf{H},$$

where $\partial_t X \doteqdot dX(\partial_t)$, and ∂_t, the *canonical vector field*, is defined via its action on functions $f \in C^\infty(M \times I)$ by

$$\partial_t|_{(x_0,t_0)} f \doteqdot \frac{d}{dt}\Big|_{t=0} f(x_0, t_0 + t).$$

The tangent bundle to $M \times I$ splits as $T(M \times I) = TM \oplus \mathbb{R}\partial_t$, where we conflate TM with the *spatial tangent bundle*, $\{\xi \in T(M \times I) : dt(\xi) = 0\}$. Here, dt is the one-form dual to ∂_t, or, equivalently, the differential of the *time projection* $(p, t) \mapsto t$ from $M \times I$ to I.

We make use of the *time-dependent connections* of Andrews–Baker [4], which extend the tangential and normal covariant derivatives to allow differentiation in space-time directions $\xi \in TM \oplus \mathbb{R}\partial_t$. That is,

$$\nabla^\top_\xi V \doteqdot \big({}^X D_\xi \big[dX(V)\big]\big)^\top \text{ and } \nabla^\perp_\xi N \doteqdot \big({}^X D_\xi N\big)^\perp,$$

where ${}^X D$ is the pullback of D to X^*TN. Observe that $\nabla^\top_\xi$ and $\nabla^\perp_\xi$ coincide with $\nabla^\top$ and $\nabla^\perp$, respectively, when $\xi \in TM$, while

$$\nabla^\top_t V = [\partial_t, V] - \mathbf{L}(V, \mathbf{H}), \tag{2.23}$$

where $[\,\cdot\,,\cdot\,]$ denotes the Lie bracket and $\nabla_t \doteqdot \nabla_{\partial_t}$.

The main advantage of working with the time-dependent connection (as opposed to the more commonly used Lie derivative) is that the tensors $g^\top$ and $g^\perp$ are ∂_t-parallel:

$$\nabla^\top_t g^\top = 0 \text{ and } \nabla^\perp_t g^\perp = 0.$$

In order to exploit this, we first derive space-time analogues of the Codazzi–Gauss–Mainardi–Ricci–Weingarten equations (following [4]).

Observe that

$$\mathbf{A}(\partial_t, u) \doteqdot ({}^X D_t[dX(u)])^\perp = \nabla^\perp_u \mathbf{H}$$

and hence

$$g^\top(\mathbf{L}(\partial_t, \nu), u) = g^\perp(\nabla^\perp_u \mathbf{H}, \nu).$$

Thus, proceeding as in the "stationary" case, we obtain the "temporal" Gauss equation

$$0 = \mathrm{Rm}^\top(\partial_t, u, v, w) + g^\perp(\nabla^\perp_w \mathbf{H}, \mathbf{A}(u, v)) - g^\perp(\nabla^\perp_v \mathbf{H}, \mathbf{A}(u, w)), \tag{2.24}$$

the temporal Codazzi–Mainardi equation

$$\nabla_t \mathbf{A}(u,v) = \nabla^\perp_u \nabla^\perp_v \mathbf{H} + \mathbf{A}(\mathbf{L}(u,\mathbf{H}),v) + g^\top(u,v)\mathbf{H}, \tag{2.25}$$

and the temporal Ricci equation

$$0 = \mathrm{Rm}^\perp(\partial_t, u, \mu, \nu) + g^\perp(\nabla^\perp_{\mathbf{L}(u,\nu)}\mathbf{H}, \mu) - g^\perp(\nabla^\perp_{\mathbf{L}(u,\mu)}\mathbf{H}, \nu)\,. \tag{2.26}$$

The Gauss and Ricci equations allow us to interchange space-time covariant derivatives. The Codazzi equation provides an evolution equation for $\mathbf{A}$.

Of course, the space-time connections $\nabla^\top$ and $\nabla^\perp$ also exhibit "spatial" Codazzi–Gauss–Mainardi–Ricci–Weingarten equations when restricted to the spatial tangent bundle.

Next, we derive a useful parabolic analogue of the Jacobi equation for minimal surfaces. It is not needed in the sequel, but it does provide some insight into the subsequent evolution equations.

Proposition 2.1 (Jacobi equation)**.** *Let $\{X_\varepsilon : M \times I \to N_K\}_{\varepsilon\in(-\varepsilon_0,\varepsilon_0)}$ be a 1-parameter family of immersed mean curvature flows with $X_0 = X$. The normal component $\sigma \doteqdot \left(\frac{d}{d\varepsilon}\big|_{\varepsilon=0} X_\varepsilon\right)^\perp$ of the variation field satisfies the **Jacobi equation***

$$(\nabla^\perp_t - \Delta^\perp)\sigma = \mathrm{tr}_{g^\top}\left(\mathbf{A}(\,\cdot\,, \mathbf{L}(\,\cdot\,,\sigma))\right) + nK\sigma\,.$$

Proof. Denote by τ the tangential component of the variation field. Fix (p,t) and consider a normal coordinate neighbourhood $(U, \{x^i\}_{i=1}^n)$ for M about p with respect to g_t. Choosing U smaller if necessary, we may identify τ with a section of TU. Using (2.5), we compute at (p,t)

$$\begin{aligned}
\nabla^\perp_i \nabla^\perp_j \sigma &= \nabla^\perp_i ({}^X D_j \sigma)^\perp \\
&= -\mathbf{A}(\partial_i, ({}^X D_j\sigma)^\top) + ({}^X D_i({}^X D_j\sigma))^\perp \\
&= \mathbf{A}(\partial_i, \mathbf{L}(\partial_j,\sigma)) + ({}^X D_i({}^X D_j(\partial_\varepsilon X - \tau)))^\perp \\
&= \mathbf{A}(\partial_i, \mathbf{L}(\partial_j,\sigma)) + \left({}^X D_i({}^X D_\varepsilon \partial_j X) - {}^X D_i(\nabla^\top_j \tau + \mathbf{A}(\partial_j,\tau))\right)^\perp \\
&= \mathbf{A}(\partial_i, \mathbf{L}(\partial_j,\sigma)) + ({}^X D_\varepsilon({}^X D_i \partial_j X) + {}^X\mathrm{Rm}(\partial_\varepsilon,\partial_i)\partial_j)^\perp \\
&\quad - \mathbf{A}(\partial_i, \nabla_j\tau) - \nabla_i\mathbf{A}(\partial_j,\tau) - \mathbf{A}(\partial_j, \nabla^\top_i\tau) \\
&= \mathbf{A}(\partial_i, \mathbf{L}(\partial_j,\sigma)) + ({}^X D_\varepsilon(dX(\nabla^\top_i\partial_j) + \mathbf{A}_{ij}) + \mathrm{Rm}(\sigma+\tau, \partial_i X)\partial_j X)^\perp \\
&\quad - \mathbf{A}(\partial_i, \nabla_j\tau) - \nabla_i\mathbf{A}(\partial_j,\tau) - \mathbf{A}(\partial_j, \nabla^\top_i\tau)\,.
\end{aligned}$$

Note that

$${}^X D_\varepsilon[dX(\nabla_i\partial_j)] = \partial_\varepsilon \Gamma_{ij}{}^k dX(\partial_k) + \Gamma_{ij}{}^{k\,X} D_\varepsilon[dX(\partial_k)]\,,$$

and hence

$$({}^X D_\varepsilon[dX(\nabla^\top_i\partial_j)])^\perp = 0$$

at (p,t). Since, by the Codazzi equation,

$$(\mathrm{Rm}(\tau,\partial_i X)\partial_j X)^\perp = 0 = \nabla_i\mathbf{A}(\tau,\partial_j) - \nabla_\tau\mathbf{A}(\partial_i,\partial_j),$$

we obtain

$$\begin{aligned}\nabla_i^\perp\nabla_j^\perp\sigma &= \mathbf{A}(\partial_i,\mathbf{L}(\partial_j,\sigma)) - (\mathrm{Rm}(\partial_i X,\sigma)\partial_j X)^\perp\\ &\quad + ({}^X D_\varepsilon(\mathbf{A}_{ij}))^\perp - \mathbf{A}(\partial_i,\nabla_j^\top\tau) - \mathbf{A}(\partial_j,\nabla_i^\top\tau) - \nabla_\tau\mathbf{A}(\partial_j,\partial_j)\,.\end{aligned}$$

On the other hand, the mean curvature flow equation yields

$$\begin{aligned}{}^X D_t\sigma &= {}^X D_t[dX(\partial_\varepsilon)-\tau]\\ &= {}^X D_\varepsilon[dX(\partial_t)] - {}^X D_t\tau\\ &= {}^X D_\varepsilon\mathbf{H} - {}^X D_t\tau\,.\end{aligned}$$

Since

$$\begin{aligned}\partial_\varepsilon g_{ij} &= \left\langle {}^X D_\varepsilon\partial_i X,\partial_j X\right\rangle + \left\langle \partial_i X, {}^X D_\varepsilon\partial_j X\right\rangle\\ &= \left\langle {}^X D_i(\tau+\sigma),\partial_j X\right\rangle + \left\langle \partial_i X, {}^X D_j(\tau+\sigma)\right\rangle\\ &= g(\nabla_i^\top\tau - \mathbf{L}(\partial_i,\sigma),\partial_j) + g(\partial_i,\nabla_j^\top\tau - \mathbf{L}(\partial_j,\sigma))\,,\end{aligned}$$

and hence

$$\begin{aligned}\partial_\varepsilon g^{ij} &= -\, g^{ip}g^{jq}\partial_\varepsilon g_{pq}\\ &= -\, g^{ip}g^{jq}\left(g(\nabla_p^\top\tau - \mathbf{L}(\partial_p,\sigma),\partial_q)\right) + g\left(\partial_p,\nabla_q^\top\tau - \mathbf{L}(\partial_q,\sigma)\right)\right)\,,\end{aligned}$$

we obtain

$$\begin{aligned}{}^X D_\varepsilon\mathbf{H} &= g^{ij\,X}D_\varepsilon(\mathbf{A}_{ij}) - 2g^{ip}g^{jq}g(\nabla_p^\top\tau - \mathbf{L}(\partial_p,\sigma),\partial_q)\mathbf{A}_{ij}\\ &= g^{ij}\left({}^X D_\varepsilon(\mathbf{A}_{ij}) - 2\mathbf{A}(\nabla_i^\top\tau - \mathbf{L}(\partial_i,\sigma),\partial_j)\right)\end{aligned}$$

and hence

$$\nabla_t^\perp\sigma = \left(g^{ij}\left({}^X D_\varepsilon\mathbf{A}_{ij} - 2\mathbf{A}(\nabla_i^\top\tau - \mathbf{L}(\partial_i,\sigma),\partial_j)\right) - {}^X D_t\tau\right)^\perp\,.$$

Finally, since

$${}^X D_t\tau - {}^X D_\tau\mathbf{H} = dX[\partial_t,\tau] = \frac{\partial\tau^i}{\partial t}dX(\partial_i)$$

has no normal component, we conclude that

$$\begin{aligned}(\nabla_t^\perp - \Delta^\perp)\sigma &= \operatorname{tr}_{g^\top}\left(\mathbf{A}(\,\cdot\,,\mathbf{L}(\,\cdot\,,\sigma)) + (\mathrm{Rm}(\,\cdot\,,\sigma)\,\cdot\,)^\perp\right)\\ &= \operatorname{tr}_{g^\top}\left(\mathbf{A}(\,\cdot\,,\mathbf{L}(\,\cdot\,,\sigma))\right) + nK\sigma\end{aligned}$$

as desired. □

Taking the trace of (2.25), or applying Proposition 2.1 with the variation $X_\varepsilon(x,t) \doteqdot X(x,t+\varepsilon)$, yields

$$(\nabla_t^\perp - \Delta^\perp)\mathbf{H} = \operatorname{tr}_{g^\top}\left(\mathbf{A}(\,\cdot\,,\mathbf{L}(\,\cdot\,,\mathbf{H}))\right) + nK\mathbf{H}. \tag{2.27}$$

It follows that

$$(\partial_t - \Delta)\frac{1}{2}|\mathbf{H}|^2 = |\mathbf{L}(\,\cdot\,,\mathbf{H})|^2 + nK|\mathbf{H}|^2 - |\nabla^\perp\mathbf{H}|^2. \tag{2.28}$$

Applying the trace Simons equation (2.15) to the temporal Codazzi equation (2.25) yields

$$\begin{aligned}(\nabla_t-\Delta)\mathbf{A}(u,v)= \operatorname{tr}_{g^\top}\Big(&\mathbf{A}\big(u,\mathbf{L}\big(\cdot,\mathbf{A}(v,\cdot)\big)+\mathbf{A}\big(v,\mathbf{L}\big(\cdot,\mathbf{A}(u,\cdot)\big)\big)-2\mathbf{A}\big(\cdot,\mathbf{L}\big(u,\mathbf{A}(v,\cdot)\big)\big)\\ &+\mathbf{A}\big(\cdot,\mathbf{L}\big(\cdot,\mathbf{A}(u,v)\big)\big)\Big)+g^\top(u,v)K\mathbf{H}-nK\Big(\mathbf{A}\big(u,v\big)-\tfrac{1}{n}g^\top(u,v)\mathbf{H}\Big).\end{aligned}\tag{2.29}$$

Tracing this of course recovers (2.27).

We also obtain

$$(\nabla_t-\Delta)\frac{1}{2}|\mathbf{A}|^2= -|\nabla\mathbf{A}|^2+|\langle\mathbf{A},\mathbf{A}\rangle^\top|^2+|\mathring{\mathbf{A}}\wedge\mathring{\mathbf{A}}|^2+nK|\mathbf{A}|^2-2nK|\mathring{\mathbf{A}}|^2,\tag{2.30}$$

where $\langle\mathbf{A},\mathbf{A}\rangle^\top\in\Gamma(NM\otimes NM)$ is formed from $\mathbf{A}\otimes\mathbf{A}$ by contracting the tangential components. That is, it is dual to the tensor $\langle\mathbf{L},\mathbf{L}\rangle^\top\in\Gamma(N^*M\otimes N^*M)$ defined by

$$\langle\mathbf{L},\mathbf{L}\rangle^\top(\mu,\nu)\doteqdot g^\top\big(\mathbf{L}(\cdot,\mu),\mathbf{L}(\cdot,\nu)\big).$$

Now consider points where $\mathbf{H}\neq 0$. By (2.28),

$$(\partial_t-\Delta)H=\big(|h|^2-|\hat{\mathbf{T}}|^2+nK\big)H\tag{2.31}$$

and, since

$$\Delta^\perp\mathbf{N}=\operatorname{div}\hat{\mathbf{T}}+\operatorname{tr}_{g^\top}\big(\mathrm{T}(\cdot,\hat{\mathbf{T}}(\cdot))\big)\mathbf{N},$$

$$\nabla_t^\perp\mathbf{N}=\operatorname{div}\hat{\mathbf{T}}+\operatorname{tr}_{g^\top}\big(\mathrm{T}(\cdot,\hat{\mathbf{T}}(\cdot))\big)\mathbf{N}+\operatorname{tr}_{g^\top}\big(\hat{\mathbf{A}}(\,\cdot\,,\mathbf{L}(\,\cdot\,,\mathbf{N}))\big)+|\hat{\mathbf{T}}|^2\mathbf{N}+2\hat{\mathbf{T}}\left(\nabla\log H\right).\tag{2.32}$$

Projecting (2.29) onto the conormal bundle yields

$$\begin{aligned}\big[(\nabla_t-\Delta)\mathbf{A}\big]^{\hat\perp}(u,v)= \operatorname{tr}_{g^\top}\Big(&\hat{\mathbf{A}}\big(u,\mathbf{L}\big(\cdot,\mathbf{A}(v,\cdot)\big)+\hat{\mathbf{A}}\big(v,\mathbf{L}\big(\cdot,\mathbf{A}(u,\cdot)\big)\big)-2\hat{\mathbf{A}}\big(\cdot,\mathbf{L}\big(u,\mathbf{A}(v,\cdot)\big)\big)\\ &+\hat{\mathbf{A}}\big(\cdot,\mathbf{L}\big(\cdot,\mathbf{A}(u,v)\big)\big)\Big)-nK\hat{\mathbf{A}}\big(u,v\big).\end{aligned}$$

On the other hand, differentiating the principal-conormal decomposition of $\mathbf{A}$ yields

$$\begin{aligned}\nabla_t\mathbf{A}&=\nabla_t^\perp\mathbf{N}\otimes h+\mathbf{N}\otimes\nabla_t^\top h+\nabla_t\hat{\mathbf{A}}\\ &=\nabla_t^\perp\mathbf{N}\otimes h+\hat\nabla_t\hat{\mathbf{A}}+\mathbf{N}\otimes\big(\nabla_t^\top h+\mathrm{T}(\cdot,\hat{\mathbf{A}})\big)\end{aligned}$$

and hence, recalling (2.22) and (2.32),

$$\begin{aligned}\big[(\nabla_t-\Delta)\mathbf{A}\big]^{\hat\perp}=(\hat\nabla_t-\hat\Delta)\hat{\mathbf{A}}&+\big[\operatorname{tr}_{g^\top}\big(\hat{\mathbf{A}}(\,\cdot\,,\mathbf{L}(\,\cdot\,,\mathbf{N}))\big)+2\hat{\mathbf{T}}\left(\nabla\log H\right)\big]\otimes h\\ &-\operatorname{tr}_{g^\top}\Big(\hat{\mathbf{T}}(\cdot)\otimes\mathrm{T}(\cdot,\hat{\mathbf{A}})+2\hat{\mathbf{T}}(\cdot)\otimes\nabla^\top_{\cdot}h\Big).\end{aligned}$$

Since

$$\begin{aligned}|\nabla\hat{\mathbf{A}}|^2&=|\hat\nabla\hat{\mathbf{A}}|^2+|\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2\\ &=|\hat\nabla\hat{\mathbf{A}}|^2+\hat g\big(\operatorname{tr}_{g^\top}(\hat{\mathbf{T}}(\cdot)\otimes\mathrm{T}(\cdot,\hat{\mathbf{A}})),\hat{\mathbf{A}}\big),\end{aligned}$$

we thus obtain, wherever $\mathbf{H} \neq 0$ (cf. [21]),

$$
\begin{aligned}
(\partial_t - \Delta)\frac{1}{2}|\hat{\mathbf{A}}|^2 &= \hat{g}((\hat{\nabla}_t - \hat{\Delta})\hat{\mathbf{A}}, \hat{\mathbf{A}}) - |\hat{\nabla}\hat{\mathbf{A}}|^2 \\
&= \big|\big\langle \mathbf{A}, \hat{\mathbf{A}}\big\rangle^\top\big|^2 - \big|\big\langle \mathbf{N}\otimes h, \hat{\mathbf{A}}\big\rangle^\top\big|^2 + |\mathbf{A}\wedge\hat{\mathbf{A}}|^2 - nK|\hat{\mathbf{A}}|^2 - |\nabla\hat{\mathbf{A}}|^2 \\
&\quad - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H}, \nabla\hat{\mathbf{A}}\Big) \\
&= \big|\big\langle \hat{\mathbf{A}}, \hat{\mathbf{A}}\big\rangle^\top\big|^2 + |\hat{\mathbf{A}}\wedge\hat{\mathbf{A}}|^2 + |\mathbf{N}\otimes h\wedge\hat{\mathbf{A}}|^2 - nK|\hat{\mathbf{A}}|^2 - |\nabla\hat{\mathbf{A}}|^2 \\
&\quad - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H}, \nabla\hat{\mathbf{A}}\Big). \qquad (2.33)
\end{aligned}
$$

Given tensor fields S and T formed from TM and NM, we denote by $S * T$ any tensor field resulting from linear combinations of contractions of $S\otimes T$ with $g^\top$ and $g^\perp$. By (2.23), (2.24), and (2.26),

$$\nabla_t(\nabla. T) = \nabla(\nabla_t T) + \mathbf{A} * \mathbf{A} * \nabla T + \mathbf{A} * \nabla\mathbf{A} * T.$$

By (2.6) and (2.10),

$$\Delta(\nabla T) = \nabla(\Delta T) + \mathbf{A} * \mathbf{A} * \nabla T + K * \nabla T + \mathbf{A} * \nabla\mathbf{A} * T\,.$$

Thus,

$$(\nabla_t - \Delta)(\nabla\mathbf{A}) = \mathbf{A} * \mathbf{A} * \nabla\mathbf{A} + K * \nabla\mathbf{A}\,,$$

and hence, by Young's inequality,

$$(\partial_t - \Delta)|\nabla\mathbf{A}|^2 \leq\, -2|\nabla^2\mathbf{A}|^2 + c(|\mathbf{A}|^2 + K)|\nabla\mathbf{A}|^2\,, \qquad (2.34)$$

where c is a constant that depends only on n and k.

Similarly,

$$(\nabla_t - \Delta)(\nabla^2\mathbf{A}) = \mathbf{A} * \mathbf{A} * \nabla^2\mathbf{A} + \mathbf{A} * \nabla\mathbf{A} * \nabla\mathbf{A} + K * \nabla^2\mathbf{A}\,,$$

and hence

$$(\partial_t - \Delta)|\nabla^2\mathbf{A}|^2 \leq\, -2|\nabla^3\mathbf{A}|^2 + c\left[(|\mathbf{A}|^2 + K)|\nabla^2\mathbf{A}|^2 + |\nabla\mathbf{A}|^4\right], \qquad (2.35)$$

where c is a constant that depends only on n and k.

Similar inequalities hold for higher derivatives of $\mathbf{A}$ since, by an induction argument,

$$(\nabla_t - \Delta)(\nabla^m\mathbf{A}) = K * \nabla^m\mathbf{A} + \sum_{i+j+k=m} \nabla^i\mathbf{A} * \nabla^j\mathbf{A} * \nabla^k\mathbf{A}\,. \qquad (2.36)$$

The following "Bernstein estimates" are a standard application of the "rough" evolution equations (2.36). For a proof in the Euclidean case (which carries over with minor modifications) see, for example, [5, Theorem 6.24].

Proposition 2.2 (Bernstein estimates)**.** *Let $X : M \times [0, \lambda K^{-1}] \to N_K^{n+\ell}$ be a solution to mean curvature flow, where $K > 0$. If*

$$\max_{M\times[0,\lambda K^{-1}]} |\mathbf{A}|^2 \leq \Lambda_0 K\,,$$

then

$$t^m|\nabla^m\mathbf{A}|^2 \le \Lambda_m K\,,$$

where Λ_m *depends only on* n, k, m, λ *and* Λ_0.

2.5. **A Poincaré-type inequality.** We need the following Poincaré-type inequality, which combines the arguments of [16, Proposition 2.2] and [18, Proposition 3.2]. Note that the hypothesis (2.37) is motivated by the codimension estimate proved in Section 3.2 below.

Proposition 2.3. *There exists* $\gamma = \gamma(n,\alpha,\eta,\delta,\Lambda) > 0$ *with the following property. Given a smoothly immersed submanifold* $X : M^n \to S^{n+\ell}_K$ *satisfying*

$$|\hat{\mathbf{A}}|^2 \le \Lambda K^\delta\big(|\mathbf{H}|^2+K\big)^{1-\delta} \;\text{ wherever }\; \mathbf{H}\neq 0\,, \tag{2.37}$$

define the "acylindrical" set $U_{\alpha,\eta}\subset M$ *by*

$$U_{\alpha,\eta} \doteqdot \left\{x\in M : |\mathbf{A}|^2 - \tfrac{1}{n-2+\alpha}|\mathbf{H}|^2 - 2(2-\alpha)K \le 0 \le |\mathbf{A}|^2 - \left(\tfrac{1}{n-1}+\eta\right)|\mathbf{H}|^2\right\}\,.$$

If $u\in W^{1,2}(M)$ *satisfies* $\operatorname{spt} u \Subset U_{\alpha,\eta}$, *then*

$$\gamma\int u^2(|\mathbf{H}|^2+K)\,d\mu \le \int u^2\left(K + \frac{|\nabla u|}{u}\frac{|\nabla\mathbf{A}|}{\sqrt{|\mathbf{H}|^2+K}} + \frac{|\nabla\mathbf{A}|^2}{|\mathbf{H}|^2+K}\right)d\mu.$$

Proof. Since $C^\infty(M)$ is dense in $W^{1,2}(M)$, it suffices to establish the estimate for smooth u. Define a tensor $E\in\Gamma(T^*M\otimes T^*M\otimes T^*M\otimes T^*M)$ by

$$\begin{aligned} E(u,v,w,z) &\doteqdot \nabla_{(u}\nabla_{v)}\mathbf{A}(w,z) - \nabla_{(w}\nabla_{z)}\mathbf{A}(u,v)\\ &\doteqdot \frac{1}{2}\big(\nabla_u\nabla_v\mathbf{A}(w,z) - \nabla_w\nabla_z\mathbf{A}(u,v) + \nabla_v\nabla_u\mathbf{A}(w,z) - \nabla_z\nabla_v\mathbf{A}(u,v)\big)\,. \end{aligned}$$

We first consider points where $\mathbf{H}\neq 0$. By Simons' equation (2.14), arguing as in [18, Lemma 3.1] yields is a constant C depending only on n such that

$$|E|^2 \ge 2\big|h\otimes h^2 - h^2\otimes h + K(g\otimes h - h\otimes g)\big|^2 - C(|h|^5 + K^{\frac{5}{2}})|\hat{\mathbf{A}}| \tag{2.38}$$

wherever $|\mathbf{H}|>0$.

Let us define

$$F \doteqdot h\otimes h^2 - h^2\otimes h + K(g\otimes h - h\otimes g)\,.$$

We claim there is a constant $\gamma=\gamma(n,\alpha,\eta)$ such that

$$|F|^2 + |\mathbf{H}|^5|\hat{\mathbf{A}}| + K^3 \ge \gamma W^3 \tag{2.39}$$

at points in $U_{\alpha,\eta}$ with $\mathbf{H}\neq 0$, where

$$W \doteqdot |\mathbf{H}|^2 + K\,.$$

In fact, we will show that (2.39) holds as an algebraic inequality. If we write λ_p for the eigenvalues of h, this inequality can be rewritten as

$$\sum_{p,q}(\lambda_p\lambda_q+K)^2(\lambda_p-\lambda_q)^2 + \operatorname{tr}(\lambda)^5|\hat{\mathbf{A}}| + K^3 \ge \gamma(\operatorname{tr}(\lambda)^2+K)^3,$$

where

$$|\lambda|^2 \doteqdot \sum_p \lambda_p^2 \ \text{ and } \ \operatorname{tr}(\lambda) \doteqdot \sum_p \lambda_p. \tag{2.40}$$

Thus, if (2.39) is false (as an algebraic inequality), then we can find a sequence of vector-valued symmetric bilinear forms $\mathbf{A}_i \in \mathbb{R}^n \odot \mathbb{R}^n \otimes \mathbb{R}^\ell$ with nonvanishing trace such that the conditions $\operatorname{tr}(\lambda^i) > 0$,

$$0 \leq |\lambda^i|^2 + |\hat{\mathbf{A}}_i|^2 - \left(\tfrac{1}{n-1} + \eta\right) \operatorname{tr}(\lambda^i)^2, \tag{2.41}$$

and

$$|\lambda^i|^2 + |\hat{\mathbf{A}}_i|^2 - \tfrac{1}{n-2+\alpha} \operatorname{tr}(\lambda^i)^2 - 2(2-\alpha)K \leq 0 \tag{2.42}$$

hold for every $i \in \mathbb{N}$, and yet

$$\frac{\sum_{p,q}(\lambda_p^i\lambda_q^i + K)^2(\lambda_p^i - \lambda_q^i)^2 + \operatorname{tr}(\lambda^i)^5|\hat{\mathbf{A}}_i| + K^3}{(\operatorname{tr}(\lambda^i)^2 + K)^3} \to 0 \tag{2.43}$$

as $i \to \infty$. It follows that $\operatorname{tr}(\lambda^i)^2 \to \infty$ as $i \to \infty$, and hence, as a consequence of (2.42),

$$\operatorname{tr}(\lambda^i)^{-2}|\lambda^i|^2 + \operatorname{tr}(\lambda^i)^{-2}|\hat{\mathbf{A}}_i|^2 \leq \tfrac{1}{n-2+\alpha} + 2(2-\alpha)K \operatorname{tr}(\lambda^i)^{-2}.$$

So (passing to a subseqence if necessary) we may extract limits

$$\operatorname{tr}(\lambda^i)^{-1}\lambda^i \to \lambda^\infty \ \text{ and } \ \operatorname{tr}(\lambda^i)^{-1}\hat{\mathbf{A}}_i \to \hat{\mathbf{A}}_\infty.$$

Passing to the limit in (2.43) we find that

$$\sum_{p,q}(\lambda_p^\infty\lambda_q^\infty)^2(\lambda_p^\infty - \lambda_q^\infty)^2 + |\hat{\mathbf{A}}_\infty| = 0,$$

and hence $\hat{\mathbf{A}}_\infty = 0$ and $\lambda^\infty = \frac{1}{m}\sum_{p=1}^m e_p$ for some $1 \leq m \leq n$, where $\{e_p\}_{p=1}^n$ is the standard basis for $\mathbb{R}^n$. Passing to limits in (2.41) and (2.42), and inserting $\hat{\mathbf{A}}_\infty = 0$ and $|\lambda^\infty|^2 = m^{-1}$, we obtain

$$\frac{1}{n-1} + \eta \leq \frac{1}{m} \leq \frac{1}{n-2+\alpha}.$$

It follows that $n-2 < m < n-1$, which is absurd. We conclude that our initial assumption was false; that is, (2.39) holds in $U_{\alpha,\eta}$ at points where $\mathbf{H} \neq 0$. Recalling (2.38) then yields

$$\gamma W^3 \leq K^3 + |E|^2 + CW^{\frac{5}{2}}|\hat{\mathbf{A}}|,$$

where $C = C(n, \alpha)$. Applying the hypothesis (2.37) and Young's inequality then yields

$$\gamma W^3 \leq K^3 + |E|^2$$

with γ taking a smaller value and depending now also on δ and Λ.

On the other hand, wherever $\mathbf{H} = 0$,

$$W^3 = K^3 \leq |E|^2 + K^3.$$

We conclude that

$$\gamma W^3 \leq |E|^2 + K^3$$

at all points of $U_{\alpha,\eta}$, where $\gamma = \gamma(n, \alpha, \eta, \delta, \Lambda)$. Thus,

$$\begin{aligned}\gamma\int u^2 W\, d\mu &\leq \int \frac{u^2}{W^2}\left(|E|^2 + K^3\right) d\mu\\ &\leq \int u^2\left(\frac{|E|^2}{W^2} + K\right) d\mu.\end{aligned}$$

We now estimate, using Simons' equation and the divergence theorem,

$$\begin{aligned}\int \frac{u^2}{W^2}|E|^2 d\mu &= \int \frac{u^2}{W^2} E * \nabla^2\mathbf{A}\, d\mu\\ &= \int \frac{u^2}{W^2}\left(\frac{\nabla u}{u} * E + \frac{\nabla W}{W} * E + \nabla E\right) * \nabla\mathbf{A}\, d\mu\\ &\leq C\int \frac{u^2}{W^2}\left(W^{\frac{3}{2}}\frac{|\nabla u|}{u} + W^{\frac{1}{2}}|\nabla W| + W|\nabla\mathbf{A}|\right)|\nabla\mathbf{A}|\, d\mu\\ &\leq C\int u^2\left(\frac{|\nabla u|}{u} + \frac{|\nabla\mathbf{A}|}{W^{\frac{1}{2}}}\right)\frac{|\nabla\mathbf{A}|}{W^{\frac{1}{2}}}\, d\mu,\end{aligned}$$

where C depends only on n, α and η. This completes the proof. □

2.6. **Notation.**

Before moving on, let us review our notation for the tensors which will appear frequently in the sequel.

$\mathbf{N}$	Principal normal vector
$\mathbf{A}$	Second fundamental form
h	Principal normal component of $\mathbf{A}$
$\hat{\mathbf{A}}$	Conormal projection of $\mathbf{A}$
$\mathbf{L}$	Weingarten tensor
$\mathbf{H}$	Mean curvature vector
H	Mean curvature scalar
$\mathring{\mathbf{A}}$	Trace-free part of $\mathbf{A}$
$\mathring{h}$	Trace-free part of h
T	Torsion (scalar valued)
$\hat{\mathbf{T}}$	Torsion (conormal valued)

3. The key estimates for smooth flows

In this section, we obtain the key estimates in the setting of smooth flows. In the following section, we will obtain extensions of appropriately modified versions of these estimates for surgically modified flows.

3.1. **Preserving quadratic pinching.**

If $X_0 : M^n \to S_K^{n+\ell}$ satisfies (1.2), then we can find $\alpha \in (\alpha_n, 1)$ such that

$$|\mathbf{A}|^2 - \frac{1}{n-2+\alpha}|\mathbf{H}|^2 - 2(2-\alpha)K \leq 0, \tag{3.1}$$

where

$$\alpha_n \doteqdot \max\{2 - \tfrac{n}{4}, 0\} = \begin{cases} \frac{3}{4} & \text{if } n = 5 \\ \frac{1}{2} & \text{if } n = 6 \\ \frac{1}{4} & \text{if } n = 7 \\ 0 & \text{if } n \geq 8. \end{cases} \tag{3.2}$$

We will show that this is preserved under mean curvature flow. In fact, we will prove (more generally) that the condition

$$|\mathbf{A}|^2 - \frac{1}{n-m+\alpha}|\mathbf{H}|^2 - 2(m-\alpha)K \leq 0 \tag{3.3}$$

is preserved for any integer $m \geq 1$ and any $\alpha \in [0, 1)$ so long as $m - \alpha \leq \frac{n}{4}$.

Proposition 3.1 (Quadratic pinching is preserved)**.** *Let $X : M^n \times [0, T) \to S_K^{n+\ell}$ be a solution to mean curvature flow such that* (3.1) *holds on $M^n \times \{0\}$ for some integer $m \geq 1$ and some $\alpha \in [0, 1)$. If $m - \alpha \leq \frac{n}{4}$, then* (3.1) *holds on $M^n \times \{t\}$ for all $t \in [0, T)$.*

Proof. Given positive numbers a and b, consider the function

$$Q \doteqdot \frac{1}{2}\left(|\mathbf{A}|^2 - a|\mathbf{H}|^2 - bK\right).$$

By (2.30) and (2.28),

$$\begin{aligned}(\partial_t - \Delta)Q = {} & |\mathring{\mathbf{A}} \wedge \mathring{\mathbf{A}}|^2 + |\langle \mathbf{A}, \mathbf{A}\rangle^\top|^2 - a|\mathbf{L}(\,\cdot\,, \mathbf{H})|^2 + nK\big(|\mathbf{A}|^2 - a|\mathbf{H}|^2\big) - 2nK|\mathring{\mathbf{A}}|^2 \\ & - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big).\end{aligned}$$

At a point where $\mathbf{H} \neq 0$, decomposing $\mathbf{A}$ into its irreducible components yields [4]

$$\begin{gathered}|\mathbf{A}|^2 = |\mathring{h}|^2 + \tfrac{1}{n}H^2 + |\hat{\mathbf{A}}|^2, \\ |\mathbf{L}(\cdot, \mathbf{H})|^2 = |\mathring{h}|^2 H^2 + \tfrac{1}{n}H^4, \\ |\mathring{\mathbf{A}} \wedge \mathring{\mathbf{A}}|^2 = 2|(\mathbf{N} \otimes \mathring{h}) \wedge \hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}} \wedge \hat{\mathbf{A}}|^2,\end{gathered}$$

and

$$\begin{aligned}|\langle \mathbf{A}, \mathbf{A}\rangle^\top|^2 &= \big|\big\langle \mathbf{N} \otimes \mathring{h} + \tfrac{1}{n}H\mathbf{N}g^\top + \hat{\mathbf{A}}, \mathbf{N} \otimes \mathring{h} + \tfrac{1}{n}H\mathbf{N}g^\top + \hat{\mathbf{A}}\big\rangle\big| \\ &= \big(|\mathring{h}|^2 + \tfrac{1}{n}H^2\big)^2 + 2\big|\big\langle \mathbf{N} \otimes \mathring{h}, \hat{\mathbf{A}}\big\rangle^\top\big|^2 + |\big\langle \hat{\mathbf{A}}, \hat{\mathbf{A}}\big\rangle^\top|^2.\end{aligned}$$

Applying the estimates [4, p. 372]

$$|\mathbf{N} \otimes \mathring{h} \wedge \hat{\mathbf{A}}|^2 + \big|\big\langle \mathring{h}, \hat{\mathbf{A}}\big\rangle^\top\big|^2 \leq 2|\mathring{h}|^2|\hat{\mathbf{A}}|^2 \tag{3.4}$$

and [2, Proposition 3] (cf. [11, Lemma 1])

$$|\hat{\mathbf{A}} \wedge \hat{\mathbf{A}}|^2 + |\big\langle \hat{\mathbf{A}}, \hat{\mathbf{A}}\big\rangle^\top|^2 \leq \tfrac{3}{2}|\hat{\mathbf{A}}|^4 \tag{3.5}$$

now yields

$$\begin{aligned}|\mathring{\mathbf{A}}\wedge\mathring{\mathbf{A}}|^2+|\langle\mathbf{A},\mathbf{A}\rangle^\top|^2 &\le 4|\mathring{h}|^2|\hat{\mathbf{A}}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^4+\big(|\mathring{h}|^2+\tfrac{1}{n}H^2\big)^2\\ &= 3|\mathring{h}|^2|\hat{\mathbf{A}}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^4+\big(|\mathring{h}|^2+\tfrac{1}{n}H^2\big)|\mathbf{A}|^2-\tfrac{1}{n}|\hat{\mathbf{A}}|^2H^2\end{aligned}$$

and hence

$$\begin{aligned}(\partial_t-\Delta)Q &= |\mathring{\mathbf{A}}\wedge\mathring{\mathbf{A}}|^2+|\langle\mathbf{A},\mathbf{A}\rangle^\top|^2-a\big(|\mathring{h}|^2+\tfrac{1}{n}H^2\big)H^2+nK\big(2Q+bK\big)-2nK|\mathring{\mathbf{A}}|^2\\ &\quad-\big(|\nabla\mathbf{A}|^2-a|\nabla^\perp\mathbf{H}|^2\big)\\ &\le 3|\mathring{h}|^2|\hat{\mathbf{A}}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^4-\tfrac{1}{n}|\hat{\mathbf{A}}|^2H^2-2nK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2\big)\\ &\quad+\big(|\mathring{h}|^2+\tfrac{1}{n}H^2+nK\big)\big(2Q+bK\big)-\big(|\nabla\mathbf{A}|^2-a|\nabla^\perp\mathbf{H}|^2\big)\\ &= \big(3|\mathring{h}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^2-\tfrac{1}{n}H^2-bK\big)|\hat{\mathbf{A}}|^2\\ &\quad+bK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2+\tfrac{1}{n}H^2+nK\big)-2nK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2\big)\\ &\quad+2Q\big(|\mathring{h}|^2+\tfrac{1}{n}H^2+nK\big)-\big(|\nabla\mathbf{A}|^2-a|\nabla^\perp\mathbf{H}|^2\big).\end{aligned}$$

Rewriting

$$|\mathbf{A}|^2=2Q+aH^2+bK\,,$$

we obtain

$$bK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2+\tfrac{1}{n}H^2+nK\big)=bK(2Q+aH^2+(b+n)K)$$

and

$$-2nK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2\big)=\,-2nK\big(2Q+\big(a-\tfrac{1}{n}\big)H^2+bK\big)\,,$$

and hence

$$\begin{aligned}&bK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2+\tfrac{1}{n}H^2+nK\big)-2nK\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2\big)\\ &\qquad=\,-2K(2n-b)Q+(a(b-2n)+2)KH^2+b(b-n)K^2.\end{aligned}$$

Similarly, using

$$\tfrac{1}{n}H^2=\tfrac{1}{an-1}\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2-2Q-bK\big)$$

we find

$$\begin{aligned}3|\mathring{h}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^2-\tfrac{1}{n}H^2-bK &= 3|\mathring{h}|^2+\tfrac{3}{2}|\hat{\mathbf{A}}|^2-\tfrac{1}{an-1}\big(|\mathring{h}|^2+|\hat{\mathbf{A}}|^2-2Q-bK\big)-bK\\ &= \big(3-\tfrac{1}{an-1}\big)|\mathring{h}|^2+\big(\tfrac{3}{2}-\tfrac{1}{an-1}\big)|\hat{\mathbf{A}}|^2+\tfrac{2}{an-1}Q-\big(1-\tfrac{1}{an-1}\big)bK\,.\end{aligned}$$

If $a\le\frac{4}{3n}$, then the first term on the right is nonpositive. Discarding this term and putting things back together, we arrive at

$$\begin{aligned}(\partial_t-\Delta)Q &\le \big(\tfrac{3}{2}-\tfrac{1}{an-1}\big)|\hat{\mathbf{A}}|^4-\big(1-\tfrac{1}{an-1}\big)bK|\hat{\mathbf{A}}|^2+b(b-n)K^2\\ &\quad+(a(b-2n)+2)KH^2+2Q\big(|\mathring{h}|^2+\tfrac{1}{an-1}|\hat{\mathbf{A}}|^2+\tfrac{1}{n}H^2+(b-n)K\big)\\ &\quad-\big(|\nabla\mathbf{A}|^2-a|\nabla^\perp\mathbf{H}|^2\big).\end{aligned}$$

Observe that if

$$a = \frac{1}{n-m+\alpha} \quad \text{and} \quad b = 2(m-\alpha)$$

for some m and $\alpha \in (0,1)$ such that $m-\alpha \leq \frac{n}{4}$ (which ensures that $a \leq \frac{4}{3n}$), then

$$a(b-2n)+2 = 0 \quad \text{and} \quad \big(1-\tfrac{1}{an-1}\big)b = 2(b-n).$$

We arrive at

$$\begin{aligned}(\tfrac{3}{2}-\tfrac{1}{an-1})|\hat{\mathbf{A}}|^4 - \big(1-\tfrac{1}{an-1}\big)bK|\hat{\mathbf{A}}|^2 + b(b-n)K^2 = (\tfrac{3}{2}-\tfrac{1}{an-1})|\hat{\mathbf{A}}|^4 - 2(b-n)K|\hat{\mathbf{A}}|^2 \\ + b(b-n)K^2.\end{aligned}$$

The discriminant of the quadratic form on the right is $3(b-n)(\frac{n}{3}-m+\alpha)$. Since $m-\alpha \leq \frac{n}{4}$ and $b = 2(m-\alpha)$, this is strictly negative. Thus,

$$(\partial_t - \Delta)Q \leq 2Q\big(|\mathring{h}|^2 + \tfrac{1}{an-1}|\hat{\mathbf{A}}|^2 + \tfrac{1}{n}H^2 + (b-n)K\big) - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big). \tag{3.6}$$

Since $n \geq 1 + \frac{2(m-\alpha)}{3}$, the gradient terms can be discarded using the Kato inequality.

On the other hand, wherever $\mathbf{H} = 0$,

$$(\partial_t - \Delta)Q = |\mathring{\mathbf{A}} \wedge \mathring{\mathbf{A}}|^2 + \big|\big\langle \mathring{\mathbf{A}}, \mathring{\mathbf{A}}\big\rangle^\top\big|^2 - nK|\mathring{\mathbf{A}}|^2 - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big).$$

So [2, Proposition 3] yields

$$\begin{aligned}(\partial_t - \Delta)Q &\leq \big(\tfrac{3}{2}|\mathring{\mathbf{A}}|^2 - nK\big)|\mathring{\mathbf{A}}|^2 - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big) \\ &= 3|\mathring{\mathbf{A}}|^2 Q + \big(\tfrac{3}{2}b - n\big)K|\mathring{\mathbf{A}}|^2 - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big).\end{aligned}$$

Since $b = 2(m-\alpha) < \frac{n}{2}$, the Kato inequality then yields

$$(\partial_t - \Delta)Q \leq 3|\mathring{\mathbf{A}}|^2 Q.$$

We conclude that

$$(\partial_t - \Delta)Q \leq fQ$$

everywhere for some locally bounded function f, at which point we may conclude that non-positivity of Q is preserved. □

We can use the preservation of pinching to obtain an estimate for the trace-free second fundamental form which improves at large times, using the maximum principle.

Proposition 3.2. *Let $X : M \times [0,T) \to S_K^{n+\ell}$ be a solution to mean curvature flow with initial condition in satisfying (3.3) with $m - \alpha < \frac{n}{4}$. There is a constant $C = C(n, m-\alpha)$ such that*

$$\frac{|\mathbf{A}|^2 - \frac{1}{n}|\mathbf{H}|^2}{|\mathbf{H}|^2 + K} \leq C\mathrm{e}^{-2Kt}. \tag{3.7}$$

Proof. Set

$$g \doteqdot \frac{1}{2}\big(|\mathbf{A}|^2 - \tfrac{1}{n}|\mathbf{H}|^2\big).$$

and

$$W \doteqdot \frac{1}{2}\big(bK + a|\mathbf{H}|^2 - |\mathbf{A}|^2\big),$$

where

$$a = \frac{4}{3n} \text{ and } b = \frac{n}{2}.$$

Applying (2.28) and (2.30) yields (cf. Proposition 3.1), wherever $\mathbf{H} \neq 0$,

$$\begin{aligned}(\partial_t - \Delta)g \leq{}& \big(3|\mathring{h}|^2 + \tfrac{3}{2}|\hat{\mathbf{A}}|^2 - \tfrac{1}{n}H^2\big)|\hat{\mathbf{A}}|^2 - 2nK\big(|\mathring{h}|^2 + |\hat{\mathbf{A}}|^2\big)\\ &+ 2g\big(|\mathring{h}|^2 + \tfrac{1}{n}H^2 + nK\big) - \big(|\nabla\mathbf{A}|^2 - \tfrac{1}{n}|\nabla^\perp\mathbf{H}|^2\big).\end{aligned}$$

Note that

$$|\mathring{h}|^2 + |\hat{\mathbf{A}}|^2 = 2g$$

so that

$$3|\mathring{h}|^2 + \tfrac{3}{2}|\hat{\mathbf{A}}|^2 - \tfrac{1}{n}H^2 = (3 - \tfrac{1}{an-1})|\mathring{h}|^2 + (\tfrac{3}{2} - \tfrac{1}{an-1})|\hat{\mathbf{A}}|^2 - \tfrac{1}{n}H^2 + \tfrac{2}{an-1}g.$$

The first three terms are non-positive. The remaining term is also non-positive by the Kato inequality. Thus,

$$\begin{aligned}(\partial_t - \Delta)g \leq{}& -4nKg + 2g\big(|\mathring{h}|^2 + \tfrac{1}{an-1}|\hat{\mathbf{A}}|^2 + \tfrac{1}{n}H^2 + nK\big)\\ ={}& 2g\big(|\mathring{h}|^2 + \tfrac{1}{an-1}|\hat{\mathbf{A}}|^2 + \tfrac{1}{n}H^2 - nK\big).\end{aligned} \tag{3.8}$$

Since

$$-(\partial_t - \Delta)W \leq -2W\big(|\mathring{h}|^2 + \tfrac{1}{an-1}|\hat{\mathbf{A}}|^2 + \tfrac{1}{n}H^2 - (n-b)K\big) - 2\gamma|\nabla\mathbf{A}|^2,$$

where

$$2\gamma = 1 - \tfrac{n+2}{3}a > 0,$$

we obtain

$$\begin{aligned}\frac{(\partial_t - \Delta)\frac{g}{W}}{\frac{g}{W}} ={}& \frac{(\partial_t - \Delta)g}{g} - \frac{(\partial_t - \Delta)W}{W} + 2\left\langle\nabla\log\frac{g}{W}, \nabla\log W\right\rangle\\ \leq{}& -2bK - 2\gamma\frac{|\nabla\mathbf{A}|^2}{W} + 2\left\langle\nabla\log\frac{g}{W}, \nabla\log W\right\rangle\end{aligned}$$

wherever $\mathbf{H} \neq 0$.

On the other hand, wherever $\mathbf{H} = 0$,

$$-(\partial_t - \Delta)W \leq -3|\mathring{\mathbf{A}}|^2W$$

and

$$(\partial_t - \Delta)g \leq \big(3|\mathring{\mathbf{A}}|^2 - 2K\big)g, \tag{3.9}$$

and hence, at such points,

$$\frac{(\partial_t - \Delta)\frac{g}{W}}{\frac{g}{W}} \leq \; -2K + 2\left\langle \nabla \log \frac{g}{W}, \nabla \log W \right\rangle .$$

Since $b \geq 1$ we conclude that

$$\frac{(\partial_t - \Delta)\frac{g}{W}}{\frac{g}{W}} \leq \; -2K + 2\left\langle \nabla \log \frac{g}{W}, \nabla \log W \right\rangle$$

everywhere. The maximum principle now implies the claim. □

Proposition 3.2 implies, in particular, that there exists $\tau = \tau(n, m - \alpha)$ such that the inequality

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}H^2 - 2K < 0$$

holds for each $t \in (0, T) \cap [\tau K^{-1}, \infty)$ on any solution initially satisfying (3.3). If $T > \tau K^{-1}$, this means that at time τK^{-1} the solution satisfies the hypotheses of [6, Main Theorem 7]. Consequently, the solution either exists forever and converges to a totally geodesic submanifold as $t \to \infty$, or else contracts to a round point in finite time.

We also find that the only quadratically pinched ancient solutions are the totally umbilic ones (cf. [2, 11, 17, 19]).

Theorem 3.3. *Let $X : M \times (-\infty, 0) \to S_K^{n+\ell}$ be a proper ancient solution to mean curvature flow. If*

$$\limsup_{t \to -\infty} \left(|\mathbf{A}|^2 - \tfrac{4}{3n}|\mathbf{H}|^2 - \tfrac{n}{2}K\right) < 0,$$

then $X(M, t)$ is totally umbilic for each t, and hence, up to a rotation, the solution is either a stationary hyperequator or a shrinking hyperparallel in $S_K^{n+1} \hookrightarrow S_K^{n+\ell}$.

3.2. **The codimension estimate.** Next, we obtain an estimate for $|\hat{\mathbf{A}}|$ which improves at high curvature scales for submanifolds satisfying (1.2). Such an estimate was obtained for high codimension mean curvature flow in Euclidean space by Naff [21]. Two new difficulties need to be overcome in our setting, however: the first is the fact that our pinching condition allows the mean curvature vector to vanish at some points (at which $\hat{\mathbf{A}}$ is not defined); the second is the presence of ambient curvature terms, which need to be controlled.

In fact, the estimate holds under the condition (3.3), so long as

$$m - \alpha < \min\left\{\frac{n}{4}, \frac{n(n-1)}{3(n+1)}\right\},$$

which when $m = 2$ is implied by (1.2). Note that this corresponds exactly to the constants in Naff's estimate in the Euclidean setting [21].

Proposition 3.4 (Codimension estimate (cf. [21]))**.** *Let $X : M^n \times [0, T) \to S_K^{n+\ell}$ be a solution to mean curvature flow such that* (3.3) *holds on $M^n \times \{0\}$ for some $m \geq 1$ and $\alpha \in [0, 1)$. If $m - \alpha < \min\left\{\frac{n}{4}, \frac{n(n-1)}{3(n+1)}\right\}$, then*

$$|\hat{\mathbf{A}}|^2 \leq CK^{\delta}(|\mathbf{H}|^2 + K)^{1-\delta} \quad \textit{wherever } \mathbf{H} \neq 0,$$

where $\delta > 0$ depends only on n and $m-\alpha$, and $C < \infty$ depends only on n, $m-\alpha$, and an upper bound for $K^{-1}\max_{M\times\{0\}}|\mathbf{A}|^2$.

Proof. Let $\tau = \tau(n, m-\alpha)$ be a positive number such that

$$m - \alpha + \tau < \min\left\{\frac{n}{4}, \frac{n(n-1)}{3(n+1)}\right\}.$$

Given $\sigma \in (0,1)$, set

$$f_\sigma \doteqdot \begin{cases} \dfrac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W} W^\sigma & \text{if } \ H > 0 \\ 0 & \text{if } \ H = 0\,, \end{cases}$$

where

$$W \doteqdot \frac{1}{2}\big(bK + aH^2 - |\mathbf{A}|^2\big)$$

with

$$a = \frac{1}{n - m + \alpha - \tau} \ \text{ and } \ b = 2(m - \alpha + \tau)\,.$$

Observe that

$$2W \geq \varepsilon b(H^2 + K)\,,$$

where $\varepsilon = \varepsilon(n, m-\alpha) > 0$. Indeed, since $(m-\alpha)$-pinching is preserved when $m - \alpha < \frac{n}{4}$,

$$\begin{aligned} -2W = &- \left(\frac{1}{n-m+\alpha-\tau} - \frac{1}{n-m+\alpha}\right) H^2 - 2\tau K + |\mathbf{A}|^2 - \tfrac{1}{n-m+\alpha} H^2 - 2(m-\alpha)K \\ \leq &- \left(\frac{1}{n-m+\alpha-\tau} - \frac{1}{n-m+\alpha}\right) H^2 - 2\tau K. \end{aligned}$$

Moreover, with this choice of a we have

$$a < \min\left\{\frac{4}{3n}, \frac{3(n+1)}{2n(n+2)}\right\}. \tag{3.10}$$

Recall that

$$\begin{aligned} (\nabla_t - \Delta)\frac{1}{2}|\hat{\mathbf{A}}|^2 = {} & |\langle \hat{\mathbf{A}}, \hat{\mathbf{A}}\rangle^\top|^2 + |h \wedge \hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}} \wedge \hat{\mathbf{A}}|^2 - nK|\hat{\mathbf{A}}|^2 \\ & - |\nabla \hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N} \otimes \nabla \frac{h}{H}, \nabla \hat{\mathbf{A}}\Big) \end{aligned}$$

and

$$\begin{aligned} -(\partial_t - \Delta)W = {} & |\langle \mathbf{A}, \mathbf{A}\rangle^\top|^2 + |\mathbf{A} \wedge \mathbf{A}|^2 - a|\mathbf{L}(\,\cdot\,, \mathbf{H})|^2 + nK\big(|\mathbf{A}|^2 - a|\mathbf{H}|^2\big) - 2nK|\mathring{\mathbf{A}}|^2 \\ & - \big(|\nabla \mathbf{A}|^2 - a|\nabla^\perp \mathbf{H}|^2\big) \end{aligned}$$

wherever $\mathbf{H} \neq 0$. We first compare the reaction terms in these two evolution equations.

Claim 3.5. There exists $\theta = \theta(n, m-\alpha) < 1$ such that

$$\frac{1}{\frac{1}{2}|\hat{\mathbf{A}}|^2}\Big(|\big\langle\hat{\mathbf{A}},\hat{\mathbf{A}}\big\rangle^\top|^2 + |h\wedge\hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}}\wedge\hat{\mathbf{A}}|^2 - nK|\hat{\mathbf{A}}|^2\Big)$$

$$\leq -\frac{\theta}{W}\Big(|\langle\mathbf{A},\mathbf{A}\rangle^\top|^2 + |\mathbf{A}\wedge\mathbf{A}|^2 - a|\mathbf{L}(\,\cdot\,,\mathbf{H})|^2 + nK\big(|\mathbf{A}|^2 - a|\mathbf{H}|^2\big) - 2nK|\mathring{\mathbf{A}}|^2\Big). \tag{3.11}$$

Proof of Claim 3.5. On the one hand, we have seen (recall (3.6)) that

$$\begin{aligned}-R_W &\doteqdot |\langle\mathbf{A},\mathbf{A}\rangle^\top|^2 + |\mathbf{A}\wedge\mathbf{A}|^2 - a|\mathbf{L}(\,\cdot\,,\mathbf{H})|^2 + nK\big(|\mathbf{A}|^2 - a|\mathbf{H}|^2\big) - 2nK|\mathring{\mathbf{A}}|^2\\ &\leq\; -2W\big(|\mathring{h}|^2 + \tfrac{1}{an-1}|\hat{\mathbf{A}}|^2 + \tfrac{1}{n}H^2 - (n-b)K\big).\end{aligned}$$

Replacing bK using

$$\big(1+\tfrac{\varepsilon}{2}\big)bK = \big(1+\tfrac{\varepsilon}{2}\big)\big(|\mathring{h}|^2 + |\hat{\mathbf{A}}|^2 + 2W - \tfrac{an-1}{n}H^2\big),$$

we obtain

$$\begin{aligned}R_W \geq 2W&\Big(\big(2+\tfrac{\varepsilon}{2}\big)|\mathring{h}|^2 + \big(1+\tfrac{\varepsilon}{2}+\tfrac{1}{an-1}\big)|\hat{\mathbf{A}}|^2 + 2\big(1+\tfrac{\varepsilon}{2}\big)W\Big)\\ &+ 2W\Big(\,\big(\tfrac{2-an}{n} - \tfrac{\varepsilon}{2}\tfrac{an-1}{n}\big)\,H^2 - \big(n+\tfrac{\varepsilon b}{2}\big)K\Big).\end{aligned}$$

The term $\frac{2-an}{n}H^2$ can be discarded since $2-an>0$.

Since $2W \geq \varepsilon b(H^2+K)$, we may estimate

$$-4W + \varepsilon bK \leq -\varepsilon b\big(2H^2 + K\big),$$

and hence

$$\begin{aligned}R_W &\geq W\Big(\big(4+\varepsilon\big)|\mathring{h}|^2 + \big(2+\varepsilon+\tfrac{2}{an-1}\big)|\hat{\mathbf{A}}|^2 + (\varepsilon b - 2n)K + 2\varepsilon W + \varepsilon(2b - \tfrac{an-1}{n})H^2\Big)\\ &\geq W\Big(\big(4+\varepsilon\big)|\mathring{h}|^2 + \big(3+\varepsilon\big)|\hat{\mathbf{A}}|^2 + (\varepsilon - 2n)K\Big).\end{aligned}$$

On the other hand, by (3.4) and (3.5),

$$\begin{aligned}R_{\frac{1}{2}|\hat{\mathbf{A}}|^2} &\doteqdot |\big\langle\hat{\mathbf{A}},\hat{\mathbf{A}}\big\rangle^\top|^2 + |h\wedge\hat{\mathbf{A}}|^2 + |\hat{\mathbf{A}}\wedge\hat{\mathbf{A}}|^2 - nK|\hat{\mathbf{A}}|^2\\ &\leq \big(4|\mathring{h}|^2 + 3|\hat{\mathbf{A}}|^2 - 2nK\big)\tfrac{1}{2}|\hat{\mathbf{A}}|^2.\end{aligned}$$

The claim follows. □

We estimate the first order terms as follows.

Claim 3.6. There exist $\theta = \theta(n, m-\alpha) < 1$ and $\Lambda = \Lambda(n, m-\alpha) < \infty$ such that

$$-|\nabla\hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big) \leq \theta\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\big(|\nabla\mathbf{A}|^2 - a|\nabla^\perp\mathbf{H}|^2\big) \tag{3.12}$$

so long as $H^2 \geq \Lambda K$.

Proof of Claim 3.6. Decomposing $\nabla\mathbf{A}$ and $\nabla\mathbf{H}$ into their irreducible components and applying the Kato inequalities (2.20) and (2.21) yields

$$\begin{aligned}|\nabla\mathbf{A}|^2-a|\nabla\mathbf{H}|^2 =&\ -\tfrac{an-1}{n}\big(H^2|\hat{\mathbf{T}}|^2+|\nabla H|^2\big)+|\hat{\mathbf{T}}\otimes\mathring{h}+\hat\nabla\hat{\mathbf{A}}|^2+|\mathrm{T}(\cdot,\hat{\mathbf{A}})+\nabla^\top\mathring{h}|^2\\ \geq&\ \big(\tfrac{2(n-1)}{n(n+2)}(1-s_2)-\tfrac{an-1}{n}\big)H^2|\hat{\mathbf{T}}|^2+\big(\tfrac{2(n-1)}{n(n+2)}(1-s_1)-\tfrac{an-1}{n}\big)|\nabla H|^2\\ &+s_1|\mathrm{T}(\cdot,\hat{\mathbf{A}})+\nabla^\top\mathring{h}|^2+s_2|\hat{\mathbf{T}}\otimes\mathring{h}+\hat\nabla\hat{\mathbf{A}}|^2\end{aligned}$$

for any $s_1,s_2\in[0,1]$. Choosing $s_2=0$ and $s_1=1-\frac{(n+2)(an-1)}{2(n-1)}$, so that

$$\tfrac{2(n-1)}{n(n+2)}(1-s_1)-\tfrac{an-1}{n}=0,$$

yields

$$|\nabla\mathbf{A}|^2-a|\nabla\mathbf{H}|^2\geq\alpha_1|\mathrm{T}(\cdot,\hat{\mathbf{A}})+\nabla^\top\mathring{h}|^2+\frac{\alpha_2}{n}H^2|\hat{\mathbf{T}}|^2, \tag{3.13}$$

where

$$\alpha_1\doteqdot 1-\tfrac{(n+2)(an-1)}{2(n-1)}\quad\text{and}\quad\alpha_2\doteqdot\tfrac{2(n-1)}{n+2}-(an-1).$$

The two components of $\nabla\mathbf{A}$ on the right of (3.13) will be sufficient to control the gradient terms in the evolution equation for $\hat{\mathbf{A}}$ (which is why we kept as much of them as possible, at the expense of the others).

So consider

$$\begin{aligned}-2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big)&=2H\,\mathrm{tr}_{g^\top}\,g\Big(\nabla^\perp_\cdot\mathbf{N}\otimes\nabla_\cdot\frac{\mathring{h}}{H},\hat{\mathbf{A}}\Big)\\ &\leq 2|\hat{\mathbf{T}}||\hat{\mathbf{A}}|\left|\nabla\mathring{h}-\frac{\nabla H}{H}\otimes\mathring{h}\right|\\ &=2|\hat{\mathbf{T}}||\hat{\mathbf{A}}|\left|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})-\mathrm{T}(\cdot,\hat{\mathbf{A}})-\frac{\nabla H}{H}\otimes\mathring{h}\right|\\ &\leq 2|\hat{\mathbf{T}}||\hat{\mathbf{A}}|\left(|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|+|\mathrm{T}(\cdot,\hat{\mathbf{A}})|+|\nabla H|\frac{|\mathring{h}|}{H}\right).\end{aligned}$$

Let $\mathring{\chi}$ be the indicator function on the support of $|\mathring{h}|^2$. Then, applying Young's inequality three times yields, for any $\gamma_1,\gamma_2>0$,

$$\begin{aligned}-2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big)\leq&\ (1+\gamma_1^{-1}+\gamma_2^{-1}\mathring{\chi})|\hat{\mathbf{T}}|^2|\hat{\mathbf{A}}|^2\\ &+\gamma_1|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2+|\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2+\gamma_2\mathring{\chi}\frac{|\mathring{h}|^2}{H^2}|\nabla H|^2.\end{aligned}$$

On the other hand, by Young's inequality and the Kato inequality (2.21),

$$\begin{aligned}|\hat\nabla\hat{\mathbf{A}}|^2+|\mathring{h}|^2|\hat{\mathbf{T}}|^2&\geq\tfrac12|\hat\nabla\hat{\mathbf{A}}+\hat{\mathbf{T}}\otimes\mathring{h}|^2\\ &\geq\frac{n-1}{n(n+2)}H^2|\hat{\mathbf{T}}|^2.\end{aligned}$$

Combining these yields

$$-|\nabla\hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big) \leq \left((1+\gamma_1^{-1}+\gamma_2^{-1}\mathring{\chi})\frac{|\hat{\mathbf{A}}|^2}{H^2}+\frac{|\mathring{h}|^2}{H^2}-\frac{n-1}{n(n+2)}\right)H^2|\hat{\mathbf{T}}|^2 + \gamma_1|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2+\gamma_2\mathring{\chi}\frac{|\mathring{h}|^2}{H^2}|\nabla H|^2.$$

Appealing to (2.20), we get

$$-|\nabla\hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big) \leq \left((1+\gamma_1^{-1}+\gamma_2^{-1}\mathring{\chi})\frac{|\hat{\mathbf{A}}|^2}{H^2}+\frac{|\mathring{h}|^2}{H^2}-\frac{n-1}{n(n+2)}\right)H^2|\hat{\mathbf{T}}|^2 + \left(\gamma_1+\gamma_2\mathring{\chi}\frac{n(n+2)}{2(n-1)}\frac{|\mathring{h}|^2}{H^2}\right)|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2.$$

Now set

$$\gamma_1=\beta_1\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\quad\text{and}\quad\gamma_2=\beta_2\frac{2(n-1)}{n(n+2)}\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\frac{H^2}{|\mathring{h}|^2}$$

at points such that $|\mathring{h}|^2>0$, where β_1 and β_2 are to be chosen. If $|\mathring{h}|^2=0$ then let γ_1 and β_1 be as before, and set $\gamma_2=\beta_2=1$. This yields

$$\left(\gamma_1+\gamma_2\mathring{\chi}\frac{n(n+2)}{2(n-1)}\frac{|\mathring{h}|^2}{H^2}\right)|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2\leq(\beta_1+\beta_2\mathring{\chi})\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2,$$

and consequently

$$\begin{aligned}-|\nabla\hat{\mathbf{A}}|^2 &- 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H},\nabla\hat{\mathbf{A}}\Big)\\ &\leq\left(\frac{|\hat{\mathbf{A}}|^2}{H^2}+\frac{|\mathring{h}|^2}{H^2}+\beta_1^{-1}\frac{2W}{H^2}+\frac{n(n+2)}{2(n-1)}\beta_2^{-1}\mathring{\chi}\frac{2W|\mathring{h}|^2}{H^4}-\frac{n-1}{n(n+2)}\right)H^2|\hat{\mathbf{T}}|^2\\ &\quad+\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\left((\beta_1+\beta_2\mathring{\chi})|\nabla\mathring{h}+\mathrm{T}(\cdot,\hat{\mathbf{A}})|^2\right).\end{aligned}\tag{3.14}$$

The terms on the second line can be controlled using the first good term in (3.13) provided $\beta_1+\beta_2\mathring{\chi}<\alpha_1$; so let β_1 and β_2 satisfy this inequality. To handle the terms on the first line we use

$$2W=bK+\tfrac{an-1}{n}H^2-|\hat{\mathbf{A}}|^2-|\mathring{h}|^2\leq bK+\tfrac{an-1}{n}H^2$$

to estimate

$$\begin{aligned}
&\frac{|\hat{\mathbf{A}}|^2}{H^2} + \frac{|\mathring{h}|^2}{H^2} + \beta_1^{-1}\frac{2W}{H^2} + \frac{n(n+2)}{2(n-1)}\beta_2^{-1}\mathring{\chi}\frac{2W|\mathring{h}|^2}{H^4} \\
&\quad = (1-\beta_1^{-1})\frac{|\hat{\mathbf{A}}|^2}{H^2} + (1-\beta_1^{-1})\frac{|\mathring{h}|^2}{H^2} + \beta_1^{-1}\left(\frac{bK}{H^2} + \frac{an-1}{n}\right) + \frac{n(n+2)}{2(n-1)}\beta_2^{-1}\mathring{\chi}\frac{2W|\mathring{h}|^2}{H^4} \\
&\quad \leq (1-\beta_1^{-1})\frac{|\hat{\mathbf{A}}|^2}{H^2} + (1-\beta_1^{-1})\frac{|\mathring{h}|^2}{H^2} + \beta_1^{-1}\Big(\frac{bK}{H^2} + \frac{an-1}{n}\Big) + \frac{n(n+2)}{2(n-1)}\beta_2^{-1}\mathring{\chi}\Big(\frac{bK}{H^2} + \frac{an-1}{n}\Big)\frac{|\mathring{h}|^2}{H^2} \\
&\quad = (1-\beta_1^{-1})\frac{|\hat{\mathbf{A}}|^2}{H^2} + \Big(1-\beta_1^{-1} + \beta_2^{-1}\mathring{\chi}\frac{(n+2)}{2(n-1)}\Big[\frac{nbK}{H^2} + an - 1\Big]\Big)\frac{|\mathring{h}|^2}{H^2} + \beta_1^{-1}\Big(\frac{bK}{H^2} + \frac{an-1}{n}\Big).
\end{aligned}$$

Since $\beta_1 < \alpha_1 < 1$, the first of these terms is non-positive. We will choose β_1 and β_2 so that the inequalities

$$1 - \beta_1^{-1} + \beta_2^{-1}\mathring{\chi}\frac{(n+2)}{2(n-1)}\left[\frac{nbK}{H^2} + an - 1\right] < 0 \tag{3.15}$$

and

$$\beta_1^{-1}\left(\frac{bK}{H^2} + \frac{an-1}{n}\right) - \frac{n-1}{n(n+2)} < 0 \tag{3.16}$$

hold whenever H^2 is sufficiently large compared to K. Together, these inequalities imply the first term on the right-hand side of (3.14) is non-positive.

To simplify notation, define $w \doteqdot \frac{(n+2)(an-1)}{2(n-1)}$. Let us set $\beta_1 = 2w + \delta$, where $\delta > 0$ is a small constant depending on n and $m - \alpha$ which we progressively refine. Note that, for all $n \geq 5$, we can indeed choose δ so that $\beta_1 < \alpha_1$. We compute

$$\beta_1^{-1}\left(\frac{bK}{H^2} + \frac{an-1}{n}\right) - \frac{n-1}{n(n+2)} = \frac{\beta_1^{-1}(n-1)}{n(n+2)}\left(\frac{n(n+2)}{n-1}\frac{bK}{H^2} - \delta\right),$$

hence there is a constant $C_0 = C_0(n)$ such that the right-hand side is negative (meaning (3.16) holds) whenever

$$\frac{H^2}{bK} \geq C_0\delta^{-1}.$$

Now we turn to (3.15). If $\mathring{\chi} = 0$ we are done, so assume we are at a point where $\mathring{\chi} = 1$. Observe that $w = 1 - \alpha_1$. Hence the inequality $\beta_1 + \beta_2 < \alpha_1$ holds if

$$\beta_2 < \alpha_1 - \beta_1 = 1 - 3w - \delta.$$

For this reason we set $\beta_2 = 1 - 3w - 2\delta$ (note this quantity is strictly positive for all $n \geq 5$ and δ suffciently small relative to n). We compute

$$\begin{aligned}
1 - \beta_1^{-1} + \beta_2^{-1}\frac{(n+2)(an-1)}{2(n-1)} &= \beta_1^{-1}\beta_2^{-1}(\beta_1\beta_2 - \beta_2 + \beta_1 w) \\
&= \beta_1^{-1}\beta_2^{-1}((4w-1)(1-w) + \delta(3 - 6w - 2\delta)).
\end{aligned}$$

Now we make use of the pinching assumption to estimate the right-hand side. In dimensions $n \geq 8$, the inequality $a \leq \frac{4}{3n}$ ensures $w < \frac{1}{4}$; in dimensions $5 \leq n \leq 7$, we need $a < \frac{3(n+1)}{2n(n+2)}$

to ensure $w < \frac{1}{4}$. Each of these conditions is ensured by (3.10), so we may indeed write $w = \frac{1}{4} - \eta$, where η is a small positive number depending on n and $m - \alpha$. Choosing δ sufficiently small relative to η and n then gives

$$1 - \beta_1^{-1} + \beta_2^{-1}\frac{(n+2)(an-1)}{2(n-1)} \leq -\beta_1^{-1}\beta_2^{-1}\alpha_1\eta,$$

and consequently,

$$1 - \beta_1^{-1} + \beta_2^{-1}\frac{(n+2)}{2(n-1)}\left[\frac{nbK}{H^2} + an - 1\right] \leq \beta_2^{-1}\left[\frac{n(n+2)}{2(n-1)}\frac{bK}{H^2} - \beta_1^{-1}\alpha_1\eta\right].$$

In particular, there is a constant $C_1 = C_1(n)$ such that, as long as δ is small enough relative to n, (3.15) holds at every point where

$$\frac{H^2}{bK} \geq C_1\eta^{-1}.$$

To recap, there are constants $\delta \in (0,1)$ and $\eta \in (0, \frac{1}{4})$ depending on n and $m - \alpha$ such that, at points where

$$\frac{H^2}{bK} \geq \Lambda \doteqdot \max\{C_0\delta^{-1}, C_1\eta^{-1}\}, \tag{3.17}$$

the inequalities (3.15) and (3.16) hold. Combining these inequalities with (3.14), we see that

$$-|\nabla\hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H}, \nabla\hat{\mathbf{A}}\Big) \leq \frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\Big((\beta_1+\beta_2)|\nabla\mathring{h} + \mathrm{T}(\cdot,\hat{\mathbf{A}})|^2\Big)$$

whenever (3.17) holds. Inserting $\beta_1 = 2w + \delta$ and $\beta_2 = 1 - 3w - 2\delta$, and combining this estimate with (3.13), we arrive at

$$-|\nabla\hat{\mathbf{A}}|^2 - 2Hg\Big(\mathbf{N}\otimes\nabla\frac{h}{H}, \nabla\hat{\mathbf{A}}\Big) \leq \frac{\alpha_1 - \delta}{\alpha_1}\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}\big(|\nabla\mathbf{A}|^2 - a|\nabla^\perp\mathbf{H}|^2\big).$$

Hence the claim holds with $\theta \doteqdot 1 - \alpha_1^{-1}\delta$. □

Since, at points where $H > 0$,

$$\frac{(\partial_t - \Delta)f_\sigma}{f_\sigma} = \frac{(\partial_t-\Delta)|\hat{\mathbf{A}}|^2}{|\hat{\mathbf{A}}|^2} - (1-\sigma)\frac{(\partial_t-\Delta)W}{W} + 2(1-\sigma)\left\langle\frac{\nabla f_\sigma}{f_\sigma}, \frac{\nabla W}{W}\right\rangle$$
$$- \sigma(1-\sigma)\frac{|\nabla W|^2}{W^2},$$

Claims 3.5 and 3.6 imply that, for $\sigma = \sigma(n, m-\alpha)$ sufficiently small,

$$\frac{(\partial_t - \Delta)f_\sigma}{f_\sigma} \leq 2(1-\sigma)\left\langle\frac{\nabla f_\sigma}{f_\sigma}, \frac{\nabla W}{W}\right\rangle \tag{3.18}$$

wherever $H^2 \geq \Lambda K$. Now, since $2W \geq \varepsilon b(H^2+K)$, where $\varepsilon = \varepsilon(n, m-\alpha)$, we can estimate

$$\frac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W} \leq \frac{\frac{1}{2}(|\hat{\mathbf{A}}|^2 + |\mathring{h}|^2 + \frac{1}{n}H^2)}{W} = \frac{aH^2 + bK - 2W}{2W} \leq C\,,$$

where $C = C(n, m-\alpha)$. This implies that H^2 is large compared to K at any point where f_σ is large compared to K^σ. Indeed, if

$$f_\sigma \geq C\,(a\Lambda + b)\,K^\sigma\,,$$

then, since $\sigma < 1$,

$$H^2 \geq \Lambda K\,.$$

We may now conclude from (3.18) that

$$f_\sigma \leq \max\left\{\max_{M\times\{0\}} f_\sigma, C\,(a\Lambda + b)\,K^\sigma\right\}.$$

The proposition follows. □

3.3. **The cylindrical estimate.** Next, we use the codimension estimate (Proposition 3.4) and the Poincaré-type inequality (Proposition 2.3) to obtain a cylindrical estimate. The argument follows the Stampacchia iteration procedure developed by Huisken [13].

Given $n \geq 5$ and $\ell \geq 2$, it will be convenient to state the estimate in terms of the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ of n-dimensional submanifolds of $S_K^{n+\ell}$ satisfying

- $|\mathbf{A}|^2 - \frac{1}{n-2+\alpha}|\mathbf{H}|^2 - 2(2-\alpha)K$, (with $\alpha \geq \alpha_n$, where α_n is defined by (3.2)),
- $\mu(M) \leq VK^{-\frac{n}{2}}$, and
- $\max_M |\mathbf{A}|^2 \leq \Theta K$.

Of course, every n-dimensional submanifold of $S_K^{n+\ell}$ satisfying (1.2) lies in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ for some $\alpha \geq \alpha_n$, $V < \infty$, and $\Theta < \infty$.

Proposition 3.7 (Cylindrical estimate)**.** *Let $X : M \times [0,T) \to S_K^{n+\ell}$ be a solution to mean curvature flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ and set $\eta_0 \doteqdot \frac{1}{n-2+\alpha_n} - \frac{1}{n-1}$. For every $\eta \in (0, \eta_0)$ there exists $C_\eta = C_\eta(n, \alpha, V, \Theta, \eta) < \infty$ so that*

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2 \leq \eta|\mathbf{H}|^2 + C_\eta K\,\mathrm{e}^{-2Kt} \quad in \quad M^n \times [0,T)\,. \tag{3.19}$$

Proof. Given $\eta > 0$ and $\sigma \in (0,1)$, consider the function

$$g_{\eta,\sigma} \doteqdot \frac{1}{2}\big[|\mathbf{A}|^2 - \big(\tfrac{1}{n-1} + \eta\big)|\mathbf{H}|^2\big]W^{\sigma-1},$$

where

$$W \doteqdot \frac{1}{2}\big(bK + a|\mathbf{H}|^2 - |\mathbf{A}|^2\big)$$

for

$$a = \tfrac{1}{n-2+\alpha_n} \;\text{ and }\; b = 2(2-\alpha_n).$$

Computing as in the proof of Proposition 3.2, we obtain, for $\eta \le \eta_0 \doteqdot \frac{1}{n-2+\alpha_n} - \frac{1}{n-1}$,

$$\frac{(\partial_t - \Delta)g_{\eta,\sigma}}{g_{\eta,\sigma}} \le -2K + \sigma n|\mathbf{A}|^2 + 2(1-\sigma)\left\langle \frac{\nabla g_{\eta,\sigma}}{g_{\eta,\sigma}}, \frac{\nabla W}{W}\right\rangle . \tag{3.20}$$

Now consider, for $k \ge 0$ and $p \ge 3$,

$$v_k \doteqdot (\mathrm{e}^{2Kt}g_{\eta,\sigma} - k)_+^{\frac{p}{2}} \text{ and } V_k(t) \doteqdot \operatorname{spt} v_k(\cdot, t).$$

By (3.20),

$$\begin{aligned}\frac{(\partial_t - \Delta)v_k^2}{v_k^2} &= p\mathrm{e}^{2Kt}\frac{(\partial_t - \Delta)g_{\eta,\sigma} + 2Kg_{\eta,\sigma}}{(\mathrm{e}^{2Kt}g_{\eta,\sigma} - k)_+} - 4\big(1 - \tfrac{1}{p}\big)\frac{|\nabla v_k|^2}{v_k^2}\\ &\le \left(\sigma np|\mathbf{A}|^2 - 2pK - 2\gamma p\frac{|\nabla \mathbf{A}|^2}{W}\right)\frac{\mathrm{e}^{2Kt}g_{\eta,\sigma}}{(\mathrm{e}^{2Kt}g_{\eta,\sigma} - kK^\sigma)_+} + 4\frac{|\nabla v_k|}{v_k}\frac{|\nabla W|}{W}\\ &\quad - 4\big(1 - \tfrac{1}{p}\big)\frac{|\nabla v_k|^2}{v_k^2}\end{aligned}$$

and hence, for p sufficiently large,

$$(\partial_t - \Delta)v_k^2 \le -2pKv_k^2 + \sigma np\mathrm{e}^{2Kt}g_{\eta,\sigma,+}(\mathrm{e}^{2Kt}g_{\eta,\sigma,+} - k)_+^{p-1}|\mathbf{A}|^2 - \gamma pv_k^2\frac{|\nabla \mathbf{A}|^2}{W} - 2|\nabla v_k|^2 .$$

It follows that

$$\begin{aligned}\frac{d}{dt}\int v_k^2\,d\mu + 2\int |\nabla v_k|^2\,d\mu + \int |\mathbf{H}|^2 v_k^2\,d\mu \le{}& -2pK\int v_k^2 d\mu - \gamma p\int v_k^2\frac{|\nabla \mathbf{A}|^2}{W}\,d\mu\\ &+ c\sigma p\int_{V_k}(\mathrm{e}^{2Kt}g_{\eta,\sigma})^p W\,d\mu ,\end{aligned} \tag{3.21}$$

where $c = c(n)$. In particular, when $k = 0$, the Poincaré inequality (Proposition 2.3) yields

$$\frac{d}{dt}\int v_0^2\,d\mu \le -2pK\int v_0^2\,d\mu \tag{3.22}$$

and hence

$$\int v_0^2\,d\mu \le C\mathrm{e}^{-2pKt}, \tag{3.23}$$

so long as $k \ge k_0$, $p \ge \ell^{-1}$ and $\sigma \le \ell p^{-\frac{1}{2}}$, where $\ell = \ell(n, \alpha, \eta)$ and $C = C(n, K, \alpha, V, \Theta, \sigma, p)$.

The L^2-estimate (3.23) can be bootstrapped to an L^∞-estimate using (3.21) by applying Huisken–Stampacchia iteration. Indeed, using (3.23) we may estimate, for $p \ge \ell$, $\sigma \le \ell^{-1}p^{-\frac{1}{2}}$,

$$V_k \le k^{-p}\int_{V_k}(\mathrm{e}^{2Kt}g_{\eta,\sigma})^p d\mu \le Ck^{-p},$$

where $C = C(n, K, \alpha, V, \Theta, \sigma, p)$, so that

$$V_k \le \frac{\omega_n}{n+1}K^{-\frac{n}{2}}$$

so long as $k \geq k_0 = k_0(n, K, \alpha, V, \Theta, \sigma, p)$. This enables us to apply the Sobolev inequality [12] (cf. [20]), which yields, at each time,

$$\frac{1}{c_S}\left(\int_{V_k} v_k^{2^*}\,d\mu\right)^{\frac{1}{2^*}} \leq \int_{V_k}\big(|\nabla v_k|^2 + v_k^2|\mathbf{H}|^2\big)\,d\mu\,,$$

where $2^* \doteqdot \frac{2n}{n-2}$ and c_S depends only on n, and consequently

$$\frac{d}{dt}\int v_k^2\,d\mu + \frac{1}{c_S}\left(\int_{V_k} v_k^{2^*}\,d\mu\right)^{\frac{1}{2^*}} \leq c\sigma p\int_{V_k}(\mathrm{e}^{2Kt}g_{\eta,\sigma})^p\big(|\mathbf{H}|^2 + K\big)\,d\mu$$

so long as $p \geq \ell$, $\sigma \leq \ell^{-1}p^{-\frac{1}{2}}$, and $k \geq k_0$. Assuming further that $k_0 \geq \max g_{\eta,\sigma}(\cdot, 0)$ (which can be achieved with dependence on n, K, α, V, Θ, σ and p only), integrating now gives

(3.24)

$$\sup_{t\in[0,T)}\int_M v_k^2(\cdot,t)\,d\mu + \int_0^T\left(\int_M v_k^{2q}\,d\mu\right)^{\frac{1}{q}}dt \leq C\sigma p\int_0^T\!\!\int_{V_k}(\mathrm{e}^{2Kt}g_{\eta,\sigma})^p\big(|\mathbf{H}|^2 + K\big)\,d\mu\,dt,$$

where $C \doteqdot c_S c$ depends only on n. Applying the Hölder, interpolation and Young inequalities, exactly as in the proof of [13, Theorem 5.1], now yields $C = C(n, \ell, \alpha, V, \Theta, \sigma, p) < \infty$ and $\gamma = \gamma(n) > 0$ such that

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

for all $h > k > k_0$, where

$$A(k) \doteqdot \int_0^T\!\!\int_{V_k} d\mu_t\,dt\,,$$

so long as $p \geq \ell^{-1}$ and $\sigma \leq \ell p^{-\frac{1}{2}}$, where $\ell = \ell(n, \alpha, \eta)$. Fixing $p \doteqdot \ell^{-1}$ and $\sigma \doteqdot \ell p^{-\frac{1}{2}}$, Stampacchia's lemma [26, Lemma 4.1] now yields

$$A(k_0 + d) = 0\,,$$

where

$$d^p \doteqdot 2^{p\gamma/(\gamma-1)}CA(k_0)^{\gamma-1}\,.$$

Estimating via (3.23)

$$A(k) \leq k_0^{-p}\int_0^T\!\!\int v_{k_0}^2\,d\mu\,dt \leq C(n, K, \alpha, V, \Theta, \eta)\,,$$

we conclude that

$$\mathrm{e}^{2Kt}g_{\eta,\sigma} \leq C(n, K, \alpha, V, \Theta, \eta)\,.$$

Young's inequality then yields

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2 \leq 2\eta|\mathbf{H}|^2 + C(n, K, \alpha, V, \Theta, \eta)\,\mathrm{e}^{-2Kt}\,.$$

The theorem now follows from the scaling covariance of the estimate. □

3.4. **The gradient estimate.** Next, we derive a suitable analogue of the "gradient estimate" [15, Theorem 6.1]. We need the following doubling estimates for solutions with initial data in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$.

Proposition 3.8. *Let $X : M \times [0,T) \to S_K^{n+\ell}$ be a maximal solution to mean curvature flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. Defining Λ_0 and λ_0 by*

$$\Lambda_0/2 \doteqdot \Theta \quad \text{and} \quad \mathrm{e}^{2n\lambda_0} \doteqdot 1 + \frac{n}{n+3\Lambda_0}, \tag{3.25}$$

we have

$$\mathrm{e}^{2nKT} \geq 1 + \frac{2n}{3\Lambda_0}\,, \tag{3.26}$$

and, for every $k \in \mathbb{N}$,

$$\max_{M\times\{\lambda_0 K^{-1}\}} |\nabla^k \mathbf{A}|^2 \leq \Lambda_k K^{k+1}\,, \tag{3.27}$$

where Λ_k depends only on n, k and Θ.

Proof. Since

$$\max_{M\times\{0\}} |\mathbf{A}|^2 \leq \Lambda_0 K/2\,,$$

a straightforward ODE comparison argument applied to the inequality

$$(\partial_t - \Delta)|\mathbf{A}|^2 \leq (3|\mathbf{A}|^2 + 2nK)|\mathbf{A}|^2$$

yields

$$\max_{M\times\{t\}} |\mathbf{A}|^2 \leq \frac{nK}{\left(3 + \frac{2n}{\Lambda_0}\right)\mathrm{e}^{-2nKt} - 3}\,.$$

We immediately obtain (3.26) and

$$|\mathbf{A}|^2(\cdot, t) \leq \Lambda_0 K \ \text{ for all } \ t \leq \lambda_0 K^{-1}\,. \tag{3.28}$$

The claim (3.27) now follows from the Bernstein estimates (Proposition 2.2). ☐

Modifying an argument of Huisken [13, Theorem 6.1] and Huisken–Sinestrari [15, Theorem 6.1], we can now obtain a pointwise estimate for the gradient of the second fundamental form which holds up to the singular time.

Proposition 3.9 (Gradient estimate (cf. [15, Theorem 6.1]))**.** *Let $X : M \times [0,T) \to S_K^{n+\ell}$, $n \geq 5$, be a solution to mean curvature flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. There exist $c = c(n, \ell, \alpha, \Theta) < \infty$, $\eta_0 = \eta_0(n) > 0$ and, for every $\eta \in (0, \eta_0)$, $C_\eta = C_\eta(n, \alpha, V, \Theta, \eta) < \infty$ such that*

$$\frac{|\nabla \mathbf{A}|^2}{|\mathbf{H}|^2 + K} + c\left(|\mathbf{A}|^2 - \frac{1}{n-1}|\mathbf{H}|^2\right) \leq \eta|\mathbf{H}|^2 + C_\eta K \mathrm{e}^{-2Kt} \tag{3.29}$$

in $M \times [\lambda_0 K^{-1}, T)$, where λ_0 is defined by (3.25).

Note that the conclusion is not vacuous since, by Proposition 3.8, the maximal existence time of a solution with initial data in the class $\mathcal{C}^n_K(\alpha, V, \Theta)$ is at least $\frac{1}{2nK}\log\left(1+\frac{2n}{3\Lambda_0}\right) > \lambda_0 K^{-1}$.

Setting $\eta = 1$, say, yields the cruder estimate

$$|\nabla \mathbf{A}|^2 \le C(|\mathbf{H}|^4 + K^2)\,, \tag{3.30}$$

where $C = C(n, \alpha, V, \Theta)$.

Proof of Proposition 3.9. We proceed as in [15, Theorem 6.1]. By (2.34),

$$(\partial_t - \Delta)|\nabla \mathbf{A}|^2 \le \, -2|\nabla^2 \mathbf{A}|^2 + c\left(|\mathbf{A}|^2 + K\right)|\nabla \mathbf{A}|^2\,.$$

We will control the bad term using the good term in the evolution equation for $|\mathbf{A}|^2$ and the Kato inequality (2.9).

By the cylindrical estimate, given any $\eta > 0$ we can find $C_\eta = C_\eta(n, \alpha, V, \Theta, \eta) > 2$ such that

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2 \le \eta|\mathbf{H}|^2 + C_\eta K \mathrm{e}^{-2Kt}\,,$$

and hence

$$G_\eta \doteqdot 2C_\eta K\mathrm{e}^{-2Kt} + \left(\eta + \tfrac{1}{n-1}\right)|\mathbf{H}|^2 - |\mathbf{A}|^2 \ge C_\eta K \mathrm{e}^{-2Kt} > 0\,.$$

By the initial pinching condition,

$$G_0 \doteqdot 4(2-\alpha_n)K + \tfrac{1}{n-2+\alpha_n}|\mathbf{H}|^2 - |\mathbf{A}|^2 \ge 2(2-\alpha_n)K > 0\,.$$

Arguing as in (3.8) and (3.9), but keeping part of the gradient terms using the Kato inequality, we obtain

$$\begin{aligned}(\partial_t - \Delta)G_\eta &\ge \, -2n|\mathbf{A}|^2(2C_\eta K\mathrm{e}^{-2Kt} - G_\eta) - 4C_\eta K^2\mathrm{e}^{-2Kt} + \beta|\nabla \mathbf{A}|^2\\ &\ge \, -2n(|\mathbf{A}|^2 + K)G_\eta + \beta|\nabla \mathbf{A}|^2,\end{aligned}$$

where $\beta \doteqdot \frac{3}{n+2} - \frac{1}{n-1} = \frac{2n-5}{(n+2)(n-1)}$, so long as $\eta \le \left(1 - \frac{3}{n+2}\right)\beta$.

Similarly,

$$(\partial_t - \Delta)G_0 \ge \, -2n(|\mathbf{A}|^2 + K)G_0\,.$$

We seek a bound for the ratio $\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}$. So suppose that $\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}$ attains a (parabolic) interior local maximum at (x_0, t_0). Then, at (x_0, t_0),

$$0 = \nabla_k \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0} = 2\frac{\langle \nabla_k \nabla \mathbf{A}, \nabla \mathbf{A}\rangle}{G_\eta G_0} - \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left(\frac{\nabla_k G_\eta}{G_\eta} + \frac{\nabla_k G_0}{G_0}\right)$$

and hence

$$4\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left\langle \frac{\nabla G_\eta}{G_\eta}, \frac{\nabla G_0}{G_0}\right\rangle \le \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left|\frac{\nabla G_\eta}{G_\eta} + \frac{\nabla G_0}{G_0}\right|^2 \le 4\frac{|\nabla^2 \mathbf{A}|^2}{G_\eta G_0}\,.$$

Furthermore, at (x_0, t_0),

$$
\begin{aligned}
0 \le (\partial_t - \Delta)\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0} &= \frac{(\partial_t - \Delta)|\nabla \mathbf{A}|^2}{G_\eta G_0} - \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left(\frac{(\partial_t - \Delta)G_\eta}{G_\eta} + \frac{(\partial_t - \Delta)G_0}{G_0}\right)\\
&\quad + \frac{2}{G_\eta G_0}\left\langle \nabla \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}, \nabla(G_\eta G_0)\right\rangle + 2\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left\langle \frac{\nabla G_\eta}{G_\eta}, \frac{\nabla G_0}{G_0}\right\rangle\\
&\le \frac{|\nabla \mathbf{A}|^2}{G_\eta G_0}\left((c+4n)(|\mathbf{A}|^2 + K) - \beta\frac{|\nabla \mathbf{A}|^2}{G_\eta}\right)
\end{aligned}
$$

and hence

$$
\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0} \le \frac{(c+4n)\beta^{-1}(|\mathbf{A}|^2 + K)}{4(2-\alpha_n)K + \frac{1}{n-2+\alpha_n}|\mathbf{H}|^2 - |\mathbf{A}|^2}.
$$

Since the solution is uniformly pinched,

$$
|\mathbf{A}|^2 \le \tfrac{1}{n-2+\alpha}|\mathbf{H}|^2 + 2(2-\alpha)K\,,
$$

we obtain, at (x_0, t_0),

$$
\frac{|\nabla \mathbf{A}|^2}{G_\eta G_0} \le C\,,
$$

where C depends only on n, k and α.

On the other hand, since $G_0 > C_0 K \doteqdot 2(2-\alpha_n)K$ and $G_\eta > C_\eta K \mathrm{e}^{-2Kt}$, if no interior local parabolic maxima are attained, then, by Proposition 3.8, we have for any $t \ge \lambda_0 K^{-1}$

$$
\begin{aligned}
\max_{\mathcal{M}\times\{t\}} \frac{|\nabla \mathbf{A}|^2}{G_0 G_\eta} &\le \max_{\mathcal{M}\times\{\lambda_0 K^{-1}\}} \frac{|\nabla \mathbf{A}|^2}{G_0 G_\eta}\\
&\le \max_{\mathcal{M}\times\{\lambda_0 K^{-1}\}} \frac{|\nabla \mathbf{A}|^2}{C_0 C_\eta K^2 \mathrm{e}^{-2\lambda_0}}\\
&\le \frac{\Lambda_1 \mathrm{e}^{2\lambda_0}}{C_0 C_\eta}\\
&\le \Lambda_1 \mathrm{e}^{2\lambda_0}\,.
\end{aligned}
$$

The theorem follows. □

3.5. **Higher order estimates.** The gradient estimate can be used to bound the first order terms which arise in the evolution equation for $\nabla^2\mathbf{A}$. A straightforward maximum principle argument exploiting this observation yields an analogous estimate for $\nabla^2\mathbf{A}$.

Proposition 3.10 (Hessian estimate (cf. [13,15]))**.** *Let $X : M \times [0,T) \to S_K^{n+1}$, $n \ge 5$, be a solution to mean curvature flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. There exists $C = C(n, \alpha, V, \Theta)$ such that*

$$
|\nabla^2\mathbf{A}|^2 \le C(|\mathbf{H}|^6 + K^3) \;\; \textit{in} \;\; M \times [\lambda_0 K^{-1}, T)\,. \tag{3.31}
$$

Proof. We proceed as in [15, Theorem 6.3]. Set $W \doteqdot |\mathbf{H}|^2 + K$. By (2.35),

$$(\partial_t - \Delta)|\nabla^2 \mathbf{A}|^2 \le c\left(W|\nabla^2\mathbf{A}|^2 + |\nabla \mathbf{A}|^4\right) - 2|\nabla^3\mathbf{A}|^2\,,$$

where c depends only on n and k. Recalling (2.28), we obtain

$$\begin{aligned}(\partial_t - \Delta)\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{5}{2}}} \le{}& \frac{c}{W^{\frac{5}{2}}}\left[W|\nabla^2\mathbf{A}|^2 + |\nabla\mathbf{A}|^4\right] - 2\frac{|\nabla^3\mathbf{A}|^2}{W^{\frac{5}{2}}}\\ &- 5\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}\left[|\mathbf{L}(\cdot,\mathbf{H})|^2 + nK|\mathbf{H}|^2 - |\nabla^\perp\mathbf{H}|^2\right]\\ &- \frac{35}{4}\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}\frac{|\nabla W|^2}{W} + \frac{5}{W^{\frac{7}{2}}}\left\langle\nabla|\nabla^2\mathbf{A}|^2, \nabla W\right\rangle.\end{aligned}$$

We can use the good third order term on the first line to absorb the ultimate term, since

$$\begin{aligned}\frac{5}{W^{\frac{7}{2}}}\left\langle\nabla|\nabla^2\mathbf{A}|^2, \nabla W\right\rangle \le{}& \frac{10}{W^{\frac{7}{2}}}|\nabla^3\mathbf{A}||\nabla^2\mathbf{A}||\nabla W|\\ \le{}& \frac{1}{W^{\frac{1}{2}}}\left(\frac{|\nabla^3\mathbf{A}|^2}{W^2} + 25\frac{|\nabla^2\mathbf{A}|^2|\nabla W|^2}{W^4}\right).\end{aligned}$$

Estimating

$$\frac{|\nabla W|^2}{W} \le 4|\nabla^\perp\mathbf{H}|^2$$

then yields

$$(\partial_t - \Delta)\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{5}{2}}} \le \frac{c}{W^{\frac{5}{2}}}\left[W|\nabla^2\mathbf{A}|^2 + |\nabla\mathbf{A}|^4\right] - \frac{|\nabla^3\mathbf{A}|^2}{W^{\frac{5}{2}}} + 100\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}|\nabla^\perp\mathbf{H}|^2.$$

Restricting to $t \ge \lambda_0 K^{-1}$ and estimating the first order terms using Proposition 3.9 (and Young's inequality) then yields

$$\begin{aligned}(\partial_t - \Delta)\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{5}{2}}} \le{}& c_1\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{3}{2}}} + C_1K^2\frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}\mathrm{e}^{-2Kt}\\ &+ \frac{c_1|\mathbf{H}|^4W^2 + C_1K^4\mathrm{e}^{-4Kt}}{W^{\frac{5}{2}}} - \frac{|\nabla^3\mathbf{A}|^2}{W^{\frac{5}{2}}}\,,\end{aligned}$$

where c_1 depends only on n, k, α and Θ, and C_1 depends also on V.

Similar arguments yield

$$(\partial_t - \Delta)\frac{|\nabla\mathbf{A}|^2}{W^{\frac{3}{2}}} \le \frac{c_2|\mathbf{H}|^2W^3 + C_2K^4\mathrm{e}^{-4Kt}}{W^{\frac{5}{2}}} - \frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{3}{2}}}$$

and

$$\begin{aligned}(\partial_t - \Delta)\frac{|\nabla\mathbf{A}|^2}{W^{\frac{7}{2}}} \le{}& c\frac{|\nabla\mathbf{A}|^2}{W^{\frac{9}{2}}}\left(W^2 + |\nabla^\perp\mathbf{H}|^2\right) - \frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}\\ \le{}& \frac{c_3|\mathbf{H}|^2W^3 + C_3K^4\mathrm{e}^{-4Kt}}{W^{\frac{9}{2}}} - \frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{7}{2}}}\,,\end{aligned}$$

where c_2 and c_3 depend only on n, α, and Θ, and C_2 and C_3 depend also on V.

Setting

$$f \doteqdot \frac{|\nabla^2\mathbf{A}|^2}{W^{\frac{5}{2}}} + c_1\frac{|\nabla\mathbf{A}|^2}{W^{\frac{3}{2}}} + C_1K^2\frac{|\nabla\mathbf{A}|^2}{W^{\frac{7}{2}}}$$

and estimating $W \geq K$, we obtain

$$\begin{aligned}(\partial_t - \Delta)f \leq{}& \frac{c_1|\mathbf{H}|^4W^2 + C_1K^4\mathrm{e}^{-4Kt}}{W^{\frac{5}{2}}} + c_1\frac{c_2|\mathbf{H}|^2W^3 + C_2K^4\mathrm{e}^{-4Kt}}{W^{\frac{5}{2}}}\\ &+ C_1K^2\frac{c_3|\mathbf{H}|^2W^3 + C_3K^4\mathrm{e}^{-4Kt}}{W^{\frac{9}{2}}}\\ \leq{}& \frac{(c_1 + c_1c_2 + c_3C_1)|\mathbf{H}|^2W^3 + (C_1 + c_1C_2 + C_1C_3)K^4\mathrm{e}^{-4Kt}}{W^{\frac{5}{2}}}\\ \leq{}& (c_1 + c_1c_2 + c_3C_1)|\mathbf{H}|^2W^{\frac{1}{2}} + (C_1 + c_1C_2 + C_1C_3)K^{\frac{3}{2}}\mathrm{e}^{-4Kt}\\ \doteqdot{}& c_4|\mathbf{H}|^2W^{\frac{1}{2}} + C_4K^{\frac{3}{2}}\mathrm{e}^{-4Kt}\,.\end{aligned}$$

On the other hand, the function G defined by

$$G^2 \doteqdot 2(2-\alpha_n)K + \tfrac{1}{n-2+\alpha_n}|\mathbf{H}|^2 - |\mathbf{A}|^2$$

satisfies

$$(\partial_t - \Delta)G \geq c_5|\mathbf{H}|^2W^{\frac{1}{2}}\,,$$

where c_5 depends only on n and α. Thus,

$$(\partial_t - \Delta)\left(f - \frac{c_4}{c_5}G + \frac{C_4}{4}K^{\frac{1}{2}}\mathrm{e}^{-4Kt}\right) \leq 0\,.$$

The maximum principle and Proposition 3.8 then yield

$$\begin{aligned}\max_{M\times\{t\}}\left(f - \frac{c_4}{c_5}G\right) \leq{}& \max_{M\times\{\lambda_0K^{-1}\}}\left(f - \frac{c_4}{c_5}G\right) + \frac{C_4}{4}K^{\frac{1}{2}}\left(\mathrm{e}^{-4\lambda_0} - \mathrm{e}^{-4Kt}\right)\\ \leq{}& C_5K^{\frac{1}{2}}\end{aligned}$$

for all $t \geq \lambda_0K^{-1}$, where C_5 depends only on n, α, V, and Θ. We conclude that

$$|\nabla^2\mathbf{A}|^2 \leq cW^3 + CK^{\frac{1}{2}}W^{\frac{5}{2}} \;\text{ in }\; M\times[\lambda_0K^{-1}, T)\,,$$

where c and C depend only on n, α, V, and Θ. The claim now follows from Young's inequality. □

Applying the Hessian estimate in conjunction with the the rough evolution equation

$$(\nabla_t - \Delta)\mathbf{A} = \mathbf{A}*\mathbf{A}*\mathbf{A} + K*\mathbf{A}$$

for $\mathbf{A}$ yields an analogous bound for $\nabla_t\mathbf{A}$, and hence, in particular, for the time derivative of $\mathbf{H}$. Thus, in high curvature regions, we obtain the following a priori bounds for $\nabla^\perp\mathbf{H}$ and $\nabla_t^\perp\mathbf{H}$.

Corollary 3.11. *Let* $X : M^n \times [0,T) \to S_K^{n+\ell}$, $n \geq 2$, *be a solution to mean curvature flow with initial condition in the class* $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. *There exist* $h_\sharp = h_\sharp(n, \alpha, V, \Theta)$ *and* $c_\sharp = c_\sharp(n, \alpha, V, \Theta)$ *such that*

$$|\mathbf{H}|(x,t) \geq h_\sharp\sqrt{K} \implies \frac{|\nabla^\perp \mathbf{H}|}{|\mathbf{H}|^2}(x,t) \leq c_\sharp \text{ and } \frac{|\nabla_t^\perp \mathbf{H}|}{|\mathbf{H}|^3}(x,t) \leq \frac{c_\sharp^2}{2}. \tag{3.32}$$

3.6. **Neck detection.** The cylindrical and gradient estimates can be used to show that regions of very high curvature which are not pinched in the sense of Andrews and Baker must form high quality 'neck' regions.

Definition 3.12. Let $X : M \to S_K^{n+\ell} \subset \mathbb{R}^{n+\ell+1}$ be an immersed submanifold of $S_K^{n+\ell}$. A point $x \in M$ *lies at the center of an* (ε, k, L)*-neck of size* r if the map $\exp^{-1}_{r^{-1}X(x)} \circ (r^{-1}X)$ is ε-cylindrical and (ε, k)-parallel at all points in the induced intrinsic ball of radius L about p in the sense of [15, Definition 3.9].

Lemma 3.13 (Neck detection (cf. [15, Lemma 7.4]))**.** *Let* $X : M \times [0,T) \to S_K^{n+\ell}$ *be a solution to mean curvature flow with initial condition in the class* $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. *Given* $\varepsilon \leq \frac{1}{100}$, *there exist parameters* $\eta_\sharp = \eta_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon) > 0$ *and* $h_\sharp = h_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon) < \infty$ *with the following property. If*

$$|\mathbf{H}|(x_0, t_0) \geq h_\sharp\sqrt{K} \quad \text{and} \quad \big(|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2\big)(x_0, t_0) \geq -\eta_\sharp |\mathbf{H}|^2(x_0, t_0)\,,$$

then

$$\Lambda_{r_0,k,\varepsilon}(x_0, t_0) \leq \varepsilon r_0^{-(k+1)}$$

for each $k = 0, \ldots, \lfloor \frac{2}{\varepsilon} \rfloor$, *where* $r_0 \doteqdot \frac{n-1}{|\mathbf{H}|(x_0,t_0)}$,

$$\Lambda_{r,0,\varepsilon}(x,t) \doteqdot \max_{\mathcal{B}_{\varepsilon^{-1}r}(x,t)\times(t-10^4 r^2, t]} \big||\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2\big|\,,$$

and, for each $k \geq 1$,

$$\Lambda_{r,k,\varepsilon}(x,t) \doteqdot \max_{\mathcal{B}_{\varepsilon^{-1}r}(x,t)\times(t-10^4 r^2, t]} |\nabla^k \mathbf{A}|\,.$$

Proof. The result can be obtained by *reductio ad absurdum*, exploiting the cylindrical, gradient and Hessian estimates for the second fundamental form. We do not include the argument since it is similar to that of [15, Lemma 7.4] (cf. [9, Theorem 7.13], [16, Lemma 4.16] and [22, Lemma 5.5]). □

By the Gauss equation and the arguments of [22, §3] (cf. [15, §3]), necks of sufficiently high quality can be integrated (after pulling up to the tangent space) to obtain "almost hypersurface" necks (in the tangent space), which can be replaced by a pair of "convex caps" in a controlled way.

3.7. **Hypersurface detection.** The codimension estimate can be used to show that, after pulling up to the tangent space, regions of high curvature almost lie in some $(n+1)$-dimensional affine subspace.

Definition 3.14. An immersed submanifold $X : M \to S_K^{n+\ell} \subset \mathbb{R}^{n+\ell+1}$ of $S_K^{n+\ell}$ is (ε, k)-*almost hypersurface* about $x_0 \in X$ if, for some $r > 0$, the map $\exp_{r^{-1}X(x_0)}^{-1} \circ (r^{-1}X)$ is (ε, k)-almost hypersurface in the sense of [22, Definition 3.1]. That is, it satisfies

$$|\nabla^m \hat{\mathbf{A}}| \leq \varepsilon \text{ for each } m = 0, \dots, k.$$

The following lemma, combined with the Gauss equation, shows that points of sufficiently large curvature have almost hypersurface neighbourhoods (of 'size' $\sim |\mathbf{H}|^{-1}$).

Lemma 3.15 (Hypersurface detection (cf. [22, Lemma 5.8]))**.** *Let $X : M \times [0, T) \to S_K^{n+\ell}$ be a solution to mean curvature flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. Given $\varepsilon \leq \frac{1}{100}$, there exist $h_\sharp = h_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon) < \infty$, $L_\sharp = L_\sharp(n, \ell, \alpha, V, \Theta) > 0$ and $\theta_\sharp = \theta_\sharp(n, \ell, \alpha, V, \Theta) > 0$ with the following property. If*

$$|\mathbf{H}|(x_0, t_0) \geq h_\sharp \sqrt{K},$$

then

$$\sup_{\mathcal{B}_{L_\sharp r_0}(x_0,t_0) \times (t_0 - \theta_\sharp r_0^2, t_0]} |\nabla^k \hat{\mathbf{A}}| \leq \varepsilon r_0^{-(k+1)}$$

for each $k = 0, \dots, \lfloor \frac{2}{\varepsilon} \rfloor$, where $r_0 \doteqdot \frac{n-1}{|\mathbf{H}|(x_0,t_0)}$.

Proof. The proof is again very similar to that of [15, Lemma 7.4]. ∎

Note that almost hypersurface regions which satisfy our pinching condition do indeed lie close to a genuine "hypersurface". This can be proved by an argument similar to [21, Proposition 2.4] (cf. [22, Theorem 6.3]). We do not require this here, however. Indeed, we only need bounds for $\hat{\mathbf{A}}$ and $\nabla \hat{\mathbf{A}}$ (note that, under the quadratic pinching condition, the torsion is controlled by $\nabla \hat{\mathbf{A}}$).

4. The key estimates for surgically modified flows

We need to show that suitable versions of the key estimates still hold in the presence of surgeries. In the following definition, surgery is performed on the middle third of a neck of size r in the obvious way:

(i) Scale by r^{-1} and precompose with $\exp_{r^{-1}X(p)}^{-1}$ to obtain a neck in $T_{r^{-1}X(p)} S_{r^2K}^{n+\ell}$.
(ii) Perform surgery on the middle third of this neck in $T_{r^{-1}X(p)} S_{r^2K}^{n+\ell}$ as described in [22, Section 3] (cf. [15, Section 3]).
(iii) Re-embed in $S_K^{n+\ell}$ by composing with $\exp_{r^{-1}X(p)}$ and scaling by r.

Definition 4.1. A *surgically modified (mean curvature) flow* in $S_K^{n+\ell}$ with *neck parameters* (ε, k, L), *surgery parameters* (τ, B), and *surgery scale* r is a finite sequence $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$ of smooth mean curvature flows $X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}$ for which the $(i+1)$-st initial datum $X_{i+1}(\cdot, T_{i+1}) : M_{i+1} \to S_K^{n+\ell}$ is obtained from the i-th final datum $X_i(\cdot, T_{i+1}) : M_i \to S_K^{n+\ell}$ by performing finitely many (τ, B)-standard surgeries, in the sense of [15, Section 3], on the middle thirds of (ε, k, L)-necks with mean curvature satisfying $\frac{n-1}{10r} \leq H \leq \frac{10(n-1)}{r}$, and then discarding finitely many connected components that are diffeomorphic either to S^n or to $S^1 \times S^{n-1}$.

4.1. **Preserving quadratic pinching.** For a suitable range of neck and surgery parameters, and surgery scales, the surgery procedure of [22] (cf [15, Section 3]) preserves the quadratic pinching condition (1.2).

In the statement of the following proposition (Proposition 4.2) and henceforth, when we refer to a surgically modified flow, it is taken for granted that the neck and surgery parameters, and the surgery scale, are fixed within a suitable range (which we progressively refine).

Proposition 4.2 (Quadratic pinching for surgically modified flows)**.** *Every surgically modified flow* $\{X_i : M^n_i \times [T_i, T_{i+1}] \to S^{n+\ell}_K\}_{i=1}^{N-1}$*,* $n \geq 5$*, with initial condition in the class* $\mathcal{C}^{n,\ell}_K(\alpha, V, \Theta)$ *(with* $\alpha > \alpha_n$ *when* $n = 5, 6, 7$*) satisfies* (3.1) *for all* $t \in [T_1, T_N]$*.*

Proof. Since $|\mathbf{A}|^2 \equiv \frac{1}{n-1}|\mathbf{H}|^2$ on a hypersurface cylinder and $|\mathbf{A}|^2 \equiv \frac{1}{n}|\mathbf{H}|^2$ on a hypersurface cap, we can ensure, for a suitable choice of neck and surgery parameters and surgery scales, that

$$|\mathbf{A}|^2 - \frac{1}{n-2+\overline{\alpha}}|\mathbf{H}|^2 < 2(2-\overline{\alpha})K$$

on regions modified or added by surgery, where $\overline{\alpha} \doteqdot \frac{\alpha+10}{11} \in (\alpha, 1)$, say. Indeed, this follows from [22, Corollary 3.20] and the Gauss equation since the surgery scale may be taken arbitrarily small. We can now proceed as in the proof of Proposition 3.4 in the time intervals (T_i, T_{i+1}). □

4.2. **The codimension estimate.** Similar reasoning shows that the codimension estimate holds for surgically modified flows for a suitable range of neck and surgery parameters, and surgery scales.

Proposition 4.3 (Codimension estimate for surgically modified flows)**.** *Let* $\{X_i : M^n_i \times [T_i, T_{i+1}] \to S^{n+\ell}_K\}_{i=1}^{N-1}$*,* $n \geq 5$*, be a surgically modified flow with initial condition in the class* $\mathcal{C}^{n,\ell}_K(\alpha, V, \Theta)$ *(with* $\alpha > \alpha_n$ *when* $n = 5, 6, 7$*). There exist* $\delta = \delta(n, \alpha) > 0$ *and* $C = C(n, \alpha, \Theta) < \infty$ *such that*

$$|\hat{\mathbf{A}}|^2 \leq CK^\delta(|\mathbf{H}|^2 + K)^{1-\delta} \quad \text{wherever } \mathbf{H} \neq 0\,.$$

Proof. As in Proposition 3.4, we seek a bound for

$$f_\sigma \doteqdot \begin{cases} \dfrac{\frac{1}{2}|\hat{\mathbf{A}}|^2}{W}W^\sigma & \text{if } \mathbf{H} \neq 0 \\ 0 & \text{if } \mathbf{H} = 0 \end{cases}$$

for some $\sigma \in (0, 1)$.

The key observation is that, since $|\hat{\mathbf{A}}|^2 \equiv 0$ on a hypersurface cylinder or cap, we can ensure, for a suitable range of neck and surgery parameters, and surgery scales, that f_σ is pointwise nonincreasing on regions modified or added by surgery. Indeed, this follows from [22, Corollary 3.20 and Corollary 3.21] and the Gauss equation. The claim follows since, recalling the proof of Proposition 3.4, f_σ is nonincreasing in the time intervals (T_i, T_{i+1}). □

4.3. **The cylindrical estimate.** Similar reasoning yields a cylindrical estimate for surgically modified flows.

Proposition 4.4 (Cylindrical estimate for surgically modified flows (Cf. [15, Theorem 5.3]))**.** *Let* $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$*,* $n \geq 5$*, be a surgically modified flow with initial condition in the class* $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ *(with* $\alpha > \alpha_n$ *when* $n = 5, 6, 7$*). For every* $\eta \in (0, \eta_0)$ *there exists* $C_\eta = C_\eta(n, \ell, \alpha, V, \Theta, \eta) < \infty$ *such that*

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2 \leq \eta|\mathbf{H}|^2 + C_\eta K \quad \text{in} \quad M_i \times [T_i, T_{i+1}] \tag{4.1}$$

for all i*.*

Proof. Since $g_{\sigma,\eta} \equiv 0$ on a neck or a cap, we can arrange, for suitable neck and surgery parameters, that $(g_{\sigma,\eta})_+$ is pointwise non-increasing in regions added or modified by surgery. Indeed, this follows from [22, Corollary 3.20] and the Gauss equation. Proceeding as in the proof of Proposition 3.7 but discarding the exponential decay term, we obtain an analogue for surgically modified flows of (3.22) on each time interval (T_i, T_{i+1}), with v_k replaced by $(g_{\sigma,\eta} - k)_+^{\frac{p}{2}}$. Since $(g_{\sigma,\eta})_+$ is pointwise non-increasing across surgery times in regions modified by surgery, this can be integrated from $T_1 = 0$ to $T_N = T$ to obtain an analogue of (3.24). The remainder of the proof of the cylindrical estimate then applies unmodified. □

4.4. **The gradient estimate.** Since the derivatives of the second fundamental form are zero on round Euclidean cylinders and spherical caps, the derivative estimates also pass to surgically modified flows.

Proposition 4.5 (Gradient estimate for surgically modified flows (Cf. [15, Theorem 6.1]))**.** *Let* $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$*,* $n \geq 5$*, be a surgically modified flow with initial condition in the class* $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ *(with* $\alpha > \alpha_n$ *when* $n = 5, 6, 7$*). There exists* $C = C(n, \ell, \alpha, V, \Theta) < \infty$ *such that*

$$|\nabla \mathbf{A}|^2 \leq C(|\mathbf{H}|^4 + K^2) \quad \text{in} \quad M_i^n \times [T_i, T_{i+1}] \tag{4.2}$$

for all i*.*

Proof. We proceed as in the proof of Proposition 3.9, but with the exponential decay term discarded and fixed $\eta = \beta$. First observe that, since $|\mathbf{A}|^2 \equiv \frac{1}{n-1}|\mathbf{H}|^2$ on a hypersurface cylinder and $|\mathbf{A}|^2 \equiv \frac{1}{n}|\mathbf{H}|^2$ on a hypersurface cap, we can ensure, for a suitable range of neck and surgery parameters, and surgery scales, that

$$|\mathbf{A}|^2 - \tfrac{1}{n-1}|\mathbf{H}|^2 \leq \tfrac{\beta}{2}|\mathbf{H}|^2$$

on regions modified or added by surgery. We may therefore arrange that

$$G_\beta \doteqdot \left(\tfrac{1}{n-1} + \beta\right)|\mathbf{H}|^2 - |\mathbf{A}|^2 + 2C_\beta K \geq \tfrac{\beta}{2}|\mathbf{H}|^2$$

and

$$G_0 \doteqdot \tfrac{3}{n+2}|\mathbf{H}|^2 - |\mathbf{A}|^2 + 2C_0 K \geq \tfrac{3}{2}\beta|\mathbf{H}|^2.$$

Furthermore, since $|\nabla \mathbf{A}|^2 \equiv 0$ on a hypersurface cylinder or cap, we can ensure, for a suitable range of neck and surgery parameters, and surgery scales, that, on regions modified or added by surgery, $|\nabla \mathbf{A}|^2 \le \mu_0 |\mathbf{H}|^4$, where μ_0 is a constant which depends only on n. Thus, in regions modified or added by surgery,

$$\frac{|\nabla \mathbf{A}|^2}{G_0 G_\beta} \le \frac{4\mu_0}{3\beta^2}\,.$$

We may now proceed as in the proof of Proposition 3.9 in the time intervals (T_i, T_{i+1}). □

4.5. **Higher order estimates.** Proceeding similarly as in Proposition 4.5 (cf. [15, Theorem 6.3]) yields estimates for higher derivatives of $\mathbf{A}$ along surgically modified flows.

Proposition 4.6 (Hessian estimate for surgically modified flows (cf. [15, Theorem 6.3]))**.** *Let $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$, $n \ge 5$, be a surgically modified flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$ (with $\alpha > \alpha_n$ when $n = 5, 6, 7$). There exists $C = C(n, \ell, \alpha, V, \Theta)$ such that*

$$|\nabla^2 \mathbf{A}|^2 \le C(|\mathbf{H}|^6 + K^3) \ \text{ in } \ M \times [\lambda_0 K^{-1}, T)\,. \tag{4.3}$$

Proof. Proceed as in Proposition 3.10 between surgeries and use the fact that, for suitable neck and surgery parameters, and surgery scales, $|\nabla^2 \mathbf{A}|^2 / |\mathbf{H}|^6$ is small in regions modified or added by surgery. □

4.6. **Neck detection.** The conclusion of the neck detection Lemma 3.13 also holds for surgically modified flows, so long as we work in regions which are not affected by surgeries (cf. [15, Lemma 7.4]).

In the following theorem, a region $U \times I$ is *free of surgeries* if at each surgery time $T_i \in I$, $i \in \{2, \dots, N-1\}$, we have $U \subset M_{i-1} \cap M_i$ and $X_{i-1}|_U(\cdot, T_i) = X_i|_U(\cdot, T_i)$ (and hence X_{i-1} and X_i may be pasted together to form a smooth mean curvature flow in $U \times I$).

Lemma 4.7 (Neck detection for surgically modified flows (cf. [15, Lemma 7.4]))**.** *Let $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$, $n \ge 5$, be a surgically modified flow with initial condition in the class $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$. Given $\varepsilon > 0$, $k \ge 2$, $L > 0$, and $\theta > 0$, there exist $\eta_\sharp = \eta_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon, k, L, \theta) > 0$ and $h_\sharp = h_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon, k, L, \theta) < \infty$ with the following property: If*

(ND1) $|\mathbf{H}(x_0, t_0)| \ge h_\sharp \sqrt{K}$ *and* $\frac{(|\mathbf{A}|^2 - \frac{1}{n-1}|\mathbf{H}|^2)(x_0,t_0)}{|\mathbf{H}(x_0,t_0)|^2} \ge -\eta_\sharp$, *and*
(ND2) *the neighbourhood $\mathcal{B}_{(L+1)r_0}(x_0, t_0) \times (t_0 - \theta r_0^2, t_0]$ is free of surgeries, where* $r_0^2 \doteqdot \frac{n-1}{|\mathbf{H}|^2(x_0,t_0)}$,

then (x_0, t_0) lies at the centre of an (ε, k, L)-neck of size r_0.

Proof. The proof of Lemma 3.13 applies using Propositions 4.4, 4.5 and 4.6 in lieu of Propositions 3.7, 3.9 and 3.10, due to the hypothesis (ND2). □

4.7. **Hypersurface detection.** Since the codimension estimate survives the surgery, we obtain an analogue of the hypersurface detection lemma in regions unaffected by surgery.

Lemma 4.8 (Hypersurface detection for surgically modified flows)**.** *Let* $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$*,* $n \geq 5$*, be a surgically modified flow in the class* $\mathcal{C}_K^{n,\ell}(\alpha, V, \Theta)$*. Given* $\varepsilon > 0$*, there exist* $h_\sharp = h_\sharp(n, \ell, \alpha, V, \Theta, \varepsilon) > 0$*,* $L_\sharp = L_\sharp(n, \ell, \alpha, V, \Theta) > 0$ *and* $\theta_\sharp = \theta_\sharp(n, \ell, \alpha, V, \Theta) > 0$ *with the following property. If*

(HD1) $|\mathbf{H}(x_0, t_0)| \geq h_\sharp \sqrt{K}$*, and*
(HD2) *the neighbourhood* $\mathcal{B}_{(L_\sharp+1)r_0}(x_0, t_0) \times (t_0 - (\theta_\sharp + 1)r_0^2, t_0]$ *is free of surgeries, where* $r_0 \doteqdot \frac{n-1}{|\mathbf{H}|(x_0,t_0)}$*,*

then

$$\sup_{\mathcal{B}_{L_\sharp r_0}(x_0,t_0)\times(t_0-\theta_\sharp r_0^2,t_0]} |\nabla^k \hat{\mathbf{A}}| \leq \varepsilon r_0^{-(k+1)}$$

for each $k = 0, \ldots, \lfloor \frac{2}{\varepsilon} \rfloor$*.*

Proof. The proof of Lemma 3.15 applies using Propositions 4.4, 4.5 and 4.6 in lieu of Propositions 3.7, 3.9 and 3.10, due to the hypothesis (HD2). □

5. Existence of terminating surgically modified flows

We say that a surgically modified flow $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$ *terminates* at the final time $T \doteqdot T_N < \infty$ if either

- each connected component of $X_{N-1}(M_{N-1}, T_N)$ is diffeomorphic to S^n or to $S^1 \times S^{n-1}$, or
- after performing surgery on $X_{N-1}(M_{N-1}, T_N)$, each connected component of the resulting hypersurface is diffeomorphic to S^n or to $S^1 \times S^{n-1}$.

Theorem 5.1 (Existence of terminating surgically modified flows)**.** *Let* $X : M \to S_K^{n+\ell}$*,* $n \geq 5$*, be a properly immersed hypersurface satisfying the quadratic pinching condition* (1.2)*. There exists a surgically modified flow* $\{X_i : M_i^n \times [T_i, T_{i+1}] \to S_K^{n+\ell}\}_{i=1}^{N-1}$ *with* $X_1(\cdot, 0) = X$ *which terminates at time* $T = T_N$*.*

Proof. Given the codimension, cylindrical and derivative estimates, and the neck and hypersurface detection lemmas, and a sufficiently small choice of the surgery scale r, we can proceed as in [22, Section 6] (cf. [15, Section 8]) using the machinery developed in [22, Section 3] (cf. [15, Sections 3 and 7]), with only minor modifications required (cf. [16]). These are:

1. In order to reconcile our data $\mathcal{C}_K^n(\alpha, V, \Theta)$ with those of [22], we replace the parameter K by introducing the scale factor $R \doteqdot 1/\sqrt{\Theta K}$. Our data α and V can then be related to the α_0 and α_2 there, respectively. The constant α_1 which appears in [22] is not needed here. Since the surgery scale may be taken as small as needed, we may then choose the surgery parameters (albeit with slightly worse values) as explained in [22, Section 6].

2. Since our ambient space is non-Euclidean, the proof of the neck continuation theorem [22, Theorem 6.3] requires modification in two places. These are explained and carried out in detail in [9, Section 8].

3. Since the maximal time is not a priori bounded in the present setting, the surgery algorithm may not terminate "on its own". This case is easily dealt with using Proposition 3.2 as in [16], however. □

Theorem 1.1 follows.

References

[1] Hilário Alencar and Manfredo do Carmo. Hypersurfaces with constant mean curvature in spheres. *Proc. Amer. Math. Soc.*, 120(4):1223–1229, 1994.
[2] Li An-Min and Li Jimin. An intrinsic rigidity theorem for minimal submanifolds in a sphere. *Arch. Math. (Basel)*, 58(6):582–594, 1992.
[3] B. Andrews. Positively curved surfaces in the three-sphere. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)*, pages 221–230. Higher Ed. Press, Beijing, 2002.
[4] Ben Andrews and Charles Baker. Mean curvature flow of pinched submanifolds to spheres. *J. Differential Geom.*, 85(3):357–395, 2010.
[5] Ben Andrews, Bennett Chow, Christine Guenther, and Mat Langford. *Extrinsic Geometric Flows.* Graduate Studies in Mathematics, Vol. 206. American Mathematical Society, 2020.
[6] Charles Baker. *The mean curvature flow of submanifolds of high codimension.* Australian National University, 2011. Thesis (Ph.D.)–Australian National University.
[7] Charles Baker and Huy The Nguyen. Evolving pinched submanifolds of the sphere by mean curvature flow. Preprint, arXiv:2004.12259.
[8] Charles Baker and Huy The Nguyen. Codimension two surfaces pinched by normal curvature evolving by mean curvature flow. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 34(6):1599–1610, 2017.
[9] Simon Brendle and Gerhard Huisken. A fully nonlinear flow for two-convex hypersurfaces in Riemannian manifolds. *Invent. Math.*, 210(2):559–613, 2017.
[10] Qing Ming Cheng and Hisao Nakagawa. Totally umbilic hypersurfaces. *Hiroshima Math. J.*, 20(1):1–10, 1990.
[11] S. S. Chern, M. do Carmo, and S. Kobayashi. Minimal submanifolds of a sphere with second fundamental form of constant length. In *Functional Analysis and Related Fields (Proc. Conf. for M. Stone, Univ. Chicago, Chicago, Ill., 1968)*, pages 59–75. Springer, New York, 1970.
[12] David Hoffman and Joel Spruck. Sobolev and isoperimetric inequalities for Riemannian submanifolds. *Commun. Pure Appl. Math.*, 27:715–727, 1974.
[13] Gerhard Huisken. Flow by mean curvature of convex surfaces into spheres. *J. Differential Geom.*, 20(1):237–266, 1984.
[14] Gerhard Huisken. Deforming hypersurfaces of the sphere by their mean curvature. *Math. Z.*, 195(2):205–219, 1987.
[15] Gerhard Huisken and Carlo Sinestrari. Mean curvature flow with surgeries of two-convex hypersurfaces. *Invent. Math.*, 175(1):137–221, 2009.
[16] Mat Langford and Huy The Nguyen. Quadratically pinched hypersurfaces of the sphere via mean curvature flow with surgery. arXiv:2006.08049, 2020.
[17] Mat Langford and Huy The Nguyen. Sharp pinching estimates for mean curvature flow in the sphere. arXiv:2009.00986, 2020.
[18] Stephen Lynch and Huy The Nguyen. Convexity estimates for high codimension mean curvature flow. Preprint, arXiv:2006.05227 [math.DG].
[19] Stephen Lynch and Huy The Nguyen. Pinched ancient solutions to the high codimension mean curvature flow. Preprint, arXiv:1709.09697.
[20] J. H. Michael and L. M. Simon. Sobolev and mean-value inequalities on generalized submanifolds of R^n. *Comm. Pure Appl. Math.*, 26:361–379, 1973.
[21] Keaton Naff. Codimension estimates in mean curvature flow. Preprint, arXiv:1906.08184.

[22] Huy The Nguyen. High codimension mean curvature flow with surgery. arXiv:2004.07163, 2020.
[23] Masafumi Okumura. Hypersurfaces and a pinching problem on the second fundamental tensor. *Amer. J. Math.*, 96:207–213, 1974.
[24] Walcy Santos. Submanifolds with parallel mean curvature vector in spheres. *Tohoku Math. J. (2)*, 46(3):403–415, 1994.
[25] James Simons. Minimal varieties in riemannian manifolds. *Ann. of Math. (2)*, 88:62–105, 1968.
[26] Guido Stampacchia. *Èquations elliptiques du second ordre à coefficients discontinus*. Séminaire de Mathématiques Supérieures, No. 16 (Été, 1965). Les Presses de l'Université de Montréal, Montreal, Que., 1966.

Department of Mathematics, University of Tennessee Knoxville, Knoxville TN, USA, 37996-1320

School of Mathematical and Physical Sciences, The University of Newcastle, Newcastle, NSW, Australia, 2308
E-mail address: mlangford@utk.edu, mathew.langford@newcastle.edu.au

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, UK, E1 4NS
E-mail address: h.nguyen@qmul.ac.uk

Eberhard Karls Universität Tübingen, Fachbereich Mathematik, Auf der Morgenstelle 10, 72076 Tübingen, Germany
E-mail address: stephen.lynch@math.uni-tuebingen.de