Invent. math. 175, 137–221 (2009)
DOI: 10.1007/s00222-008-0148-4

Inventiones mathematicae

Mean curvature flow with surgeries of two–convex hypersurfaces

Gerhard Huisken[1], **Carlo Sinestrari**[2]

[1] Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1, 14476 Golm, Germany (e-mail: gerhard.huisken@aei.mpg.de)
[2] Dipartimento di Matematica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italy (e-mail: sinestra@mat.uniroma2.it)

Oblatum 8-VI-2006 & 5-VII-2008
Published online: 17 September 2008 –

1 Introduction

Let $F_0 : \mathcal{M} \to \mathbb{R}^{n+1}$ be a smooth immersion of an n–dimensional hypersurface in Euclidean space, $n \geq 2$. The evolution of $\mathcal{M}_0 = F_0(\mathcal{M})$ by mean curvature flow is the one–parameter family of smooth immersions $F : \mathcal{M} \times [0, T[\to \mathbb{R}^{n+1}$ satisfying

$$\frac{\partial F}{\partial t}(p,t) = -H(p,t)\nu(p,t), \qquad p \in \mathcal{M}, t \geq 0, \tag{1.1}$$

$$F(\cdot,0) = F_0, \tag{1.2}$$

where $H(p,t)$ and $\nu(p,t)$ are the mean curvature and the outer normal respectively at the point $F(p,t)$ of the surface $\mathcal{M}_t = F(\cdot,t)(\mathcal{M})$. The signs are chosen such that $-H\nu = \overrightarrow{H}$ is the mean curvature vector and the mean curvature of a convex surface is positive.

For closed surfaces the smooth solution of (1.1)–(1.2) exists on a maximal time interval $[0, T[$, with $T > 0$ finite, and the curvature of the surfaces becomes unbounded for $t \to T$. In the last decades several different notions of a weak solution have been introduced to define a flow after the singular time T, see among others [3,4,8,1]. The purpose of this paper is to define a flow after singularities by a new approach based on a surgery procedure. Compared with the notions of weak solutions existing in the literature, the flow with surgeries has the advantage that it keeps track of the changes of topology of the evolving surface and thus can be applied to classify all geometries that are possible for the initial manifold. Our surgery construction is inspired by the procedure originally introduced by Hamilton in [14] for a Ricci flow with surgery deforming metrics on a Riemannian manifold, see also [13]. In three dimensions Hamilton's Ricci flow is the fundamental approach to Thurston's geometrization conjecture and was employed by Perelman in conjunction with a different surgery procedure in [25,26].

The intuitive idea of the present surgery construction consists of stopping the smooth mean curvature flow at a certain time $T_1 < T$ which is very close to the first singular time T. One then removes the regions in the surface which have large curvature and replaces them by more regular ones such that the maximum curvature drops by a certain fixed factor. One also removes connected components of the surface that are recognized as being diffeomorphic to $\mathbb{S}^n$ or $\mathbb{S}^{n-1} \times \mathbb{S}^1$. After this the smooth flow is restarted until a new singular time is approached and the whole procedure is repeated. The central claim to be proven is that the whole procedure can be quantitatively controlled in terms of a few parameters depending only on the initial data and terminates after finitely many steps when all remaining components are recognized as being diffeomorphic to copies of $\mathbb{S}^n$ or $\mathbb{S}^{n-1} \times \mathbb{S}^1$. Once this is shown we can conclude that the initial manifold is diffeomorphic either to $\mathbb{S}^n$ or to a finite connected sum of $\mathbb{S}^{n-1} \times \mathbb{S}^1$.

The required procedure can only be defined when one has a detailed knowledge of the possible profiles of the surface near the singularities. Namely, one needs to show that the nearly singular parts of the surface are either uniformly convex and cover a whole connected component of the surface or they contain regions which are very close to cylinders $\mathbb{S}^{n-1} \times [a, b]$; these regions are called *necks* and are the ones which will be removed by the surgery. We emphasize that our procedure is modelled on the original construction by Hamilton in [14], while Perelman's approach [25] has some differences (for example, in Perelman's approach surgery is performed exactly at the singular time rather than a short time before). Our approach benefits from new a priori estimates on the second fundamental form and its gradient that we establish for mean curvature flow with surgery in case $n \geq 3$. These estimates allow a detailed quantitative control of necks and at this stage have no analogue yet in 2-dimensional mean curvature flow or in 3-dimensional Ricci flow.

The main result of this paper is that a procedure for mean curvature flow with surgery having the desired properties can be constructed if the initial surface is closed of dimension at least 3 and is *two-convex*. A surface is called two-convex if the sum of the two smallest principal curvatures $\lambda_1 + \lambda_2$ is nonnegative everywhere.

Theorem 1.1 *Let $F_0 : \mathcal{M} \to \mathbb{R}^{n+1}$ be a smooth immersion of a closed n–dimensional hypersurface, with $n \geq 3$. Suppose that $\mathcal{M}_0 = F_0(\mathcal{M})$ is two–convex, i.e. that $\lambda_1 + \lambda_2 \geq 0$ everywhere on $\mathcal{M}_0$. Then there exists a mean curvature flow with surgeries starting from $\mathcal{M}_0$ which terminates after a finite number of steps.*

This result has the following immediate topological consequences.

Corollary 1.2 *Any smooth closed n-dimensional two–convex immersed surface $F_0 : \mathcal{M} \to \mathbb{R}^{n+1}$ with $n \geq 3$ is diffeomorphic either to $\mathbb{S}^n$ or* *to a finite connected sum of* $\mathbb{S}^{n-1} \times \mathbb{S}^1$*. Furthermore there exists a handlebody* Ω*, see e.g.* [22, Ch. 3] *, and an immersion* $G : \bar{\Omega} \to \mathbb{R}^{n+1}$ *such that* $\partial\Omega \simeq \mathcal{M}$ *is diffeomorphic to the initial hypersurface* $\mathcal{M}$ *and such that* $G|_{\partial\Omega} = F_0$.

We will see in Theorem 3.26 that embedded initial surfaces remain embedded under mean curvature flow with surgery. This leads to the following Schoenflies type theorem for simply connected two-convex surfaces.

Corollary 1.3 *Any smooth closed simply connected n-dimensional two–convex embedded surface* $\mathcal{M} \subset \mathbb{R}^{n+1}$ *with* $n \geq 3$ *is diffeomorphic to* $\mathbb{S}^n$ *and bounds a region whose closure is diffeomorphic to a smoothly embedded* $(n+1)$*-dimensional standard closed ball.*

We note that the condition of two-convexity is slightly weaker than other natural curvature conditions: For 3-dimensional hypersurfaces it is implied by nonnegative scalar curvature, for 4-dimensional hypersurfaces it is implied by nonnegative isotropic curvature. Previous results related to Corollary 1.2 have been obtained by other authors [9,23,27,30] using different approaches. Also, an analysis of a flow through singularities of two-dimensional mean convex surfaces has been announced in [5], while geometric information on the behavior of weak solutions of mean curvature flow can be found in [6,28,29].

Since the present construction of mean curvature flow with surgeries appears to be rather robust we believe that apart from the specific application of Corollary 1.2 there will be further applications in different contexts, e.g. in the case where the ambient space is not euclidean. We also hope to remove the assumption $n \geq 3$ in future work.

We now give an outline of the current paper. In Sect. 2 we introduce notation and collect basic facts about mean curvature flow as well as two-convex surfaces. We then introduce the principal parameters controlling two-convex initial surfaces and give a precise definition of mean curvature flow with surgeries.

Section 3 investigates necks and their quantitative description. We recall that Hamilton [14] has introduced different equivalent notions of necks for abstract Riemannian manifolds. Here we adapt his theory to the setting of immersed surfaces. It turns out to be possible to give extrinsic notions of necks for hypersurfaces consistent with Hamilton's intrinsic ones in [14], allowing us to make use of some of the results obtained there. The different equivalent definitions of necks are discussed at the beginning of Sect. 3. Among other possible descriptions, the reader can keep in mind the characterization of a neck as a region of our surface which can be represented, up to a homothety, as a graph over a cylinder $\mathbb{S}^{n-1} \times [a, b]$ with small C^k norm for a suitable k. In the rest of Sect. 3 we introduce the standard surgery procedure for necks, which replaces a piece of a neck close to some cylinder by two convex spherical caps, while keeping track of all curvature quantities.

Section 4 first revisits the convexity estimates established in [18, 19] for smooth mean curvature flow and establishes that they are retained by the surgery procedure if the surgery parameters are chosen appropriately. We state the estimates in a non-technical form which is made precise in Theorem 4.13 (see also Lemma 5.1):

Theorem 1.4 (Convexity estimates) *For a given smooth closed n-dimensional initial hypersurface $\mathcal{M}_0$ in $\mathbb{R}^{n+1}$ of positive mean curvature the parameters of standard surgery can be chosen in such a way that the solution $\mathcal{M}_t$, $t \in [0, T[$ of mean curvature flow with surgery satisfies the following estimate on the smallest eigenvalue λ_1 of the second fundamental form: For any $\delta > 0$, there exists $C_\delta = C_\delta(\mathcal{M}_0) > 0$ such that*

$$\lambda_1 \geq -\delta H - C_\delta \text{ on } \mathcal{M}_t, \quad \forall t \in [0, T[. \tag{1.3}$$

This result implies that any surface obtained by rescaling our flow around a singularity is convex. Roughly speaking, the nearly singular regions of the surface become asymptotically convex as a singular time is approached. We note that B. White [29] has obtained convexity estimates for level-set solutions of mean curvature flow using methods from geometric measure theory.

In Sect. 5 we prove new a priori estimates for two-convex surfaces moving by mean curvature with surgery which imply that points where λ_1 is small have curvature close to the curvature of a cylinder. The precise result is given in Theorem 5.3, but it can be stated in non-technical form as follows:

Theorem 1.5 (Cylindrical estimate) *For a given smooth closed two-convex initial hypersurface $\mathcal{M}_0$ in $\mathbb{R}^{n+1}$, $n \geq 3$, the parameters of standard surgery can be chosen in such a way that the solution $\mathcal{M}_t$, $t \in [0, T[$ of mean curvature flow with surgery satisfies the following estimate: for any $\eta > 0$, there exists $C_\eta = C_\eta(\mathcal{M}_0) > 0$ such that at every point we have the property*

$$|\lambda_1| \leq \eta H \implies |\lambda_i - \lambda_j|^2 \leq c(n)\eta H^2 + C_\eta, \quad \forall i, j \geq 2, \tag{1.4}$$

where $c(n)$ only depends on n.

To compare the curvature at different points of the surface we prove derivative estimates for the curvature in Sect. 6, making use of the convexity estimates and the cylindrical estimate in the previous sections.

Theorem 1.6 (Gradient estimate) *For a given smooth closed two-convex initial hypersurface $\mathcal{M}_0$ in $\mathbb{R}^{n+1}$, $n \geq 3$, the parameters of standard surgery can be chosen in such a way that the solution $\mathcal{M}_t$, $t \in [0, T[$ of mean curvature flow with surgery satisfies the following estimate: there exist $c = c(n)$ and $C = C(\mathcal{M}_0) > 0$ such that the inequality*

$$|\nabla A|^2 \leq c|A|^4 + C$$

holds everywhere on $\mathcal{M} \times [0, T[$.

We recall that gradient estimates for mean curvature flow are already available in the literature. The novelty of the above result is that it has a pointwise character and the constants do not depend on the maximum of the curvature on the surface. It is fundamental for our construction that the estimates hold for the flow with surgeries, with the same constants across all surgeries.

In the last two sections we demonstrate the existence of an algorithm governing our flow with surgeries and prove our main theorem. Roughly speaking, this requires two main steps. First, we have to show that before a singular time it is possible to do surgery unless the surface is diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$. For this we have to show that nearly singular regions necessarily contain necks unless they are uniformly positively curved. The basic quantitative tool in this analysis is given in Lemma 7.4, which shows that any point where H is large enough while the ratio λ_1/H is small enough necessarily lies at the center of a neck. Further results are needed to treat the case of a point at which H is large but λ_1/H is not small, which may occur in a so called degenerate neckpinch. In this case we can use our gradient estimate for the curvature to prove that a neck can be found within a controlled distance from the given point of high curvature.

The second step is to show that the procedure ends after a finite number of surgeries. To establish this we prove in Theorem 8.2 that, provided the parameters involved are chosen appropriately, any neck can be continued until the curvature at the ends is much smaller than in the middle. Therefore, the surgeries can be done in such a way that the curvature of the surface decreases by a fixed factor. This ensures that we can maintain a fixed upper bound on the curvature throughout our mean curvature flow with surgery, depending only on the initial data. In particular it implies that the time between two surgeries is uniformly bounded from below, and that only finitely many of them can occur. This control also allows us to prove that for embedded initial data surgeries can be done in such a way that they do not destroy embeddedness, completing the proof of the main theorems.

2 Preliminaries and statement of main results

In this section we collect basic facts and notation about mean curvature flow and 2-convex surfaces and state the main results. We also describe mean curvature flow with surgeries together with the key a priori estimates controlling it.

Let $F : \mathcal{M} \times [0, T[\to \mathbb{R}^{n+1}$ be a solution of mean curvature flow (1.1)–(1.2) with closed, smoothly immersed evolving surfaces $\mathcal{M}_t = F(\cdot, t)(\mathcal{M})$. We denote the induced metric by $g = \{g_{ij}\}$, the surface measure by $d\mu$, the second fundamental form by $A = \{h_{ij}\}$ and the Weingarten operator by $W = \{h^i_j\}$. We then denote by $\lambda_1 \leq \cdots \leq \lambda_n$ the principal curvatures, i.e. the eigenvalues of W, and by $H = \lambda_1 + \cdots + \lambda_n$ the mean curvature. In addition, $|A|^2 = \lambda_1^2 + \cdots + \lambda_n^2$ will denote the squared norm of A. All these quantities depend on $(p, t) \in \mathcal{M} \times [0, T[$ and satisfy the following equations computed in [16].

Lemma 2.1 *If $\mathcal{M}_t$ evolves by mean curvature flow, the associated quantities introduced above satisfy the following equations (here ∇ and Δ denote respectively the covariant derivative and the Laplace–Beltrami operator induced by the metric on $\mathcal{M}_t$):*

$$\text{(i) } \frac{\partial}{\partial t} g_{ij} = -2Hh_{ij}, \quad \text{(ii) } \frac{\partial}{\partial t} d\mu = -H^2 d\mu, \quad \text{(iii) } \frac{\partial}{\partial t} h^i_j = \Delta h^i_j + |A|^2 h^i_j,$$

$$\text{(iv) } \frac{\partial}{\partial t} H = \Delta H + |A|^2 H, \quad \text{(v) } \frac{\partial}{\partial t} |A|^2 = \Delta |A|^2 - 2|\nabla A|^2 + 2|A|^4.$$

Let us also recall the following well known result (see e.g. [7]).

Theorem 2.2 *Let $\mathcal{M}_0 = F_0(\mathcal{M})$ be smooth and closed. Then the mean curvature flow* (1.1)–(1.2) *has a unique smooth solution, which is defined in a maximal time interval* $[0, T[$ *and satisfies* $\max_{\mathcal{M}_t} |A|^2 \to \infty$ *as $t \uparrow T$.*

The key assumption of this paper is that the dimension n of the hypersurfaces is at least 3 and that the initial hypersurface is 2-convex, i.e. that the sum of the smallest two principal curvatures is non-negative: $\lambda_1 + \lambda_2 \geq 0$ everywhere on $\mathcal{M}_0$. Let us give a sufficient condition for 2-convexity in terms of the elementary symmetric polynomials. In the following we set

$$S_k = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_k}$$

to denote the k-th elementary symmetric polynomial of the principal curvatures. In particular, $S_1 = H$, and S_2 is the scalar curvature.

Lemma 2.3 *Let $\mathcal{M}$ be a smooth n-dimensional hypersurface such that $S_1 > 0, \ldots, S_{n-1} > 0$. Then $\lambda_1 + \lambda_2 > 0$.*

Proof. We recall a property of the symmetric polynomials (see Lemma 2.4 in [19]): if $S_1 > 0, \ldots, S_k > 0$, then $S_{h,i} > 0$ for any $h = 1, \ldots, k-1$ and $i = 1, \ldots, n$. Here $S_{h,i} > 0$ denotes the sum of the terms of S_h not containing λ_i. We argue by induction. In the case $n = 3$ we have $\lambda_1 + \lambda_2 = S_{1,3} > 0$ by the property recalled above. For a general $n > 3$, we observe that $S_{1,n} > 0, S_{2,n} > 0, \ldots, S_{n-2,n} > 0$. This means that $\lambda_1, \ldots, \lambda_{n-1}$ satisfy the assumptions of the lemma and so, by the induction hypothesis, $\lambda_1 + \lambda_2 > 0$. □

In particular, a 3-dimensional manifold with positive mean curvature and positive scalar curvature is 2-convex. It is also easy to see that, for $n \geq 4$, a hypersurface with positive isotropic curvature is 2-convex (see [23, §3]).

Remark 2.4 (i) The converse implication is not true, namely $\lambda_1 + \lambda_2 > 0$ does not imply the positivity of $S_1, \dots, S_{n-1}$. For example, in three dimensions, if $\lambda_1 = -2, \lambda_2 = \lambda_3 = 3$ we have $\lambda_1 + \lambda_2 > 0$ and $S_2 < 0$. More precisely, it can be checked that, if $\lambda_1 + \lambda_2 > 0$, then all S_k with $2k \leq n$ are positive, while the remaining ones can be negative.

(ii) To maintain quantitative control of 2-convexity we consider in the following uniformly 2-convex surfaces, i.e. surfaces that satisfy $\lambda_1 + \lambda_2 \geq \alpha_0 H$ for some $\alpha_0 > 0$. We may assume that $0 < \alpha_0 < 1/(n-1)$ since any surface with $\lambda_1 + \lambda_2 \geq H/(n-1)$ is convex: In fact, suppose that $\lambda_1 + \lambda_2 \geq H/(n-1)$. Then, keeping into account that $H \geq \lambda_1 + (n-1)\lambda_2$, we find that $\lambda_1 \geq \lambda_1/(n-1)$, which implies that $\lambda_1 \geq 0$.

In order to deal with mean curvature flow and surgery on 2-convex surfaces in a transparent quantitative way, we introduce a class of surfaces which is controlled by just a few parameters and will be invariant under both mean curvature flow and surgery:

Definition 2.5 *For a set positive of constants $R, \alpha_0, \alpha_1, \alpha_2$ we denote by $\mathcal{C}(R, \alpha)$ with $\alpha = (\alpha_0, \alpha_1, \alpha_2)$ the class of all smooth and closed hypersurface immersions $F : \mathcal{M} \to \mathbb{R}^{n+1}$ satisfying the estimates*

$$\text{(i) } \lambda_1 + \lambda_2 \geq \alpha_0 H, \quad \text{(ii) } H \geq \alpha_1 R^{-1}, \quad \text{(iii) } |\mathcal{M}| \leq \alpha_2 R^n. \tag{2.1}$$

Here R plays the role of a scaling parameter, making α_0, α_1 and α_2 scaling invariant: if $F \in \mathcal{C}(R, \alpha)$, then $(r_0 F) \in \mathcal{C}(r_0 R, \alpha)$. We will choose R such that $|A|^2 \leq R^{-2}$ holds on the initial surface $\mathcal{M}_0$.

The class $\mathcal{C}(R, \alpha)$ is well adapted to mean curvature flow:

Proposition 2.6 (i) *Given any smooth, closed, weakly 2-convex hypersurface immersion $\mathcal{M}_0$, the solution $\mathcal{M}_t$ of mean curvature flow is strictly 2-convex for each $t > 0$.*

(ii) *For every strictly 2-convex, smooth closed hypersurface $\mathcal{M}$ we can choose R and α such that $\mathcal{M} \in \mathcal{C}(R, \alpha)$ and $|A|^2 \leq R^{-2}$ holds everywhere on $\mathcal{M}$.*

(iii) *Each class $\mathcal{C}(R, \alpha)$ is invariant under smooth mean curvature flow.*

Proof. If we denote by $W(v_1, v_2)$ the Weingarten operator applied to two tangent vectors v_1, v_2 at any point, we have

$$\lambda_1 + \lambda_2 = \min\{W(e_1, e_1) + W(e_2, e_2) \, : \, |e_1| = |e_2| = 1, \; e_1 \perp e_2\}.$$

This shows that $\lambda_1 + \lambda_2$ is a concave function of the Weingarten operator, being the infimum of a family of linear maps. Therefore the inequality $\lambda_1 + \lambda_2 \geq \alpha_0 H$ describes a convex cone of matrices for $\alpha_0 \geq 0$. On the other hand, the components h_j^i of the Weingarten operator of a surface evolving by mean curvature flow satisfy the equation

$$\partial_t h_j^i = \Delta h_j^i + |A|^2 h_j^i$$

in view of Lemma 2.1(iii). Observe that the system of ODEs obtained dropping the Δh^i_j term above changes any matrix by a homothety and thus it leaves invariant any cone of matrices. Hence, we can apply the maximum principle for tensors (see [10, Sect. 4]) to obtain that the inequality is preserved for any $\alpha_0 \geq 0$. If $\lambda_1 + \lambda_2$ were not strictly positive for positive times, the strong maximum principle in [10, Sect. 4] would imply that $\lambda_1 + \lambda_2 = 0$ everywhere on $\mathcal{M}_0$. But this is impossible since there is at least one point on $\mathcal{M}_0 \to \mathbb{R}^{n+1}$ where all eigenvalues are strictly positive, proving (i). To see (ii), define for given $\mathcal{M}$ the constant R by $\sup |A|^2 = R^{-2}$, the existence of the other constants α_i then follows from compactness of $\mathcal{M}$. The first inequality of assertion (iii) was shown already, the other two inequalities follow from the evolution equations for H and $d\mu$ in Lemma 2.1(iv) and (ii). □

The inequality $|A|^2 \leq R^{-2}$, instead, is clearly not invariant under the flow; notice that this property does not appear in the definition of the class $\mathcal{C}(R, \alpha)$ and is only requested on the initial surface $\mathcal{M}_0$. In view of the previous proposition we will always assume from now on that our initial surface $\mathcal{M}_0$ satisfies $|A|^2 \leq R^{-2}$ and is in some class $\mathcal{C}(R, \alpha)$. Later in Proposition 3.22 we also show that $\mathcal{C}(R, \alpha)$ is preserved by the surgery procedure described there. We now note some easy properties of surfaces in this class:

Proposition 2.7 (i) *On any 2-convex surface of dimension $n \geq 3$ we have the estimate $\frac{1}{n}H^2 \leq |A|^2 \leq nH^2$.*
(ii) *The inequality $\lambda_1 + \lambda_2 \geq \alpha_0 H$ implies that $\lambda_i \geq \frac{\alpha_0}{2} H$, $i = 2, \ldots, n$.*
(iii) *The inequality $\lambda_1 + \lambda_2 \geq \alpha_0 H$ implies that*

$$|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 \geq \frac{\alpha_0^2}{8} H^2 |\nabla H|^2.$$

(iv) *For any initial surface $\mathcal{M}_0 \in \mathcal{C}(R, \alpha)$ there is an upper bound on the maximal time of existence given by $T \leq \frac{n}{2}\alpha_1^{-2} R^2$. An alternative upper bound is given by $T \leq D_0^2/2n$, where D_0 is the diameter of $\mathcal{M}_0$.*
(v) *If the initial surface has normalised curvature $|A|^2 \leq R^{-2}$ then the maximal time T of existence is bounded below by $T \geq (1/2)R^2$.*

Proof. To prove the first assertion, note the inequalities

$$-\lambda_1 = \lambda_1 - 2\lambda_1 \leq \lambda_1 + 2\lambda_2 \leq \lambda_1 + \lambda_2 + \lambda_3 \leq H$$

and also

$$\lambda_n \leq \lambda_1 + \lambda_2 + \lambda_n \leq H.$$

Thus, any 2-convex surface satisfies

$$-H \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq H, \tag{2.2}$$

which also implies

$$|A|^2 \le n(\max|\lambda_i|)^2 \le nH^2 \tag{2.3}$$

as required. The second claim is immediate from

$$\lambda_i \ge \frac{\lambda_1+\lambda_2}{2} \ge \frac{\alpha_0}{2}H, \quad i = 2, \dots, n. \tag{2.4}$$

To see (iii) notice that by the proof of Lemma 3.2 in [18]

$$|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 \ge \frac{1}{2}\sum_{i=1}^{n-1}\lambda_i^2|\nabla H|^2.$$

Moreover we have, by (2.4),

$$\sum_{i=1}^{n-1}\lambda_i^2 \ge \sum_{i=2}^{n-1}\lambda_i^2 \ge (n-2)\frac{\alpha_0^2}{4}H^2.$$

The first upper bound for the maximal time of existence follows from the evolution equation for the mean curvature in Lemma 2.1(iv) and comparison with the corresponding ODE. The second bound is the well known comparison principle with an enclosing round sphere. Similarly the lower bound for the time of existence follows from the evolution equation for $|A|^2$ in Lemma 2.1(v) and comparison with the corresponding ODE, in conjunction with standard regularity theory. □

To extend mean curvature flow beyond singularities we introduce in Sect. 3 a scaling invariant standard surgery procedure for necks which is inspired by Hamilton's work in [14]. A neck is a region of the hypersurface that is geometrically close to a cylinder in a quantitative way that is made precise in Sect. 3. The standard surgery replaces a portion of a neck by two spherical caps. We show in Sects. 3–6 that the parameters controlling this standard surgery can be chosen once and for all in such a way that the classes $\mathcal{C}(R, \alpha)$ are preserved and all other relevant a priori estimates remain valid. This allows us to combine smooth mean curvature flow with standard surgery as follows:

Mean curvature flow with surgeries is determined by an algorithm that assigns to each initial smooth closed two-convex hypersurface immersion $F_0 : \mathcal{M}_1 \to \mathbb{R}^{n+1}$ in some class $\mathcal{C}(R, \alpha)$ a sequence of intervals $[0, T_1], [T_1, T_2], \dots, [T_{N-1}, T_N]$, a sequence of manifolds $\mathcal{M}_i$, $1 \le i \le N$ and a sequence of smooth mean curvature flows $F_t^i : \mathcal{M}_i \to \mathbb{R}^{n+1}$, $t \in [T_{i-1}, T_i]$ such that the following is true:

(i) The initial hypersurface for the family F^1 is given by $F_0 : \mathcal{M}_1 \to \mathbb{R}^{n+1}$.
(ii) The initial hypersurface for the flow $F_t^i : \mathcal{M}_i \to \mathbb{R}^{n+1}$ on $[T_{i-1}, T_i]$ for $2 \le i \le N$ is obtained from $F_{T_{i-1}}^{i-1}$ by the following 2-step procedure

using the standard surgery from Sect. 3. In the first step a hypersurface $\hat{F}^{i-1}_{T_{i-1}} : \mathcal{M}_i \to \mathbb{R}^{n+1}$ is obtained from $F^{i-1}_{T_{i-1}} : \mathcal{M}_{i-1} \to \mathbb{R}^{n+1}$ by standard surgery replacing finitely many disjoint necks by two spherical caps. In the second step finitely many disconnected components are removed from the surface $\hat{F}^{i-1}_{T_{i-1}}$ constructed in the first step that are recognised as being diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$, resulting in the initial surface $F^i_{T_{i-1}}$ for smooth mean curvature flow on $[T_{i-1}, T_i]$.

We say that the mean curvature flow with surgeries *terminates after finitely many steps* at time T_N if either at time T_N all connected components of $F^N_{T_i}$ are recognised as being diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$, or if this is the case on $\hat{F}^N_{T_i}$ after carrying out the surgeries in the first step in (ii) above.

To make sure that the flow terminates after finitely many steps we have to carefully control the choice of times T_i and the scale of the surgery at these times. As we explain in detail in Sect. 8 we will show that for each class $\mathcal{C}(R, \alpha)$ we can choose three values $\omega_1, \omega_2, \omega_3 > 1$ depending only on α such that if we take any $H_1 \geq \omega_1 R^{-1}$ and then set $H_2 = \omega_2 H_1$, $H_3 = \omega_3 H_2$, the mean curvature flow with surgeries can be constructed with the following properties:

- The surgery times T_i, $1 \leq i \leq N$ are determined as the earliest times in $[T_{i-1}, T_i]$ where the mean curvature H has maximum value H_3.
- The maximum curvature $H^i_{max}(T_{i-1})$ of $F^i_{T_{i-1}}$ is reduced to H_2 by the surgery in step 2 of (ii) for each $1 \leq i \leq N$. In particular, the mean curvature on $\mathcal{M}_t$ will be uniformly bounded by H_3 throughout mean curvature flow with surgery.
- All surgeries in step 1 of (ii) are performed in regions of the surface where the mean curvature is approximately H_1, say $H_1/2 \leq H \leq 2H_1$.

We note that the algorithm described above is completely determined firstly by the choice of the parameters of the standard surgery in Sect. 3 depending only on n and secondly by the choice of the parameters H_1, H_2, H_3 above, which only have to lie above the threshold determined by the constants ω_i depending on α.

With these definitions our main theorem can then be stated as follows.

Theorem 2.8 *For any given initial surface $\mathcal{M}_0$ in some class $\mathcal{C}(R, \alpha)$ there exists a mean curvature flow with surgeries starting from $\mathcal{M}_0$ which terminates after a finite number of steps. All surfaces of the flow satisfy a uniform curvature bound determined by R, α and all time intervals have length bounded from below by a uniform constant depending only on R and α.*

3 Necks and surgery

We develop the notion of an (ε, k, L)-neck for regions of a hypersurface $\mathcal{M}^n \to \mathbb{R}^{n+1}$ which are ε-close to a standard cylinder of length $2L$ in

C^{k+2}-norm after appropriate rescaling. In a first step we show that the presence of such necks can be deduced from appropriate local estimates on the second fundamental form and its derivatives. We then show that necks can be parametrised in a canonical way with the help of constant mean curvature slices and harmonic mappings, leading to the concept of a maximally extended (ε, k, L)-neck. Finally we show that the class of surfaces $\mathcal{C}(R, \alpha)$ introduced in the previous section is invariant under the surgery provided they are performed on (ε, k, L)-necks with parameters in a suitable range of values.

The approach in this section is inspired by the work of Hamilton in [14]. In particular we recall that Hamilton defines on one hand *curvature necks*, which are regions with intrinsic curvature resembling that of a cylinder, and on the other hand *geometric necks*, which have an actual cylindrical parametrisation with metric close to the one a standard cylinder. At a first sight the two notions seem different, since the former depends on the pointwise behaviour of the curvature, while the latter is also related to the global properties of the region. However, Hamilton proves that they are basically equivalent; in particular, a large enough curvature neck possesses a suitable subset which can be parametrized as a geometric neck. It is useful to have these different notions at disposal. In fact, the results about the detection of necks are more easily stated for curvature necks, because they employ the a priori estimates on curvature quantities satisfied by the solutions of the flow. On the other hand, the surgery construction is only possible in a region known to be diffeomorphic to a cylinder, and therefore requires the notion of geometric neck.

In our case, the setting in an ambient manifold allows an easier recognition of a neck from curvature properties and suggests a different approach via normal coordinates to the specific surgery construction. For this reason, it is natural to give an extrinsic analogue of Hamilton's definitions, both of curvature necks (Definition 3.1) and of geometric necks (Definition 3.9). In addition, the property that any large enough curvature neck also admits a cylindrical parametrization (see Proposition 3.5) has an easier and more direct proof than in the intrinsic case.

Definition 3.1 (Extrinsic curvature necks) *Let $\mathcal{M}^n \to \mathbb{R}^{n+1}$ be a smooth hypersurface and $p \in \mathcal{M}^n$.*

(i) *We say that the extrinsic curvature is ε-spherical at p if the Weingarten map $W(p) : T_p\mathcal{M}^n \to T_p\mathcal{M}^n$ satisfies*

$$|W(p) - Id(p)| \leq \varepsilon, \tag{3.1}$$

where $Id(p)$ is the identity map on $T_p\mathcal{M}^n$.

(ii) *We say the extrinsic curvature is ε-cylindrical at p if there exists an orthonormal frame at p such that*

$$|W(p) - \bar{W}| \leq \varepsilon, \tag{3.2}$$

where $\bar{W}$ is the Weingarten map on the tangent space to $\mathbb{S}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n+1}$ in a standard frame.

(iii) *We say the extrinsic curvature is (ε, k)-parallel at p if*

$$|\nabla^l W(p)| \leq \varepsilon \quad \textit{for } 1 \leq l \leq k. \tag{3.3}$$

(iv) *We say the extrinsic curvature is (ε, k)-spherical on $\mathcal{M}^n$ if it is ε-spherical and (ε, k)-parallel at every point $p \in \mathcal{M}^n$. We say the extrinsic curvature is (ε, k, L)-cylindrical around p if it is ε-cylindrical and (ε, k)-parallel at every point in the intrinsic ball of radius L around p.*

(v) *We say the extrinsic curvature is (ε, k)-homothetically spherical or (ε, k, L)-homothetically cylindrical around p if there exists a scaling constant σ such that $\sigma\mathcal{M}$ satisfies the corresponding property in* (iv). *As in* [14] *we say that p lies at the center of an (ε, k, L)-extrinsic curvature neck if the extrinsic curvature is (ε, k, L)-homothetically cylindrical around p.*

In a first step we show that extrinsically spherical or cylindrical surfaces are also intrinsically spherical or cylindrical in the sense of Hamilton in [14]. In the following we assume the reader to be familiar with Sect. 3 of that article while trying to keep the exposition of this paper selfcontained. We will add the word *intrinsic* to mean that we are referring to Hamilton's definitions concerning the Riemann curvature tensor, applied to our induced metric.

Proposition 3.2 *For every $\varepsilon > 0$ there is $\varepsilon' > 0$ depending only on ε and n such that every $p \in \mathcal{M}^n$ which has (ε', k)-extrinsic spherical curvature also has (ε, k)-intrinsic spherical curvature with respect to the induced metric g.*

Proof. This is an immediate consequence of the Gauss equations, $R_{ijkl} = h_{ik}h_{jl} - h_{il}h_{jk}$. □

Proposition 3.3 *For every $\varepsilon > 0$ there is $\varepsilon' > 0$ depending only on ε and n such that every point p at the center of an (ε', k, L)-extrinsic curvature neck also lies at the center of an (ε, k, L)-intrinsic curvature neck.*

Proof. In a standard frame on a cylinder $\mathbb{S}^{n-1} \times \mathbb{R} \subset \mathbb{R}^{n+1}$ the second fundamental form equals $\{h_{ij}\} = \operatorname{diag}[0, 1, \ldots, 1]$. In view of the Gauss equations the Riemann curvature operator of g diagonalises in the same frame with eigenvalues equal to the sectional curvatures $\sigma_{ij} = R_{ijij} = h_{ii}h_{jj}$ for $i \neq j$ and the result follows easily. □

In the next step we show that (ε, k)-extrinsically spherical or cylindrical surfaces are positioned close to spheres or cylinders in $\mathbb{R}^{n+1}$.

Proposition 3.4 *For $k \geq 1$ and $0 < \varepsilon \leq \varepsilon(n)$ small enough any surface which is (ε, k)-spherical everywhere in a ball $B_d(p)$ of diameter $d \geq 4$ is a uniformly convex closed surface and can be written as a graph of a function $u : \mathbb{S}^n \to \mathbb{R}$ over the standard sphere of radius 1 in $\mathbb{R}^{n+1}$ with $\|u\|_{C^{k+2}} \leq c(n)\varepsilon$.*

Proof. The result is well-known even under much weaker assumptions, e.g. it follows from the exponential convergence of uniformly convex surfaces to spheres under mean curvature flow, see [16]. In the case $k \geq 1$ we sketch an elementary proof for the convenience of the reader. Choosing $0 < \varepsilon \leq \varepsilon(n)$ small enough we can ensure that all eigenvalues λ_i of the second fundamental form are between $9/10$ and $11/10$, ensuring uniform convexity. This also ensures that $\mathcal{M}$ is diffeomorphic to an embedded sphere satisfying a diameter bound, say $\operatorname{diam}(\mathcal{M}) \leq 4$, in view of Myers' theorem if $n \geq 2$. (In case $n = 1$ this argument is still true for simple closed curves). Now let

$$z(p) = p - \frac{n}{H(p)}\nu(p) \tag{3.4}$$

be an approximate center of the surface as seen from p. In view of $k \geq 1$ we get

$$|\nabla_i z(p)| = \left| e_i - \frac{n}{H(p)} W(p)(e_i) + \frac{n}{H^2(p)} \nabla_i H(p)\nu(p) \right| \leq c(n)\varepsilon \tag{3.5}$$

in view of the assumptions $|H(p) - n| \leq \varepsilon$, $|W(e_i) - e_i| \leq \varepsilon$ and $|\nabla H| \leq n\varepsilon$. Hence $z(p)$ varies at most by a fixed multiple of ε on $\mathcal{M}$ and we may choose an approximate sphere $\mathbb{S}^n_1(z(p_0))$ for some arbitrary but fixed $p_0 \in \mathcal{M}$. Then it follows easily that $\mathcal{M}$ is a graph over this sphere with the desired properties since the extrinsic curvature and its derivatives up to order k control the function u and its derivatives up to order $k + 2$. □

Proposition 3.5 *Let $k \geq 1$. For all $L \geq 10$ there exists $\varepsilon(n, L) > 0$ and $c(n, L)$ such that any point $p \in \mathcal{M}$ which lies at the center of an (ε, k, L)-extrinsic curvature neck with $0 < \varepsilon \leq \varepsilon(n, L)$ has a neighbourhood which after appropriate rescaling can be written as (cylindrical) graph of a function $u : \mathbb{S}^{n-1} \times [-(L-1), (L-1)] \to \mathbb{R}$ over some standard cylinder in $\mathbb{R}^{n+1}$, satisfying*

$$\|u\|_{C^{k+2}} \leq c(n, L)\varepsilon. \tag{3.6}$$

Proof. As in the previous result we first choose $\varepsilon(n)$ so small that $|\lambda_1| \leq 1/10$ and $|\lambda_i - 1| \leq 1/10$ for $i = 2, 3, \ldots, n$ everywhere in the (intrinsic) ball $B_L(p)$. Then in particular the eigenvector $e_1(p)$ corresponding to $\lambda_1(p)$ is well defined and we may consider the unique cylinder of radius 1, $\bar{F} : \mathbb{S}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n+1}$ touching $\mathcal{M}$ at p with equal orientation of its mean curvature vector and axis parallel to $e_1(p)$. In this alignment we can choose an orthonormal frame $e_1(p), \ldots, e_n(p), \nu(p) \in \mathbb{R}^{n+1}$ for $T_p\mathcal{M}$ agreeing with a standard frame for the cylinder.

Now consider normal coordinates for $(\mathcal{M}, g)$ at p,

$$G : \mathbb{R}^n \cong T_p\mathcal{M} \longrightarrow \mathcal{M} \xrightarrow{F} \mathbb{R}^{n+1} \tag{3.7}$$

with $G(x) = \exp_p^g(x)$ for $x \in \mathbb{R}^n \cong T_p\mathcal{M}$. Using radial parallel transport in these normal coordinates on the $e_i(p)$ we obtain an adapted orthonormal frame $e_i(r\xi)$, $\nu(r\xi) = \nu(F \circ G(r\xi))$, for $x = r\xi$, $|\xi| = 1$, in a neighbourhood of p. Abusing notation we denote by ξ both the direction in $\mathbb{S}^{n-1} \subset T_p\mathcal{M}$ as well as its parallel transport. Then for each fixed direction ξ the frames $e_i(r\xi)$, $\nu(r\xi)$ satisfy a system of ODEs in radial direction controlled by the second fundamental form:

$$\begin{aligned} \frac{d}{dr} e_i(r\xi) &= -\langle W(F \circ G(r\xi))(e_i(r\xi)), \xi\rangle \nu(r\xi), \\ \frac{d}{dr} \nu(r\xi) &= \langle W(F \circ G(r\xi))(\xi), e_i(r\xi)\rangle e_i(r\xi). \end{aligned} \tag{3.8}$$

The analogous procedure on the standard cylinder $\bar{F} : \mathbb{S}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n+1}$ attached at p leads to the same system of ODEs, only with $\bar{W}$ instead of $W(F \circ G(r\xi))$, for the standard adapted orthonormal frame $\bar{e}_i(r\xi)$, $\bar{\nu}(r\xi)$ there. In view of the gradient estimate $|\nabla W| \le \varepsilon$ we have

$$|W(p) - W(F \circ G(r\xi))| \le c(n)\varepsilon r. \tag{3.9}$$

Since the initial frame is the same by construction and $|W(p) - \bar{W}| \le \varepsilon$, standard results from ODE theory yield the estimate

$$\begin{aligned} \max_{\xi \in S^{n-1} \subset T_p\mathcal{M}} \left\{ \sum_{i=1}^n |e_i(r\xi) - \bar{e}_i(r\xi)| + |\nu(r\xi) - \bar{\nu}(r\xi)| \right\} \\ \le \varepsilon c(n, L)(\exp(c(n)r) - 1). \end{aligned} \tag{3.10}$$

The surface $\mathcal{M}$ can now be recovered from the frame by integration starting at p, leading for $r \le L$ to a corresponding estimate for the parametrisation

$$\|F \circ G(r\xi) - \bar{F} \circ \bar{G}(r\xi)\|_{C^{2+k}} \le c(n, L)\varepsilon. \tag{3.11}$$

To show that $\mathcal{M}$ closes up as a cylinder consider the y_1-coordinate in $\mathbb{R}^{n+1}$ associated with the direction $e_1(p) =: w$ and its levelsets $\Sigma_z := \{q \in \mathcal{M} \mid y_1(q) = z\}$. We may assume that $y_1(p) = 0$. For $r \le L$ we derive from (3.10) that $\sup_{|\xi|=1} |e_1(r\xi) - w| \le c(n, L)\varepsilon$. This implies, in particular, that for $\varepsilon \le \varepsilon(n, L)$ small enough we have $w = \nabla y_1 \neq 0$ and thus $\Sigma_z \cap B_L(p)$ is a regular smooth surface. Moreover, in view of (3.10) and (3.11) we may use the $e_i(r\xi)$, $\nu(r\xi)$ to construct an adapted orthonormal frame for Σ_z as a hypersurface in the plane $\{y_1 = z\}$, say $\tilde{e}_i(q)$, $\tilde{\nu}(q)$, $i = 2, \ldots, n$, with

$$\sup_{q \in \Sigma_z \cap B_L(p)} \left\{ \sum_{i=2}^n |\tilde{e}_i(q) - e_i(q)| + |\tilde{\nu}(q) - \nu(q)| \right\} \le c(n, L)\varepsilon. \tag{3.12}$$

Using again the assumptions on W and its derivatives and observing that in view of (3.12) the metric on $\mathcal{M}$ is close to the metric on the standard cylinder we see that for $|z| \le L-1$, $L \ge 10$, the surface $\Sigma_z \cap B_L(p) \subset \{x_1 = z\}$ contains an $(n-1)$-ball of diameter at least 4 and is $(c(n)\varepsilon, k)$-spherical in the sense of Proposition 3.4. Noting that $n \ge 3$, hence $(n-1) \ge 2$, we conclude that Σ_z is a closed uniformly convex surface close to a round $(n-1)$-sphere in the plane $\{y_1 = z\}$ which can be represented as the graph of some function $u(\cdot, z) : \mathbb{S}^{n-1} \to \mathbb{R}$ satisfying

$$\|u(\cdot, z)\|_{C^{k+2}} \le c(n, L)\varepsilon. \tag{3.13}$$

Noting $L \ge 10$ and the assumptions on W in $B_L(p)$ it is then easy to see that u can be extended to a map, say $u : \mathbb{S}^{n-1} \times [-(L-1), (L-1)] \to \mathbb{R}$ with the desired properties. □

Remark 3.6 We could have used Proposition 3.3 together with Hamilton's Theorem C3.2 in [14] to conclude immediately that $B_{(L-1)}(p) \subset \mathcal{M}$ has the structure of an intrinsic geometric neck. Note that in our proof y_1 replaces the function f constructed in Hamilton's proof of his result.

We will now follow [14] to construct normal parametrisations for maximally extended extrinsic necks. For a given topological neck, that is for a local diffeomorphism $N : \mathbb{S}^{n-1} \times [a, b] \to \mathcal{M}$ we denote by $r : [a, b] \to \mathbb{R}$ the average radius of the cross sections $\Sigma_z = N(\mathbb{S}^{n-1} \times \{z\})$ with respect to the pullback of the metric g on $\mathcal{M}$, i.e. $|\Sigma_z|_g = \sigma_{n-1} r(z)^{n-1}$, where σ_{n-1} is the area of the standard $(n-1)$- sphere of radius 1. We also denote by $\bar{g}$ the standard metric on the cylinder $\mathbb{S}^{n-1} \times [a, b]$. We recall

Definition 3.7 *The local diffeomorphism $N : \mathbb{S}^{n-1} \times [a, b] \to (\mathcal{M}, g)$ is called an (intrinsic) (ε, k)-cylindrical* geometric neck *if it satisfies the following conditions:*

(i) *The conformal metric $\hat{g} = r^{-2}(z)g$ satisfies the estimates*

$$|\hat{g} - \bar{g}|_{\bar{g}} \le \varepsilon, \quad |\bar{D}^j \hat{g}|_{\bar{g}} \le \varepsilon \quad \textit{for } 1 \le j \le k, \tag{3.14}$$

uniformly on $\mathbb{S}^{n-1} \times [a, b]$.

(ii) *The mean radius function $r : [a, b] \to \mathbb{R}$ satisfies the estimates*

$$\left| \left(\frac{d}{dz} \right)^j \log r(z) \right| \le \varepsilon \tag{3.15}$$

for all $1 \le j \le k$ everywhere on $[a, b]$.

Hamilton then proceeds to show that away from the boundary, for a suitable range of parameters (ε, k), each geometrical (ε, k)-cylindrical neck $N : \mathbb{S}^{n-1} \times [a, b] \to (\mathcal{M}, g)$ can be changed by diffeomorphism to a normal neck $\tilde{N} : \mathbb{S}^{n-1} \times [\tilde{a}, \tilde{b}] \to (\mathcal{M}, g)$ which is unique up to isometries of the standard cylinder, see [14, Theorem 2.2]. We recall the notion of an (intrinsic) normal neck.

Definition 3.8 *A local diffeomorphism* $N : \mathbb{S}^{n-1} \times [a, b] \to (\mathcal{M}, g)$ *is called* normal *if it satisfies the following conditions:*

(i) *Each cross section* $\Sigma_z = N(\mathbb{S}^{n-1} \times \{z\}) \subset (\mathcal{M}, g)$ *has constant mean curvature.*

(ii) *The restriction of* N *to each* $\mathbb{S}^{n-1} \times \{z\}$ *equipped with the standard metric is a harmonic map to* Σ_z *equipped with the metric induced by* g*, and*

(iia) *in case* $n = 3$ *only, the center of mass of the pull-back of* g *on* $\mathbb{S}^2 \times \{z\}$ *considered as a subset of* $\mathbb{R}^3 \times \{z\}$ *lies at the origin* $\{0\} \times \{z\}$.

(iii) *The volume of any subcylinder with respect to the pullback of* g *is given by*

$$Vol(S^{n-1} \times [v, w], g) = \sigma_{n-1} \int_v^w r(z)^n dz. \tag{3.16}$$

(iv) *For any Killing vector field* $\bar{V}$ *on* $\mathbb{S}^{n-1} \times \{z\}$ *we have that*

$$\int_{S^{n-1}\times\{z\}} \bar{g}(\bar{V}, U) d\mu = 0, \tag{3.17}$$

where U *is the unit normal vector field to* Σ_z *in* $(\mathcal{M}, g)$ *and* $d\mu$ *is the measure of the metric* $\bar{g}$ *on the standard cylinder.*

Now we adapt these definitions to include the extrinsic curvature.

Definition 3.9 *Let* $N : \mathbb{S}^{n-1} \times [a, b] \to (\mathcal{M}, g) \subset \mathbb{R}^{n+1}$ *be an (intrinsic)* (ε, k)*-cylindrical geometric neck in a smooth hypersurface* $F : \mathcal{M} \to \mathbb{R}^{n+1}$ *with induced metric* g *and Weingarten map* W*. We say that* N *is an* (ε, k)*-cylindrical* hypersurface neck *if in addition to the assumptions in Definition 3.7 it is true that*

$$|W(q) - r(z)^{-1}\bar{W}| \le \varepsilon r(z)^{-1} \qquad \textit{and}$$
$$|\nabla^l W(q)| \le \varepsilon r(z)^{-l-1}, \quad 1 \le l \le k, \tag{3.18}$$

for all $q \in \mathbb{S}^{n-1} \times \{z\}$ *and all* $z \in [a, b]$.

Remark 3.10 (i) Every point sufficiently far from the boundary of an (ε, k)-hypersurface neck lies at the center of some (ε', k', L)-extrinsic curvature neck.

(ii) As in the intrinsic case [14], in order to check that we have an (ε, k)-curvature neck, it suffices to find a fixed scaling constant r_0 such that the hypersurface $\tilde{F} = r_0^{-1} F : \mathcal{M} \to \mathbb{R}^{n+1}$ satisfies the conditions

$$\begin{aligned} &|\tilde{g} - \bar{g}| \le \varepsilon', \qquad |\bar{D}^j \tilde{g}| \le \varepsilon', \qquad 1 \le j \le k, \\ &|\tilde{W} - \bar{W}| \le \varepsilon', \quad |\bar{\nabla}^l \tilde{W}|_{\bar{g}} \le \varepsilon', \quad 1 \le l \le k, \end{aligned} \tag{3.19}$$

for suitable (ε', k'). The dependence of the scaling factor r on z allows to combine necks of different scales to long necks as in [14].

(iii) The parameter z along the cylinder determined by condition (iii) in Definition 3.8 has the advantage of being invariant under rescaling which simplifies the exposition of the standard surgery procedure. This is the reason why the exponent which appears in (3.16) is n, rather than $n-1$ which would appear more natural at first sight.

With the help of the above definitions we can now combine overlapping necks into long necks and define the notion of maximal normal hypersurface necks analogous to [14].

Definition 3.11 *We call an* (ε, k)*-cylindrical hypersurface neck* N *a* maximal *normal* (ε, k)*-cylindrical hypersurface neck if* N *is normal and if whenever* N^* *is another such neck with* $N = N^* \circ G$ *for some diffeomorphism* G *then the map* G *is onto.*

Since our definition of a (maximal) normal (ε, k)-hypersurface neck is an extension of the corresponding notion for an intrinsic neck in [14] we can now proceed exactly as in Lemma C2.1, Theorem C2.2 and Theorem C2.4 of that paper to obtain uniqueness, existence and overlapping properties for (maximal) normal parametrisations on (ε, k)-cylindrical hypersurface necks. We summarise these results in the following theorem.

Theorem 3.12 *Let* $F : \mathcal{M} \to \mathbb{R}^{n+1}$ *be a smooth closed hypersurface with* $n \geq 3$.

(i) *For any* $\delta > 0$ *we can choose* $\varepsilon > 0$ *and* k *so that if* $N : \mathbb{S}^{n-1} \times [a, b] \to \mathcal{M}$ *is an* (ε, k)*-cylindrical hypersurface neck with* $b - a \geq 3\delta$ *then we can find a normal neck* N^* *and a diffeomorphism* G *of the domain cylinder of* N^* *onto a region in the domain cylinder of* N *containing all points at least* δ *from the ends, such that* $N^* = N \circ G$.
(ii) *For any* $\delta > 0$ *and any* (ε', k') *we can choose* (ε, k) *so that the normal neck* N^* *in (i) is an* (ε', k')*-cylindrical hypersurface neck.*
(iii) *For* $0 < \varepsilon \leq \varepsilon(n)$ *sufficiently small and* $k \geq 1$*, if* N_1 *and* N_2 *are both normal necks which are* (ε, k)*-cylindrical hypersurface necks, and if there is a diffeomorphism* G *of the corresponding cylinders such that* $N_2 = G \circ N_1$*, then* G *is an isometry in the standard metrics on the cylinders.*
(iv) *For* $k \geq 1$ *and any* $\Lambda > 0$ *there is* $\tilde{\varepsilon}(\Lambda, n) > 0$ *such that any two normal* (ε, k)*-hypersurface necks* N_1, N_2 *with* $0 < \varepsilon \leq \tilde{\varepsilon}(n, \Lambda)$ *that overlap on some collar* $\mathbb{S}^{n-1} \times [z_0, z_0 + \Lambda]$ *agree there up to isometries of the standard cylinder and can be combined into a common normal* (ε, k)*-hypersurface neck.*
(v) *The normal neck* N^* *constructed in* (i) *is contained in a maximal normal* (ε, k)*- hypersurface neck unless the target hypersurface* $\mathcal{M}$ *is diffeomorphic to* $\mathbb{S}^{n-1} \times \mathbb{S}^1$.

Remark 3.13 (i) Different from [14, Theorem C2.5] we have $\mathbb{S}^{n-1} \times \mathbb{S}^1$ as the only possibility in (v) in view of the ambient geometry.

(ii) As in [14] the theorem implies that two different maximal normal (ε, k)-hypersurface necks with lengths at least 3δ in the standard metrics cannot overlap except within distance δ of their ends, if $0 < \varepsilon \leq \varepsilon(\delta, n)$ is small enough and $k \geq 1$, since they could otherwise be combined into a longer neck.

Combining Proposition 3.5 and Theorem 3.12 we can now prove that points at the center of an (ε, k, L)-extrinsic curvature neck have a neighbourhood contained in a suitable hypersurface neck.

Theorem 3.14 *For every (ε, k, L) with $L \geq 10$ there exist (ε', k') such that if the extrinsic curvature is (ε', k', L)-cylindrical around $p \in \mathcal{M}$, then p lies at the center of a normal (ε, k)-cylindrical hypersurface neck $N : \mathbb{S}^{n-1} \times [-(L-1), (L-1)] \to \mathcal{M}$, which is contained in a maximal normal (ε, k)-hypersurface neck unless the target hypersurface $\mathcal{M}$ is diffeomorphic to $\mathbb{S}^{n-1} \times \mathbb{S}^1$.*

Proof. From Proposition 3.5 we see that p has a neighbourhood which after rescaling can be written as a graph over the standard cylinder $\mathbb{S}^{n-1} \times [-(L-1), (L-1)]$ which is C^{k+2}-close to the standard cylinder. Then Theorem 3.12(i) yields a normal parametrisation and part (v) of the same theorem yields the extension to a maximal normal hypersurface neck. $\square$

Now suppose that $F : \mathcal{M} \to \mathbb{R}^{n+1}$ is a smooth closed hypersurface of dimension $n \geq 3$ as before and let

$$N : \mathbb{S}^{n-1} \times [a, b] \to \mathcal{M} \tag{3.20}$$

be a maximal normal (ε, k)-hypersurface neck, where (ε, k) is in a range where the conclusions of Theorem 3.12 hold. Let $z_0 \in [a, b]$ have sufficient distance to the ends of the neck, i.e. $[z_0 - 4\Lambda, z_0 + 4\Lambda] \subset [a, b]$ for some standard length $\Lambda > 0$ to be determined later.

For each such pair (N, z_0) and given parameters $0 < \tau < 1$, $B > 10\Lambda$ we now define the *standard surgery with parameters* τ, B at the cross section $\Sigma_{z_0} = N(\mathbb{S}^{n-1} \times \{z_0\})$, replacing the cylindrical image of $\mathbb{S}^{n-1} \times [z_0 - 4\Lambda, z_0 + 4\Lambda]$ smoothly by two properly adapted spherical caps. First let us denote by $\bar{C}_{z_0} : \mathbb{S}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n+1}$ the straight cylinder best approximating $\mathcal{M}$ at the cross section Σ_{z_0}: the radius of $\bar{C}$ is chosen as the mean radius $r(z_0) = r_0$, a point on its axis if given by the center of mass of Σ_{z_0} with its induced metric, and its axis is parallel to the average of the unit normal field to $\Sigma_{z_0} \subset (\mathcal{M}, g)$, taken again with respect to the induced metric. Then the standard surgery with parameters τ, B is performed as follows:

a) The surgery leaves the two collars $\mathbb{S}^{n-1} \times [a, z_0 - 3\Lambda]$ and $\mathbb{S}^{n-1} \times [z_0 + 3\Lambda, b]$ unchanged.
b) It replaces the two cylinders $N(\mathbb{S}^{n-1} \times [z_0 - 3\Lambda, z_0])$ and $N(\mathbb{S}^{n-1} \times [z_0, z_0 + 3\Lambda])$ by two n-balls attached smoothly to $\Sigma_{z_0 - 3\Lambda}$ and $\Sigma_{z_0 + 3\Lambda}$ respectively. In the following we will only describe the procedure at

the left end $[z_0 - 4\Lambda, z_0]$, the procedure on the right hand interval $[z_0, z_0+4\Lambda]$ is analogous. For convenience let $z_0-4\Lambda = 0$ and consider a normal parametrisation $N : \mathbb{S}^{n-1} \times [0, 4\Lambda] \to \mathcal{M}$ in the following.

c) Motivated by Hamilton [14] we use the function $u(z) \equiv r_0 \exp(-\frac{B}{z-\Lambda})$ on $[\Lambda, 3\Lambda]$ for $B > 10\Lambda$ in Gaussian normal coordinates to bend the surface inwards into a surface which is strictly convex on $\mathbb{S}^{n-1} \times [2\Lambda, 3\Lambda]$, for a parameter $0 < \tau < 1$:

$$\tilde{N}(\omega, z) := N(\omega, z) - \tau u(z)\nu(\omega, z). \tag{3.21}$$

d) To blend the resulting surface into an axially symmetric one, we choose a fixed smooth transition function $\varphi : [0, 4\Lambda] \to \mathbb{R}^+$ with $\varphi = 1$ on $[0, 2\Lambda]$, $\varphi = 0$ on $[3\Lambda, 4\Lambda]$ with $\varphi' \leq 0$. Denoting by $\tilde{C}_{z_0} : \mathbb{S}^{n-1} \times [0, 4\Lambda] \to \mathbb{R}^{n+1}$ the bending of the approximating cylinder defined earlier, $\tilde{C}_{z_0} = \bar{C}_{z_0}(\omega, z) - \tau u(z)\nu_{\bar{C}}(\omega, z)$, we then interpolate to obtain

$$\hat{N}(\omega, z) := \varphi(z)\tilde{N}(\omega, z) + (1 - \varphi(z))\tilde{C}_{z_0}(\omega, z). \tag{3.22}$$

We note that the function φ only depends on Λ, and that it can be defined in such a way that all its derivatives are smaller if Λ is larger. In particular, if we assume $\Lambda \geq 10$, each derivative of φ is bounded by some fixed constant.

e) At last we suitably change u on $[3\Lambda, 4\Lambda]$ to a function $\hat{u}$ to ensure that $\tau\hat{u}(z) \to r(z_0) = r_0$ as z approaches some $z_1 \in (3\Lambda, 4\Lambda]$, such that $\tilde{C}_{z_0}([3\Lambda, 4\Lambda])$ is a smoothly attached axially symmetric and uniformly convex cap. Since this last deformation on $[3\Lambda, 4\Lambda]$ only concerns the axisymmetric case, it can be made for each pair τ, B of parameters in such a way that on the attached convex cap there is some fixed upper bound for the curvature and each of its derivatives, independent of $\Lambda \geq 10$ and the surgery parameters τ, B.

We now need to compute how the bending of a neck affects its curvature. Given a neck $N : \mathbb{S}^{n-1} \times [a, b] \to \mathcal{M} \to \mathbb{R}^{n+1}$ we note the Weingarten relations

$$\begin{aligned} \frac{\partial^2 N}{\partial x^i \partial x^j} - \Gamma^k_{ij} \frac{\partial N}{\partial x^k} &= -h_{ij}\nu, \\ \frac{\partial \nu}{\partial x^i} &= h_{il} g^{lk} \frac{\partial N}{\partial x^k}, \end{aligned} \tag{3.23}$$

for $1 \leq i, j \leq n$. We assume the neck N to be in normal parametrisation and consider a positive scalar function $u : [a, b] \to \mathbb{R}$ of the scaling invariant standard length $z = x_1$ along the neck as well as a parameter $\tau \geq 0$ such that $\tau(|u(z)| + |u'(z)|) \leq r(z)$. Then the deformed neck is given by

$$\tilde{N}_\tau(p) = N(p) - \tau u(z)\nu(p) \tag{3.24}$$

for all $p = (\omega, z) \in \mathbb{S}^{n-1} \times [a, b]$ and the metric and normal of the new surface can be computed as follows.

Lemma 3.15 *The induced metric $\tilde g^\tau_{ij}$ and the normal $\tilde\nu_\tau$ of the deformed neck satisfy*

(i) $\tilde g^\tau_{ij} = g_{ij} + \delta^1_i\delta^1_j\tau^2(u')^2 + \tau^2u^2h_{il}h^l_j - 2\tau u h_{ij}$,
(ii) $\frac{d}{d\tau}\tilde g^\tau_{ij} = -2uh_{ij} + 2\tau\delta^1_i\delta^1_j(u')^2 + 2\tau u^2h_{il}h^l_j$,
(iii) $\frac{d}{d\tau}\sqrt{\det\tilde g^\tau} = \sqrt{\det\tilde g^\tau}\,\tilde g^{ij}(-uh_{ij} + \tau\delta^1_i\delta^1_j(u')^2 + \tau u^2h_{il}h^l_j)$,
(iv) $\frac{d}{d\tau}\tilde\nu_\tau = \langle\tilde\nu_\tau, \delta^1_l u'\nu + uh_{lk}g^{kj}\frac{\partial N}{\partial x^j}\rangle\tilde g^{lm}\frac{\partial\tilde N_\tau}{\partial x^m}$.

Proof. The first three identities are immediate consequences of $\tilde g^\tau_{ij} = \langle\frac{\partial\tilde N_\tau}{\partial x^i}, \frac{\partial\tilde N_\tau}{\partial x^i}\rangle$ and the Weingarten relations. To derive (iv) we compute

$$\frac{d}{d\tau}\tilde\nu_\tau = \left\langle\frac{d}{d\tau}\tilde\nu_\tau, \frac{\partial\tilde N_\tau}{\partial x^l}\right\rangle\tilde g^{lm}\frac{\partial\tilde N_\tau}{\partial x^m} = -\left\langle\tilde\nu_\tau, \frac{\partial}{\partial x^l}(-u\nu)\right\rangle\tilde g^{lm}\frac{\partial\tilde N_\tau}{\partial x^m} \tag{3.25}$$

and the result follows from the Weingarten relations since u only depends on z. □

Corollary 3.16 *For $k \geq 1$ and $0 \leq \varepsilon < \varepsilon_0$ suitably small there is a fixed constant $c > 0$ such that for all deformations $\tilde N_\tau$ given as in* (3.24) *of an (ε, k)-cylindrical neck in normal parametrisation we have the estimates*

(i) $|\tilde g^\tau_{ij} - (g_{ij} - 2\tau uh_{ij})| \leq c\tau^2(|u|^2 + |u'|^2)$,
(ii) $|\tilde g^{ij}_\tau - (g^{ij} + 2\tau uh^{ij})| \leq c\tau^2r^{-4}(|u|^2 + |u'|^2)$,
(iii) $|\sqrt{\det\tilde g^\tau} - \sqrt{\det g}(1 - \tau uH)| \leq c\tau^2r^{n-2}(|u|^2 + |u'|^2)$,
(iv) $|\tilde\nu_\tau - (\nu + \tau u'g^{1m}\frac{\partial N}{\partial x^m})| \leq c\tau^2r^{-2}(|u'|^2 + |u|^2)$,

everywhere on $\mathbb{S}^{n-1}\times[a, b]$.

Proof. Taking Definitions 3.7–3.9 into account, the relations (i)–(iii) are immediate consequences of Lemma 3.15(i)–(iii). To see the estimate for $\tilde\nu_\tau$ observe first that by Lemma 3.15(iv) we have an easy estimate $|\tilde\nu_\tau - \nu| \leq c\tau r^{-1}(|u'| + |u|)$. Using the lemma again we then derive that

$$\begin{aligned}\tilde\nu_\tau - \nu &= \int_0^\tau\left\langle\tilde\nu_\sigma, \delta^1_l u'\nu + uh_{kl}g^{kj}\frac{\partial N}{\partial x^j}\right\rangle\tilde g^{lm}\frac{\partial\tilde N_\sigma}{\partial x^m}d\sigma\\ &= \int_0^\tau\langle\tilde\nu_\sigma, \nu\rangle u'\tilde g^{1m}\frac{\partial\tilde N_\sigma}{\partial x^m}d\sigma + \int_0^\tau\left\langle\tilde\nu_\sigma, \frac{\partial N}{\partial x^j}\right\rangle uh^j_l\tilde g^{lm}\frac{\partial\tilde N_\sigma}{\partial x^m}d\sigma.\end{aligned} \tag{3.26}$$

The second term on the RHS can then be estimated by $c\tau^2r^{-2}|u|(|u| + |u'|)$ and the result follows from $|\tilde g^{1m} - g^{1m}| \leq c\tau r^{-3}(|u| + \tau r^{-1}(|u|^2 + |u'|^2)$ as well as $\frac{\partial\tilde N}{\partial x^m} = \frac{\partial N}{\partial x^m} - \tau\delta^n_m u'\nu - \tau u\frac{\partial\nu}{\partial x^m}$ since $\tau(|u| + |u'|) \leq r$. □

We are now ready to control the second fundamental form of the bent surfaces.

Proposition 3.17 *We can choose* $\varepsilon_0 > 0$ *small enough and a fixed constant* $c > 0$ *such that for all deformations* $\tilde{N}_\tau$ *of* (ε, k)*-cylindrical necks in normal parametrisation with* $0 < \varepsilon \le \varepsilon_0$, $k \ge 1$, *we have the estimates*

(i) $|\tilde{h}_{ij} - (h_{ij} + \tau\delta_i^1\delta_j^1 u'' - \tau u h_{il} h_j^l)| \le c\tau\varepsilon|u'| + c\tau^2 r^{-1}(|u''|^2 + |u'|^2 + |u|^2)$,

(ii) $|\tilde{h}_j^i - (h_j^i + \tau g^{1i}\delta_j^1 u'' + \tau u h_l^i h_j^l)| \le c\tau r^{-2}\varepsilon|u'| + c\tau^2 r^{-3}(|u''|^2 + |u'|^2 + |u|^2)$

everywhere on $\mathbb{S}^{n-1} \times [a, b]$.

Proof. From the Weingarten relations we compute

$$\begin{aligned} \tilde{h}_{ij}^\tau &= -\left\langle \tilde{\nu}_\tau, \frac{\partial^2 \tilde{N}_\tau}{\partial x^i \partial x^j} \right\rangle = -\left\langle \tilde{\nu}_\tau, \frac{\partial}{\partial x^i}\left(\frac{\partial N}{\partial x^j} - \tau\delta_j^1 u'\nu - \tau u \frac{\partial \nu}{\partial x^j} \right) \right\rangle \\ &= -\left\langle \tilde{\nu}_\tau, \frac{\partial^2 N}{\partial x^i \partial x^j} - \tau\delta_i^1\delta_j^1 u''\nu - \tau u' \left(\delta_i^1 \frac{\partial \nu}{\partial x^j} + \delta_j^1 \frac{\partial \nu}{\partial x^i} \right) \right\rangle \\ &\quad + \tau u \left\langle \tilde{\nu}_\tau, \frac{\partial}{\partial x^i}\left(h_j^l \frac{\partial N}{\partial x^l} \right) \right\rangle. \end{aligned} \tag{3.27}$$

Replacing $\tilde{\nu}$ by ν we arrive at

$$\tilde{h}_{ij} = h_{ij} + \tau\delta_i^1\delta_j^1 u'' - \tau u h_{il} h_j^l + \langle \tilde{\nu} - \nu, B_{ij} \rangle \tag{3.28}$$

where

$$\begin{aligned} B_{ij} &= -\frac{\partial^2 N}{\partial x^i \partial x^j} + \tau\delta_i^1\delta_j^1 u''\nu + \tau u' \left(\delta_i^1 \frac{\partial \nu}{\partial x^j} + \delta_j^1 \frac{\partial \nu}{\partial x^i} \right) \\ &\quad + \tau u \left(\frac{\partial}{\partial x^i} h_j^l \frac{\partial N}{\partial x^l} + h_j^l \frac{\partial^2 N}{\partial x^i \partial x^l} \right). \end{aligned} \tag{3.29}$$

In view of Corollary 3.16 the first term on the RHS leads to

$$-\left\langle \frac{\partial^2 N}{\partial x^i \partial x^j}, \frac{\partial N}{\partial x^m} \right\rangle g^{1m} u'(z)\tau, \tag{3.30}$$

up to terms which can be estimated by $\tau^2 r^{-1}(|u'(z)|^2 + |u(z)|^2)$ as claimed. Now

$$\left\langle \frac{\partial^2 N}{\partial x^i \partial x^j}, \frac{\partial N}{\partial x^m} \right\rangle g^{1m} = \Gamma_{ij}^1, \tag{3.31}$$

a Christoffel symbol for the metric in the normal parametrisation which can be estimated by $c\varepsilon$. All remaining terms involving B_{ij} can be estimated as claimed in view of Corollary 3.16(i).

For the second estimate we use again that $\tau(|u(z)| + |u'(z)|) \le r(z)$ as well as the relation

$$\left| \tilde{g}_\tau^{ij} - (g^{ij} + 2\tau u h^{ij}) \right| \le c\tau^2 r^{-4}(|u|^2 + |u'|^2) \tag{3.32}$$

from Corollary 3.16(ii). This immediately implies the desired relation for $\tilde{h}_j^i = \tilde{g}^{ik}\tilde{h}_{kj}$ in view of (i). □

For a neck $N : \mathbb{S}^{n-1} \times [0, 4\Lambda] \to \mathcal{M} \to \mathbb{R}^{n+1}$ in normal parametrisation we now choose the function

$$u(z) = r_0 f(z) \equiv r_0 \exp\left(-\frac{B}{z-\Lambda}\right), \quad z \in [\Lambda, 4\Lambda], \tag{3.33}$$

where $r_0 = r(z_0) = r(4\lambda)$ is the scale of the neck and the constant $B > 1$ is to be chosen. We will use the following properties of the function f, see also Hamilton [10, Lemma D2.4]:

Lemma 3.18 (i) *For $z \in [\Lambda, 4\Lambda]$ the function $f(z) = \exp(-\frac{B}{z-\Lambda})$ satisfies the estimates*

$$f(z) \leq 1, \quad f''(z) \leq \frac{5}{B^2}, \quad f'(z) = \frac{B}{(z-\Lambda)^2} f(z), \quad |f'(z)|^2 \leq \frac{5}{B^2} f(z), \tag{3.34}$$

as well as, for $B \geq 12\Lambda$,

$$f''(z) \geq \frac{B^2}{2(z-\Lambda)^4} f(z). \tag{3.35}$$

(ii) *In particular, for any Λ and any $\delta > 0$ we can choose B large enough that*

$$f(z) \leq \delta f''(z), \quad |f'(z)|(1 + |f'(z)|) \leq \delta f''(z), \quad |f''(z)| \leq \delta \tag{3.36}$$

hold everywhere on $[\Lambda, 4\Lambda]$.

We then have the following result.

Theorem 3.19 *For any $\theta > 0$ and any $\Lambda \geq 10$ we may choose $k \geq 1$ and $0 \leq \varepsilon < \varepsilon_0$, and then fix $0 < \tau_0 < 1$ small enough and B large enough such that the second fundamental form of the deformed surface $\tilde{N}^{\tau_0}(p) = N(p) - \tau_0 r_0 f(z)\nu(p)$ satisfies*

(i) $|\tilde{h}_{ij}^{\tau_0} - (h_{ij} + \tau_0 r_0 \delta_i^1 \delta_j^1 f'' - \tau_0 r_0 f h_{il} h_j^l)| \leq \theta \tau_0 r_0 f''$,

(ii) $|\tilde{h}_j^{\tau_0 i} - (h_j^i + \tau_0 r_0 g^{1i} \delta_j^1 f'' + \tau_0 r_0 f h_l^i h_j^l)| \leq \theta \tau_0 r^{-1} f''$

on $[\Lambda, 4\Lambda]$ for any (ε, k)-cylindrical neck $N : \mathbb{S}^{n-1} \times [0, 4\Lambda] \to \mathcal{M} \subset \mathbb{R}^{n+1}$ with normal parametrisation.

Proof. Given the standard length Λ we may choose for $k \geq 1$ the parameter $0 < \varepsilon \leq \varepsilon_0$ sufficiently small to make the mean radius r as close to r_0 on $[0, 4\Lambda]$ as we like, i.e. $|r/r_0 - 1| \leq \theta/10$, in view of Definition 3.7(ii). The result then is an immediate consequence of Proposition 3.17 and Lemma 3.18. ◻

Remark 3.20 As in [10, Theorem 2.6] we can also state the estimates in terms of an orthonormal frame diagonalising the second fundamental form at some point of the neck, $e_1, e_2, \ldots, e_n$, such that e_1 corresponds to the smallest eigenvalue. Then we see as in [10] that the deformation function f satisfies

$$|D_i f| \leq \frac{\eta}{r_0} f'', \quad |D_i D_j f| \leq \frac{\eta}{r_0^2} f'', \quad 1 \leq i \leq n,\ 2 \leq j \leq n,$$
$$\left| D_1 D_1 f - \frac{1}{r_0^2} f''(z) \right| \leq \frac{\eta}{r_0^2} f'' \tag{3.37}$$

for any $\eta > 0$ and Λ, provided (ε, k) are suitable and B is large enough. Thus, in terms of the eigenvalues $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_n$ we get

$$\left|\tilde{\lambda}_i - \left(\lambda_i + \tau_0 u \lambda_i^2\right)\right| \leq \theta \tau_0 D_1 D_1 u, \quad 2 \leq i \leq n,$$
$$\left|\tilde{\lambda}_1 - \left(\lambda_1 + \tau_0 D_1 D_1 u + \tau_0 u \lambda_1^2\right)\right| \leq \theta \tau_0 D_1 D_1 u, \tag{3.38}$$

everywhere on $[\Lambda, 4\Lambda]$. Similarly we get for the induced volume element

$$\left|\sqrt{\det \tilde{g}} - \sqrt{\det g}(1 - \tau_0 u H)\right| \leq \theta \tau_0 r_0^{n-1} u \tag{3.39}$$

as well as

$$\left|\sqrt{\det \tilde{g}} - \sqrt{\det g}(1 - \tau_0 u H)\right| \leq \theta \tau_0^2 r_0^{n+1} D_1 D_1 u, \tag{3.40}$$

everywhere on $[\Lambda, 4\Lambda]$.

Corollary 3.21 *For any $\Lambda \geq 10$ we may choose $k_0 \geq 1$, $0 < \varepsilon_0$, $0 < \tau_0 < 1$ and B large enough such that for all $0 < \varepsilon \leq \varepsilon_0$, $k \geq k_0$ large enough the deformed surface $\tilde{N}_{\tau_0}$ satisfies:*

(i) $\tilde{H} \geq H$, $\tilde{\lambda}_1 + \tilde{\lambda}_2 \geq \lambda_1 + \lambda_2$, $\sqrt{\det \tilde{g}} \leq \sqrt{\det g}$ *on* $[\Lambda, 4\Lambda]$,
(ii) $\tilde{\lambda}_1 \geq \frac{1}{2}\tau_0 D_1 D_1(r_0 u)$, $\tilde{\lambda}_1 + \tilde{\lambda}_2 \geq \lambda_1 + \lambda_2 + \frac{1}{2}\tau_0 D_1 D_1(r_0 u)$ *on* $[2\Lambda, 3\Lambda]$,
(iii) $(\tilde{\lambda}_1 + \tilde{\lambda}_2)/\tilde{H} \geq (\lambda_1 + \lambda_2)/H$ *on* $[\Lambda, 4\Lambda]$,
(iv) $\tilde{H} \geq H + \frac{1}{2}\tau_0 D_1 D_1(r_0 u)$, $\sqrt{\det \tilde{g}} \leq \sqrt{\det g}(1 - \frac{1}{2}\tau_0 u H)$ *on* $[2\Lambda, 3\Lambda]$.

Proof. This is an immediate consequence of the previous remark. □

Having shown in Proposition 2.6 that the class $\mathcal{C}(R, \alpha)$ of surfaces is invariant under smooth mean curvature flow we are now ready to extend this to mean curvature flow with surgeries:

Theorem 3.22 *For any $\Lambda \geq 10$ we may choose $k_0 \geq 1$, $0 < \varepsilon_0$, $0 < \tau_0$ small enough and B large enough such that for all $\alpha = (\alpha_0, \alpha_1, \alpha_2)$ and $R > 0$ the class $\mathcal{C}(R, \alpha)$ is invariant under standard surgery with parameters τ_0, B on a normal (ε, k)-hypersurface neck $N : \mathbb{S}^{n-1} \times [-4\Lambda, 4\Lambda] \to \mathcal{M}$, for all $0 < \varepsilon \leq \varepsilon_0$, and $k \geq k_0$.*

Proof. Note that in the region $[\Lambda, 2\Lambda]$ the claim is immediate from Corollary 3.21(i) and the fact that on an approximate cylinder λ_2 roughly equals $\frac{1}{(n-1)} H > \frac{1}{n} H$. In the region $[2\Lambda, 3\Lambda]$ we can ensure that the interpolated surface

$$\hat{N}(\omega, z) := \varphi(z)\tilde{N}(\omega, z) + (1 - \varphi(z))\tilde{C}_{z_0}(\omega, z) \tag{3.41}$$

is arbitrarily close to $\tilde{N}$ in any norm if $k \geq 1$ and ε are chosen appropriately. Hence $\hat{N}$ will also satisfy the estimates on $H, \lambda_1 + \lambda_2, \sqrt{\det g}$ of Corollary 3.21(ii) and (iii) implying the desired inequalities there. Finally we can smoothly attach the strictly convex cap in $[3\Lambda, 4\Lambda]$ increasing the curvature and decreasing the area further by making an appropriate choice of the function $\hat{u}$ there. □

At this stage we can already fix the parameters $\Lambda = 10$ and $k_0 = 2$ once and for all. The parameters B and τ_0 will be fixed depending only on n such that Theorem 3.22 is valid and the algebraic conditions implying (4.19) and (5.13) in the next sections are satisfied. In the following chapters we show that the crucial estimates controlling convexity, roundness of necks and gradient of curvature can be established for both smooth mean curvature flow and mean curvature flow with standard surgeries with these parameters, provided $0 < \varepsilon \leq \varepsilon_0$ is chosen small.

We conclude this section by showing that topological properties of $\mathcal{M}$ before the surgery can be recovered from the properties of the surface $\tilde{\mathcal{M}}$ after the surgery. We first note that surgery amounts to no more than "cutting handles".

Proposition 3.23 *There is a range of parameters $\Lambda \geq 10$, $0 < \varepsilon \leq \varepsilon_0$ and $k \geq k_0$ depending only on n, such that the following is true. Suppose standard surgery is performed on a normal (ε, k)-hypersurface neck $N : \mathbb{S}^{n-1} \times [-4\Lambda, 4\Lambda] \to \mathcal{M}$ in some connected smooth closed immersed hypersurface $F : \mathcal{M} \to \mathbb{R}^{n+1}$ resulting in a new smooth hypersurface $\tilde{\mathcal{M}}$. If $\tilde{\mathcal{M}}$ is connected then $\mathcal{M}$ is diffeomorphic to the manifold obtained from $\tilde{\mathcal{M}}$ by a standard connected sum with itself. If $\tilde{\mathcal{M}}$ is disconnected with the two components $\tilde{\mathcal{M}}_1$ and $\tilde{\mathcal{M}}_2$ then $\mathcal{M}$ is diffeomorphic to the standard connected sum of $\tilde{\mathcal{M}}_1$ and $\tilde{\mathcal{M}}_2$. In particular, if $\tilde{\mathcal{M}}$ is disconnected and $\tilde{\mathcal{M}}_2$ is diffeomorphic to $\mathbb{S}^n$, then $\tilde{\mathcal{M}}_1$ is diffeomorphic to $\mathcal{M}$.*

Proof. This is clear from the construction: the two open n-discs attached by the surgery are diffeomorphic to the standard disc and the two collars $[-4\Lambda, 0]$, $[0, 4\Lambda]$ are (ε, k)-close to the standard cylinder. □

We now show that embedded 2-convex surfaces in $\mathbb{R}^{n+1}$ are still embedded after surgery. It is well known that such surfaces separate $\mathbb{R}^{n+1}$ into a bounded and an unbounded region. We show that the diffeomorphism type of the region enclosed before the surgery can be recovered from the region enclosed after the surgery. For possible future use we state the following lemma in case of general k-convex surfaces, i.e. surfaces where the sum of the smallest k principal curvatures is positive everywhere.

Lemma 3.24 *Let $\mathcal{M}^n = \mathcal{M} \subset \mathbb{R}^{n+1}$, $n \geq 3$, be a connected closed hypersurface which is smoothly embedded in $\mathbb{R}^{n+1}$. Suppose that $\mathcal{M}$ is strictly k-convex for some $1 \leq k \leq n-1$. Let $E^n = E \subset \mathbb{R}^{n+1}$ be a hyperplane transverse to $\mathcal{M}$ such that $\emptyset \neq \Sigma^{n-1} = \Sigma = E \cap \mathcal{M}$ is a smooth closed hypersurface of E. Then each component of Σ is strictly k-convex and bounds a region in E that does not contain another component of Σ.*

Proof. Let $e_1, \dots, e_{n-1}, \gamma_E$ be a local adapted orthonormal frame of $\Sigma = E \cap \mathcal{M}$ in E. Since E is transverse to $\mathcal{M}$ we may choose γ_E such that $\langle \nu, \gamma_E \rangle > 0$ everywhere. The second fundamental form A of $\mathcal{M} \subset \mathbb{R}^{n+1}$ and A^E of $\Sigma \subset E$ are then given by

$$A(e_i, e_j) = -\langle \bar{\nabla}_{e_i} e_j, \nu \rangle, \quad A^E(e_i, e_j) = -\langle \bar{\nabla}_{e_i} e_j, \gamma_E \rangle, \quad 1 \leq i, j \leq n-1.$$

If $w \in \mathbb{R}^{n+1}$, $|w| = 1$, is a unit normal to E we have $\nu = \langle \nu, \gamma_E \rangle \gamma_E + \langle \nu, w \rangle w$ such that

$$A(e_i, e_j) = -\langle \nu, \gamma_E \rangle \langle \bar{\nabla}_{e_i} e_j, \gamma_E \rangle - \langle \nu, w \rangle \langle \bar{\nabla}_{e_i} e_j, w \rangle = \langle \nu, \gamma_E \rangle A^E(e_i, e_j).$$

It follows immediately that Σ is again k-convex, since the sum of the smallest k eigenvalues can be characterised by

$$\lambda_1 + \cdots + \lambda_k = \min \left\{ \sum_{i=1}^{k} A(X_i, X_i) \,\middle|\, X_i \perp X_j, |X_i| = 1, 1 \leq i, j \leq k \right\}.$$

It also follows that the mean curvature vectors $\overrightarrow{H} = -H\nu$ of $\mathcal{M} \subset \mathbb{R}^{n+1}$ and $\overrightarrow{H}^E = -H^E \gamma_E$ of $\Sigma \subset E^n$ have a positive angle since $\langle \overrightarrow{H}, \overrightarrow{H_E} \rangle = HH_e \langle \nu, \gamma_E \rangle > 0$. To prove the second claim of the lemma, recall a standard result from topology [15, Theorems 4.4.4 and 4.4.6], that $\mathcal{M}$ separates $\mathbb{R}^{n+1}$ into exactly two components U^1, U^2, such that $\mathcal{M}$ is the topological boundary of each component. Let U^1 be the bounded component and U^2 be the unbounded component. Similarly, each of the finitely many connected components Σ_l, $1 \leq l \leq N$, of $\Sigma \subset E$ separates E into exactly two components Ω_l^1, Ω_l^2 with common topological boundary Σ_l; let Ω_l^1 be bounded and Ω_l^2 be unbounded. Now suppose that for some component Σ_l the region Ω_l^1 contains some other Σ_k, $k \neq l$. Among these choose k such that $d = \operatorname{dist}(\Sigma_k, \Sigma_l)$ is smallest and pick $p \in \Sigma_l$, $q \in \Sigma_k$ such that this distance is attained on the straight line segment $[p, q]$ from p to q: $d = \operatorname{dist}(p, q)$. It follows that the open line segment (p, q) satisfies $(p, q) \subset (\Omega_l^1 \cap \Omega_k^2)$. Furthermore, $(p, q) \subset U^1$, since the mean curvature vector $\overrightarrow{H}(p)$ of $\mathcal{M}$ at p has a positive angle with the mean curvature vector $\overrightarrow{H}^E(p)$ of Σ at p and since the mean curvature vector points into the bounded region. But this leads to a contradiction to the direction of the mean curvature vectors $\overrightarrow{H}(q)$ and $\overrightarrow{H}^E(q)$ at q: $\overrightarrow{H}^E(p)$ must point in direction $(q-p)$ into the bounded region Ω_k^1 while $\overrightarrow{H}(q)$ points into the bounded region U^1 just found to contain (p, q). Then the projection of $\overrightarrow{H}(q)$ into E points into direction $(p-q)$ contradicting the positive angle with $\overrightarrow{H}^E(q)$. □

We are now ready to control the solid tube enclosed by a normal hypersurface neck.

Proposition 3.25 *Given a normal (ε, k)-hypersurface neck $N : \mathbb{S}^{n-1} \times [-4\Lambda, 4\Lambda] \to \mathcal{M}$ in some smooth hypersurface $\mathcal{M} \subset \mathbb{R}^{n+1}$ with sufficiently good parameters $\Lambda \geq 10$, $0 < \varepsilon \leq \varepsilon_0$ and $k \geq k_0$ depending only on n, there is a unique local diffeomorphism*

$$G : \bar{B}_1^n \times [-4\Lambda, 4\Lambda] \to \mathbb{R}^{n+1} \tag{3.42}$$

with the following properties:

(i) *The restriction of G to the boundary $\mathbb{S}^{n-1} \times [-4\Lambda, 4\Lambda]$ coincides with N.*
(ii) *Each disc $G(\bar{B}_1^n \times \{z_0\}) \subset \mathbb{R}^{n+1}$ is an embedded area minimizing hypersurface.*
(iii) *The restriction of G to each disc $\bar{B}_1^n \times \{z_0\}$ is a harmonic diffeomorphism.*
(iv) *The diffeomorphism G is ε-close in the C^{k+1}-norm to the standard isometric embedding of some solid cylinder $\bar{B}_1^n \times [-4\Lambda, 4\Lambda]$ in $\mathbb{R}^{n+1}$.*

We call G the normal solid tube associated with the normal neck N.

Proof. Since the boundary data $N(\mathbb{S}^{n-1} \times \{z_0\})$ are very close to some round $(n-1)$-sphere in $\mathbb{R}^{n+1}$, they have a convex projection onto a suitable hyperplane E_{z_0}. It is well known that the minimal surface equation can then be solved with these boundary data to produce the unique area minimizing disc as a graph over E_{z_0}. By regularity theory this surface will be smooth in the interior and of class $C^{k+1,\alpha}$ up to the boundary for all $0 < \alpha < 1$. We can then turn the graphical representation into a harmonic mapping with the claimed C^{k+1}-regularity by means of the implicit function theorem since the boundary data are close to the data of a standard $(n-1)$-sphere. □

Theorem 3.26 *There is a range of parameters $\Lambda \geq 10$, $0 < \varepsilon \leq \varepsilon_0$ and $k \geq k_0$ depending only on n, such that the following is true. Suppose $\mathcal{M} \subset \mathbb{R}^{n+1}$, $n \geq 3$ is a connected, smooth, closed and embedded hypersurface which is strictly 2-convex. Let U be the closed bounded region enclosed by $\mathcal{M}$.*

(i) *If standard surgery is performed on a normal (ε, k)-hypersurface neck $N : \mathbb{S}^{n-1} \times [-4\Lambda, 4\Lambda] \to \mathcal{M}$ then the resulting hypersurface $\tilde{\mathcal{M}}$ is again embedded.*

(ii) *If $\tilde{\mathcal{M}}$ is connected with the resulting enclosed bounded region $\tilde{U}$, then the region U is diffeomorphic to a connected sum of $\tilde{U}$ with itself. If $\tilde{U}$ is disconnected consisting of two disjoint bounded regions $\tilde{U}^1$ and $\tilde{U}^2$, then U is diffeomorphic to the connected sum of $\tilde{U}^1$ and $\tilde{U}^2$. In particular, if $\tilde{U}^2$ is diffeomorphic to a standard closed disc $\bar{B}_1^n \subset \mathbb{R}^{n+1}$ then U is diffeomorphic to $\tilde{U}^1$.*

Proof. Claim (i) follows from the fact that for a given normal (ε, k)-hypersurface neck N the interior of the solid tube G associated with N does not intersect $\mathcal{M}$ and hence the surgery construction does not destroy the embeddedness of $\mathcal{M}$. Indeed, suppose that there is some $z_0 \in [-3\Lambda, 3\Lambda]$ such that $\mathcal{M} \cap G(B_1^n \times \{z_0\})$ is not empty. Since $\mathcal{M}$ is smooth and the neck is an approximate cylinder we can then also find a plane $E^n = E$ close to this disc which intersects the 2-convex surface $\mathcal{M}$ transversely both in an approximate sphere close to $G(\mathbb{S}^{n-1} \times \{z_0\})$ and in another component in the interior of that sphere, contradicting Lemma 3.24. To prove the second statement we notice that in our surgery construction the convex caps attached in the neck intervals $[3\Lambda, 4\Lambda]$ and $[4\Lambda, 5\Lambda]$ bound convex regions that are diffeomorphic to the standard half ball $(\bar{B}_1^n)^+ \subset \mathbb{R}^{n+1}$. They are attached along a solid collar with boundary given by the solid normal tube G associated with N which is (ε, k)-close to the standard solid cylinder by Proposition 3.25. This ensures that the inverse operation to standard surgery is the construction of a connected sum by attaching a solid handle to $\tilde{U}$ from the outside. □

4 Convexity estimates in the presence of surgery

In this chapter we prove that the a priori convexity estimates of our previous paper [19] are still valid in a mean curvature flow with surgeries provided standard surgery is done on (ε, k)-necks with $k \geq 2$ and $0 < \varepsilon \leq \varepsilon_0$ small depending only on n.

In our previous paper [19] we proved the following lower bound for the elementary symmetric functions S_m of the principal curvatures.

Theorem 4.1 *Let $\mathcal{M}_t$, $t \in [0, T[$ be a family of smooth closed n-dimensional surfaces immersed in $\mathbb{R}^{n+1}$ evolving by mean curvature. Suppose that $\mathcal{M}_0$ has positive mean curvature. Then, for any $\delta > 0$, there exists $C_\delta = C_\delta(\mathcal{M}_0) > 0$ such that, for all $m = 2, \ldots, n$, we have*

$$S_m \geq -\delta H^m - C_\delta \quad \text{on } \mathcal{M}_t, \ \forall t \in [0, T[. \tag{4.1}$$

Here we want to show that the estimates (4.1) still hold for mean curvature flow with surgeries in a class $\mathcal{C}(R, \alpha)$ with constants depending only on R and α, provided the surgery parameters are chosen appropriately for this class. In order to deal with the presence of surgeries and in order to explicitly state the precise dependence of all constants on the main parameters R, α it will be necessary to review the original proof of these estimates in [19].

Remark 4.2 The convexity estimates in this chapter do not need the 2-convexity and hold in the class of mean convex surfaces. They will not depend on α_0 but only on α_1, α_2, the initial curvature bound $|A|^2 \leq R^{-2}$ as well as an upper bound on the ratio $|A|^2/H^2$ which is well known to be preserved for mean convex surfaces. The case of 2-convex surfaces with $n \geq 3$ is just a special case with $|A|^2 \leq nH^2$, compare Proposition 2.7(i).

Let us start with the case $m = 2$ which was first done in [18] and is technically simpler. We introduced the following function on surfaces with positive mean curvature.

$$f_{\sigma,\eta} = \frac{|A|^2 - (1+\eta)H^2}{H^{2-\sigma}}, \tag{4.2}$$

where σ, η are small positive parameters. An important property of this function is that suitable L^p-norms of its positive part are nonincreasing under the flow, as shown in [18, Prop. 3.6]):

Theorem 4.3 *Let $\mathcal{M}_0$ be a closed n-dimensional immersed surface such that*

$$H \geq \beta_1 |A| > 0 \quad \text{on } \mathcal{M}_0 \tag{4.3}$$

for some $1 \geq \beta_1 > 0$, and let $\mathcal{M}_t$ be the smooth evolution of $\mathcal{M}_0$ by mean curvature. Then there exist $c_1, c_2 > 0$ depending only on n, η, β_1 such that, for any $\sigma \leq 1/c_1$ and $p \geq c_2/\sigma^2$, the integral

$$\int_{\mathcal{M}_t} (f_{\sigma,\eta})_+^p d\mu$$

is a decreasing function of t.

Remark 4.4 By choosing c_1, c_2 suitably larger (but still only depending on n, η, β_1), one obtains that the following integrals are also decreasing in time

$$\int_{\mathcal{M}_t} H^2 (f_{\sigma,\eta})_+^p d\mu, \qquad \int_{\mathcal{M}_t} H^n (f_{\sigma,\eta})_+^p d\mu.$$

In fact, the integrand can be written as $(f_{\sigma',\eta})_+^p$, with $\sigma' = \sigma + 2/p$ and $\sigma' = \sigma + n/p$ respectively. Another integral which is important in the following is

$$\int_{\mathcal{M}_t} H^{2r} (f_{\sigma,\eta})_+^{pr} d\mu,$$

where $r > 1$ is a suitable constant depending only on n. Again, this integral is nonincreasing if σ, p satisfy bounds of the above form.

We now show that the integrals above cannot increase under the standard surgery with the surgery parameters as in Sect. 3.

Proposition 4.5 (i) *We can choose* $\varepsilon_0 > 0$*,* $\eta_0 > 0$ *and* $\sigma_0 > 0$ *small enough such that* $(f_{\sigma,\eta})_+$*, with* $f_{\sigma,\eta}$ *defined as in* (4.2)*, is nonincreasing under standard surgery with surgery parameters as in Sect. 3 on a normal* (ε, k)*-hypersurface neck for any* $0 < \sigma < \sigma_0, 0 < \eta < \eta_0$ *and any* $0 < \varepsilon < \varepsilon_0, k \geq 2$*. By this we mean in the notation of Sect. 3 that* $(f_{\sigma,\eta})_+$ *is nonincreasing in regions such as* $[0, 3\Lambda]$ *of the surface which are modified by the surgery and it is zero on the regions such as* $[3\Lambda, 4\Lambda]$ *which are added by the surgery.*

(ii) *The statement of Theorem 4.3 holds for mean curvature flow in a class* $\mathcal{C}(R, \alpha)$ *with surgeries determined as in* (i).

Proof. Notice that on an approximate cylinder $|A|^2$ is close to $\frac{1}{n-1}H^2$. Since $n \geq 3$ it follows immediately that for sufficiently small ε_0 the function $(f_{\sigma,\eta})_+$ vanishes everywhere on the regions affected by surgery, proving (i). The second statement is an immediate consequence of (i) since the inequality $H \geq \beta_1|A|$ for $1 \geq \beta_1 > 0$ is not affected by the surgery and therefore the constants c_1, c_2 of Theorem 4.3 never change. □

We are now ready to prove the case $m = 2$ of Theorem 4.1 for a flow with standard surgeries.

Theorem 4.6 *Let* $\mathcal{M}_0$ *be a closed n-dimensional immersed surface satisfying* $|A|^2 \leq R^{-2}$*,* $H \geq \beta_1|A|$ *for* $1 \geq \beta_1 > 0$ *as well as* $H \geq \alpha_1 R^{-1}$*,* $|\mathcal{M}_0| \leq \alpha_2 R^n$*. Then for any* $\delta > 0$ *there is a constant* β_2 *depending on* $n, \delta, \alpha_1, \alpha_2, \beta_1$ *such that a solution* $\mathcal{M}_t, t \in [0, T[$*, of mean curvature flow with initial data* $\mathcal{M}_0$ *and with surgeries as in Proposition 4.5 satisfies the estimate* $|A|^2 - H^2 \leq \delta H^2 + \beta_2 R^{-2}$ *on* $\mathcal{M}_t$ *for all* $t \in [0, T[$.

Corollary 4.7 *Let* $\mathcal{M}_0 \in \mathcal{C}(R, \alpha)$ *be a surface satisfying* $|A|^2 \leq R^{-2}$*. Then for any* $\delta > 0$ *there is a constant* β_2 *depending on* n, δ *and* α *such that a solution* $\mathcal{M}_t, t \in [0, T[$*, of mean curvature flow with initial data* $\mathcal{M}_0$ *and with surgeries as in Proposition 4.5 satisfies the estimate* $|A|^2 - H^2 \leq \delta H^2 + \beta_2 R^{-2}$ *on* $\mathcal{M}_t$ *for all* $t \in [0, T[$.

Proof. The above statement is proved in [18, Theor. 3.1] (see also [16, Theorem 5.1]) in the case of a smooth flow without surgeries. Here we are going to review the main steps of the proof to show that in view of Proposition 4.5 the same argument applies also to a flow with surgeries.

Let us set $\eta = \delta/2$ and fix any σ, p satisfying the restriction of Theorem 4.3. For any $k > 0$, we consider the truncated function

$$v = v_{k,\sigma,\eta,p} = (f_{\sigma,\eta} - k)_+^p$$

and set $A(k, t) = \{x \in \mathcal{M}_t \;:\; v > 0\}$. Observe that by Proposition 4.5 v is pointwise decreasing with surgeries specified there. From the evolution equation for v one obtains, for a smooth flow

$$\frac{d}{dt}\int_{\mathcal{M}_t} v^2 d\mu + \int_{\mathcal{M}_t} |\nabla v|^2 d\mu \leq \sigma p \int_{A(k,t)} |A|^2 (f_{\sigma,\eta})_+^p d\mu. \tag{4.4}$$

We are going to exploit the $|\nabla v|^2$ term using the Sobolev-type inequality (see [24])

$$\Big(\int_{\mathcal{M}_t} v^{2q} d\mu\Big)^{1/q} \leq c_3 \int_{\mathcal{M}_t} |\nabla v|^2 d\mu + c_3 \Big(\int_{A(k,t)} H^n d\mu\Big)^{2/n} \Big(\int_{\mathcal{M}_t} v^{2q} d\mu\Big)^{1/q}. \tag{4.5}$$

Here $c_3 = c_3(n)$, while $q = n/(n-2)$ if $n > 2$ and an arbitrary number greater than 1 if $n = 2$. Suppose that $k \geq k_1$, where $k_1 = k_1(\eta, \sigma, p, n)$ is given by

$$k_1 = (2c_3)^{\frac{n}{2p}} \Big(\int_{\mathcal{M}_0} H^n (f_{\sigma,\eta})_+^p\Big)^{\frac{1}{p}}. \tag{4.6}$$

Recalling Remark 4.4, we obtain

$$\begin{aligned}\Big(\int_{A(k,t)} H^n d\mu\Big)^{2/n} &\leq k^{-2p/n} \Big(\int_{A(k,t)} H^n (f_{\sigma,\eta})_+^p d\mu\Big)^{2/n} \\ &\leq k^{-2p/n} \Big(\int_{\mathcal{M}_0} H^n (f_{\sigma,\eta})_+^p d\mu\Big)^{2/n} \leq \frac{1}{2c_3}.\end{aligned}$$

Thus, we obtain from (4.4) and (4.5)

$$\frac{d}{dt}\int_{\mathcal{M}_t} v^2 d\mu + \frac{1}{2c_3}\Big(\int_{\mathcal{M}_t} v^{2q} d\mu\Big)^{1/q} \leq \sigma p \int_{A(k,t)} |A|^2 (f_{\sigma,\eta})_+^p d\mu. \tag{4.7}$$

We deduce, for any $s \in]0, T[$,

$$\begin{aligned}\int_{\mathcal{M}_s} v^2 d\mu - \int_{\mathcal{M}_0} v^2 d\mu + \frac{1}{2c_3}\int_0^s \Big(\int_{\mathcal{M}_t} v^{2q} d\mu\Big)^{1/q} ds \\ \leq \sigma p \int_0^s \int_{A(k,t)} |A|^2 (f_{\sigma,\eta})_+^p d\mu ds. \end{aligned}\tag{4.8}$$

Until now we have assumed that $\mathcal{M}_t$ is a smooth flow. However, it is easy to see that (4.8) holds also if $\mathcal{M}_t$ is a flow with surgeries satisfying the conditions of Proposition 4.5. In fact, suppose that there are surgery times $t_1, t_2, \ldots, t_h$ with $0 < t_1 < t_2 < \cdots < t_h \leq s$. We observe that (4.7) holds on each interval $[0, t_1]$, $[t_1, t_2], \ldots, [t_h, s]$; in fact it only requires condition (4.3) and the monotonicity of the L^p norm of $f_{s,\eta}$, and these properties are preserved by the surgeries. Thus, we can integrate (4.7) on each interval $[0, t_1], [t_1, t_2], \ldots, [t_h, s]$ and sum up the contributions, taking into account that the integral of v^2 is decreasing after each surgery in view of Proposition 4.5. Equivalently, we can say that inequalities (4.4), (4.7) hold in a weak sense for a flow with surgeries.

The rest of the proof requires no special discussion in the presence of surgeries because the regularity in time of the functions involved plays no longer any role. We define $k_2 = k_2(\sigma, \eta)$ as

$$k_2 = \sup_{\mathcal{M}_0} f_{\sigma,\eta}. \tag{4.9}$$

Then $v \equiv 0$ at $t = 0$ if $k \geq k_2$ and so the contribution of the initial data in (4.8) vanishes. Let us also set $q_0 = 2 - 1/q$ and choose $r > 1$ such that $1 - 1/q_0 - 1/r > 0$. In addition, let us set

$$\|A(k)\| = \int_0^T \int_{A(k,t)} d\mu.$$

After some computations we obtain from (4.8)

$$\int_0^T \int_{A(k,t)} v^2 d\mu dt \leq c_4 \sigma p \|A(k)\|^{2-\frac{1}{q_0}-\frac{1}{r}} \Big(\int_0^T \int_{A(k,t)} H^{2r} (f_{\sigma,\eta})_+^{pr} d\mu dt \Big)^{\frac{1}{r}},$$

where $c_4 = c_4(n, \beta_1)$. If we set

$$k_3 = \int_{\mathcal{M}_0} H^{2r} (f_{\sigma,\eta})_+^{pr} d\mu, \quad \gamma = 1 - \frac{1}{q_0} - \frac{1}{r} \tag{4.10}$$

then we conclude, by Remark 4.4 and by the definition of v,

$$\int_0^T \int_{A(k,t)} (f_{\sigma,\eta} - k)_+^p d\mu dt \leq c_4 \sigma p \|A(k)\|^{1+\gamma} T^{1/r} k_3^{1/r}. \tag{4.11}$$

We recall that this estimate holds provided $k \geq \max\{k_1, k_2\}$ and p, σ satisfy the bounds in Theorem 4.3 (see also Remark 4.4) which are of the form $\sigma \leq 1/c_1$, $p \geq c_2/\sigma^2$, for suitable constants c_1, c_2 only depending on n, β_1, η. Then we also have, for any $h > k \geq \max\{k_1, k_2\}$,

$$|h - k|^p \|A(h)\| \leq c_4 \sigma p \|A(k)\|^{1+\gamma} T^{1/r} k_3^{1/r}.$$

Then, since $\gamma > 0$, a well known result by Stampacchia (see e.g. [20, Lemma II.B.1]) implies that

$$\|A(k)\| = 0, \quad \forall k > \max\{k_1, k_2\} + k_4,$$

where

$$k_4^p = c_4 \sigma p (k_3 T)^{\frac{1}{r}} 2^{\frac{p(\gamma+1)}{\gamma}} \|A(\max\{k_1, k_2\})\|^\gamma.$$

Setting $k_5 = \max\{k_1, k_2\} + k_4$, we obtain that $|A|^2 \leq (1+\eta) H^2 + k_5 H^{2-\sigma} \leq (1 + 2\eta) H^2 + C$, for a suitable $C = C(\eta, \mathcal{M}_0)$. Since we have chosen $\eta = \delta/2$ we only need to show that the constant C has the desired form.

If $\mathcal{M}_0$ is in $\mathcal{C}(R,\alpha)$ we can estimate the constants introduced in (4.6), (4.9) and (4.10) as follows:

$$k_2 = \sup_{\mathcal{M}_0} f_{\sigma,\eta} \le \beta_1 2(\sup_{\mathcal{M}_0} H)^\sigma \le \beta_1 2(\sqrt{n})^\sigma R^{-\sigma} \le c(n,\beta_1)R^{-\sigma},$$
$$k_1 \le c(n)|\mathcal{M}_0|^{1/p}(\sup_{\mathcal{M}_0} H)^{n/p} k_2 \le c(n,\alpha,\beta_1)R^{-\sigma},$$
$$k_3 \le |\mathcal{M}_0|(\sup_{\mathcal{M}_0} H)^{2r} k_2^{pr} \le c(n,\alpha,\beta_1)R^{n-2r-\sigma pr}.$$

From Proposition 2.7 we recall the estimate $T \le \frac{n}{2}\alpha_1^{-2}R^2$, which is unaffected by surgeries since the mean curvature is always increasing there. If we take into account that $(n+2)(\gamma + \frac{1}{r}) = (n+2)(1-\frac{1}{q_0}) = 2$ we obtain

$$\begin{aligned} k_4^p &\le c(n,\alpha,\beta_1,\eta)(k_3 T)^{\frac{1}{r}}(T|\mathcal{M}_0|)^\gamma \\ &\le c(n,\alpha,\beta_1,\eta)R^{\frac{n+2}{r}-2-\sigma p}R^{(n+2)\gamma} = c(n,\alpha,\beta_1,\eta)R^{-\sigma p}. \end{aligned}$$

Thus, $k_5 = \max\{k_1,k_2\} + k_4 \le c(n,\alpha,\beta_1,\eta)R^{-\sigma}$. We conclude, using Young's inequality,

$$|A|^2 - H^2 \le \eta H^2 + c(n,\alpha,\beta_1,\eta)R^{-\sigma}H^{2-\sigma} \le 2\eta H^2 + c'(n,\alpha,\beta_1,\eta)R^{-2}.$$

Since $\eta = \delta/2$, we obtain the conclusion. □

The case where $m > 2$ in Theorem 4.1 is technically more complicated. We use an induction procedure. We suppose that we have proved estimate (4.1) for the polynomials S_l with $l = 2,\ldots,m$ for some $m \ge 2$ and we want to deduce from this the estimate for S_{m+1}. The basic idea is to consider the quotient $Q_{m+1} := S_{m+1}/S_m$, which is a concave function of the principal curvatures. Such a quotient, however, need not even be well defined, since estimate (4.1) does not prevent S_m from being zero somewhere. However, we can define a perturbation of the second fundamental form after which S_m becomes positive, as we now show. For given $0 < \rho < 1/n$, $D > 0$ we set

$$b_{ij;\rho,D} = h_{ij} + (\rho H + D)g_{ij}.$$

We denote by λ_i^b the eigenvalues of b_{ij}, which are given by $\lambda_i^b = \lambda_i + \rho H + D$, and denote by S_m^b, Q_m^b the symmetric polynomials and their quotients evaluated at λ_i^b instead of λ_i.

A key property of the perturbation introduced above is that a *lower bound* of the form $S_m \ge -\delta H^m - C_\delta R^{-m}$ on the unperturbed polynomials implies the *positivity* of the perturbed polynomials S_m^b, as shown by the following purely algebraic property of the elementary symmetric functions (see [19, Lemma 2.7]).

Lemma 4.8 *Suppose that for given $m \in \{2, \ldots, n-1\}$ and fixed $R > 0$ the elementary symmetric polynomials $S_l, l = 2, \ldots, m$, satisfy the following property $(E)_m$ on some subset T of the upper halfplane $\{(\lambda_1, \ldots, \lambda_n) \mid H = \lambda_1 + \cdots + \lambda_n > 0\}$:*

$(\mathbf{E})_{\mathbf{m}}$ *For any $0 < \delta \leq 1/2$ there exist constants $C_{2,\delta}, \ldots, C_{m,\delta} > 0$ such that everywhere on T*

$$S_l \geq -\delta H^l - C_{l,\delta} R^{-l}, \quad l = 2, \ldots, m. \tag{4.12}$$

Then, for any $0 < \rho \leq 1/n$ there exists $c_{hskip-1pt\rho} > 0$ such that, for any $D \geq c_\rho R^{-1}$, we have everywhere on T

$$S^b_{m;\rho,D} \geq \frac{\rho(n-m+1)}{m(1+n\rho)} S^b_{m-1;\rho,D} H^b_{\rho,D} > 0.$$

The constant c_ρ only depends on $C_{2,\delta}, \ldots, C_{m,\delta}$ for a suitable $\delta = \delta(\rho)$.

Now our induction procedure goes as follows. Suppose that we have proved estimates (4.12) for the elementary symmetric functions of the principal curvatures in our flow up to a certain integer m (the first step, $m = 2$, is given by Theorem 4.6). By the previous lemma, for every $\rho > 0$, we can find D_ρ as in Lemma 4.8 such that for $D \geq D_\rho$ the quotient $Q^b_{m+1;\rho,D}$ is well defined. We consider the function

$$f_{\sigma,\eta,\rho,D} = \frac{-Q^b_{m+1;\rho,D} - \eta H^b_{\rho,D}}{\left(H^b_{\rho,D}\right)^{1-\sigma}}. \tag{4.13}$$

The next result (see [19, Lemma 2.8]) is again purely algebraic and shows that an upper bound on the function $f_{\sigma,\eta,\rho,D_\rho}$ implies property $(\mathbf{E})_{\mathbf{m+1}}$ for the elementary symmetric functions.

Lemma 4.9 *Suppose that some subset $T \subset \{(\lambda_1, \ldots, \lambda_n) | H > 0\}$ satisfies $(\mathbf{E})_{\mathbf{m}}$ for a given $m \in \{2, \ldots, n-1\}$ and $R > 0$. Suppose that, for any $0 < \rho \leq 1/n$, $0 < \eta \leq 1/2$, there exist $c_\rho = c(\rho)$, $\sigma = \sigma(\rho, \eta)$, $K = K(\rho, \eta)$ such that, setting $D_\rho = c_\rho R^{-1}$, the function $f_{\sigma,\eta,\rho,D_\rho}$ is well defined and satisfies*

$$f_{\sigma,\eta,\rho,D_\rho} \leq K R^{-\sigma}.$$

Then, for any $0 < \delta \leq 1/2$, there exists $C_{m+1,\delta} > 0$ such that

$$S_{m+1} \geq -\delta |A|^{m+1} - C_{m+1,\delta} R^{-m-1}$$

holds everywhere on T. In addition, $C_{m+1,\delta}$ only depends on the value of c, σ, K corresponding to a suitable pair $\rho(\delta), \eta(\delta)$.

Thus, if we can bound from above the function $f_{\sigma,\eta,\rho,D_\rho}$, we can iterate our procedure with m replaced by $m+1$ and obtain Theorem 4.1 after a finite number of steps. The purely algebraic nature of the iteration lemmata just stated ensures that all constants will only depend on the scaling invariant parameters $\alpha_1, \alpha_2, \beta_1$ and the scale R of the initial data. The proof of the estimate of $f_{\sigma,\eta,\rho,D_\rho}$ is rather long and technical (see [19]). Here we only mention that, using the special algebraic properties of the quotients Q_m, one can show the function $f_{\sigma,\eta,\rho,D_\rho}$ satisfies an evolution equation which allows to follow a procedure similar to the case $m = 2$ seen before. The perturbation in the second fundamental form induces the presence of lower order terms in the estimates, which do not affect substantially the procedure. The following result is analogous to Theorem 4.3 and was shown for smooth mean curvature flow in [19, Corollary 3.7].

Theorem 4.10 *Let $\mathcal{M}_0$ be a closed n-dimensional immersed surface satisfying $|A|^2 \le R^{-2}$, $H \ge \beta_1 |A|$ for $1 \ge \beta_1 > 0$ as well as $H \ge \alpha_1 R^{-1}$, $|\mathcal{M}_0| \le \alpha_2 R^n$. Let $\mathcal{M}_t$, for $t \in [0, T[$, be the smooth evolution of $\mathcal{M}_0$ by mean curvature flow. Suppose in addition that assumption $(\mathbf{E})_\mathbf{m}$ already holds on $\mathcal{M}_t$ for some $m \in \{2, \dots, n-1\}$ with constants $C_{2,\delta}, \dots, C_{l,\delta}$ depending only on $\delta, n, \alpha_1, \alpha_2$ and β_1. Let $0 < \rho \le 1/n$, $0 < \eta \le 1/2$ and set $D_\rho = c_\rho R^{-1}$, with c_ρ as in Lemma 4.8. Then there exist $c_1, c_2, c_3 > 0$ (depending on n, η, ρ, β_1) and $K_1 > 0$ (depending on $n, \eta, \rho, \beta_1, c_\rho$) such that, for any $\sigma \le 1/c_1$ and $p \ge c_2/\sigma^2$ we have the estimate*

$$\int_{\mathcal{M}_t} (f_{\sigma,\eta,\rho,D_\rho})_+^p d\mu \le e^{c_3 t/R^2} \int_{\mathcal{M}_0} (f_{\sigma,\eta,\rho,D_\rho})_+^p d\mu + K_1 R^{-\sigma p} |\mathcal{M}_0| (e^{c_3 t/R^2} - 1).$$

Remark 4.11 In view of Lemmata 4.8 and 4.9 for each $2 \le m \le n$ and for each set of parameters $\alpha_1, \alpha_2, \beta_1$ the constant c_ρ in $D_\rho = c_\rho R^{-1}$ can be fixed as a universal function of $0 < \rho \le 1/n$.

As in the case $m = 2$ we now show that the integrals above cannot increase under surgery.

Theorem 4.12 (i) *We can choose surgery parameters B and τ_0 depending only on n and we can choose $\varepsilon_0 > 0$ as well as parameters $\rho_0 > 0$, $\eta_0 > 0$, $\sigma_0 > 0$, $D_0 > 0$ such that $(f_{\sigma,\eta,\rho,D})_+$ is nonincreasing under standard surgery with parameters B, τ_0 on a normal (ε, k)-hypersurface neck for any $0 < \rho \le \rho_0$, $0 < \sigma \le \sigma_0$, $0 < \eta \le \eta_0$, $D \ge D_{\rho_0}$ and any $0 < \varepsilon \le \varepsilon_0$, $k \ge 2$.*

(ii) *The statement of Theorem 4.10 holds for mean curvature flow in a class $\mathcal{C}(R, \alpha)$ with surgeries determined as in* (i).

Proof. Notice that on an approximate cylinder the quotient Q_{m+1} is close to $\frac{(n-m-1)}{(n-1)(m+1)} H$ since $\lambda_2, \dots, \lambda_n$ are close to r_0^{-1} and $|\lambda_1| \le \varepsilon r_0^{-1}$. Also notice that at points where $D \ge D_{\rho_0} \ge |\lambda_1|$ we have $\lambda_1^b = \lambda_1 + \rho + D > 0$ and hence $Q_{m+1}^b > 0$, such that $f_{\sigma,\eta,\rho,D}$ is negative at such points for all $2 \le m \le n$.

So we only need to consider points where $D \leq \varepsilon r_0^{-1}$. Choosing then $\rho_0 > 0$ small depending on n we immediately see that in case $2 \leq m < n-1$ for sufficiently small ε_0 depending only on n the function $f_{\sigma,\eta,\rho,D}$ is negative everywhere on the regions affected by surgery, proving (i) in this case. The critical case is $m = n-1$. We compute the purely algebraic identity

$$\begin{aligned}\frac{\partial}{\partial\lambda_i}\big(-Q^b_{n;\rho,D} - \eta H^b_{\rho,D}\big) &= -\sum_j \frac{\partial Q^b_{n;\rho,D}}{\partial\lambda^b_j}\big(\delta^i_j + \rho\big) - \eta(1+n\rho)\\ &= -\frac{\partial Q^b_{n;\rho,D}}{\partial\lambda^b_i} - \rho\sum_j \frac{\partial Q^b_{n;\rho,D}}{\partial\lambda^b_j} - \eta(1+n\rho),\end{aligned} \tag{4.14}$$

as well as

$$\frac{\partial Q^b_{n;\rho,D}}{\partial\lambda^b_i} = (S^b)^{-2}_{n-1;\rho,D}\prod_{j\neq i}\big(\lambda^b_j\big)^2 \geq 0. \tag{4.15}$$

We then compute for the bent surface

$$\tilde{N}_\tau(p) = N(p) - \tau u(z)\nu(p), \qquad p = (\omega, z) \in \mathbb{S}^{n-1}\times[0, 4\Lambda] \tag{4.16}$$

the change of our functions with respect to τ: first we get from (3.38) the change for the eigenvalues λ_i

$$\frac{d\tilde{\lambda}_i}{d\tau} = \delta^1_i D_1D_1u + u\lambda_i^2 + O(\theta D_1D_1u) \tag{4.17}$$

and then from (4.14)

$$\begin{aligned}&\frac{d}{d\tau}\big(-\tilde{Q}^b_{n,\rho,D} - \eta\tilde{H}^b_{\rho,D}\big)\\ &\quad= -\sum_i\big(\tilde{S}^b_{n-1}\big)^{-2}\prod_{j\neq i}\big(\lambda^b_j\big)^2\big(\delta^1_i D_1D_1u + u\lambda_i^2 + O(\theta D_1D_1u)\big)\\ &\qquad -(\eta + n\eta\rho)\big(D_1D_1u + |A|^2u\big) + O(\theta D_1D_1u)\\ &\quad\leq -\big(\tilde{S}^b_{n-1}\big)^{-2}\Big(\prod_{j\neq 1}\big(\tilde{\lambda}^b_j\big)^2 D_1D_1u + \sum_{i\neq 1}\prod_{j\neq i}\big(\tilde{\lambda}^b_j\big)^2 O(\theta D_1D_1u)\Big)\\ &\qquad + O(\theta D_1D_1u).\end{aligned} \tag{4.18}$$

Now notice that $\lambda_j = r_0^{-1}$ for $j = 2, \ldots, n$ on the standard cylinder while $\lambda_1 = 0$. Choosing again $0 < \rho < \rho_0$ small depending on n and recalling that we may assume $D \leq \varepsilon r_0^{-1}$ at the point under consideration we conclude for small enough $0 < \theta$ depending on n (with corresponding choice of B and τ_0) and $0 < \varepsilon \leq \varepsilon_0$ depending on n that on the bent neck $\tilde{N}$ the estimate

$$\big(-\tilde{Q}^b_{n,\rho,D} - \eta\tilde{H}^b_{\rho,D}\big) \leq \big(-Q^b_{n,\rho,D} - \eta H^b_{\rho,D}\big) - \frac{1}{2}\tau_0 r_0^{-1}D_1D_1u. \tag{4.19}$$

everywhere on $[0, 3\Lambda]$. Thus the positive region of $f_{\sigma,\eta,\rho,D}$ is shrinking and the numerator is decreasing there. Since the mean curvature is increasing we conclude that $(f_{\sigma,\eta,\rho,D})_+$ is nonincreasing everywhere on $[0, 3\Lambda]$. Notice that both the deformed surface and its interpolation with the axially symmetric surface is convex in $[2\Lambda, 3\Lambda]$ such that $(f_{\sigma,\eta,\rho,D})_+$ vanishes in $[2\Lambda, 4\Lambda]$ after standard surgery is completed. This completes the proof of (i). The second statement is then an immediate consequence of (i) and Theorem 4.10 since the inequality $H \geq \beta_1|A|$ for $1 \geq \beta_1 > 0$ is not affected by the surgery and therefore the constants c_1, c_2, c_3 of the theorem never change. □

In contrast to Theorem 4.3, the L^p norms appearing in Theorem 4.10 are not decreasing in time; however, the result still gives a uniform bound on the L^p norms of f which allows us to derive an L^∞ bound with a procedure analogous to the one of Theorem 4.6, compare [19]. Starting from initial data satisfying $|A|^2 \leq R^{-2}$ and keeping carefully track of the scaling parameter R we derive that for any $0 < \eta \leq \eta_0$ small, $0 < \rho \leq \rho_0$ small with corresponding $D_\rho = c_\rho R^{-1}$ as in Lemma 4.9 and Remark 4.11, there is $0 < \sigma$ depending also on $n, \alpha_1, \alpha_2, \beta_1$ such that $f_{\sigma,\eta,\rho,D_\rho} \leq K_2 R^{-\sigma}$, with $K_2 > 0$ depending on $n, \eta, \alpha_1, \alpha_2, \beta_1, \rho, D_\rho$. This allows the next step in the iteration Lemma 4.9 thereby finally leading to the following result.

Theorem 4.13 *Let $\mathcal{M}_0$ be a closed n-dimensional immersed surface satisfying $|A|^2 \leq R^{-2}$, $H \geq \beta_1|A|$ for $1 \geq \beta_1 > 0$ as well as $H \geq \alpha_1 R^{-1}$, $|\mathcal{M}_0| \leq \alpha_2 R^n$. Then for any $2 \leq m \leq n$ and any $\delta > 0$ there is a constant β_m depending on $n, m, \delta, \alpha_1, \alpha_2, \beta_1$ such that a solution $\mathcal{M}_t$, $t \in [0, T[$, of mean curvature flow with initial data $\mathcal{M}_0$ and with surgeries as in Theorem 4.12 satisfies the estimates*

$$S_m \geq -\delta H^m - \beta_m R^{-m} \quad \text{on } \mathcal{M}_t \text{ for all } t \in [0, T[.$$

For an initial surface in our class $\mathcal{C}(R, \alpha)$ we obtain:

Corollary 4.14 *Let $\mathcal{M}_0 \in \mathcal{C}(R, \alpha)$ be a surface satisfying $|A|^2 \leq R^{-2}$. Then for any $\delta > 0$ there is a constant β_m depending on n, δ and α such that a solution $\mathcal{M}_t$, $t \in [0, T[$, of mean curvature flow with initial data $\mathcal{M}_0$ and with surgeries as in Proposition 4.12 satisfies the estimates*

$$S_m \geq -\delta H^m - \beta_m R^{-m} \quad \text{on } \mathcal{M}_t \text{ for all } t \in [0, T[.$$

Remark 4.15 We emphasize that we were able to choose the surgery parameters B, τ_0 depending only on n, independent of the scale of the neck and the class of initial data $\mathcal{C}(R, \alpha)$.

While we have shown that all estimates in this section work without the assumption of 2-convexity for sufficiently good surgery parameters, 2-convexity is crucial for the cylindrical estimates in the following section that in turn lead to curvature gradient estimates and ultimately allow the detection and control of necks.

5 Cylindrical estimates

In this section we want to prove an estimate showing, roughly speaking, that for the surfaces we are considering any rescaling near a singularity which is not strictly convex must be cylindrical. We will assume as before that we have a solution of mean curvature flow with surgery in some class $\mathcal{C}(R, \alpha)$ with normalised initial data $\mathcal{M}_0$ satisfying $|A|^2 \leq R^{-2}$. We will also assume that the surgery parameters have been chosen such that $\mathcal{C}(R, \alpha)$ is preserved and the conclusions of Theorem 4.12 and Corollary 4.14 hold. The strategy is to combine 2-convexity with the information provided by the estimates of Corollary 4.14.

Lemma 5.1 *Let $\mathcal{M}_t$ in $\mathcal{C}(R, \alpha)$ be a solution of mean curvature flow with surgery and normalised initial data as above. Then for any $\delta > 0$ there exists $K_\delta = K_\delta(n, \alpha)$ such that $\lambda_1 \geq -\delta H - K_\delta R^{-1}$ on $\mathcal{M}_t$ for any $t > 0$.*

Proof. It suffices to consider the case when $\lambda_1 < 0$. From Corollary 4.14 we know that for any $\delta > 0$ there exists $C_\delta = C_\delta(n, \alpha)$ such that

$$S_n \geq -\left(\frac{\alpha_0}{2}\right)^{n-1} \delta H^n - C_\delta R^{-n}.$$

Thus we find, by (2.4),

$$\left(\frac{\alpha_0}{2}\right)^{n-1} \delta H^n + C_\delta R^{-n} \geq (-\lambda_1)\lambda_2\lambda_3 \ldots \lambda_n \geq (-\lambda_1)\left(\frac{\alpha_0}{2} H\right)^{n-1},$$

which implies

$$-\lambda_1 \leq \delta H + \frac{2^{n-1} C_\delta}{(\alpha_0 H)^{n-1} R^n}.$$

Taking into account that $-\lambda_1 \leq H$, and considering separately the cases $H \leq 1/R$ and $H \geq 1/R$ we conclude

$$-\lambda_1 \leq \delta H + \max\left\{\frac{1}{R}, \frac{2^{n-1} C_\delta}{\alpha_0^{n-1} R}\right\}.$$

□

Actually, it is not difficult to check that the above result can be proved also without the hypothesis of 2-convexity. We have assumed this property because it is required in the rest of our analysis and it allows for a more explicit proof.

As in [16] we recall Simons' identity

$$\frac{1}{2}\Delta |A|^2 = \langle h_{ij}, \nabla_i \nabla_j H \rangle + |\nabla A|^2 + Z, \tag{5.1}$$

where $Z = H\mathrm{tr}(A^3) - |A|^4 = (\sum \lambda_i)(\sum \lambda_i^3) - (\sum \lambda_i^2)^2$. Let us derive an estimate for Z.

Lemma 5.2 *Under the same assumptions as in the previous lemma there exists a constant $\gamma_1 > 0$ depending only on n, α_0 such that for any $\delta > 0$ there exists $K_\delta = K_\delta(n, \alpha)$ such that*

$$Z \geq \gamma_1 H^2 \left(|A|^2 - \frac{1}{n-1} H^2 - \delta H^2 \right) - K_\delta R^{-1} H^3$$

on $\mathcal{M}_t$ for any $t > 0$.

Proof. We first observe that

$$|A|^2 - \frac{1}{n-1} H^2 = \frac{1}{n-1} \Big(\sum_{1<i<j} (\lambda_i - \lambda_j)^2 + \lambda_1 (n\lambda_1 - 2H) \Big). \quad (5.2)$$

Therefore, by (2.4),

$$\begin{aligned} Z &= \sum_{j=2}^n \lambda_1 \lambda_j (\lambda_1 - \lambda_j)^2 + \sum_{1<i<j} \lambda_i \lambda_j (\lambda_i - \lambda_j)^2 \\ &\geq \sum_{j=2}^n \lambda_1 \lambda_j (\lambda_1 - \lambda_j)^2 + \frac{(\alpha_0 H)^2}{4} \sum_{1<i<j} (\lambda_i - \lambda_j)^2 \\ &= \frac{(\alpha_0 H)^2}{4} ((n-1)|A|^2 - H^2) \\ &\quad + \lambda_1 \left(\sum_{j=2}^n \lambda_j (\lambda_1 - \lambda_j)^2 + \frac{(\alpha_0 H)^2}{4} (2H - n\lambda_1) \right). \end{aligned} \quad (5.3)$$

We now estimate the second term in the right–hand side. It suffices to consider the case when $\lambda_1 < 0$ since otherwise this term is nonnegative and the assertion of the theorem is immediate. From (2.2) we deduce that

$$\begin{aligned} &\sum_{j=2}^n \lambda_j (\lambda_1 - \lambda_j)^2 + \frac{(\alpha_0 H)^2}{4} (2H - n\lambda_1) \\ &\quad \leq (n-1) H (2H)^2 + \frac{(\alpha_0 H)^2}{4} (n+2) H = \left(4(n-1) + (n+2) \frac{\alpha_0^2}{4} \right) H^3. \end{aligned}$$

Therefore (5.3) implies

$$Z \geq \frac{(\alpha_0 H)^2}{4} ((n-1)|A|^2 - H^2) + \lambda_1 \left(4(n-1) + (n+2) \frac{\alpha_0^2}{4} \right) H^3,$$

from which we easily obtain the assertion, using Lemma 5.1. $\square$

To derive the cylindrical estimate let us consider, for $\eta \in \mathbb{R}$ and $\sigma \in [0, 2]$, the function

$$f_{\sigma,\eta} = \frac{|A|^2 - \left(\frac{1}{n-1} + \eta\right) H^2}{H^{2-\sigma}}. \tag{5.4}$$

Although we use the same notation, the function $f_{\sigma,\eta}$ here is slightly different than the one introduced in (4.2). The factor in front of H^2 is chosen here in such a way that, if $\eta = 0$, the function vanishes on a cylinder. In addition, if $\eta = 0$ and $\lambda_1 = 0$, the numerator is nonnegative and vanishes if and only if $\lambda_2 = \cdots = \lambda_n$. The aim of the rest of the section is to prove the following estimate. Observe that, in view of (5.2), the result easily implies the estimate of Theorem 1.5.

Theorem 5.3 (i) *Let $\mathcal{M}_t$, $t \in [0, T[$ be a smooth solution of mean curvature flow in $\mathcal{C}(R, \alpha)$ with $n \geq 3$ and initial data satisfying $|A|^2 \leq R^{-2}$. Then, for any $\eta > 0$ there exists a constant $C_\eta = C_\eta(n, \alpha) > 0$ such that*

$$|A|^2 - \frac{H^2}{n-1} \leq \eta H^2 + C_\eta R^{-2}$$

on $\mathcal{M}_t$ for any $t \in [0, T[$.

(ii) *For all $\Lambda \geq 10$ we can choose $k_0 \geq 2$, $\varepsilon_0 > 0$, surgery parameters B, τ_0, as well as parameters $\eta_0 > 0$, $\sigma_0 > 0$ such that $(f_{\sigma,\eta})_+$ is nonincreasing under standard surgery on a normal (ε, k)-hypersurface neck for any $0 < \sigma \leq \sigma_0, 0 < \eta \leq \eta_0$ and any $0 < \varepsilon \leq \varepsilon_0, k \geq k_0$. For mean curvature flow with such surgeries and parameters $0 < \eta < \eta_0$ we then have the same estimate as in* (i).

In the proof of Theorem 5.3, we consider values of σ, η in $]0, 1[$ and we write for simplicity $f_{\sigma,\eta} = f$ as long as σ, η are kept fixed. Let us observe that (2.3) implies

$$f_{\sigma,\eta} \leq n H^\sigma. \tag{5.5}$$

Lemma 5.4 *There exist constants $c_1, c_2 > 1$, depending only on n, α_0, such that for solutions of smooth mean curvature flow we have the estimate*

$$\frac{d}{dt}\int_{\mathcal{M}_t} f_+^p d\mu \leq -\frac{p(p-1)}{2}\int_{\mathcal{M}_t} f_+^{p-2}|\nabla f|^2 d\mu - \frac{p}{c_1}\int_{\mathcal{M}_t} \frac{f_+^p}{H2}|\nabla H|^2 d\mu$$
$$-p\int_{\mathcal{M}_t} \frac{f_+^{p-1}}{H^{4-\sigma}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 d\mu + p\sigma \int_{\mathcal{M}_t} |A|^2 f_+^p d\mu$$

for any $p \geq c_2$.

Proof. We recall (see [16, 18]) that f satisfies the equation

$$\frac{\partial f_{\sigma,\eta}}{\partial t} = \Delta f_{\sigma,\eta} + \frac{2(1-\sigma)}{H}\langle \nabla H, \nabla f_{\sigma,\eta}\rangle - \frac{\sigma(1-\sigma)}{H^2} f_{\sigma,\eta}|\nabla H|^2 \qquad (5.6)$$
$$- \frac{2}{H^{4-\sigma}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 + \sigma |A|^2 f_{\sigma,\eta}.$$

Thus we have, for $p \geq 2$,

$$\begin{aligned}\frac{d}{dt}\int f_+^p d\mu &= \int \left(\frac{\partial f_+^p}{\partial t} - H^2 f_+^p\right) d\mu = \int \left(p f_+^{p-1}\frac{\partial f}{\partial t} - H^2 f_+^p\right) d\mu \\ &\leq -p(p-1)\int f_+^{p-2}|\nabla f|^2 d\mu \\ &\quad + 2(1-\sigma)p \int \frac{f_+^{p-1}}{H}\langle \nabla H, \nabla f\rangle d\mu \\ &\quad -2p\int \frac{f_+^{p-1}}{H^{4-\sigma}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 d\mu + p\sigma \int |A|^2 f_+^p d\mu.\end{aligned} \qquad (5.7)$$

From Lemma 2.7(iii) and inequality (5.5) we deduce

$$\frac{f_+^{p-1}}{H^{4-\sigma}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 \geq \frac{\alpha_0^2}{8n}\frac{f_+^p}{H^2}|\nabla H|^2. \qquad (5.8)$$

Therefore, if $p - 1 \geq 32n/\alpha_0^2$ we obtain

$$\begin{aligned}2(1-\sigma)p\frac{f_+^{p-1}}{H}\langle \nabla H, \nabla f\rangle &\leq 2p\frac{f_+^{p-1}}{H}|\nabla H||\nabla f| \\ &\leq \frac{\alpha_0^2 p}{16n}\frac{f_+^p}{H^2}|\nabla H|^2 + \frac{16np}{\alpha_0^2} f_+^{p-2}|\nabla f|^2 \\ &\leq p\frac{f_+^{p-1}}{H^{4-\sigma}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 \\ &\quad - \frac{\alpha_0^2 p}{16n}\frac{f_+^p}{H^2}|\nabla H|^2 + \frac{p(p-1)}{2} f_+^{p-2}|\nabla f|^2.\end{aligned}$$

Substituting in inequality (5.7) we obtain the conclusion in the case of a smooth flow. □

Lemma 5.5 *There exist constants $c_3 = c_3(n, \alpha_0), C = C(n, \eta, \alpha) > 1$, such that*

$$\frac{\eta}{c_3}\int_{\mathcal{M}_t} |A|^2 f_+^p d\mu \leq \frac{p}{\beta}\int_{\mathcal{M}_t} f_+^{p-2}|\nabla f|^2 d\mu + (1+\beta p)\int_{\mathcal{M}_t}\frac{f_+^p}{H^2}|\nabla H|^2 d\mu$$
$$+ \frac{C}{R}\int_{\mathcal{M}_t} H f_+^p d\mu$$

for any $\beta > 0$, $p > 2$.

Proof. Let us set $a = 2 - \sigma$. We compute

$$\begin{aligned}\Delta f &= \Delta\left(\frac{|A|^2}{H^a}\right) - \left(\frac{1}{n-1} + \eta\right)\Delta H^{2-a} \\ &= \frac{\Delta|A|^2}{H^a} - a\frac{|A|^2}{H^{a+1}}\Delta H - a(a-1)\frac{|A|^2}{H^{a+2}}|\nabla H|^2 - \frac{2a}{H}\left\langle \nabla H, \nabla \frac{|A|^2}{H^a}\right\rangle \\ &\quad + \left(\frac{1}{n-1} + \eta\right)(2-a)\left(\frac{a-1}{H^a}|\nabla H|^2 - \frac{1}{H^{a-1}}\Delta H\right).\end{aligned}$$

Using (5.1) can rewrite the above equality as

$$\begin{aligned}\Delta f = &\frac{2}{H^a}\langle h_{ij}, \nabla_i\nabla_j H\rangle + \frac{2}{H^a}Z + \frac{2}{H^{a+2}}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 \\ &- \frac{2(a-1)}{H}\langle \nabla H, \nabla f\rangle - \left(\frac{af}{H} + 2\left(\frac{1}{n-1} + \eta\right)H^{1-a}\right)\Delta H \\ &+ \frac{(2-a)(a-1)}{H^2} f|\nabla H|^2.\end{aligned}$$

We multiply this equality by $f_+^p H^{a-2}$ and integrate on $\mathcal{M}_t$. Integrating by parts and taking into account the Codazzi equation we obtain

$$\begin{aligned}\int \frac{2Z}{H^2} f_+^p d\mu = &-p\int \frac{f_+^{p-1}}{H^{2-a}}|\nabla f|^2 d\mu + 2p\int \frac{f_+^{p-1}}{H^2}\langle h_{ij}, \nabla_i f \nabla_j H\rangle d\mu \\ &- 4\int \frac{f_+^p}{H^3}\langle h_{ij}, \nabla_i H \nabla_j H\rangle d\mu \\ &- 2\int \frac{f_+^p}{H^4}|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2 d\mu \\ &- \int \left(ap\frac{f_+^p}{H^{3-a}} + 2p\left(\frac{1}{n-1} + \eta\right)\frac{f_+^{p-1}}{H}\right)\langle \nabla f, \nabla H\rangle d\mu \\ &+ 2\int \left(\frac{f_+^{p+1}}{H^{4-a}} + \left(2 + \frac{1}{n-1} + \eta\right)\frac{f_+^p}{H^2}\right)|\nabla H|^2 d\mu.\end{aligned}$$

Using repeatedly inequality (5.5) and the assumptions $a \in]1, 2[$, $\eta \in]0, 1[$ we obtain

$$\int \frac{2Z}{H^2} f_+^p d\mu \leq 8np\int \frac{f_+^{p-1}}{H}|\nabla H||\nabla f| d\mu + 10n\int \frac{f_+^p}{H^2}|\nabla H|^2 d\mu.$$

In addition we have, for any $\beta > 0$

$$2\frac{f_+^{p-1}}{H}|\nabla H||\nabla f| \leq \frac{f_+^{p-2}}{\beta}|\nabla f|^2 + \beta\frac{f_+^p}{H^2}|\nabla H|^2. \tag{5.9}$$

On the other hand, choosing $\delta = \eta/2$ in Lemma 5.2 we obtain that

$$f \geq 0 \implies Z \geq \frac{\gamma_1 \eta}{2} H^4 - \frac{K_\delta}{R} H^3.$$

Therefore, by (2.3),

$$\frac{2Z}{H^2} f_+^p \geq \gamma_1 \eta H^2 f_+^p - \frac{2K_\delta H f_+^p}{R} \geq \frac{\gamma_1}{n} \eta |A|^2 f_+^p - \frac{2K_\delta H f_+^p}{R}$$

and the assertion follows. □

Proposition 5.6 *There exist constants c_4, c_5 (depending on n, α_0, η) such that for any $p \geq c_4$, $0 < \sigma \leq 1/c_5\sqrt{p}$ and some constant $K_2 > 0$ depending on $n, \alpha, \eta, \sigma, p$ the following estimate holds:*

$$\int_{\mathcal{M}_t} (f_{\sigma,\eta})_+^p d\mu \leq \int_{\mathcal{M}_0} (f_{\sigma,\eta})_+^p d\mu + tK_2 R^{-2-\sigma p} |\mathcal{M}_0|.$$

Proof. Let c_1, c_2, c_3, C be as in Lemmas 5.4, 5.5. We recall that all these constants are greater than 1. For a fixed $\eta \in]0, 1]$, suppose that p, σ satisfy

$$p \geq \max\{2, c_2\}, \quad \sigma \leq \frac{\eta}{8c_1 c_3 \sqrt{p}}$$

and set $\beta = \frac{1}{\sqrt{p}}$. Then

$$\begin{cases} \dfrac{p2\sigma c_3}{\eta\beta} \leq \dfrac{p2}{8c_1} \leq \dfrac{p(p-1)}{4} \\ \dfrac{p\sigma c_3}{\eta}(1+\beta p) \leq \dfrac{\sqrt{p}}{8c_1}(1+\sqrt{p}) < \dfrac{p}{2c_1}. \end{cases}$$

Thus, by Lemma 5.5,

$$p\sigma \int |A|^2 f_+^p d\mu \leq \frac{p(p-1)}{4} \int f_+^{p-2} |\nabla f|^2 d\mu + \frac{p}{2c_1} \int \frac{f_+^p}{H^2} |\nabla H|^2 d\mu + \frac{C\sqrt{p}}{8c_1 R} \int H f_+^p d\mu.$$

Let

$$q = \frac{2+\sigma p}{1+\sigma p}, \quad q' = 2 + \sigma p, \quad B = \left(\frac{pq\sigma}{2n}\right)^{\frac{1}{q}}.$$

Then $1/q + 1/q' = 1$ and so, by Young's inequality

$$\begin{aligned}\frac{C\sqrt{p}}{8c_1 R} H &= H^{-\sigma p} B H^{1+\sigma p} \frac{C\sqrt{p}}{8Bc_1 R} \\ &\leq H^{-\sigma p}\left(\frac{B^q H^{q(1+\sigma p)}}{q} + KR^{-q'}\right) \\ &\leq \frac{p\sigma}{2n} H^2 + KR^{-2-\sigma p} H^{-\sigma p}\end{aligned}$$

for some $K = K(\alpha_0, \eta, n, C, \sigma, p)$. Therefore, by (5.5),

$$\begin{aligned}\frac{C\sqrt{p}}{8c_1 R}\int H f_+^p d\mu &\leq \frac{p\sigma}{2n}\int H^2 f_+^p d\mu + KR^{-2-\sigma p}\int H^{-\sigma p} f_+^p d\mu \\ &\leq \frac{p\sigma}{2}\int |A|^2 f_+^p d\mu + KR^{-2-\sigma p} n^p |\mathcal{M}_t|\end{aligned}$$

and so

$$\begin{aligned}p\sigma\int |A|^2 f_+^p d\mu \leq \frac{p(p-1)}{2}\int f_+^{p-2}|\nabla f|^2 d\mu + \frac{p}{c_1}\int \frac{f_+^p}{H2}|\nabla H|^2 d\mu \\ + 2KR^{-2-\sigma p} n^p |\mathcal{M}_t|.\end{aligned}$$

By Lemma 5.4 we conclude

$$\frac{d}{dt}\int f_+^p d\mu \leq K_2(n, \alpha, \eta, \sigma, p) R^{-2-\sigma p}|\mathcal{M}_0|,$$

which implies the assertion. $\square$

Proof of Theorem 5.3. In the smooth case the result can be obtained with the same procedure as in Theorems 4.6 and 4.13. It remains to prove that the function $(f_{\sigma,\eta})_+$ is decreasing under surgery with appropriate parameters, leaving the integral estimates above unchanged. To see this we compute first the algebraic identity

$$\frac{\partial}{\partial\lambda_i}\left(|A|^2 - \left(\frac{1}{n-1} + \eta\right)H^2\right) = 2\left(\lambda_i - \left(\frac{1}{n-1} + \eta\right)H\right). \tag{5.10}$$

Then the bent surface

$$\tilde{N}_\tau(p) = N(p) - \tau u(z)\nu(p), \quad p = (\omega, z) \in \mathbb{S}^{n-1} \times [0, 4\Lambda] \tag{5.11}$$

satisfies the following rate of change in view of (3.38) for the change of eigenvalues:

$$\begin{aligned}
\frac{d}{d\tau}\left(|A|^2 - \left(\frac{1}{n-1}+\eta\right)H^2\right)
&= 2\sum_i \left[\left(\lambda_i - \left(\frac{1}{n-1}+\eta\right)H\right)\left(u\lambda_i^2 + \delta_i^1 D_1 D_1 u\right)\right] \\
&\quad + O(\theta H D_1 D_1 u) \\
&= 2\left[u\sum_i \lambda_i^3 - \left(\frac{1}{n-1}+\eta\right)uH|A|^2 + \lambda_1 D_1 D_1 u\right. \\
&\qquad \left. - \left(\frac{1}{n-1}+\eta\right)HD_1D_1u\right] + O(\theta H D_1 D_1 u). \qquad (5.12)
\end{aligned}$$

Then notice that for $\varepsilon < \varepsilon_0$ sufficiently small the eigenvalues $\lambda_2, \dots, \lambda_n$ are approximately equal to $\frac{1}{n-1}H$ while λ_1 and u are small such that the dominating term on the RHS is $-2(\frac{1}{n-1}+\eta)HD_1D_1u$. Hence we get for $0 < \eta < \eta_0$ sufficiently small that

$$|\tilde{A}|^2 - \left(\frac{1}{n-1}+\eta\right)\tilde{H}^2 \le |A|^2 - \left(\frac{1}{n-1}+\eta\right)H^2 - \frac{\tau_0}{n-1}HD_1D_1u \tag{5.13}$$

on $[0, 3\Lambda]$. Hence the positive region of $f_{\sigma,\eta}$ is shrinking, and, since the denominator is increasing, $(f_{\sigma,\eta})_+$ is nonincreasing everywhere on $[0, 3\Lambda]$. Since on $[2\Lambda, 3\Lambda]$ the term D_1D_1u is bounded from below, the integrand is decreasing by some definite amount there, allowing us to blend with an axially symmetric surface as in (3.22), creating the surface $\hat{N}$. Then $\hat{N}$ will still satisfy the estimate (5.13) with $\frac{1}{2}D_1D_1u$ replaced by $\frac{1}{4}D_1D_1u$. Thus after the surgery the integrand $(f_{\sigma,\eta})_+$ will be zero everywhere on $[2\Lambda, 4\Lambda]$, completing the argument. □

6 Derivative estimates for the curvature

In this section we derive a pointwise derivative estimate for the curvature in mean curvature flow of 2-convex surfaces, depending only on the mean curvature at a given point rather than on some maximum of curvature as in the more general derivative estimates of [7].

The proof makes essential use of the assumption $n \ge 3$ ensuring that the constant κ_n,

$$\kappa_n = \frac{1}{2}\left(\frac{3}{n+2} - \frac{1}{n-1}\right), \tag{6.1}$$

satisfies $\kappa_n > 0$. This enables us to combine the estimate of the previous section with the inequality $|\nabla A|^2 \geq 3|\nabla H|^2/(n+2)$ obtained from the Codazzi equations in [16, Lemma 2.1]. Notice that a corresponding a priori derivative estimate for $n = 2$ cannot be true on general immersed mean convex surfaces moving by mean curvature since it is wrong on the one-dimensional translating grim reaper curve modelling possible singularities of the flow in that case. We also note that at this stage there is no analogous direct a priori estimate known for Ricci flow, the corresponding estimate for curvature derivatives in [25,26] is obtained via contradiction arguments.

Theorem 6.1 *Let $\mathcal{M}_t$ in $\mathcal{C}(R, \alpha)$ be a solution of mean curvature flow with surgery and normalised initial data. Then there is a constant $\gamma_2 = \gamma_2(n)$ and a constant γ_3 depending only on n, α such that for suitable surgery parameters as in the previous two sections the flow satisfies the uniform estimate*

$$|\nabla A|^2 \leq \gamma_2 |A|^4 + \gamma_3 R^{-4}$$

for every $t \geq (1/4)R^2$.

Proof. The main idea will be to estimate by the maximum principle the function

$$\frac{|\nabla A|^2}{g_1 g_2}$$

where $g_i = a_i H^2 - |A|^2 + C_i R^{-2}$ for $i = 1, 2$. Here a_1, a_2 are suitable constants greater than $1/(n-1)$ which will be specified below, while C_1, C_2 will be chosen in such a way that the corresponding g_i is strictly positive.

We start by defining g_1 and deriving some estimate from its evolution equation. Let us fix any $\eta \in (0, \kappa_n]$. We know from Theorem 5.3 that there exists $C_\eta = C(\eta, n, \alpha)$ such that

$$\left(\frac{1}{n-1} + \eta\right) H^2 - |A|^2 + C_\eta R^{-2} \geq 0.$$

Thus, if we set

$$g_1 = \left(\frac{1}{n-1} + \eta\right) H^2 - |A|^2 + 2C_\eta R^{-2},$$

we have $g_1 \geq C_\eta R^{-2}$ and $g_1 - 2C_\eta R^{-2} \geq g_1 - 2g_1 = -g_1$. Using the evolution equations for $|A|^2, H^2$ [16, Cor. 3.5] and the inequality

$|\nabla A|^2 \geq 3|\nabla H|^2/(n+2)$ [16, Lemma 2.1] we find

$$\begin{aligned}
\frac{\partial}{\partial t}g_1 - \Delta g_1 &= -2\left(\left(\frac{1}{n-1}+\eta\right)|\nabla H|^2 - |\nabla A|^2\right) + 2|A|^2\left(g_1 - \frac{2C_\eta}{R^2}\right) \\
&\geq 2\left(1 - \frac{n+2}{3}\left(\frac{1}{n-1}+\eta\right)\right)|\nabla A|^2 - 2|A|^2 g_1 \\
&\geq 2\kappa_n \frac{n+2}{3}|\nabla A|^2 - 2|A|^2 g_1. \qquad (6.2)
\end{aligned}$$

Let us now define g_2. We know that there is $C_0 = C_0(n,\alpha)$ such that

$$\left(\frac{1}{n-1}+\kappa_n\right)H^2 - |A|^2 + C_0 R^{-2} \geq 0.$$

Let us set

$$g_2 = \frac{3}{n+2}H^2 - |A|^2 + 2C_0 R^{-2}.$$

Then we have $g_2 \geq \kappa_n H^2 + C_0 R^{-2} \geq \kappa_n |A|^2/n + C_0 R^{-2}$. In addition

$$\begin{aligned}
\frac{\partial}{\partial t}g_2 - \Delta g_2 &= -2\left(\frac{3}{n+2}|\nabla H|^2 - |\nabla A|^2\right) + 2|A|^2\left(g_2 - \frac{2C_0}{R^2}\right) \\
&\geq -2|A|^2 g_2. \qquad (6.3)
\end{aligned}$$

We now want to derive a differential inequality for $|\nabla A|^2/g_1$. We have the general formula

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{f}{g}\right) - \Delta\left(\frac{f}{g}\right) \\
&\quad = \frac{2}{g}\left\langle \nabla g, \nabla\left(\frac{f}{g}\right)\right\rangle + \frac{1}{g}\left(\frac{\partial f}{\partial t} - \Delta f\right) - \frac{f}{g^2}\left(\frac{\partial g}{\partial t} - \Delta g\right). \qquad (6.4)
\end{aligned}$$

In addition (see [16, Th. 7.1])

$$\frac{\partial}{\partial t}|\nabla A|^2 - \Delta|\nabla A|^2 \leq -2|\nabla^2 A|^2 + c_n|A|^2|\nabla A|^2. \qquad (6.5)$$

Observe that, by the Schwarz inequality, we have for any function g

$$\begin{aligned}
\left\langle \nabla g, \nabla|\nabla A|^2\right\rangle &= 2\nabla_i g \nabla^i \nabla^k h^{lm} \nabla_k h_{lm} \\
&\leq 2|\nabla g||\nabla A||\nabla^2 A| \leq \frac{1}{g}|\nabla g|^2|\nabla A|^2 + g|\nabla^2 A|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
&-\frac{2}{g}|\nabla^2 A|^2 + \frac{2}{g}\left\langle \nabla g, \nabla\left(\frac{|\nabla A|^2}{g}\right)\right\rangle \\
&\quad = -\frac{2}{g}|\nabla^2 A|^2 - \frac{2}{g^3}|\nabla g|^2|\nabla A|^2 + \frac{2}{g^2}\langle \nabla g, \nabla|\nabla A|^2\rangle \leq 0.
\end{aligned}$$

It follows, using (6.4), (6.2) and (6.5),

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{|\nabla A|^2}{g_1}\right) - \Delta\left(\frac{|\nabla A|^2}{g_1}\right) \\
&\qquad\le \frac{2}{g_1}\left\langle \nabla g_1, \nabla\left(\frac{|\nabla A|^2}{g_1}\right)\right\rangle + \frac{1}{g_1}\big(-2|\nabla^2 A|^2 + c_n|A|^2|\nabla A|^2\big) \\
&\qquad\quad - 2\frac{|\nabla A|^2}{g_1^2}\left(\kappa_n \frac{n+2}{3}|\nabla A|^2 - |A|^2 g_1\right) \\
&\qquad\le (c_n+2)|A|^2\frac{|\nabla A|^2}{g_1} - 2\kappa_n\frac{n+2}{3}\frac{|\nabla A|^4}{g_1^2}.
\end{aligned}$$

Now we apply again (6.4) with $f = |\nabla A|^2/g_1$ and $g = g_2$. Taking into account (6.3) we obtain

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{|\nabla A|^2}{g_1 g_2}\right) - \Delta\left(\frac{|\nabla A|^2}{g_1 g_2}\right) - \frac{2}{g_2}\left\langle \nabla g_2, \nabla\left(\frac{|\nabla A|^2}{g_1 g_2}\right)\right\rangle \\
&\quad\le \frac{1}{g_2}\left((c_n+2)|A|^2\frac{|\nabla A|^2}{g_1} - 2\kappa_n\frac{n+2}{3}\frac{|\nabla A|^4}{g_1^2}\right) + 2\frac{|\nabla A|^2}{g_1 g_2^2}|A|^2 g_2 \\
&\quad\le \frac{1}{g_2}\left((c_n+4)|A|^2\frac{|\nabla A|^2}{g_1} - 2\kappa_n\frac{n+2}{3}\frac{|\nabla A|^4}{g_1^2}\right) \\
&\quad\le \frac{|\nabla A|^2|A|^2}{g_1 g_2}\left((c_n+4) - 2\kappa_n^2\frac{n+2}{3n}\frac{|\nabla A|^2}{g_1 g_2}\right).
\end{aligned}$$

Let us first consider the case without surgeries. In view of Proposition 2.7(v) and the interior a priori estimate for the derivatives of curvature in [7] we have at the (smooth) time $t_0 = (1/4)R^2$ an upper bound $|\nabla A|^2 \le m_0 R^{-4}$ with a constant m_0 depending only on n, α. Applying the maximum principle and recalling that $g_1 g_2 \ge R^{-4}$, we obtain

$$\frac{|\nabla A|^2}{g_1 g_2} \le \max\left\{m_0, \frac{3n(c_n+4)}{2\kappa_n 2(n+2)}\right\}$$

as desired. To derive the estimate in the case of a flow with surgeries, let us be more specific in our choice of test function by choosing $\eta = \kappa_n$ in the definition of g_1. Since $|A|^2 - H^2/(n-1) = 0$ on an exact cylinder it is clear from our construction in Sect. 3 that for all reasonable surgery parameters we have the estimate $H^2/(n-1) - |A|^2 \ge -\kappa_n H^2/2$ in the region of the surface affected by a surgery. Hence, in such a region we have $g_1 \ge \kappa_n H^2/2$ as well as $g_2 \ge 3\kappa_n H^2/2$. Furthermore, notice that $|\nabla A|^2 = 0$ on a standard cylinder, which implies that for any (ε, k)-neck with $k \ge 1$ the quantity $|\nabla A|^2$ is very small compared to H^4. For a given choice of transition function φ and choice of a convex cap in steps d) and e) of the surgery construction of Sect. 3 there is a fixed constant μ_0 depending only on n such that for all surgery parameters considered in the previous Sects. 4 and 5 we have the uniform estimate $|\nabla A|^2 \leq \mu_0 H^4$ on the region altered by surgery and hence

$$\frac{|\nabla A|^2}{g_1 g_2} \leq \frac{4\mu_0}{3\kappa_n^2}.$$

Iterating the argument in every time interval between two surgeries we find

$$\frac{|\nabla A|^2}{g_1 g_2} \leq \max\left\{m_0, \frac{3n(c_n+4)}{2\kappa_n^2(n+2)}, \frac{4\mu_0}{3\kappa_n^2}\right\}.$$

Since we have chosen $\eta = \kappa_n$ in the definition of g_1, the corresponding constant C_η only depends on n, α. Thus,

$$g_1 g_2 \leq H^4 + C(\alpha, n) R^{-4}$$

and so the above estimate implies

$$|\nabla A|^2 \leq c(n)|A|^4 + C(\alpha, n) R^{-4}.$$

□

Remark 6.2 Observe that on a neck we have, up to lower order terms, $g_1 \approx \eta H^2$. Hence, on a neck the gradient estimate $|\nabla A|^2 \leq c g_1 g_2$ implies $|\nabla A|^2 \leq c\eta H^4 + C$ and so it can be interesting to choose η small in the definition of g_1. Since this property will not be used in the sequel, we have just chosen a fixed η (not small) in the proof.

Let us also observe that, if we consider an ancient solution obtained as limit of rescalings, the lower order terms vanish and we can let $\eta \to 0$ in the definition of g_1 to obtain that $|\nabla A|^2 \leq cH^2(\frac{1}{n-1}H^2 - |A|^2)$.

In order to control the time derivative of curvature with an explicit a priori estimate we derive estimates for the second derivatives of curvature.

Theorem 6.3 *Let $\mathcal{M}_t$ in $\mathcal{C}(R, \alpha)$ be a solution of mean curvature flow with surgery and normalised initial data. Then there is a constant $\gamma_4 = \gamma_4(n)$ and a constant γ_5 depending only on n, α such that for suitable surgery parameters as in the previous two sections the flow satisfies the uniform estimate*

$$|\nabla^2 A|^2 \leq \gamma_4 |A|^6 + \gamma_5 R^{-6}$$

for every $t \geq (1/4)R^2$.

Proof. We assume as in the proof of the previous theorem that a fixed choice has been made for the transition function φ and the convex cap in steps d) and e) of the surgery construction in Sect. 3 such that the estimate of the previous theorem holds for all surgery parameters under consideration. We infer from the interior regularity theory in [7] that at time

$t_0 = (1/4)R^2$ there is a bound $|\nabla A|^2 \le m_0 R^{-4}$ as well as some upper bound $|\nabla^2 A|^2 \le m_1 R^{-6}$ with constants m_i depending only on n, α. In the following computations we denote by $k_1, k_2, \dots$ any constant depending only on n and by $C_1, C_2, \dots$ the constants depending also on α. We have (see [16, Th. 7.1])

$$\begin{aligned}\frac{\partial}{\partial t}|\nabla^2 A|^2 - \Delta|\nabla^2 A|^2 &\le -2|\nabla^3 A|^2 + k_1|A|^2|\nabla^2 A|^2 + k_2|A||\nabla A|^2|\nabla^2 A| \\ &\le -2|\nabla^3 A|^2 + (k_1 + k_2/2)|A|^2|\nabla^2 A|^2 \\ &\quad + (k_2/2)|\nabla A|^4.\end{aligned}$$

Therefore

$$\begin{aligned}&\frac{\partial}{\partial t}\frac{|\nabla^2 A|^2}{H^5} - \Delta\frac{|\nabla^2 A|^2}{H^5} \\ &\qquad \le \frac{1}{H^5}\left(-2|\nabla^3 A|^2 + \left(k_1 + \frac{k_2}{2}\right)|A|^2|\nabla^2 A|^2 + \frac{k_2}{2}|\nabla A|^4\right) \\ &\qquad\quad - 5|A|^2\frac{|\nabla^2 A|^2}{H^5} - 30\frac{|\nabla^2 A|^2}{H^7}|\nabla H|^2 + \frac{10}{H^6}\langle \nabla H, \nabla|\nabla^2 A|^2\rangle.\end{aligned}$$

Now if we estimate

$$\frac{10}{H^6}\langle \nabla H, \nabla|\nabla^2 A|^2\rangle \le \frac{1}{H^5}|\nabla^3 A|^2 + \frac{100}{H^7}|\nabla H|^2|\nabla^2 A|^2$$

we obtain, using also Theorem 6.1,

$$\begin{aligned}&\frac{\partial}{\partial t}\frac{|\nabla^2 A|^2}{H^5} - \Delta\frac{|\nabla^2 A|^2}{H^5} \\ &\qquad \le -\frac{|\nabla^3 A|^2}{H^5} + k_3\frac{|\nabla^2 A|^2}{H^3} + 70\frac{|\nabla^2 A|^2}{H^7}|\nabla H|^2 + \frac{k_2}{2}\frac{|\nabla A|^4}{H^5} \\ &\qquad \le -\frac{|\nabla^3 A|^2}{H^5} + k_4\frac{|\nabla^2 A|^2}{H^3} + \frac{C_1}{R^4}\frac{|\nabla^2 A|^2}{H^7} + \frac{k_5 H^8 + C_2 R^{-8}}{H^5}.\end{aligned}$$

Similarly we find,

$$\begin{aligned}\frac{\partial}{\partial t}\frac{|\nabla A|^2}{H^3} - \Delta\frac{|\nabla A|^2}{H^3} &\le -\frac{|\nabla^2 A|^2}{H^3} + \frac{k_6 H^8 + C_3 R^{-8}}{H^5}, \\ \frac{\partial}{\partial t}\frac{|\nabla A|^2}{H^7} - \Delta\frac{|\nabla A|^2}{H^7} &\le -\frac{|\nabla^2 A|^2}{H^7} + \frac{k_7 H^8 + C_4 R^{-8}}{H^9}.\end{aligned}$$

Let us now set

$$f = \frac{|\nabla^2 A|^2}{H^5} + N\frac{|\nabla A|^2}{H^3} + \frac{M}{R^4}\frac{|\nabla A|^2}{H^7} - \kappa H$$

where $N, M, \kappa > 0$ will be chosen below. We have

$$\begin{aligned}\frac{\partial}{\partial t} f - \Delta f \le{}& k_4 \frac{|\nabla^2 A|^2}{H^3} + k_5 H^3 + \frac{C_1}{R^4} \frac{|\nabla^2 A|^2}{H^7} + \frac{C_2}{R^8 H^5} \\ &- N \frac{|\nabla^2 A|^2}{H^3} + N k_6 H^3 + \frac{N C_3}{R^8 H^5} \\ &- \frac{M}{R^4} \frac{|\nabla^2 A|^2}{H^7} + \frac{k_7 M}{R^4 H} + \frac{C_4 M}{R^{12} H^9} - \kappa |A|^2 H.\end{aligned}$$

If we choose

$$N > k_4, \quad \kappa > n(N k_6 + k_5), \quad M > C_1,$$

then we find, taking into account that $H \ge \alpha_1 R^{-1}$,

$$\frac{\partial}{\partial t} f - \Delta f \le C_5 R^{-3},$$

which implies

$$\max_{\mathcal{M}_t} f \le \max_{\mathcal{M}_{t_0}} f + C_5 R^{-3}(t - t_0).$$

Since $\mathcal{M}_t$ is in $C(R, \alpha)$ we have at $t = t_0 = (1/4)R^2$ the estimate

$$f \le \left(\frac{m_1}{\alpha_1^5} + \frac{N m_0}{\alpha_1^3} + \frac{M m_0}{\alpha_1^7} - \kappa \alpha_1 \right) R^{-1}.$$

In addition, in a region added or modified by a surgery we have estimates $|\nabla A|^2 \le \mu_0 H^4$ and $|\nabla^2 A|^2 \le \mu_1 H^6$ with constants μ_0, μ_1 for each n only depending on the fixed choices in steps d) and e) of the surgery construction in Sect. 3, uniformly in all other surgery parameters. Choosing then κ somewhat larger,

$$\kappa > n(N k_6 + k_5) + \mu_1 + N \mu_0,$$

we conclude that on a region added or modified by surgery

$$f \le (\mu_1 + N\mu_0 - \kappa) H + M \mu_0 \alpha_1^{-3} R^{-1} \le M \mu_0 \alpha_1^{-3} R^{-1}.$$

Recalling the upper bound on the time interval in Proposition 2.7(iv) we conclude that at any time we have $f \le C(n, \alpha, \mu_0, \mu_1) R^{-1}$, which implies that

$$|\nabla^2 A|^2 \le c(n, \mu_0, \mu_1) H^6 + C(n, \alpha, \mu_0, \mu_1) H^5 R^{-1}.$$

□

In a completely analogous way we can estimate the derivatives $|\nabla^m A|$ of any order $m \le k_0$, where k_0 is the parameter counting the number of derivatives in the necks where we do the surgeries. In addition, the estimates on the space derivatives immediately yield estimates on the time derivatives, e.g. $|\partial_t h^i_j| = |\Delta h^i_j + |A|^2 h^i_j| \le |\nabla^2 A| + |A|^3 \le c_1|A|^3 + c_2 R^{-3}$. Strictly speaking, such an argument cannot be applied at a surgery time T_i, since the flow in this case is not even continuous with respect to t. However, the flow is smooth and satisfies the estimates before and after T_i; thus, the right and left time derivatives $\partial_t A(p, T_i-)$ and $\partial_t A(p, T_i+)$ exist and satisfy the same estimate as above. With an abuse of notation, we will simply write $\partial_t A$, with the convention that at the surgery times we mean the one sided time derivatives. Summarizing, we have the following result.

Corollary 6.4 *Under the same hypotheses as in the previous theorem there are constants $\gamma' = \gamma'(n)$ and $\gamma'' = \gamma''(n, \alpha)$ such that*

$$\left|\partial_t^h \nabla^m A\right|^2 \le \gamma' |A|^{4h+2m+2} + \gamma'' R^{-(4h+2m+2)}$$

for all $h, m \ge 0$ such that $2h + m \le k_0$.

To estimate the derivatives of order higher than k_0 we can apply the results of [7]. Such results, however, are based on interior parabolic regularity, and therefore can only be applied away from the surgery times. The estimates of Corollary 6.4, instead, hold with the same constants, regardless of how close we are to a surgery time.

As a special case of our estimates, we obtain the following statement, which is convenient for the analysis of the regions where the curvature is large. It is interesting to compare this result with estimate (1.3) in [26] for ancient solutions to the Ricci flow, which is obtained by a completely different method.

Corollary 6.5 *Let $\mathcal{M}_t$ be a mean curvature flow with surgeries starting from a surface in $C(R, \alpha)$. Then we can find $c^\sharp > 0$, $H^\sharp > 0$ such that, for all $p \in \mathcal{M}$ and $t > 0$,*

$$H(p,t) \ge H^\sharp \implies |\nabla H(p,t)| \le c^\sharp H^2(p,t), \quad |\partial_t H(p,t)| \le c^\sharp H^3(p,t), \tag{6.6}$$

where $c^\sharp$ only depends on the dimension n, and $H^\sharp = h_0 R^{-1}$ with h_0 depending on α, n.

It is well known that gradient estimates like the above ones allow to control the size of the curvature in a neighbourhood of a given point. We conclude the section by giving one of such applications, which will be needed in the following.

Lemma 6.6 *Let $F : \mathcal{M} \to \mathbb{R}^{n+1}$ be an n-dimensional immersed surface. Suppose that there are $c^\sharp$, $H^\sharp > 0$ such that $|\nabla H(p)| \le c^\sharp H^2(p)$ for any $p \in \mathcal{M}$ such that $H(p) \ge H^\sharp$. Let $p_0 \in \mathcal{M}$ satisfy $H(p_0) \ge \gamma H^\sharp$ for some*

$\gamma > 1$. *Then*

$$H(q) \geq \frac{H(p_0)}{1 + c^\sharp d(p_0, q) H(p_0)} \geq \frac{H(p_0)}{\gamma} \quad \textit{for all } q$$
$$\textit{s.t.} \quad d(p_0, q) \leq \frac{\gamma - 1}{c^\sharp} \frac{1}{H(p_0)}.$$

Proof. Consider first the case where there are points $q \in \mathcal{M}$ such that $H(q) < H(p_0)/\gamma$. Let q_0 be a point with this property having minimal distance from p_0. Then let us set $d_0 = d(p_0, q_0) H(p_0)$ and $\theta_0 = \min\{d_0, \frac{\gamma-1}{c^\sharp}\}$.

Now let $q \in \mathcal{M}$ be any point with $d(q, p_0) \leq \theta_0 / H(p_0)$, and let ζ : $[0, d(p_0, q)] \to \mathcal{M}$ be a geodesic from p_0 to q, parametrised by arc length. By definition of θ_0, we have

$$H(\zeta(s)) \geq H(p_0)/\gamma \geq H^\sharp$$

for any $s \in [0, d(p_0, q)]$. Then we have $|\nabla H(\zeta(s))| \leq c^\sharp H^2(\zeta(s))$ and

$$\frac{d}{ds} H(\zeta(s)) \geq -c^\sharp H^2(\zeta(s))$$

for all $s \in [0, d(p_0, q)]$. Integrating we obtain

$$H(\zeta(s)) \geq \frac{H(p_0)}{1 + c^\sharp s H(p_0)}, \quad s \in [0, d(p_0, q)],$$

which implies

$$H(q) \geq \frac{H(p_0)}{1 + c^\sharp d(p_0, q) H(p_0)} \geq \frac{H(p_0)}{1 + c^\sharp \theta_0}. \tag{6.7}$$

This holds for all q such that $d(p_0, q) \leq \frac{\theta_0}{H(p_0)}$. Suppose now that $d_0 < \frac{\gamma-1}{c^\sharp}$. Then $d_0 = \theta_0$ and (6.7) holds also with $q = q_0$. This yields $H(q_0) > H(p_0)/\gamma$, in contradiction with the definition of q_0. Therefore we have $d_0 \geq \frac{\gamma-1}{c^\sharp}$, which implies that $\theta_0 = \frac{\gamma-1}{c^\sharp}$. Then (6.7) becomes our assertion.

In the case where $H(q) \geq H(p_0)/\gamma$ for all $q \in \mathcal{M}$, then we have $|\nabla H| \leq c^\sharp H^2$ everywhere, and so the result follows from the same argument in a more direct way. □

7 Neck detection

We want to show that the surgery procedure introduced in the previous sections can be used to alter mean curvature flow before a singular time in such a way that the mean curvature stays bounded, unless we already recognize the surface to be convex or of type $\mathbb{S}^{n-1} \times \mathbb{S}^1$. To this purpose we need results ensuring that, if we are close enough to a singular time and the surface is not yet uniformly convex everywhere, the regions of our manifold with largest curvature contain a neck where we can do the surgery. In this section we will do this preliminary analysis, collecting the results which will be needed for the surgery algorithm of the next section.

We start by introducing some notation. Let us first consider a smooth mean curvature flow (without surgeries) $F : \mathcal{M} \times [0, T_0] \to \mathbb{R}^{n+1}$. We denote by $g(t)$ the metric on $\mathcal{M}$ induced by the immersion at time t. Given $p \in \mathcal{M}$ and $r > 0$, we let $\mathcal{B}_{g(t)}(p, r) \subset \mathcal{M}$ be the closed ball of radius r around p with respect to the metric $g(t)$. In addition, if t, θ are such that $0 \leq t - \theta < t \leq T_0$, we set

$$\mathcal{P}(p, t, r, \theta) = \{(q, s) \ : \ q \in \mathcal{B}_{g(t)}(p, r), s \in [t - \theta, t]\}. \tag{7.1}$$

Such a set will be called a (backward) *parabolic neighbourhood* of (p, t). This definition of parabolic neighbourhood agrees with the one used in [26].

Let us extend these definitions to the case of a flow with surgeries. In this case we have a family of flows $F^i : \mathcal{M}_i \times [T_{i-1}, T_i] \to \mathbb{R}^{n+1}$, where $T_0 = 0$ and $0 < T_1 < T_2 < \ldots$ are the surgery times. We simply write $\mathcal{M}$, F instead of $\mathcal{M}_i$, F^i if there is no risk of confusion. For a flow with surgeries, we can define $\mathcal{B}_{g(t)}(p, r)$ like in the smooth case. The neighbourhood $\mathcal{B}_{g(t)}(p, r)$ will belong to the manifold $\mathcal{M}_i$ corresponding to the interval $[T_{i-1}, T_i]$ containing t. We have to be more specific when t coincides with one of the surgery times T_i, since the definition becomes ambiguous; in this case we write $g(t-)$ (resp. $g(t+)$) to denote the manifold before (resp. after) the surgery. We also use the convention that $g(t) = g(t-)$, that is, at the surgery times we take our flow to be continuous from the left. If we want to consider parabolic neighbourhoods in the case of a flow with surgeries, we must take into account that $\mathcal{P}(p, t, r, \theta)$ might be not well defined if there are surgeries between time $t - \theta$ and t. Therefore, we give the following definition.

Definition 7.1 *Let* $F^i : \mathcal{M}_i \times [T_{i-1}, T_i] \to \mathbb{R}^{n+1}$, $i = 1, 2, \ldots$ *be a mean curvature flow with surgeries. Let* $(p, t) \in \mathcal{M}_i \times [T_{i-1}, T_i]$ *for some* i*, let* $\theta \in (0, t]$ *and* $r > 0$. *We say that* $\mathcal{B}_{g(t)}(p, r)$ *has not been* changed by the surgeries *in the interval* $[t - \theta, t]$ *if there are no points of* $\mathcal{B}_{g(t)}(p, r)$ *which belong to a region changed by a surgery occurred at a time* $s \in (t - \theta, t]$. *In this case, we define the parabolic neighbourhood* $\mathcal{P}(p, t, r, \theta)$ *according to* (7.1) *as in the smooth case. We also describe this behaviour by saying that* $\mathcal{P}(p, t, r, \theta)$ *does not* contain surgeries.

Observe that in the above definition we allow the presence of surgeries in the time interval $(t - \theta, t]$, provided they are performed on parts of the surface disjoint from our domain $\mathcal{B}_{g(t)}(p, r)$. In this case $\mathcal{B}_{g(t)}(p, r)$ is a subset of different $\mathcal{M}_i$'s before and after the surgery times; however, since it is not changed by the surgeries, we can consider it as a fixed domain during the whole time interval $[t - \theta, t]$. Observe also that we allow $\mathcal{B}_{g(t)}(p, r)$ to be changed by a surgery at the initial time $t - \theta$; in this case, at time $t - \theta$ we consider the manifold $\mathcal{M}_{t-\theta+}$ after the surgery, so that the parabolic neighbourhood does not keep track of the surgery at the initial time.

Let us introduce some notation which simplifies some formulas in the analysis of necks. Given a point (p, t), we set

$$\hat{r}(p,t) := \frac{n-1}{H(p,t)}, \quad \hat{\mathcal{P}}(p,t,L,\theta) := \mathcal{P}(p,t,\hat{r}(p,t)L,\hat{r}(p,t)^2\theta). \quad (7.2)$$

Observe that, if (p, t) lies on a neck, then $\hat{r}(p,t)$ is approximately equal to the radius of the neck. In addition, if we rescale the flow in space and time in order to have $\hat{r}(p,t) = 1$, then $\hat{\mathcal{P}}(p,t,L,\theta)$ becomes $\mathcal{P}(p,t,L,\theta)$.

Throughout this section, we consider a mean curvature flow with surgeries where the initial manifold belongs to $\mathcal{C}(R,\alpha)$ for suitable R, α, and where the surgeries are done according to the procedure of the previous sections for a fixed choice of the parameters.

The following lemma is a simple consequence of the gradient estimates. Particularly relevant for our later applications is part (ii), stating that if a point (p, t) has a curvature which is much larger than the curvature in the regions modified by the previous surgeries, then a suitable parabolic neighbourhood of (p, t) does not contain surgeries. This property depends in a crucial way on the fact that our gradient estimates are not obtained using interior parabolic regularity and therefore they hold in every smooth part of the flow with the same constants, regardless of how close we are to the surgeries.

Lemma 7.2 *Let $c^\sharp, H^\sharp$ be the constants of Corollary 6.5. Define $d^\sharp = (8(n-1)^2c^\sharp)^{-1}$. Then the following properties hold.*

(i) *Let (p,t) satisfy $H(p,t) \geq 2H^\sharp$. Then, given any $r, \theta \in (0, d^\sharp]$ such that $\hat{\mathcal{P}}(p,t,r,\theta)$ does not contain surgeries, we have*

$$\frac{H(p,t)}{2} \leq H(q,s) \leq 2H(p,t) \quad (7.3)$$

for all $(q,s) \in \hat{\mathcal{P}}(p,t,r,\theta)$.

(ii) *Suppose that, for any surgery performed at time less than t, the regions modified by the surgery have mean curvature less than K, for some $K \geq H^\sharp$. Let (p,t) satisfy $H(p,t) \geq 2K$. Then the parabolic neighbourhood*

$$\mathcal{P}\left(p, t, \frac{1}{8c^\sharp K}, \frac{1}{8c^\sharp K^2}\right) \quad (7.4)$$

does not contain surgeries. In particular, the neighbourhood $\hat{\mathcal{P}}(p,t,d^\sharp,d^\sharp)$ does not contain surgeries and all points (q,s) contained there satisfy (7.3).

Proof. Both parts are obtained integrating the gradient estimates of Corollary 6.5 in space and time with the same procedure of Lemma 6.6. Let us just show the proof of (ii). Suppose that the neighbourhood in (7.4) is modified by surgeries. Let us take a point (q, s) in there which is modified by a surgery, with s the maximal time at which we can find such a point. By our assumptions we have $H(q, s+) \leq K$. Integrating the estimate on $\partial_t H$ from Corollary 6.5 we obtain

$$\frac{1}{H^2(q,t)} \geq \frac{1}{H^2(q,s)} - 2c^\sharp(t-s) \geq \frac{3}{4K^2}.$$

Then we can integrate along a geodesic from q to p at time t and use the estimate on ∇H to obtain

$$\frac{1}{H(p,t)} \geq \frac{1}{H(q,t)} - c^\sharp d_{g(t)}(p,q) \geq \frac{\sqrt{3} - 1/4}{2K} > \frac{1}{2K},$$

in contradiction with our assumptions. In this computation we have assumed that $H \geq H^\sharp$ along the integration paths in order to apply the gradient estimates of Corollary 6.5. If it is not so, we can choose the last point along the path with $H \leq H^\sharp$ and integrate from that point on, obtaining a contradiction in the same way. Finally, from the definitions of $d^\sharp$ we find

$$\frac{(n-1)}{H(p,t)} d^\sharp \leq \frac{1}{16(n-1)c^\sharp K} < \frac{1}{8c^\sharp K}, \quad \frac{(n-1)^2}{H(p,t)^2} d^\sharp \leq \frac{1}{64c^\sharp K^2} < \frac{1}{8c^\sharp K^2}.$$

Thus, by definition, $\hat{\mathcal{P}}(p, t, d^\sharp, d^\sharp)$ is contained in the neighbourhood (7.4). We deduce that $\hat{\mathcal{P}}(p, t, d^\sharp, d^\sharp)$ does not contain surgeries and that part (i) can be applied to this neighbourhood. □

In the following we will refer to the notions of neck introduced in Sect. 3. We say that a point (p, t) lies at the center of a (curvature, geometric, ...) neck if $p \in \mathcal{M}$ lies at the center of a neck with respect to the immersion $F(\cdot, t)$. Let us give a time-dependent version of the notion of curvature neck. We first set, for $s \leq 0$,

$$\rho(r,s) = \sqrt{r^2 - 2(n-1)s} \tag{7.5}$$

i.e. $\rho(r, s)$ is the radius at time s of a standard n-dimensional cylinder evolving by mean curvature flow and having radius r at time $s = 0$. Let us observe that, if $d^\sharp$ is the constant of Lemma 7.2, we have

$$r \leq \rho(r,s) \leq 2r, \quad \forall s \in [-d^\sharp r^2, 0], \tag{7.6}$$

otherwise the standard cylinder would violate Lemma 7.2(i).

Definition 7.3 *We say that a point* (p_0, t_0) *lies at the center of an* $(\varepsilon, k, L, \theta)$*-shrinking curvature neck if, after setting* $r_0 = \hat{r}_0(p_0, t_0)$ *and* $\mathcal{B}_0 = \mathcal{B}_{g(t_0)}(p_0, r_0 L)$*, the following properties hold:*

(i) *the parabolic neighbourhood* $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$ *does not contain surgeries;*

(ii) *for every* $t \in [t_0 - r_0^2\theta+, t_0]$, *the region* $\mathcal{B}_0$, *with respect to the immersion* $F(\cdot, t)$ *multiplied by the scaling factor* $\rho(r_0, t - t_0)^{-1}$, *is* ε*-cylindrical and* (ε, k)*-parallel at every point.*

The notation $t_0 - r_0^2\theta+$ in requirement (ii) means the limit from the right and keeps into account the case that $t_0 - r_0^2\theta$ is a surgery time. The definition says that at any point of $\mathcal{P}\left(p_0, t_0, r_0 L, r_0^2\theta\right)$ the Weingarten operator of our surface and its space derivatives, up to order k, after appropriate rescaling, are ε-close to the corresponding ones of the standard shrinking cylinder. Using the evolution equation for the Weingarten operator, one can check that also the derivatives of the form $\partial_t^i \nabla^h W$, with $2i + h \leq k$, are close to the ones on the cylinder up to an order $O(\varepsilon)$. Similarly to the above definition, one could give a time-dependent version of the notions of geometric necks, but it will not be necessary for our later analysis.

To be able to define a flow beyond the singular time using our surgery procedure, we have to show that the surface develops necks in the regions with large curvature as the singular time is approached. The next result (which will be called in the following the *neck detection lemma*) provides the first basic step in this direction.

Lemma 7.4 (Neck detection) *Let* $\mathcal{M}_t$, $t \in [0, T[$ *be a mean curvature flow with surgeries as in the previous sections, starting from an initial manifold* $\mathcal{M}_0 \in C(R, \alpha)$ *for some* R, α. *Let* $\varepsilon, \theta, L > 0$, *and* $k \geq k_0$ *be given (where* $k_0 \geq 2$ *is the parameter measuring the regularity of the necks where surgeries are performed). Then we can find* η_0, H_0 *with the following property. Suppose that* $p_0 \in \mathcal{M}$ *and* $t_0 \in [0, T[$ *are such that*

(ND1) $H(p_0, t_0) \geq H_0$, $\frac{\lambda_1(p_0, t_0)}{H(p_0, t_0)} \leq \eta_0$,

(ND2) *the neighbourhood* $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$ *does not contain surgeries.*

Then

(i) *the neighbourhood* $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$ *is an* $(\varepsilon, k_0 - 1, L, \theta)$*-shrinking curvature neck;*
(ii) *the neighbourhood* $\hat{\mathcal{P}}(p_0, t_0, L - 1, \theta/2)$ *is an* $(\varepsilon, k, L - 1, \theta/2)$*-shrinking curvature neck.*

The constant η_0 *only depends on* α *and on* $\varepsilon, k, L, \theta$, *while* H_0 *can be written as* $h_0 R^{-1}$, *with* h_0 *also depending only on* $\alpha, \varepsilon, k, L, \theta$.

Proof. We use a contradiction argument based on a rescaling procedure like the ones which are often used e.g. in [14,25]. Let us prove assertion (i). Suppose that for some values of ε, L, θ the conclusion is not true, no matter how η_0, H_0 are chosen. Then we can find a sequence $\{\mathcal{M}_t^j\}_{j \geq 1}$ of solutions to the flow, a sequence of times t_j, a sequence of points $p_j \in \mathcal{M}^j$ such that, denoting by $H_j, \lambda_{1,j}$ the values of H, λ_1 at the point $F_j(p_j, t_j) \in \mathcal{M}_{t_j}^j$, and setting $\hat{r}_j = (n-1)/H_j$, we have

(a) each flow starts from a manifold belonging to the same class $C(R, \alpha)$ and therefore satisfies the estimates of the previous sections with the same constants;
(b) the parabolic neighbourhood $\mathcal{P}^j(p_j, t_j, \hat{r}_j L, \hat{r}_j^2\theta)$ is not changed by surgeries (the superscript in $\mathcal{P}^j$ denotes that it belongs to $\mathcal{M}^j$);
(c) $H_j \to \infty$, $\lambda_{1,j}/H_j \to 0$ as $j \to \infty$;
(d) (p_j, t_j) does not lie at the center of an $(\varepsilon, k_0 - 1, L, \theta)$-shrinking neck.

Actually, to ensure that $\mathcal{P}^j(p_j, t_j, \hat{r}_j L, \hat{r}_j^2\theta)$ is well-defined, one should also require that $t_j - \hat{r}_j^2\theta > 0$. However, this certainly holds for j large enough, as a consequence of the other assumptions. In fact, we have that $H_j \to \infty$. Since the curvature of the flows is uniformly bounded at time $t = 0$ and attains the value H_j at time t_j, the sequence t_j must be bounded away from zero. Therefore $t_j > \theta(n-1)^2 H_j^{-2}$ for j large enough, which is equivalent to $t_j - \hat{r}_j^2\theta > 0$.

We now perform a parabolic rescaling of each flow $\mathcal{M}_t^j$ in such a way that the mean curvature at (p_j, t_j) becomes $n-1$; in addition, we translate in space and time so that (p_j, t_j) is mapped to the origin $0 \in \mathbb{R}^{n+1}$ and t_j becomes 0. More precisely, if F_j is the parametrization of the original flow $\mathcal{M}_t^j$, then we denote the rescaled flow is by $\bar{\mathcal{M}}_\tau^j$ and we define it as

$$\bar{F}_j(p, \tau) = \frac{1}{\hat{r}_j}\left[F_j\left(p, \hat{r}_j^2\tau + t_j\right) - F_j(p_j, t_j)\right].$$

For simplicity, we choose for every flow a local coordinate system centered at p_j, so that in these coordinates we can write 0 instead of p_j. The neighbourhood $\mathcal{P}^j(p_j, t_j, \hat{r}_j L, \hat{r}_j^2\theta)$ in the original flow becomes $\bar{\mathcal{P}}_j(0, 0, \theta, L)$ in the rescaled flow (we use the notation $\bar{\mathcal{P}}_j$ for a neighbourhood belonging to the flow $\bar{\mathcal{M}}_\tau^j$). By (b), such a neighbourhood contains no surgeries. Our aim is to show that the restrictions of the rescaled flows to $\bar{\mathcal{P}}_j(0, 0, \theta, L)$ converge, up to a subsequence, to a limit flow which is a portion of a shrinking cylinder, and this will yield a contradiction with (d).

By construction, each rescaled flow satisfies $\bar{F}_j(0, 0) = 0$, $\bar{H}_j(0, 0) = n - 1$. The gradient estimates of the previous section allow us to obtain bounds (uniform in j) on $|A|$ and its derivatives up to order k_0 at least in a neighbourhood of the form $\bar{\mathcal{P}}_j(0, 0, d, d)$, for a suitable $d > 0$ (in general smaller that L, θ). This implies uniform estimates for the immersions $\bar{F}_j$ in the C^{k_0+2}-norm, and therefore compactness in the C^{k_0+1} norm. Thus, in this neighbourhood a subsequence of the flows converges in the C^{k_0+1} norm to some limit flow, which we denote by $\tilde{\mathcal{M}}_\tau^\infty$.

Let us analyze the properties of the limit flow. We denote with a tilde the geometric quantities associated with $\tilde{\mathcal{M}}_\tau^\infty$. Passing to the limit in the convexity estimates we find that $\tilde{S}_i \geq 0$ for all $i = 1, \ldots, n$, which implies that $\tilde{\lambda}_i \geq 0$ for all $i = 1, \ldots, n$. On the other hand, we know that $\tilde{\lambda}_1 + \tilde{\lambda}_2 \geq \alpha_0 \tilde{H}$. We deduce that $\tilde{\lambda}_i > 0$ for $i \geq 2$, and so $\tilde{S}_i > 0$ for $i \leq n-1$. In addition, by property (c) above we have that $\tilde{\lambda}_1(0,0) = 0$.

We consider now the quotient $\tilde{Q}_n = \tilde{S}_n/\tilde{S}_{n-1}$. It is nonnegative everywhere and, by the strong maximum principle (the evolution equation satisfied by Q_n can be found in (3.2) in [19]), if it positive somewhere in the interior of $\tilde{\mathcal{P}}^\infty(0,0,d,d)$, it is positive everywhere at all later times. But $\tilde{Q}_n(0,0) = 0$ and therefore $\tilde{Q}_n \equiv 0$ on $\tilde{\mathcal{P}}^\infty(0,0,d,d)$. This shows that $\tilde{\lambda}_1 \equiv 0$ in this set. An alternative proof of this claim can be obtained by applying Hamilton's maximum principle for tensors to the Weingarten operator $\tilde{h}^i_j$ of the limit flow.

From the property that $\tilde{\lambda}_1 = 0$ and $\tilde{\lambda}_i > 0$ for $i \geq 2$ we deduce that $|\tilde{A}|^2 - \frac{1}{n-1}\tilde{H}^2 \geq 0$. On the other hand, passing to the limit in the roundness estimates (Theorem 5.3), we deduce that $|\tilde{A}|^2 - \frac{1}{n-1}\tilde{H}^2 \leq 0$. Hence the quantity $|\tilde{A}|^2 - \frac{1}{n-1}\tilde{H}^2$ vanishes identically. Recalling the evolution equation for $|A|^2/H^2$, see (5.6), we obtain that the tensor $|H\nabla_i h_{kl} - \nabla_i H h_{kl}|^2$ vanishes identically on $\tilde{\mathcal{M}}^\infty_\tau$. Arguing as in [17, Theorem 5.1] (see also [21, Theorem 4]), we deduce that $\tilde{\mathcal{M}}^\infty_\tau$ on $\tilde{\mathcal{P}}^\infty(0,0,d,d)$ is a portion of a shrinking cylinder.

Now we iterate the procedure to show that the whole neighbourhoods $\bar{\mathcal{P}}^\infty_j(0,0,L,\theta)$ of the rescaled flows converge to a cylinder. From the first step we know that, for j large enough, the curvature on $\bar{\mathcal{P}}^\infty_j(0,0,d,d)$ is close to the curvature of a unit cylinder, and therefore satisfies e.g. $\bar{H}_j \leq 2(n-1)$. Then, using the gradient estimates, we have uniform bounds on $\bar{H}_j$ also on larger neighbourhoods, say on $\bar{\mathcal{P}}^\infty_j(0,0,2d,2d)$. We repeat the previous argument to prove convergence to a cylinder there. After a finite number of iterations we obtain convergence of the neighbourhoods $\bar{\mathcal{P}}^\infty_j(0,0,L,\theta)$. The immersions converge in the C^{k_0-1}-norm and this ensures that, for j large enough, the neighbourhoods are $(\varepsilon, k_0-1, L, \theta)$-shrinking necks. This contradicts assumption (d), and thus proves part (i) of the lemma.

The argument for assertion (ii) is very similar. Again we argue by contradiction and take a sequence of rescaled flows. As in the previous part, we have uniform C^2-bounds on the Weingarten operator. If we consider the smaller parabolic neighbourhoods $\bar{\mathcal{P}}_j(0,0,L-1,\theta/2)$, we can apply the interior regularity results of [7] to find bounds in the C^{k+1} norm as well. This yields compactness in the C^k-norm, which allows us to conclude also in this case. □

Remark 7.5 In assumption (ND1) the quantity λ_1/H is only required to be bounded from above. We do not require explicitly a corresponding bound from below because it follows from the convexity estimates together with the assumption that H is large.

We stress that the parameters ε, k, L in the above result are the ones describing the neck we are willing to find, and are in general different from the ones related to the necks where the surgeries are done. In fact, it will be useful at some stages to find necks which are more regular than what is needed for the standard surgery.

We also remark the difference between statements (i) and (ii) of the lemma. Part (i) concerns the whole parabolic neighbourhood $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$, which is surgery-free, but can be arbitrarily close to a surgery; the points of the neighbourhood are even allowed to be modified by a surgery at the initial time $t_0 - \theta r_0^2$. Therefore the description goes up to $k_0 - 1$ derivatives, that is, the regularity that is preserved by the surgeries minus one derivative which is lost in the compactness argument (this could be easily refined by the use of Hölder spaces, but it is not necessary for our purposes). Part (ii) of the statement, instead, is concerned with a smaller parabolic neighbourhood, where we can use interior parabolic regularity and have as many derivatives as we wish.

Our neck detection lemma is similar to Theorem C.5.1 by Hamilton [14]. However, in contrast to that result, we do not have to assume a priori a bound on the curvature in a neighbourhood of the point under consideration. In fact such a bound can be obtained from our gradient estimates, as we have seen in the proof.

Remark 7.6 It is easy to see that the constants η_0, H_0 of the neck detection lemma depend continuously on the parameters L, θ measuring the size of the parabolic neighbourhood. Thus, if $L_2 > L_1 > 0$ and $\theta_2 > \theta_1 > 0$, it is possible to find η_0, H_0 which apply to any $L \in [L_1, L_2]$ and $\theta \in [\theta_1, \theta_2]$.

We now derive some consequences and refinements of the neck detection lemma. In the first one we combine it with Proposition 3.5 and Theorem 3.14 to obtain that the point (p_0, t_0) lies also at the center of a cylindrical graph and of a normal hypersurface neck. We only consider the structure of the hypersurface at the final time t_0, which will be enough for most of our later arguments.

In what follows, we define the *length* of a hypersurface neck $\mathbb{S}^{n-1} \times [a, b] \to \mathbb{R}^{n+1}$ equal to $b - a$. The length defined in this way is a scale invariant quantity. In contrast with this, when we talk about the *distance* between points on the neck or on the manifold, we mean it with respect to the metric $g(t)$, and thus it depends on the scale. For instance, if a hypersurface neck has a length L and a radius approximately r_0, then the distance between the two components of its boundary is approximately $r_0 L$. Similarly, we use the term *cylindrical graph of length* $2L$ *and* C^{k+2}*-norm less than* ε to denote a region which, after appropriate rescaling, can be written as a graph of a function $u : \mathbb{S}^{n-1} \times [-L, L] \to \mathbb{R}$ satisfying $\|u\|_{C^{k+2}} \leq \varepsilon$. Like in the case of a hypersurface neck, the length is a scale invariant quantity; the actual distance between the two planes containing the boundary of the cylindrical graph is equal to $2L$ multiplied by the scaling factor.

Corollary 7.7 *Given* $\varepsilon, \theta > 0$*,* $L \geq 10$ *and* $k > 0$ *integer, we can find* $\eta_0, H_0 > 0$ *such that the following holds. Let* p_0, t_0 *satisfy* (ND1) *and*

(ND2) *of Lemma 7.4. Then*

(i) *the point* (p_0, t_0) *lies at the center of a cylindrical graph of length* $2(L-2)$ *and* C^{k+2}*-norm less than* ε*;*
(ii) *the point* (p_0, t_0) *lies at the center of a normal* $(\varepsilon, k, L-2)$*-hypersurface neck.*

Proof. By Proposition 3.5 and Theorem 3.14, both properties are true if (p_0, t_0) lies at the center of an $(\varepsilon', k', L-1)$-curvature neck for suitable ε', k'. A fortiori, the properties hold if (p_0, t_0) lies at the center of an $(\varepsilon', k', L-1, \theta/2)$-shrinking curvature neck. Thus, it suffices to apply part (ii) of the neck detection lemma with parameters $(\varepsilon', k', L, \theta)$ and use the corresponding values of η_0, H_0. □

The following lemma has a somehow more technical statement, but it is also an immediate consequence of the neck detection lemma and of its proof. It shows that the shrinking curvature neck given by that lemma can be represented at each time as a hypersurface neck; for our future applications, it is important to observe that this representation can be done at all times under consideration, even the initial one which could coincide with a surgery time. Before giving the result, let us remark an elementary property.

Remark 7.8 Let $C = \mathbb{S}^{n-1} \times \mathbb{R}$ be a standard cylinder in $\mathbb{R}^{n+1}$, and let $p_0 \in C$ be a point on the $x_{n+1} = 0$ section. Let $B_L \subset C$ be the set of points of C having (intrinsic) distance less than L from p_0. Clearly, B_L cannot be written in the form $\mathbb{S}^{n-1} \times [a, b]$ for any a, b. However, it is easy to see that, if $L \geq (\pi^2 + 1)/2$, then

$$\mathbb{S}^{n-1} \times [-(L-1), L-1] \subset B_L \subset \mathbb{S}^{n-1} \times [-L, L].$$

Roughly speaking, for large L the ball B_L differs little from a subcylinder of length $2L$.

Lemma 7.9 *In Lemma 7.4, we can choose the constants* η_0, H_0 *so that the additional following property holds. Suppose that* $L \geq 10$ *and that* $\theta \leq d^\sharp$*. Denote as usual*

$$r_0 = \frac{n-1}{H(p,t)}, \qquad \mathcal{B}_0 = \mathcal{B}_{g(t_0)}(p_0, r_0 L).$$

Then, for any $t \in [t_0 - \theta r_0^2 +, t_0]$*, the point* (p_0, t) *lies at the center of a* $(\varepsilon, k_0 - 1)$*-hypersurface neck* $\mathcal{N}_t \subset \mathcal{B}_0$*, satisfying the following properties:*

(i) *the mean radius* $r(z)$ *of every cross section of* $\mathcal{N}_t$ *is equal to* $\rho(r_0, t-t_0)(1 + O(\varepsilon))$*;*
(ii) *the length of* $\mathcal{N}_t$ *is at least* $L-2$*;*
(iii) *there exists a unit vector* $\omega \in \mathbb{R}^{n+1}$ *such that* $|\nu(p,t) \cdot \omega| \leq \varepsilon$ *for any* $p \in \mathcal{N}_t$*.*

Proof. The above result can be proved exactly in the same way as part (i) of Lemma 7.4. By a contradiction argument, we show that, for a suitable choice of η_0, H_0, our parabolic neighbourhood is as close as we wish to a portion of an exact cylinder evolving by mean curvature flow over the same time interval. The cylinder has radius r_0 at the final time, hence it has radius $\rho(r_0, t - t_0)$ at the previous times.

At the final time, C_{t_0} is a neighbourhood of radius $r_0 L$ of p_0; by Remark 7.8, it contains a subcylinder of length $2(L-1)$ (we recall that the length of a neck is computed after a homothety which makes the radius equal to 1). The same subcylinder is contained in C_t for $t < t_0$; however, since the scaling factor is given by $\rho(r_0, t-t_0)$ rather than r_0, the length of the subcylinder becomes $2r_0(L-1)/\rho(r_0, t-t_0)$. Recalling (7.6), we see that the subcylinder has length at least $2(L-1)$ for the times under consideration. Since we can make our parabolic neighbourhood as close as we wish in the $(k_0 - 1)$-norm to the cylinder C_t, we can find a geometric neck parametrizing the part of the neighbourhood corresponding to the subcylinder found above, and this neck will satisfy properties (i) and (ii) of our statement. Property (iii) simply follows choosing ω to be the axis of the cylinder C_t. $\square$

We have seen that assumption (ND2), giving the existence of a parabolic neighbourhood not modified by surgeries, is crucial in the proof of the neck detection lemma. It may seem that such an assumption can be easily implied by a suitable lower bound on the time between two surgeries. As we will discuss later in Remark 7.17, there is actually no easy way of doing this. Therefore it is useful to have arguments which ensure the validity of (ND2) in certain special cases. The next result shows that (ND2) follows from the other assumptions of the neck detection lemma, provided the curvature at (p_0, t_0) is large enough compared to the curvature of the regions changed in the previous surgeries.

Lemma 7.10 *Consider a flow with surgeries satisfying the same assumptions of Lemma 7.4. Let $d^\sharp$ be the value given from Lemma 7.2. Let $\varepsilon, k, L, \theta$ be given, with $\theta \le d^\sharp$. Then we can find η_0, H_0 with the following property. Let (p_0, t_0) be any point satisfying*

$$H(p_0, t_0) \ge \max\{H_0, 5K\}, \qquad \frac{\lambda_1(p_0, t_0)}{H(p_0, t_0)} \le \eta_0, \tag{7.7}$$

where K is the maximum of the curvature at the points changed in the surgeries at times before t_0. Then (p_0, t_0) satisfies hypothesis (ND2) *and the conclusions* (i)–(ii) *of Lemma 7.4. In addition, the neighbourhood*

$$\mathcal{P}\left(p_0, t_0, \frac{n-1}{H(p_0, t_0)} L, \frac{(n-1)^2}{K^2} d^\sharp\right)$$

(which is larger in time than the one in (ND2) *) does not contain surgeries.*

Proof. Let $\varepsilon, k, L, \theta$ be given, with $\theta \leq d^\sharp$. As observed in Remark 7.6, we can find values η_0, H_0 such that the conclusions of the neck detection lemma hold, not only for this choice of $(\varepsilon, k, L, \theta)$, but also if we replace L with any $L' \in [d^\sharp, L]$. In addition, we can assume that $H_0 \geq 2H^\sharp$, where $H^\sharp$ is the constant in Corollary 6.5. We claim that such values of η_0, H_0 satisfy the conclusions of the present lemma.

By our choice of η_0, H_0, the conclusions of the lemma can fail only if (ND2) is not satisfied, that is, if $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$ contains surgeries. Let us observe that, by Lemma 7.2(ii), at least the neighbourhood $\hat{\mathcal{P}}(p_0, t_0, d^\sharp, \theta)$ does not contain surgeries. Therefore, if (ND2) is violated, there exists a maximal $L' \in [d^\sharp, L)$ such that $\hat{\mathcal{P}}(p_0, t_0, L', \theta)$ does not contain surgeries. We can apply the neck detection lemma to this neighbourhood and deduce that it is an $(\varepsilon, k_0 - 1, L', \theta)$-shrinking neck. In particular, $H(p, t_0) = H(p_0, t_0)(1 + O(\varepsilon)) \geq 4K$ for all p such that $d_{g(t_0)}(p_0, p) \leq \hat{r}(p_0, t_0)d^\sharp$. But then Lemma 7.2(ii) shows that the larger neighbourhood $\hat{\mathcal{P}}(p_0, t_0, L' + d^\sharp, \theta)$ does not contain surgeries as well, contradicting the maximality of L'. This proves that (ND2) holds and that the neck detection lemma can be applied to the whole neighbourhood $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$.

To obtain the last assertion of our statement, consider any q such that

$$d_{t_0}(q, p_0) \leq \frac{n-1}{H(p_0, t_0)} L.$$

By the previous part of the statement, $H(q, t_0) = H(p_0, t_0)(1+O(\varepsilon)) > 2K$. Then, Lemma 7.2(ii) implies that q has not been affected by any surgery between time $t_0 - (n-1)^2 d^\sharp / K^2$ and t_0. Since this holds for any q in the neighbourhood, the statement is proved. □

If the assumptions of the previous lemma are not satisfied, then we may not be able to exclude the presence of surgeries in the parabolic neighbourhood. In such cases we can exploit the fact that we have a precise knowledge of the structure of the region changed by a surgery. Actually, for our purposes it suffices to consider the limiting situation, where we have a parabolic neighbourhood which does not contain surgeries, but where we find surgeries as soon as we enlarge the neighbourhood. Let us give a definition to describe this case.

Definition 7.11 *We say that the parabolic neighbourhood $\mathcal{P}(p_0, t_0, r, \tau)$ is* adjacent to a surgery region *if it has not been changed by surgeries, but there exists $p \in \mathcal{M}$ such that $d_{g(t_0)}(p, p_0) = r$, and which belongs to the boundary of a region changed by a surgery at a time $s \in [t_0 - \tau, t_0]$. We say that a hypersurface neck $\mathcal{N} \subset \mathcal{M}$ is* bordered on one side by a disk *if one of the two components of $\partial \mathcal{N}$ is also the boundary of a closed domain $\mathcal{D} \subset \mathcal{M}$, which is diffeomorphic to a disk and has no interior points in common with $\mathcal{N}$.*

In the next result we assume that our flow with surgeries satisfies certain properties, which we list below. We recall that the full surgery procedure consists first in removing some portion of necks and replacing them by a cap on each side, and then in removing some connected components of the surface which are recognized as diffeomorphic to spheres or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$. We denote these components as "components removed afterwards" in the following properties.

(s1) The surgeries take place in a region where the mean curvature is approximately equal to a fixed value K^* (the same for all surgeries). More precisely, there is $K^* > 2H^\sharp$ (where $H^\sharp$ is the value of Corollary 6.5) such that each surgery is performed at a cross section Σ_{z_0} of a normal neck with $r(z_0) = r^*$, where $r^* = (n-1)/K^*$ (see Sect. 3).

(s2) The two portions of a normal neck resulting from a surgery have the following properties. One portion belongs to a component of the surface which will be removed after the surgery. In the other portion, the part of the neck which has been left unchanged by the surgery has the following structure: on the first cross section (the one which coincides with the boundary of the region changed by the surgery) the mean radius satisfies $r(z) \leq (11/10)r^*$, on the last section $r(z) \geq 2r^*$ and in the sections in between $r^* \leq r(z) \leq 2r^*$.

(s3) Each surgery is essential for removing a region of the surface with curvature greater than $10K^*$. That is, if we consider any of the surgeries performed at a given surgery time, we can find a component removed afterwards which contains some point with mean curvature $H \geq 10K^*$, and which would not have been disconnected from the rest of the surface without the surgery we are considering.

Property (s1) easily implies that all points modified by a surgery have mean curvature between $K^*/2$ and $2K^*$ after the surgery, if the neck parameter ε_0 in the surgery procedure is chosen small enough. Also, the property that $r(z) \leq (11/10)r^*$ on the first cross section in (s2) is actually implied by (s1).

The above properties will be consequences of the way the surgery algorithm is defined in the next section. Property (s3) is a natural assumption since the aim of the surgery procedure is to reduce the curvature by a certain amount. Properties (s1) and (s3) together show that the regions with largest curvature are not the ones removed by the surgery itself (which have curvature close to K^*) but are the ones which become disconnected from the rest of the surface and are removed because they have a known topology. Property (s2) means intuitively that the surgeries are not performed at the very end of a neck, but at a certain distance so that there is a final part left untouched where the radius becomes twice as large. The presence of this long part of the neck in the surface after the surgery will be helpful in the proof of the next lemma.

Lemma 7.12 *Consider a flow with surgeries satisfying our usual assumptions, and in addition properties* (s1)–(s3) *above. Let $L, \theta > 0$ be such that $\theta \leq d^\sharp$, where $d^\sharp$ is the constant of Lemma 7.2, and that $L \geq 20$. Then there exist η_0, H_0 such that the following property holds. Let (p_0, t_0) satisfy properties* (ND1), (ND2) *of the neck detection Lemma 7.4. Suppose in addition that the parabolic neighbourhood $\hat{\mathcal{P}}(p_0, t_0, L, \theta)$ is adjacent to a surgery region. Then (p_0, t_0) lies at the center of a hypersurface neck $\mathcal{N}$ of length at least $L - 3$, which is bordered on one side by a disc $\mathcal{D}$. The mean curvature on $\mathcal{N} \cup \mathcal{D}$ at time t_0 is less than $5K^*$, where K^* is the constant in* (s1).

Proof. We first apply the neck detection Lemma 7.4(i) to find η_0, H_0 such that any point (p_0, t_0) satisfying (ND1), (ND2) lies at the center of an $(\varepsilon, k_0 - 1, L, \theta)$-shrinking curvature neck. By possibly refining the choice of η_0, H_0, we can also obtain that this neck can be parametrized as a geometric neck for all times under consideration as described in Proposition 7.9. Now, let (p_0, t_0) satisfy the hypotheses of the present lemma for such values of η_0, H_0. As usual, let us set

$$r_0 = \frac{n-1}{H(p_0, t)}, \qquad \mathcal{B}_0 = \{p \in \mathcal{M} \,:\, d_{g(t_0)}(p, p_0) \leq r_0 L\}.$$

Our assumptions are that $\mathcal{B}_0$ is not modified by any surgery for $t \in [t_0 - \theta r_0^2, t_0]$, but that there is a point $q_0 \in \partial\mathcal{B}_0$ and a time $s_0 \in [t_0 - \theta r_0^2, t_0]$ such that q_0 lies in the closure of a region modified by a surgery at time s_0. In the rest of the proof, we simply write s_0 instead of s_0^+, that is, we refer to the surface after the surgery. As a first step, we will show that our region has the desired topological structure at the surgery time s_0; then, we will show that the structure is not affected by the other surgeries which may occur between time s_0 and t_0.

Let us denote by $\mathcal{D}^*$ the region modified by the surgery which includes q_0 in its closure, and let $\mathcal{N}^*$ be the part of the neck left unchanged with the properties described in (s2). Let us denote by Σ_1^* and Σ_2^* the two components of $\partial\mathcal{N}^*$ having mean radius less than $(11/10)r^*$ and greater than $2r^*$ respectively. By (s2), $\Sigma_1^* = \partial\mathcal{D}^*$, and so $q_0 \in \Sigma_1^*$. It follows that the mean radius of Σ_1^* is equal to $(n-1)/H(q_0, s_0)$ up to an error of order $O(\varepsilon)$. Then we know that $H(p, s_0) \geq \frac{n-1}{r^*}(10/11 + O(\varepsilon)) = K^*(10/11 + O(\varepsilon))$ for all $p \in \mathcal{B}_0$, because the mean curvature is constant up to $O(\varepsilon)$ on $\mathcal{B}_0$ at any fixed time.

We claim that $\mathcal{B}_0$ must be contained in $\mathcal{N}^*$. In fact, we know that $\mathcal{B}_0$ has not been changed by the surgery at time s_0, and so it has no common points with $\mathcal{D}^*$. If $\mathcal{B}_0$ were not contained in $\mathcal{N}^*$, then it would intersect the other component Σ_2^* of $\partial\mathcal{N}^*$. But this is impossible, since at time s_0 the points in $\mathcal{B}_0$ and in Σ_2^* have mean curvature respectively greater than $(10/11)K^*$ and less than $K^*/2$ up to $O(\varepsilon)$.

Let $z \in [z_1, z_2]$ be the parameter describing the cross sections of $\mathcal{N}^*$ (where $z = z_i$ corresponds to Σ_i^*). Then we can find a maximal interval

$[a, b] \subset [z_1, z_2]$ such that the neck corresponding to $z \in [a, b]$ is centered at p_0 and is contained in $\mathcal{B}_0$. Let us denote by $\mathcal{N}_0$ this neck. Arguing as in Lemma 7.9, it is easy to see that $\mathcal{N}_0$ has a length at least $L - 2$.

Let us now denote with $\mathcal{N}'$ the part of $\mathcal{N}^*$ corresponding to $z \in [z_1, a]$. Then we have that p_0 belongs to $\mathcal{N}_0$, which is a normal k_0-hypersurface neck of length at least $L - 2$, and which is bordered on one side by the region $\mathcal{N}' \cup \mathcal{D}^*$, which is diffeomorphic to a disc. This is the statement of our theorem, except that it holds at the surgery time s_0 rather than the final t_0.

It remains to show that, if there are any surgeries between time s_0 and t_0, they do not affect the region $\mathcal{N}_0 \cup \mathcal{N}' \cup \mathcal{D}^*$. To this purpose, first observe that $H(p, s_0) \le 2K^*$ for any p in this region. By our choice of $d^\sharp$, $H^\sharp$ (see Lemma 7.2) we have that $H(p, t) \le 4K^*$ for any $p \in \mathcal{N}_0 \cup \mathcal{N}' \cup \mathcal{D}^*$ and t between s_0 and either t_0 or the first surgery time affecting this region, if there is any. But this shows that there cannot be any such surgery. In fact, let us first observe that $\mathcal{N}_0$ is contained in $\mathcal{B}_0$, which by assumption is not changed by surgeries in $[s_0, t_0]$. The neck $\mathcal{N}_0$ disconnects the region $\mathcal{D}^* \cup \mathcal{N}'$ from the rest of the manifold. By (s3), if a surgery changes this part, it must disconnect a region entirely contained in $\mathcal{D}^* \cup \mathcal{N}' \cup \mathcal{N}_0$ where the maximum of the curvature is at least $10K^*$. But this contradicts the bound on the curvature found just above. This proves that the topology of the region does not change up to time t_0, and that the curvature remains below the value $5K^*$ in this region. To conclude the proof, it suffices to parametrize the geometric neck $\mathcal{N}_0$ in normal form at the final time t_0, using the property that $\mathcal{N}_0 \subset \mathcal{B}_0$ which is an $(\varepsilon, k_0 - 1)$ curvature neck at any fixed time. □

Remark 7.13 As for Lemma 7.4, in the previous result it is easy to see that it is possible to choose the constants η_0, H_0 so that the conclusion applies to a whole interval of $L \in [L_1, L_2]$ and $\theta \in [\theta_1, \theta_2]$, provided $L_1 \ge 20$ and $0 < \theta_1 \le \theta_2 \le d^\sharp$.

In the neck detection lemma we assume that at the point under consideration the quantity λ_1/H is small. The next result will enable us to deal with the case where λ_1/H is not small, showing that it can be reduced in some sense to the former one. The result has some analogies with the main theorem in [11], but has a more elementary proof because we can use here our gradient estimates. It is a general property of hypersurfaces and is not related to mean curvature flow.

Theorem 7.14 *Let $F : \mathcal{M} \to \mathbb{R}^{n+1}$, with $n > 1$, be a smooth connected immersed hypersurface (not necessarily closed). Suppose that there exist $c^\sharp, H^\sharp > 0$ such that $|\nabla H(p)| \le c^\sharp H^2(p)$ for all $p \in \mathcal{M}$ such that $H(p) \ge H^\sharp$. Then, for any $\eta_0 > 0$ we can find $\alpha_0 > 0$ and $\gamma_0 > 1$ (depending only on $c^\sharp, \eta_0$) such that the following holds. Let $p \in \mathcal{M}$ satisfy $\lambda_1(p) > \eta_0 H(p)$ and $H(p) \ge \gamma_0 H^\sharp$. Then either $\mathcal{M}$ is closed with*

$\lambda_1 > \eta_0 H > 0$ *everywhere, or there exists a point* $q \in \mathcal{M}$ *such that*

(i) $\lambda_1(q) \le \eta_0 H(q)$,
(ii) $d(p, q) \le \alpha_0/H(p)$,
(iii) $H(q') \ge H(p)/\gamma_0$ *for all* $q' \in \mathcal{M}$ *such that* $d(p, q') \le \alpha_0/H(p)$*; in particular,* $H(q) \ge H(p)/\gamma_0$.

Proof. Given $\alpha_0 > 0$, let us set $\gamma_0 = 1 + c^\sharp \alpha_0$. For a given point $p \in \mathcal{M}$, let us set $\mathcal{M}_{p,\alpha_0} = \{q \in \mathcal{M} \ : \ d(p, q) \le \alpha_0/H(p)\}$. By Lemma 6.6, we obtain that, if $H(p) \ge \gamma_0 H^\sharp$, then

$$H(q) \ge \frac{H(p)}{1 + c^\sharp d(p, q) H(p)} \ge \frac{H(p)}{\gamma_0} \quad \text{for all } q \in \mathcal{M}_{p,\alpha_0}.$$

Suppose now that $p \in \mathcal{M}$ is such that $H(p) \ge \gamma_0 H^\sharp$ and that $\lambda_1(q) > \eta_0 H(q)$ for all $q \in \mathcal{M}_{p,\alpha_0}$. We claim that, if α_0 is suitably large, these properties imply that $\mathcal{M}$ coincides with $\mathcal{M}_{p,\alpha_0}$ and is therefore compact with $\lambda_1 > \eta_0 H$ everywhere. This will prove the theorem.

To prove this, we will show that the Gauss map $\nu : \mathcal{M}_{p,\alpha_0} \to \mathbb{S}^n$ is surjective. Let us take any $\omega \in \mathbb{S}^n$, such that $\omega \ne \pm\nu(P)$. We consider the curve γ solution of the ODE

$$\begin{cases} \dot\gamma = \frac{\omega^T(\gamma)}{|\omega^T(\gamma)|} \\ \gamma(0) = p \end{cases} \tag{7.8}$$

where, for any $q \in \mathcal{M}$, $\omega^T(q) = \omega - \langle \omega, \nu(q)\rangle \nu(q)$ is the component of ω tangential to $\mathcal{M}$ at q. Since $|\dot\gamma| = 1$, the curve γ will be parametrized by arclength. The curve can be continued until $|\omega^T(\gamma)| \ne 0$, i.e. $\nu(\gamma) \ne \pm\omega$. As long as $\gamma(s)$ is contained in $\mathcal{M}_{p,\alpha_0}$ (which is surely the case if $s \in [0, \alpha_0/H(p)]$), we can use the property $\lambda_1 > \eta_0 H$ to derive some estimate. Namely, if we denote by $e_1, \dots, e_n$ an orthonormal basis of the tangent space to $\mathcal{M}$ at a point $\gamma(s)$, we have

$$\begin{aligned} \frac{d}{ds}\langle \nu, \omega\rangle &= \sum_{i=1}^n \langle \dot\gamma, e_i\rangle \langle \nabla_{e_i}\nu, \omega\rangle = \frac{1}{|\omega^T|} \sum_{i,j=1}^n h_{ij} \langle \omega, e_i\rangle \langle \omega, e_j\rangle \\ &\ge \frac{1}{|\omega^T|} \eta_0 H |\omega^T|^2 = \eta_0 H \sqrt{1 - \langle \nu, \omega\rangle^2}, \end{aligned}$$

which implies

$$\frac{d}{ds} \arcsin\langle \nu, \omega\rangle \ge \eta_0 H.$$

Now suppose that $\gamma(s)$ exists for $s \in [0, \alpha_0/H(p)]$. Then we have

$$\begin{aligned} \pi &> \arcsin\left\langle \nu\left(\gamma\left(\frac{\alpha_0}{H(p)}\right)\right), \omega\right\rangle - \arcsin\langle \nu(p), \omega\rangle \ge \eta_0 \int_0^{\alpha_0/H(p)} H(\gamma(s)) ds \\ &\ge \eta_0 \int_0^{\alpha_0/H(p)} \frac{1}{H(P)^{-1} + c^\sharp s} ds = \frac{\eta_0}{c^\sharp} \ln\left(1 + c^\sharp \alpha_0\right). \end{aligned}$$

Thus, if α_0 is large enough to have

$$\alpha_0 > \frac{1}{c^\sharp}\left(\exp\left(\frac{c^\sharp \pi}{\eta_0}\right) - 1\right) \tag{7.9}$$

we obtain a contradiction. Therefore there exists $s^* \in\,]0, \alpha_0/H(p)[$ such that either $\nu(\gamma(s)) \cdot \omega \to 1$ or $\nu(\gamma(s)) \cdot \omega \to -1$ as $s \to s^*$. Since $\arcsin\langle \nu, \omega\rangle$ is increasing, only the first possibility can occur. This shows that $\gamma(s)$ converges, as $s \to s^*$, to some point $q^* \in \mathcal{M}_{p,\alpha_0}$ such that $\nu(q^*) = \omega$, as desired.

It remains to consider the case when $\omega = \pm\nu(p)$. If $\omega = \nu(p)$, then ω trivially belongs to the image of the Gauss map. If instead we have $\omega = -\nu(p)$, it suffices to replace p with another point p' sufficiently close to p; by convexity, we have $\nu(p') \neq \nu(p) = -\omega$ and the previous argument can be applied. Thus, we have proved that the Gauss map is surjective from $\mathcal{M}_{p,\alpha_0}$ to $\mathbb{S}^n$. Since $\lambda_1 > 0$ on $\mathcal{M}_{p,\alpha_0}$, the Gauss map is also a local diffeomorphism. Then, since $\mathbb{S}^n$ is simply connected for $n > 1$, well known results imply that the map is a global diffeomorphism. □

The results obtained until now suffice to obtain a rough result about the existence of necks before the first singular time is approached.

Corollary 7.15 *Let $\mathcal{M}_t$ be a smooth mean curvature flow of closed 2-convex hypersurfaces. Given neck parameters ε, k, L, there exists H^* (depending on the initial data) such that, if $H_{max}(t_0) \geq H^*$, then the hypersurface at time t_0 either contains an (ε, k, L)-hypersurface neck or it is convex.*

Proof. It suffices to combine Corollary 7.7 of the neck detection lemma with Theorem 7.14. Since we assume that the flow is smooth, the parabolic neighbourhood in hypothesis (ND2) trivially does not contain surgeries. □

Thus, unless the surface becomes convex, it is possible to perform a surgery before the singular time. At this stage it is not clear that after the surgery the surface will be in some sense less singular, e.g. that the maximum of the curvature will decrease. Also, proving the existence of necks after the first surgery requires additional arguments because one has to check condition (ND2) in the neck detection lemma. To deal with these issues we need to define a suitable algorithm determining when and where we shall perform our surgeries; this will be done in the next section.

Remark 7.16 As we already mentioned, the neck detection Lemma 7.4 is in the spirit of Hamilton's results in the Ricci flow. Our analysis of the complementary case where λ_1/H is not small (Theorem 7.14) is instead quite different both from Hamilton's and Perelman's one. Their approach consists of rescaling the flow to obtain an ancient solution and of studying the structure of ancient solutions using the Harnack estimate (which has an analogue for mean curvature flow, see [12]). We are not using any of these tools here. Also, we do not need to analyze the fine asymptotic structure of the singularities usually denoted as *degenerate neckpinches*, which is quite complicate even in the rotationally symmetric case [2].

Remark 7.17 We can now make some informal comment on hypothesis (ND2) in the neck detection lemma to explain why we need results like Lemmas 7.10 and 7.12 in our later analysis. As already mentioned, our aim is to define a flow with surgeries such that, after each surgery procedure, the maximum of the mean curvature drops from MH^* to H^*, where H^*, M are suitably large values, the same for all surgeries. Then, it may appear that there is a simple argument which ensures that hypothesis (ND2) is satisfied, not involving the results mentioned above. The argument is as follows. Between two consecutive surgery times T_1 and T_2 the curvature must increase from H^* to MH^*. Then, the inequality $\partial_t H - \Delta H = |A|^2 H \leq nH^3$ (see (2.3)) and a standard comparison argument yield

$$T_2 - T_1 \geq \frac{1}{2n}\frac{M^2-1}{M^2}\frac{1}{(H^*)^2}.$$

Thus, hypothesis (ND2) surely holds provided

$$\theta \leq \frac{H(p_0,t_0)^2}{2n(n-1)^2}\frac{M^2-1}{M^2}\frac{1}{(H^*)^2}. \tag{7.10}$$

Since the parameter θ can be chosen as small as we wish, it appears that (ND2) is always satisfied if the parameters are chosen appropriately. However, this is not quite true. The problem is that we need to apply the neck detection lemma at points (p_0, t_0) where the curvature can be much smaller than H^*. This happens when we want to remove by a surgery a point (p', t') such that $H(p', t') \geq H^*$ and $\lambda_1(p', t') > \eta_0 H(p', t')$. In this case we use Theorem 7.14 to find another point (p_0, t_0) at controlled distance from (p', t') such that $\lambda_1(p_0, t_0) \leq \eta_0 H(p_0, t_0)$. The theorem says that $H(p_0, t_0) \geq H(p', t')/\gamma_0 \geq H^*/\gamma_0$. Thus, in order to satisfy (7.10), we need

$$\theta \leq \frac{1}{2n(n-1)^2}\frac{M^2-1}{M^2}\frac{1}{\gamma_0^2}.$$

However, there is no clear way to obtain this. The problem is that choosing θ small influences the value of γ_0. In fact, the values of η_0, H_0 given by the neck detection lemma depend on θ. Thus, the smaller we choose θ, the smaller will be η_0. On the other hand, the smaller is η_0, the larger is the corresponding value of γ_0 given by Theorem 7.14. Since the dependence of γ_0 on θ is not explicit, there is no way to obtain the inequality we desire. Thus, it is not clear how to ensure a priori that all the points where we need to apply the neck detection lemma have a surgery-free parabolic neighbourhood. □

We will now give some definitions and prove some auxiliary results which will be used in the main theorem of next section and employ techniques similar to those of Theorem 7.14. Let $\mathcal{N}$ be an (ε, k)-hypersurface neck contained in a closed hypersurface $\mathcal{M}$, with $k \geq 1$. As usual, we denote by z the parameter along the neck. We know that $\mathcal{N}$ can be locally represented as a cylindrical graph; let $p_0 \in \mathcal{N}$ be at the center of a cylindrical graph $\mathcal{N}_1 \subset \mathcal{N}$ of C^1-norm less than ε_1 for some $\varepsilon_1 > 0$. Let ω be a unit vector parallel to the axis of $\mathcal{N}_1$. For simplicity, we assume that ω is parallel to the x_{n+1}-axis and points in the direction of the increasing x_{n+1}, we set $y = x_{n+1}$ and assume that p_0 lies on the $y = 0$ plane. We will call "vertical" the direction of the y-axis and "horizontal" any direction orthogonal to this one.

We have two different parametrizations for $\mathcal{N}_1$: the one as a cylindrical graph and the one induced by the normal parametrization of $\mathcal{N}$. The two representations differ little from each other, in the sense that on a cross section where z is constant, the coordinate y is constant up to $O(\varepsilon)$, and viceversa. We recall that the coordinate z is scale invariant, while y is not; thus, an increase Δy in the y-coordinate corresponds approximately to an increase $r(z)\Delta z$ in the z-coordinate. We assume that the y–axis is oriented in such a way that the directions of the increasing y and z agree.

We want to study the behaviour of the quantity $\omega \cdot \nu$ on $\mathcal{N}$, where ν is the normal to $\mathcal{N}$. Intuitively, if $\omega \cdot \nu$ is very small in some part of $\mathcal{N}$, this means that the axis of $\mathcal{N}$ is very close to ω in that part. Also, if $\omega \cdot \nu$ has constant sign in some region, e.g. it is positive, this means that the radius of the neck is decreasing. We want to study how these properties are related to the convexity of the neck. To do this, we will consider an ODE similar to (7.8).

Let Σ_0 be the intersection of the cylindrical graph $\mathcal{N}_1$ with the $y = 0$ plane. Then, by construction, we have $|\omega \cdot \nu(p)| \leq \varepsilon_1$ for all $p \in \Sigma_0$. Let us consider, for any $p \in \Sigma_0$, the curve $\gamma(p, \tau)$ satisfying the equation

$$\begin{cases} \dot{\gamma} = \dfrac{\omega^T(\gamma)}{|\omega^T(\gamma)|^2}, & \tau \geq 0 \\ \gamma(0) = p, \end{cases} \tag{7.11}$$

where the dot means derivative with respect to τ. It is easy to see that $y(\gamma(p, \tau)) = \tau$ for all $p \in \Sigma_0$; thus we can write $\gamma(p, y)$ instead of $\gamma(p, \tau)$ since τ and y coincide along γ. We consider $\gamma(p, y)$ for $y \geq 0$ since the analysis for $y \leq 0$ is analogous. We do not require the trajectories to remain inside $\mathcal{N}_1$ or $\mathcal{N}$; we follow them until they are well defined, that is, until ω is not orthogonal to γ. This cannot hold for arbitrarily large y by the compactness of our surface. Thus, there exists some finite $y_{max} > 0$ such that $\gamma(p, y)$ is defined for all $p \in \Sigma_0$ and $y \in [0, y_{max})$, and such that $\omega^T(\gamma(p, y)) \to 0$ as $y \to y_{max}$ at least for some p.

For $\bar{y} \in (0, y_{max})$, let us set $\Sigma_{\bar{y}} = \{\gamma(p, \bar{y}) \ : \ p \in \Sigma_0\}$. Clearly $\Sigma_{\bar{y}}$ is a smooth $(n-1)$-dimensional surface contained in the $y = \bar{y}$ hyperplane and diffeomorphic to Σ_0 under the flow (hence diffeomorphic to $\mathbb{S}^{n-1}$). We can compare two different surfaces by considering their projections on a fixed horizontal n-dimensional hyperplane; in particular, we will say that the surfaces Σ_y are *shrinking* if the projection of Σ_{y_2} is contained in the subset of the hyperplane enclosed by the projection of Σ_{y_1} for any $y_2 \geq y_1$.

Proposition 7.18 *Under the above hypotheses, suppose in addition that* $\lambda_1 \geq \alpha \geq 0$ *everywhere on* $\mathcal{N}$. *Then*

(i) *For any* $p \in \Sigma_0$, *we have that* $|\omega^{\perp}(\gamma(p, y))|$ *is bounded away from zero as long as* $\gamma(p, y) \in \mathcal{N}$; *therefore, any curve* $\gamma(p, y)$ *is well defined as long as it is contained in* $\mathcal{N}$.

(ii) *Along any trajectory* $\gamma(p, y)$ *we have* $\frac{d}{dy}\langle \nu, \omega \rangle \geq \alpha$ *as long as* γ *is contained in* $\mathcal{N}$.

(iii) *The axis of the neck* $\mathcal{N}$ *is approximately equal to* ω *everywhere. More precisely, any representation of a subset of* $\mathcal{N}$ *as a cylindrical graph of* C^1*-norm of size* $O(\varepsilon)$ *has an axis* $\tilde{\omega}$ *such that* $1 - \langle \omega, \tilde{\omega} \rangle = O(\varepsilon)$.

(iv) *If for some* $y_1 \geq 0$ *we have* $\nu(q) \cdot \omega \geq 0$ *for all* $q \in \Sigma_{y_1}$, *then the surfaces* Σ_y *are shrinking as long as they are contained in* $\mathcal{N}$.

Proof. As in the proof of Theorem 7.14, we find that

$$\frac{d}{dy}\langle \nu, \omega \rangle = \sum_{i=1}^{n} \langle \dot{\gamma}, e_i \rangle \langle \nabla_{e_i} \nu, \omega \rangle = \frac{1}{|\omega^T|^2} \sum_{i,j=1}^{n} h_{ij} \langle \omega, e_i \rangle \langle \omega, e_j \rangle \geq \lambda_1, \tag{7.12}$$

which proves (ii). Next, we recall that $|\langle \nu, \omega \rangle| \leq \varepsilon_1$ on Σ_0 by construction. By (ii), $\langle \nu, \omega \rangle$ is nondecreasing and therefore we have $\langle \nu, \omega \rangle \geq -\varepsilon_1$ along any trajectory $\gamma(p, y)$ as long as it stays inside $\mathcal{N}$. Suppose now that $\tilde{\omega}$ is the axis of any cylindrical graph representation of a subset $\tilde{\mathcal{N}} \subset \mathcal{N}$. Then $|\nu(q) \cdot \tilde{\omega}| = O(\varepsilon)$ for every $q \in \tilde{\mathcal{N}}$. If $\omega \neq \tilde{\omega}$, let us define

$$v = \omega - \langle \omega, \tilde{\omega} \rangle \tilde{\omega}. \tag{7.13}$$

Then $|v| = \sqrt{1 - \langle \omega, \tilde{\omega} \rangle^2} \neq 0$ and v is orthogonal to $\tilde{\omega}$. On an exact cylinder with axis $\tilde{\omega}$ we can find points where the normal is $\pm v/|v|$. Since $\tilde{\mathcal{N}}$ is close to a cylinder, we can find $q \in \tilde{\mathcal{N}}$ such that $|\nu(q) + \frac{v}{|v|}| = O(\varepsilon)$. Then we have

$$-\varepsilon_1 \leq \nu(q) \cdot \omega = \left(\nu(q) + \frac{v}{|v|} \right) \cdot \omega - \frac{v}{|v|} \cdot \omega \leq -\sqrt{1 - \langle \omega, \tilde{\omega} \rangle^2} + O(\varepsilon),$$

which shows that $\sqrt{1 - \langle \omega, \tilde{\omega} \rangle^2} = O(\varepsilon)$. We can choose the orientation of $\tilde{\omega}$ such that $\langle \omega, \tilde{\omega} \rangle \geq 0$; then the above estimate shows that $\langle \omega, \tilde{\omega} \rangle = 1 - O(\varepsilon)$, which implies (iii).

Property (i) is an easy consequence of (iii). To prove (iv), let us consider the projections of Σ_y on a horizontal hyperplane. The exterior normal (up to a normalizing factor) is given by $\nu - \langle \omega, \nu \rangle \omega$. Observe that

$$\begin{aligned} \langle \dot{\gamma}, \nu - \langle \omega, \nu \rangle \omega \rangle &= |\omega^T|^{-2} \langle \omega^T, \nu - \langle \omega, \nu \rangle \omega \rangle \\ &= |\omega^T|^{-2} (\langle \omega^T, \nu \rangle - \langle \omega, \nu \rangle \langle \omega^T, \omega \rangle) = -\langle \omega, \nu \rangle. \end{aligned} \tag{7.14}$$

This shows that the horizontal component of $\dot{\gamma}$ points towards the interior of Σ_y provided $\langle \omega, \nu \rangle > 0$. On the other hand, if $\langle \omega, \nu \rangle > 0$ for some value of y, the same holds for all greater values of y as a consequence of (ii). This concludes the proof. $\square$

We give one more elementary lemma which will be useful to study the behaviour of the trajectories $\gamma(p, y)$ once they leave the neck $\mathcal{N}$. Roughly speaking, the result says that if all submanifolds Σ_y have a small diameter, than the whole surface foliated by the Σ_y's has large mean curvature. We will use here the gradient estimate of Corollary 6.5.

Lemma 7.19 *Let $c^\sharp$, $H^\sharp$ be as in Corollary 6.5, and let us set $\Theta = 1 + (2 + \pi)(n - 1)c^\sharp$. Let us define the trajectories $\gamma(p, y)$ as before Lemma 7.18. Suppose that, for some $0 \le y_1 < y_2 < y_{max}$, we have $\lambda_1(\gamma(p, y)) > 0$ for all $y \in [y_1, y_2]$, $p \in \Sigma_0$ and that $\omega \cdot \nu(p) \ge 0$ for all $p \in \Sigma_{y_1}$. Suppose also that Σ_{y_1} has a diameter equal to $2(n - 1)/K$ for some $K \ge \Theta H^\sharp$, and that $H(p) \ge K$ for all $p \in \Sigma_{y_1}$. Then we have $H(\gamma(p, y)) \ge K/\Theta$ for all $y \in [y_1, y_2]$, $p \in \Sigma_0$.*

Proof. Using our assumptions and (7.12) we deduce that $\omega \cdot \nu > 0$ along all trajectories γ for $y \in [y_1, y_2]$. Then Proposition 7.18(iv) shows that for $y \in [y_1, y_2]$ the surfaces Σ_y are shrinking. By assumption, Σ_{y_1} is enclosed by an $(n - 1)$-dimensional sphere of mean curvature K, that is, of radius $R := (n - 1)/K$. Therefore, we can find a round cylinder with radius R and axis ω which encloses $\cup_{y \in [y_1, y_2]} \Sigma_y$.

Let us first consider the case where $y_2 - y_1 < R$. Then it is easy to see that, given any $p \in \Sigma_y$, we can find $p' \in \Sigma_{y_1}$ such that $d(p, p') \le 2R$. We obtain from Corollary 6.5

$$H(p) \ge \frac{K}{1 + 2(n - 1)c^*}.$$

Suppose now that $y_2 - y_1 \ge R$. Given any $y \in [y_1, y_2]$, let y' be such that $y \in [y', y' + R] \subset [y_1, y_2]$. Now let us take a portion of a cone C having circular section, axis ω, lower and upper basis in the $y = y'$ and $y = y' + R$ hyperplanes respectively. By a suitable choice of the radii $R_1, R_2 \le R$ of the upper and lower basis we can obtain that C touches $\cup_{y \in [y', y' + R]} \Sigma_y$ from the outside at some point q not lying in the $y = y'$, $y = y' + R$-planes. Then $H(q)$ is greater than the mean curvature of C at q, which is greater than K. Now, given any $p \in \Sigma_y$, it is easy to see that $d(p, q) \le (2 + \pi)R$. It follows that

$$H(p) \ge \frac{K}{1 + (2 + \pi)(n - 1)c^*}.$$ $\square$

8 The flow with surgeries

This final section is devoted to the proof of the following result.

Theorem 8.1 *Let $\mathcal{M}_0 \in \mathcal{C}(R, \alpha)$ be a smooth closed two-convex hypersurface immersed in $\mathbb{R}^{n+1}$, with $n \ge 3$, satisfying $|A|^2 \le R^{-2}$. Then there* *exists constants $H_1 < H_2 < H_3$ and a mean curvature flow with surgeries starting from $\mathcal{M}_0$ with the following properties:*

- *each surgery takes place at a time T_i such that* $\max H(\cdot, T_i-) = H_3$
- *after the surgery, all the components of the manifold (except the ones diffeomorphic to spheres or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$ which are neglected afterwards) satisfy* $\max H(\cdot, T_i+) \leq H_2$
- *each surgery starts from a cross section of a normal hypersurface neck with mean radius $r(z_0) = (n-1)/H_1$.*
- *the flow with surgeries terminates after finitely many steps.*

The constants H_i can be any values such that $H_1 \geq \omega_1 R^{-1}$, $H_2 = \omega_2 H_1$ and $H_3 = \omega_3 H_2$, with $\omega_i > 1$ depending only on the parameters α.

We start with the definition of the constants H_1, H_2, H_3. As the reader will see, this will require several steps in which other auxiliary parameters are introduced. The motivation of some details of the definition may be unclear at this stage, but it will be seen during the proof of the theorem. However, we give here some intuitive idea lying behind the definition.

One reason of our choice of parameters is that we want to apply the neck detection lemma in an iterative way. We can roughly explain it as follows. Given ε, k, L, the neck detection lemma gives η_0, H_0 such that any point (p_0, t_0) with $H(p_0, t_0) \geq H_0$ and $\lambda_1(p_0, t_0) \leq \eta_0 H(p_0, t_0)$ lies at the center of an (ε, k, L)-neck. In particular, any point p in the neck satisfies $H(p, t_0) \approx H(p_0, t_0)$ and $\lambda_1(p, t_0) \leq \varepsilon H(p, t_0)$. Although the proof of the neck detection lemma is not constructive, it is clear that in general η is much smaller than ε; thus the information on λ_1 in a general point of the neck is weaker than the hypothesis at the center p_0.

However, we can let the parameter η_0 play the role of ε in a further application of the lemma. Namely, we can find η_0', H_0' such that any point (p_0, t_0) with $H(p_0, t_0) \geq H_0'$ and $\lambda_1(p_0, t_0) \leq \eta_0' H(p_0, t_0)$ lies at the center of an $(\eta_0, 1, L)$-neck. We can choose H_0' larger than H_0. Then any point p of the $(\eta_0, 1, L)$-neck centered at p_0 will satisfy $H(p, t_0) \geq H_0$ and $\lambda_1(p, t_0) \leq \eta_0 H(p, t_0)$; thus it is in turn the center of an (ε, k, L)-neck.

In the proof of Theorem 8.1 we will need to iterate the neck detection lemma in the way just described. In addition, we will combine it with Theorem 7.14 as in Corollary 7.15. This is the reason why we need to introduce many different constants measuring the smallness of λ_1 and the size of H in the following definition.

Choice of the parameters

(P0) (Choice of the neck parameters) In Sects. 3–6 we have defined a surgery procedure on (ε_0, k_0)–hypersurface necks in normal form of length L, where ε_0 must be suitably small (with bounds depending only on the dimension), $k_0 \geq 2$ is any integer, and $L \geq 10 + 8\Lambda$, where Λ is the length parameter in the surgery. We also assume that $L \geq 20 + 8\Lambda$ and that ε_0 is small enough so that, if $\mathcal{N}$ is a normal

$(\varepsilon_0, 1)$-hypersurface neck of length $2L$ then the mean curvature at any two points of $\mathcal{N}$ can differ by a factor at most 2.

(P1) (Summary of known parameters) We define $c^\sharp$, $H^\sharp$ as in Corollary 6.5, $d^\sharp$ as in Lemma 7.2 and Θ as in Lemma 7.19.

(P2) (First application of the neck detection lemma) We choose η_0, K_0 such that, if (p, t_0) satisfies

$$H(p, t_0) \geq K_0, \quad \lambda_1(p, t_0) \leq \eta_0 H(p, t_0), \tag{8.1}$$

and if $\hat{\mathcal{P}}(p, t_0, L', \theta')$ does not contain surgeries for some $L' \in [L/4, L]$, $\theta' \in [d^\sharp/1600, d^\sharp]$, then $\hat{\mathcal{P}}(p, t_0, L', \theta')$ is a shrinking neck and (p, t_0) lies at the center of a normal (ε_0, k_0)-hypersurface neck of length at least $2L' - 2$ (see Corollary 7.7(i) and Remark 7.6). We also require that η_0, K_0 are such that, if (p, t_0) satisfies (8.1) and in addition $H(p, t_0) \geq 5K$, where K is the maximum of the mean curvature in the regions inserted in the surgeries, then the conclusions of Lemma 7.10 apply. Finally, we also require that η_0, K_0 are such that Proposition 7.12 can be applied to the parabolic neighbourhood $\hat{\mathcal{P}}(p, t_0, L', \theta')$ for the values of θ', L' above (see also Remark 7.13).

(P3) (Second application of the neck detection lemma) Next we set $\varepsilon_1 = (n-1)\eta_0/2$. We apply Corollary 7.7(ii) to find η_1, K_1 such that, if (p, t_0) satisfies

$$H(p, t_0) \geq K_1, \quad \lambda_1(p, t_0) \leq \eta_1 H(p, t_0), \tag{8.2}$$

and the parabolic neighbourhood $\hat{\mathcal{P}}(p, t_0, 10, d^\sharp/1600)$ does not contain surgeries, then (p, t_0) lies at the center of a cylindrical graph of length 5 and C^1-norm less than ε_1. We will choose η_1, K_1 such that $K_1 \geq K_0$, $K_1 \geq H^\sharp$ and $\eta_1 \leq \eta_0$.

(P4) (Application of the pinching Theorem 7.14) Now we choose γ_0 such that, if $H(p, t_0) > \gamma_0 H^\sharp$ and $\lambda_1(p, t_0) > \eta_1 H(p, t_0)$ then either $\lambda_1 > \eta_1 H$ everywhere on $\mathcal{M}_{t_0}$ or there exists q such that $\lambda_1(q, t_0) \leq \eta_1 H(q, t_0)$ and such that $H(q', t_0) \geq H(p, t_0)/\gamma_0$ for all q' with $d_{t_0}(q', p) \leq d_{t_0}(q, p)$ (see Theorem 7.14).

(P5) (Third application of the neck detection lemma) Let us set $\theta_2 = (10^4 n^3 \Theta^2 \gamma_0^2)^{-1}$. Then let us choose K_2, η_2 such that, if $H(p, t_0) \geq K_2$, if $\lambda_1(p, t_0) \leq \eta_2 H(p, t_0)$ and if $\hat{\mathcal{P}}(p, t_0, 10, \theta_2)$ does not contain surgeries, then (p, t_0) lies on a cylindrical graph of length 5 and C^1-norm less than ε_1. We also require $K_2 \geq K_1$ (while no comparison is needed for η_2).

(P6) We finally define H_1 to be any value such that $H_1 \geq 4\Theta K_2$, and then H_2, H_3 by

$$H_2 = 10\gamma_0 H_1, \quad H_3 = 10 H_2.$$

To have a definite value of these constants, one can simply pick $H_1 = 4\Theta K_2$. However, it is useful to remark that the H_i's can be also chosen arbitrarily large.

All the parameters introduced above only depend on the parameters α, R describing the initial surface. More precisely, the curvature parameters H_i, K_i, $H^\sharp$ can be written as constants depending only on α multiplied by R^{-1}, while the other parameters only depend on α.

In the proof of Theorem 8.1, we will define the surgery algorithm in such a way that the following properties are satisfied.

(S) Each surgery is performed on a normal (ε_0, k_0)-hypersurface neck. The surgeries are performed at times T_i such that $\max H(\cdot, T_i) = H_3$. After the surgeries are performed, and suitable components whose topology is known are removed, we have $\max H(\cdot, T_i+) \leq H_2$. In addition, all surgeries satisfy properties (s1)–(s3) (which are stated before Proposition 7.12) with $K^* = H_1$.

The proof of Theorem 8.1 will consist of a finite induction procedure. Namely, we suppose to have a mean curvature flow starting from $\mathcal{M}_0$, either smooth or with surgeries satisfying (S), defined up to some time t_0 such that $\max_{\mathcal{M}_{t_0}} H = H_3$. We then show that we can perform a finite number of surgeries at time t_0 which also satisfy (S). It is then easy to conclude that such a flow must terminate after a finite number of steps.

Let us observe that property (S), together with a standard comparison argument (see Remark 7.17) implies that the difference between two consecutive surgery times satisfies the uniform lower bound

$$T_{k+1} - T_k \geq \frac{10^2 - 1}{10^2} \frac{1}{2nH_2^2} > \frac{49}{10^4 n \gamma_0^2 H_1^2}. \tag{8.3}$$

The crucial step for the proof of Theorem 8.1 is provided by the next result, which we call the *neck continuation theorem*, and which we state and prove separately. Roughly speaking, the result says that the neck given by the neck detection Lemma 7.4 can be continued until either the curvature has decreased by a certain amount or the surface ends with a convex cap.

Theorem 8.2 (Neck continuation) *Suppose that $\mathcal{M}_t$, with $t \in [0, t_0]$, is a mean curvature flow with surgeries satisfying* (S)*, and let* $\max_{\mathcal{M}_{t_0}} H \geq H_3$. *Let p_0 be such that*

$$H(p_0, t_0) \geq 10 H_1, \quad \lambda_1(p_0, t_0) \leq \eta_1 H(p_0, t_0), \tag{8.4}$$

where η_1, H_1 are defined as in (P0)–(P6). *Then (p_0, t_0) lies on some (ε_0, k_0)-hypersurface neck $\mathcal{N}_0$ in normal form (with (ε_0, k_0) as in* (P0)*), which either covers the whole component of $\mathcal{M}_{t_0}$ including p_0, or has a boundary consisting of two cross sections Σ_1, Σ_2, each of which satisfies one of the two following properties:*

(i) *the mean radius of Σ is $2(n-1)/H_1$, or*
(ii) *the cross section Σ is the boundary of a region $\mathcal{D}$ diffeomorphic to a disc, where the curvature is at least H_1/Θ. The region $\mathcal{D}$ lies "after" the cross section Σ, that is, it is disjoint from $\mathcal{N}_0$.*

Proof of Theorem 8.2. Let us take p_0 such that (8.4) is satisfied. Recalling our definitions, we have

$$H(p_0, t_0) \geq 10K_1 \geq 10K_0, \quad \lambda_1(p_0, t_0) \leq \eta_1 H(p_0, t_0) \leq \eta_0 H(p_0, t_0).$$

Thus, at (p_0, t_0) we can apply neck detection both on the "finer" ε_1-level (see (P3)) and on the "rougher" ε_0-level (see (P2)).

We first consider the ε_0-level. Recall that all previous surgeries are peformed on necks with curvature close to H_1 and thus $K = 2H_1$ is a bound from above for the curvature in the regions modified by the surgeries. It follows that Lemma 7.10 can be applied with $K = 2H_1$. Since $H(p_0, t_0) \geq 10H_1$, definition (P2) and Lemma 7.10 ensure that the parabolic neighbourhood $\hat{\mathcal{P}}(p_0, t_0, L, d^\sharp)$ does not contain surgeries, and that the point (p_0, t_0) lies at the center of a normal (ε_0, k_0)-hypersurface neck of length at least $2L - 2$. Let us denote by $\mathcal{N}_0$ the maximal normal (ε_0, k_0)-hypersurface neck containing p_0. If $\mathcal{N}_0$ covers the whole manifold, we are done. Otherwise we have to show that in both directions, starting from p_0, we find a cross section of $\mathcal{N}_0$ which satisfies either property (i) or (ii) of our statement.

Let z be the parameter along the neck in its normal parametrization. We choose it in such a way that the cross section containing p_0 corresponds to $z = 0$. We follow the neck in the direction of the increasing z; the same analysis can be then repeated in the other direction. If there is a cross section where the average radius is $r(z) = 2(n-1)/H_1$, we are done. Therefore we assume that no such section exists, i.e. that $r(z) < 2(n-1)/H_1$ for all $z \in [0, z_{max}]$, where z_{max} is the value of z corresponding to the last section of the neck. This also implies that $H > H_1/4$ everywhere until the last section of the neck. We need to show that in this case the neck is bordered by a disc.

Let us outline the strategy of our proof. The neck detection lemma ensures that the neck $\mathcal{N}_0$ can be continued as long as we can find points enjoying the following three properties:

- the curvature H is large
- the ratio λ_1/H is small
- a suitable parabolic neighbourhood of the point is surgery-free.

Now, since the neck $\mathcal{N}_0$ ends somewhere, one of these three properties must fail. The first one, however, is ensured by the inequality $H > H_1/4$. Hence it must be one of the other two. We will show that if the second one fails (λ_1/H no longer small) then the neck starts closing up until it ends with a convex cap. If the third one is violated, instead, we will use Proposition 7.12 to conclude that the neck is bordered by the disc inserted in a previous surgery.

To make the above argument precise, we define a closed subset $\Omega \subset \mathcal{N}_0$ of our neck as follows. We say that $p \in \Omega$ if

(Ω1) $\lambda_1(p, t_0) \le \eta_0 H(p, t_0)$
(Ω2) the parabolic neighbourhood

$$\mathscr{P}\left(p, t_0, \frac{n-1}{H(p,t_0)} L, \frac{(n-1)^2}{(10H_1)^2} d^\sharp\right)$$

does not contain surgeries.

We are going to show that the points of Ω satisfy the hypotheses of the neck detection lemma, and therefore the neck $\mathcal{N}_0$ cannot end as long as it contains such points. It will follow that the last part of $\mathcal{N}_0$ does not contain points of Ω, and this will be exploited to deduce consequences on the last part of the neck.

First let us remark that, by Lemma 7.10, a point p which satisfies (Ω1) but not (Ω2) is necessarily such that $H(p, t_0) < 10H_1$. In particular, our starting point p_0 belongs to Ω. We also recall that all points $p \in \mathcal{N}_0$ on the side where $z \ge 0$ satisfy $H(p, t_0) \ge H_1/4$. Therefore, we have

$$\frac{(n-1)^2}{(40)^2 H(p,t_0)^2} \le \frac{(n-1)^2}{(10H_1)^2},$$

which implies

$$\hat{\mathscr{P}}\left(p, t_0, L, d^\sharp/40^2\right) \subset \mathscr{P}\left(p, t_0, \frac{n-1}{H(p,t_0)} L, \frac{(n-1)^2}{(10H_1)^2} d^\sharp\right).$$

Therefore, we know from (P2) that any $p \in \Omega$ lies at the center of a normal (ε_0, k_0)-hypersurface neck of length $2L - 2$. Thus, since the neck ends when $z = z_{max}$, at least the sections with $z \in (z_{max} - L + 1, z_{max}]$ do not contain any point of Ω. Let us define z^* to be the maximal value of z with the following property: the cross section of $\mathcal{N}_0$ with coordinate z^* contains a point $p_1 \in \Omega$, while there are no points of Ω for $z \in (z^*, z^* + 10)$. We consider two different cases.

(a) There is at least one point p_2 with $z \in (z^*, z^* + 10)$ satisfying (Ω1).
(b) All points with $z \in (z^*, z^* + 10)$ do not satisfy (Ω1).

Let us consider case (a). This will be the case where we find points which have been modified by a previous surgery, and where we can apply Proposition 7.12. To show that the hypotheses of that proposition are satisfied, we need some preliminary work. By definition, p_2 does not satisfy (Ω2), that is, the parabolic neighbourhood

$$\mathscr{P}\left(p_2, t_0, \frac{n-1}{H(p_2,t_0)} L, \frac{(n-1)^2}{(10H_1)^2} d^\sharp\right) \tag{8.5}$$

is modified by some surgery. We recall that, by (P0), the mean curvature can vary at most by a factor 2 in the part of the neck containing p_1 and p_2. Therefore we have

$$H(p_2, t_0) \geq H(p_1, t_0)/2,$$
$$d_{g(t_0)}(p_1, p_2) < 2(\pi + 10)\frac{n-1}{H(p_2, t_0)} < \frac{(n-1)L}{4H(p_2, t_0)}.$$

This implies

$$\mathcal{P}\left(p_2, t_0, \frac{n-1}{H(p_2, t_0)}\frac{L}{4}, \frac{(n-1)^2}{(10H_1)^2}d^\sharp\right) \subset \mathcal{P}\left(p_1, t_0, \frac{n-1}{H(p_1, t_0)}L, \frac{(n-1)^2}{(10H_1)^2}d^\sharp\right).$$

The neighbourhood at the right-hand side does not contain surgeries, because $p_1 \in \Omega$, so the one at the left-hand side does the same. By continuity, we deduce that we can replace the L in (8.5) with a suitable $L' \in [L/4, L]$ to obtain a neighbourhood which is not modified by surgeries, but is adjacent to a surgery, on the side of the increasing z. If we set

$$\theta' = \frac{H(p_2, t_0)^2}{(10H_1)^2}d^\sharp$$

we can write such a neighbourhood as $\hat{\mathcal{P}}(p_2, t_0, L', \theta')$. Since $H_1/4 \leq H(p_2, t_0) \leq 10H_1$, we have that $d^\sharp/(40)^2 \leq \theta' \leq d^\sharp$. By (P2), we are able to apply Lemma 7.12 and conclude that (p_2, t_0) lies in a hypersurface neck $\mathcal{N}$ bounded on one side by a disk $\mathcal{D}$. The same lemma tells us that the mean curvature on $\mathcal{N} \cup D$ is strictly less than $10H_1$. The hypersurface neck $\mathcal{N}$ can be combined with $\mathcal{N}_0$ to form a unique neck. The side bordered by $\mathcal{D}$ must be in the direction of the increasing z; otherwise, $\mathcal{N}$ should include all the neck $\mathcal{N}_0$, and this is impossible, because $\mathcal{N}_0$ contains the point p_0 which satisfies $H(p_0, t_0) \geq 10H_1$. Thus, the theorem is proved in this case.

We then turn to case (b). We assume therefore that all points in $\mathcal{N}_0$ with $z \in (z^*, z^* + 10)$ satisfy $\lambda_1 > \eta_0 H$. We will show that this convexity property is enough to ensure that the neck starts closing up, that is, that its radius starts decreasing at a strictly positive rate. This is the part where we use the property that the starting point p_0 lies on an ε_1-neck. After this, we will prove that the $(z^* + 10)$-cross section bounds a region which is convex and diffeomorphic to a disc.

Before using our information on the region $z \in (z^*, z^* + 10)$, we have to go back to the starting point p_0 of our neck on the $z = 0$ section. We use the property that $\lambda_1(p_0, t_0) \leq \eta_1 H(p_0, t_0)$ and thus p_0 lies on a cylindrical region with parameter ε_1 much finer than ε_0. Rather than p_0, it is convenient to consider the last point of the neck with this property (i.e., the one with the largest z); then we will have the additional information that $\lambda_1 > \eta_1 H$ after that point.

More precisely, let $\bar{z} \in [0, z^*]$ be the largest value of z such that the corresponding cross section contains a point $\bar{q}$ with $\lambda_1 \leq \eta_1 H$. We claim that $\hat{\mathcal{P}}(\bar{q}, t_0, 10, d^\sharp/1600)$ does not contain surgeries. In fact, from our definitions we deduce that there is a point $q \in \Omega$ with z coordinate in $[\bar{z}-10, \bar{z}]$. Then it is easy to check that

$$\hat{\mathcal{P}}\left(\bar{q}, t_0, 10, \frac{d^\sharp}{1600}\right) \subset \mathcal{P}\left(q, t_0, \frac{n-1}{H(q, t_0)}L, \frac{(n-1)^2}{(10H_1)^2}d^\sharp\right),$$

which does not contain surgeries, by definition of Ω. Thus, we know from (P3) that there exists a region $\mathcal{G} \subset \mathcal{N}_0$ centered at $\bar{q}$ which can be written as a cylindrical graph with C^1-norm less than ε_1.

We now apply the analysis of the last part of the previous section. We let ω be a unit vector parallel to the axis of $\mathcal{G}$. We assume that ω is parallel to the y axis, where we have set $y = x_{n+1}$ to denote the $(n+1)$-coordinate. We normalize y so that $F(\bar{q}, t_0)$ lies on the $\{y = 0\}$-hyperplane. We call Σ_0 the intersection of $\mathcal{G}$ with the $\{y = 0\}$-hyperplane. For any $p \in \Sigma_0$, we consider the curve $y \to \gamma(y, p)$ which solves (7.11) for $y \geq 0$. We denote by y_{max} the supremum of the values for which $\gamma(y, p)$ is defined defined for all $p \in \Sigma_0$ and we set $\Sigma_y =: \{\gamma(y, p) : p \in \Sigma_0\}$, for $0 \leq y < y_{max}$. In addition, given $0 \leq y_1 < y_2 < y_{max}$, we set

$$\Sigma(y_1, y_2) = \cup\{\Sigma_y, y_1 \leq y \leq y_2\}.$$

Let us denote by $\mathcal{N}_0'$ the part of $\mathcal{N}_0$ corresponding to $z \in [\bar{z}, z^* + 10]$. The $z = \bar{z}$ cross section contains the point $\bar{q}$ and so it is very close to Σ_0. By definition of $\bar{z}$, we have $\lambda_1 \geq \eta_1 H > 0$ on the part of $\mathcal{N}_0'$ with $z \in [\bar{z}, z^*]$; in the part with $z \in [z^*, z^* + 10]$ we have the stronger convexity property $\lambda_1 \geq \eta_0 H > 0$. In any case, $\mathcal{N}_0'$ is a convex region. Then, by Proposition 7.18, the axis of $\mathcal{N}_0'$ is approximately ω everywhere. In addition, the trajectories of (7.11) are defined at least as long they are contained in $\mathcal{N}_0'$. It follows that there exists a smallest value $y' < y_{max}$ such that $\gamma(y', p) \in \partial\mathcal{N}_0'$ for some $p \in \Sigma_0$. By construction, we have $|\nu(p) \cdot \omega| \leq \varepsilon_1$ for all $p \in \Sigma_0$, since Σ_0 is contained in the cylindrical graph $\mathcal{G}$. Recalling (7.12), we see that along all curves γ we have $\frac{d}{dy}\langle \nu, \omega\rangle \geq \lambda_1 > 0$, which implies in particular that $\nu(p) \cdot \omega \geq -\varepsilon_1$ for all $p \in \Sigma(0, y')$.

Now we exploit the property that $\lambda_1 \geq \eta_0 H$ on the cross sections of $\mathcal{N}_0$ corresponding to $z \in [z^*, z^* + 10]$. Let us set $r^* = r(z^*)$ to denote the mean radius of the z^*-section and let $H^* = (n-1)/r^*$. By assumption, $H^* > H_1/2$. The y-coordinate is almost constant on each cross section, because the axis of the neck is close to ω. Thus, the y-coordinates on the $z = z^*$ section and on the $z = z^* + 10$ section differ by approximately $10r^*$. It follows that at least the points of $\Sigma(y' - 5r^*, y')$ have a z coordinate such that $z \in [z^*, z^* + 10]$. Since H varies slowly on a neck, we have that $H \geq H^*/2$ on $\Sigma(y' - 5r^*, y')$. Thus we deduce from (7.12) that, along any curve γ, we have, for $y \in [y' - 5r^*, y']$,

$$\frac{d}{dy}\langle \nu, \omega \rangle = \frac{1}{|\omega^T|^2} \sum_{i,j=1}^{n} h_{ij} \langle \omega, e_i \rangle \langle \omega, e_j \rangle \geq \frac{1}{|\omega^T|^2} \eta_0 H |\omega^T|^2 \geq \eta_0 \frac{H^*}{2}.$$

Thus, for any $p' \in \Sigma_{y'}$, i.e. $p = \gamma(y', p)$ for some $p \in \Sigma_0$, we have

$$\begin{aligned} \langle \nu(p), \omega \rangle &= \langle \nu(\gamma(y' - 5r^*, p)), \omega \rangle + \int_{y'-5r^*}^{y'} \frac{d}{dy} \langle \nu, \omega \rangle dy \\ &\geq -\varepsilon_1 + 5r^* \frac{\eta_0 H^*}{2} > 4\varepsilon_1. \end{aligned}$$

The positivity of $\langle \nu, \omega \rangle$ on $\Sigma_{y'}$ means, roughly speaking, that the neck is closing up as z increases. We want to show that the part of the surface coming "after" $\Sigma_{y'}$ is a convex cap. To show this, we will continue the analysis of the curves $\gamma(y, p)$ for $y > y'$. Observe that the region swept by these curves is in general no longer a neck as y grows. Nevertheless, the curves are well defined until a value y_{max}, which is the first value such that $\nu(\gamma(y, p)) \to \pm\omega$ for some p as $y \to y_{max}$. Such a value y_{max} exists, by the compactness of our surface. We will prove that the region covered by the curves is convex, and that all curves converge to the same point as $y \to y_{max}$.

To this purpose, we will show that for all $y \in [y', y_{max})$ the following properties hold along all trajectories of (7.11):

$$\text{(i) } |\nu \cdot \omega| < 1, \quad \text{(ii) } \lambda_1 > 0, \quad \text{(iii) } H > H_1/4\Theta, \quad \text{(iv) } \nu \cdot \omega > \varepsilon_1. \tag{8.6}$$

Property (i) holds for all $y \in [y', y_{max})$ by definition. Since the other inequalities hold for y close to y', if they do not hold in the whole interval there must be a smallest $y \in (y', y_{max})$, which we denote by $y^\sharp$, where one of them becomes an equality. Let us show that this is impossible.

We first observe that (iv) necessarily holds also at $y = y^\sharp$. In fact, since (ii) holds for $y \in [y', y^\sharp)$, we know from (7.12) that $\nu \cdot \omega$ is increasing along any trajectory of (7.11) for y in this interval. Thus, (iv) still holds at $y = y^\sharp$. On the other hand, the fact that (iv) holds in $y \in [y', y^\sharp]$ implies that (iii) holds at $y^\sharp$. In fact, $\Sigma_{y'}$ has diameter less than $(n-1)8/H_1$, while by (P3) and (P6) we have $H_1 > 4\Theta H^\sharp$ and so we can apply Lemma 7.19.

Suppose now that (ii) fails for $y = y^\sharp$, that is, there exists $p^\sharp \in \Sigma_{y^\sharp}$ such that $\lambda_1(p^\sharp) = 0$. By the definition of θ_2 in (P5) we have

$$\theta_2 \frac{(n-1)^2}{H(p^\sharp, t_0)^2} \leq \theta_2 \frac{16(n-1)^2\Theta^2}{H_1^2} < \frac{16}{10^4 n \gamma_0^2 H_1^2}.$$

Recalling the estimate (8.3) between two surgery times, we see that the parabolic neighbourhood centered at $\hat{\mathcal{P}}(p^\sharp, t_0, 10, \theta_2)$ does not contain surgeries. By (P5), we deduce that a portion of the surface around $p^\sharp$ can be written as a cylindrical graph with C^1-norm less than ε_1. Let $\tilde{\omega}$ be the axis of this graph; then $\tilde{\omega}$ must be different from ω otherwise we have a contradiction with (iv). However, we find a contradiction in any case. Namely, let us define $v = \omega - \langle \omega, \tilde{\omega} \rangle \tilde{\omega}$ as in (7.13). Since v is orthogonal to $\tilde{\omega}$, we can find a point $q^\sharp$ close to $p^\sharp$ such that $|\nu(q^\sharp) + \frac{v}{|v|}| \leq \varepsilon_1$. But then, by (iv),

$$4\varepsilon_1 < \nu(q^\sharp) \cdot \omega = \left(\nu(q^\sharp) + \frac{v}{|v|} \right) \cdot \omega - \frac{v}{|v|} \cdot \omega \leq \varepsilon_1 - \sqrt{1 - \langle \omega, \tilde{\omega} \rangle^2},$$

which gives a contradiction. Therefore all properties (i)–(iv) hold for any $y < y_{max}$.

Now, we know that there exists at least a trajectory γ^* of (7.11) such that $\gamma^*(y) \to p^*$ as $y \to y_{max}$ for some $p^* \in \mathcal{M}_{t_0}$ such that $\nu(p^*) \cdot \omega = 1$. In fact, $\nu \cdot \omega$ cannot tend to -1 by property (iv). Let us set

$$\Sigma_{y_{max}} = \{ \lim_{y \to y_{max}} \gamma(y, p) : p \in \Sigma_0 \}.$$

We claim that $\Sigma_{y_{max}}$ reduces to the single point p^*. This implies that all trajectories γ of (7.11) tend to the same p^* as $y \to y_{max}$, and shows that the region after the neck is a convex cap.

Our claim follows from some standard arguments of Morse theory, observing that p^* is a critical point of the height function y, and that the Hessian has all negative eigenvalues because of convexity. However, it can be easily obtained by a direct argument. Since $\nu(p^*) \cdot \omega = 1$, the tangent plane to $\mathcal{M}_{t_0}$ at p^* is the plane $y = y_{max}$. Since the second fundamental form of $\mathcal{M}_{t_0}$ is positive definite at p^*, locally $\mathcal{M}_{t_0}$ lies below the plane $y = y_{max}$; this shows that p^* is an isolated point of $\Sigma_{y_{max}}$. On the other hand, $\Sigma_{y_{max}}$ is the limit of the convex surfaces Σ_y, and so it also convex. This is a contradiction unless $\Sigma_{y_{max}}$ consists uniquely of the point p^*. This completes the proof. □

Proof of Theorem 8.1. We use an iterative argument. We consider a flow defined in $[0, t_0]$, which either is smooth or has surgeries satisfying (S) at times smaller than t_0. We assume that t_0 is the first time after the last surgery (or the first time at all, if the flow is smooth), such that $H_{max}(t_0) = H_3$. We then want to show that we can perform a finite number of surgeries on $\mathcal{M}_{t_0}$, which satisfy (S), and that after these surgeries the maximum of the curvature satisfies $H_{max} \leq H_2$ except on the regions diffeomorphic to spheres or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$ which are neglected afterwards.

Let us consider any point p_0 such that $H(p_0, t_0) \geq H_2$. We consider first the case where $\lambda_1(p_0, t_0) \leq \eta_1 H(p_0, t_0)$. Then we apply the neck continuation Theorem 8.2 to find that p_0 belongs to a neck $\mathcal{N}_0$ satisfying the properties described there. Let us denote by $\mathcal{A}$ the region consisting of the neck $\mathcal{N}_0$ together possibly with the one or two regions diffeomorphic to discs which occur in case (ii) of the theorem. The region $\mathcal{A}$ contains the point p_0 and has one of the three following possible structures: (a) it has two boundary components, and it is diffeomorphic to $\mathbb{S}^{n-1} \times [-1, 1]$, or (b) it has one boundary component and it is diffeomrphic to a disc, or

(c) it has no boundary, it coincides with the connected component of $\mathcal{M}$ containing p_0 and it is diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$. In any case, the boundary $\partial\mathcal{A}$ (if non empty) consists of one or two cross sections of the neck $\mathcal{N}_0$ with mean radius equal to $2(n-1)/H_1$ and hence with mean curvature approximately equal to $H_1/2$.

If we have instead $\lambda_1(p_0, t_0) > \eta_1 H(p_0, t_0)$, we proceed as follows. We apply Theorem 7.14 to find a point q_0 such that $\lambda_1(q_0, t_0) \leq \eta_1 H(q_0, t_0)$ and that $H(q, t_0) \geq H(p_0, t_0)/\gamma_0$ for all q such that $d_{t_0}(q, p_0) \leq d_{t_0}(q_0, p_0)$. In particular we have that $H(q_0, t_0) \geq H_2/\gamma_0 \geq 10H_1$. Then we proceed as in the first case and define a region $\mathcal{A}$ containing q_0 and consisting of a neck with the possible union of one or two discs at the ends. We claim that $\mathcal{A}$ includes the point p_0 too. If this is not the case, this means that any path from p_0 to q_0 must intersect the boundary of $\mathcal{A}$. At the points of $\partial\mathcal{A}$, however, the curvature H is close to $H_1/2$. On the other hand, we know that along the geodesic from p_0 to q_0 we have $H \geq H(p_0, t_0)/\gamma_0 \geq 10H_1$, and thus we obtain a contradiction.

In both cases we have defined a region $\mathcal{A}$ including the point p_0 with the structure described above. We then repeat the analysis until we cover all points with curvature larger than H_2 by similar regions. That is, suppose that there is $p_0' \notin \mathcal{A}$ such that $H(p_0', t_0) > H_2$. We proceed as above to define a region $\mathcal{A}'$ including the point p_0'. We have to ensure that $\mathcal{A}$ and $\mathcal{A}'$ are disjoint, otherwise the surgeries we are going to perform would interfere with each other. To show this, we recall that $\partial\mathcal{A}$ consists of cross sections of a neck with mean radius equal to $2(n-1)/H_1$. This means that, if we meet one such cross section in the application of the neck continuation theorem 8.2 we stop there because we have achieved property (i) of the theorem. This implies that $\mathcal{A}$ and $\mathcal{A}'$ can overlap at most at boundary points.

Observe that the area of any region defined in this way is bounded from below by a fixed multiple of H_2^{-n}. Therefore, we find a finite collection $\mathcal{A}, \mathcal{A}', \ldots, \mathcal{A}^{(k)}$ which covers all points of $\mathcal{M}_{t_0}$ with mean curvature greater than H_2.

After having identified the regions with large curvature, we proceed with the surgeries. The $\mathcal{A}^{(i)}$'s with no boundary components are diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$ and can be discarded for later times. In the other ones we do a surgery near to each boundary component in the following way. We know that any such component is a cross section of a neck with mean radius $2(n-1)/H_1$. We consider the cross section $\Sigma^{(i)}$ closest to the boundary with mean radius $(n-1)/H_1$. Such a cross section surely exists by continuity because the neck contains the point p_0 (respectively q_0) where the curvature is at least $H_2 \geq 10H_1$. We then perform the standard surgery centered at the cross section $\Sigma^{(i)}$. If the boundary $\partial\mathcal{A}^{(i)}$ has two components we apply this procedure on both sides. Notice that the surgeries performed on different regions are independent from each other because the $\mathcal{A}^{(i)}$'s can touch only at boundary points, while the surgeries are performed at cross sections well inside the interior of the $\mathcal{A}^{(i)}$'s, where the mean radius is half of the one on the boundary.

In both cases, the surgeries (whether one or two) create a connected component diffeomorphic to a sphere which includes all points of $\mathcal{A}^{(i)}$ with curvature larger than H_2. Such a component will be neglected in the later evolution, and so we see that the maximum of the curvature has decreased under H_2 at the end of the procedure.

It is easy to check that the surgeries defined in this way satisfy all requirements listed in (S), including properties (s1)–(s3) before Proposition 7.12 with $K^* = H_1$ and $r^* = (n-1)/H_1$. In particular, our construction ensures that each surgery takes place on a cross section with mean radius r^* and leaves unchanged a collar at whose end the mean radius is equal to $2r^*$. In addition, each surgery is essential for removing a component of the surface where the maximum of the mean curvature is larger than H_2.

We then restart the flow, until we reach again a time where $H_{max} = H_3$ and we repeat the procedure. There can only be a finite number of surgery times, because the area of the surface is decreasing under the smooth flow and each surgery decreases the area of the surface by a fixed multiple of $(H_1)^{-n}$. This implies that eventually the whole surface is removed in the surgery procedure because the remaining pieces will be identified as diffeomorphic either to $\mathbb{S}^n$ or $\mathbb{S}^{n-1} \times \mathbb{S}^1$. □

Proof of Corollaries 1.2 and 1.3. Having completed the proof of the main Theorem 1.1 we notice that at the termination of the mean curvature flow with surgeries we are left with finitely many disjoint smooth closed surfaces which are either diffeomorphic to $\mathbb{S}^n$ or to $\mathbb{S}^{n-1} \times \mathbb{S}^1$. This includes finitely many surfaces of this type that were discarded during the flow at various surgery times.

In the case of positive mean curvature considered here, smooth mean curvature flow $F : \mathcal{M} \times [t_0, t_1] \to \mathbb{R}^{n+1}$ is a smooth immersion of $\mathcal{M} \times [t_0, t_1]$ into $\mathbb{R}^{n+1}$ such that all restrictions $F : \mathcal{M} \times [a, b] \to \mathbb{R}^{n+1}$, $t_0 \le a < b \le t_1$, are isotopic to each other via rescaling of time. In particular, the surfaces $\mathcal{M}_t = F(\cdot, t)(\mathcal{M})$ are all diffeomorphic to each other between surgery times, such that the first statement of Corollary 1.2 is then an immediate consequence of Proposition 3.23.

To control the bounded region swept out by $\mathcal{M}_t$ let us assume first that the initial surface is embedded, which is a preserved property by Theorem 3.26(i). To begin we examine the finitely many pieces left by the flow case by case. A surface of type $\mathbb{S}^{n-1} \times \mathbb{S}^1$ can only occur if it is recognised as a maximal normal (ε, k)-hypersurface neck N without boundary. The solid tube G associated with N then provides a smooth diffeomorphism of the interior region with the standard solid tube $\bar{B}_1^n \times \mathbb{S}^1$ equipped with the flat metric. Another simple case occurs when during neck detection at some surgery time T_i or at the final time T we recognise a convex closed surface which is well known to bound a smooth convex region diffeomorphic to $\bar{B}_1^n$, compare Theorem 7.14. Two more cases where the interior region has to be carefully controlled occur during "neck continuation", see Theorem 8.2 and Proposition 7.12. While following along a normal neck at a surgery time T_i along the z-coordinate we may encounter a backward parabolic neighbourhood containing a surgery at some previous time T_k, $k < i$. From Proposition 7.12 we concluded that at time T_k there is a normal (ε_0, k_0)-hypersurface neck capped by a convex disc attached by surgery at this earlier time. The convex disc then bounds a smooth half ball which is smoothly glued to the solid tube associated with the normal (ε_0, k_0)-hypersurface neck at time T_k. From Proposition 7.12 we also infer that this region does not encounter any further surgeries till time T_i while the normal (ε_0, k_0)-hypersurface neck collaring this region persists. It follows that the mean curvature flow is an isotopy in this regime between times T_k and T_i such that also at the later time T_i the solid tube associated with the neck ends in a smooth half ball attached in the standard way. The last case we need to consider concerns the situation where during "neck continuation" we encounter a convex cap satisfying the estimates (8.6). It is then easy to see that this convex cap can be written as the graph of a strictly convex function over a hyperplane orthogonal to the vector ω occurring in (8.6). Hence the region bounded by this cap is also diffeomorphic to a standard half ball. In summary we conclude that all the finitely many disjoint regions created by mean curvature flow with surgery are diffeomorphic either to $\bar{B}_1^{n+1}$ or to $\bar{B}_1^n \times \mathbb{S}^1$. Since smooth mean curvature flow with positive mean curvature provides an isotopy of the enclosed regions at successive times it then follows from Theorem 3.26 that the initial closed region is diffeomorphic to a solid handlebody, i.e. the connected sum of finitely many copies of $\bar{B}_1^n \times \mathbb{S}^1$. In particular, if the initial surface $\mathcal{M}$ is simply connected, it must have bounded a region which is diffeomorphic to a standard closed ball.

If the initial surface is only immersed, the conclusion of Corollary 1.2 is still valid: Each surgery is a local construction inside one particular solid tube associated with a normal (ε_0, k_0)-hypersurface neck and is not affected by other parts of the surface $\mathcal{M}$ which may intersect this tube. So we can apply Proposition 3.23 locally at each such surgery, creating a new smoothly immersed surface from which the previous surface can be reconstructed via standard connected sums. Since a positive lower bound on the mean curvature is preserved during surgeries, all $F_i : \mathcal{M}_i \times [T_{i-1}, T_i] \to \mathbb{R}^{n+1}$ are smooth immersions of $\mathcal{M}_i \times [T_{i-1}, T_i]$ into $\mathbb{R}^{n+1}$. Using the normal coordinates for solid tubes in the surgery region constructed in Proposition 3.25 these can be explicitly spliced together into the immersion of a handlebody into $\mathbb{R}^{n+1}$ with boundary $F_0 : \mathcal{M} \to \mathbb{R}^{n+1}$. $\square$

Acknowledgements. We thank Richard Hamilton and Klaus Ecker for many stimulating discussions on this work. We also thank the referees for the valuable comments. Parts of this project were carried out while the authors were visiting the Isaac Newton Institute (Cambridge, UK), the IPAM (Los Angeles, USA) and the BIRS (Banff, Canada). The authors would like to thank each of these institutions for their hospitality and support. The first author thanks the Department of Mathematics of University of Rome "Tor Vergata" for its hospitality. He acknowledges support by the SFB382 and SFB647 of the DFG. The second author thanks the AEI Institute in Golm for its hospitality and acknowledges supported by MIUR projects Cofin2002 and PRIN2005.

References

1. Altschuler, S., Angenent, S.B., Giga, Y.: Mean curvature flow through singularities for surfaces of rotation. J. Geom. Anal. **5**, 293–358 (1995)
2. Angenent, S.B., Velazquez, J.J.L.: Degenerate neckpinches in mean curvature flow. J. Reine Angew. Math. **482**, 15–66 (1997)
3. Brakke, K.A.: The Motion of a Surface by its Mean Curvature. Princeton University Press, Princeton (1978)
4. Chen, Y.G., Giga, Y., Goto, S.: Uniqueness and existence of viscosity solutions of generalized mean curvature flow equation. J. Differ. Geom. **33**, 749–786 (1991)
5. Colding, T., Kleiner, B.: Singularity structure in mean curvature flow of mean-convex sets. Electron. Res. Announc. Am. Math. Soc. **9**, 121–124 (2003) (electronic)
6. Ecker, K.: Regularity Theory for Mean Curvature Flow. Birkhäuser, Boston (2004)
7. Ecker, K., Huisken, G.: Interior estimates for hypersurfaces moving by mean curvature. Invent. Math. **105**, 547–569 (1991)
8. Evans, L.C., Spruck, J.: Motion of level sets by mean curvature, I. J. Differ. Geom. **33**, 635–681 (1991)
9. Fraser, A.M.: Minimal disks and two-convex hypersurfaces. Am. J. Math. **124**, 483–493 (2002)
10. Hamilton, R.S.: Four–manifolds with positive curvature operator. J. Differ. Geom. **24**, 153–179 (1986)
11. Hamilton, R.S.: Convex hypersurfaces with pinched second fundamental form. Commun. Anal. Geom. **2**, 167–172 (1994)
12. Hamilton, R.S.: The Harnack estimate for the mean curvature flow. J. Differ. Geom. **41**, 215–226 (1995)
13. Hamilton, R.S.: Formation of singularities in the Ricci flow. Surv. Differ. Geom. **2**, 7–136 (1995)
14. Hamilton, R.S.: Four–manifolds with positive isotropic curvature. Commun. Anal. Geom. **5**, 1–92 (1997)
15. Hirsch, M.W.: Differential Topology. Grad. Texts Math., vol. 33. Springer, Berlin (1976)
16. Huisken, G.: Flow by mean curvature of convex surfaces into spheres. J. Differ. Geom. **20**, 237–266 (1984)
17. Huisken, G.: Local and global behaviour of hypersurfaces moving by mean curvature. Proc. Symp. Pure Math. **54**, 175–191 (1993)
18. Huisken, G., Sinestrari, C.: Mean curvature flow singularities for mean convex surfaces. Calc. Var. Partial Differ. Equ. **8**, 1–14 (1999)
19. Huisken, G., Sinestrari, C.: Convexity estimates for mean curvature flow and singularities of mean convex surfaces. Acta Math. **183**, 45–70 (1999)
20. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, San Diego (1980)
21. Lawson, H.B.: Local rigidity theorems for minimal hypersurfaces. Ann. Math. **89**, 187–197 (1969)
22. Matsumoto, Y.: An introduction to Morse theory. Trans. Math. Monogr., vol. 208. American Math. Soc., Providence (2002)
23. Mercuri, F., Noronha, M.H.: Low codimensional submanifolds of Euclidean space with nonnegative isotropic curvature. Trans. Am. Math. Soc. **348**, 2711–2724 (1996)
24. Michael, J.H., Simon, L.M.: Sobolev and mean value inequalities on generalized submanifolds of $\mathbb{R}^n$. Commun. Pure Appl. Math. **26**, 361–379 (1973)
25. Perelman, G.: The entropy formula for the Ricci flow and its geometric applications. Preprint (2002). arXiv:math/0211159v1

26. Perelman, G.: Ricci flow with surgery on three-manifolds. Preprint (2003). arXiv:math/0303109v1
27. Sha, J.-P.: *p*-convex Riemannian manifolds. Invent. Math. **83**, 437–447 (1986)
28. White, B.: The topology of hypersurfaces moving by mean curvature. Commun. Anal. Geometry **3**, 317–333 (1995)
29. White, B.: The nature of singularities in mean curvature flow of mean-convex sets. J. Am. Math. Soc. **16**, 123–138 (2002)
30. Wu, H.: Manifolds of partially positive curvature. Indian. Univ. Math. J. **36**, 525–548 (1987)