

# DIAMETER AND CURVATURE CONTROL UNDER MEAN CURVATURE FLOW

PANAGIOTIS GIANNIOTIS AND ROBERT HASLHOFER

ABSTRACT. We prove that for the mean curvature flow of two-convex hypersurfaces the intrinsic diameter stays uniformly controlled as one approaches the first singular time. We also derive sharp  $L^{n-1}$ -estimates for the regularity scale of the level set flow with two-convex initial data. Our proof relies on a detailed analysis of cylindrical regions ( $\varepsilon$ -tubes) under mean curvature flow. The results are new even in the most classical case of mean convex surfaces evolving by mean curvature flow in  $\mathbb{R}^3$ .

## 1. INTRODUCTION

A family of hypersurfaces  $\{M_t^n \subset \mathbb{R}^{n+1}\}_{t \in [0, T)}$  evolves by mean curvature flow if the normal velocity at each point is given by the mean curvature vector. By classical theory (see e.g. [14, 16]), given any closed embedded initial hypersurface  $M_0^n \subset \mathbb{R}^{n+1}$ , there exists a unique smooth solution defined on a maximal time interval  $[0, T)$ . The maximal existence time  $T < \infty$  is characterized by

$$(1.1) \quad \lim_{t \nearrow T} \max_{M_t} |A| = \infty,$$

where  $|A|$  denotes the norm of the second fundamental form.

One naturally wonders to what extent one can control the geometry of the hypersurfaces – curvature integrals, intrinsic diameter, etc – as one approaches the first singular time (more generally, one can pose these questions also beyond the first singular time in the setting of mean curvature flow with surgery and level set flow, respectively):

**Question 1.1** (Curvature control). *Can one control the curvature integrals  $\int_{M_t} |A|^p d\mu$  as the flow approaches the first singular time?*

**Question 1.2** (Diameter control). *Can one control the intrinsic diameter as the flow approaches the first singular time?*

The two questions are in fact tightly related. For example, by a result of Topping [24, Thm. 1.1] we have the estimate

$$(1.2) \quad \text{diam}(M_t, d_t) \leq C_n \int_{M_t} |H|^{n-1} d\mu.$$

To obtain diameter control, one thus might try to derive uniform  $L^{n-1}$ -bounds for the mean curvature.

For  $n > 2$ , the situation where there is some hope to get uniform  $L^{n-1}$ -bounds concerns the evolution of two-convex hypersurfaces, i.e. when the sum of the smallest two principal curvatures  $\lambda_1 + \lambda_2$  is positive.<sup>1</sup> This curvature condition arises naturally in the work on mean curvature flow with surgery by Huisken-Sinestrari [17] (see also [12, 3, 2]), and its main feature is that it excludes generalized cylinders  $\mathbb{R}^j \times S^{n-j}$  with more than one  $\mathbb{R}$ -factor (i.e.  $j \geq 2$ ) as potential blowup limits.

For  $n = 2$ , the notion of two-convexity boils down to the simpler notion of mean convexity, i.e. positive mean curvature, or in other words the assumption that the flow is moving inwards.

It has been proved by Head [13] and Cheeger-Haslhofer-Naber [5] that for the mean curvature flow of two-convex hypersurfaces one has

$$(1.3) \quad \int_{M_t} |A|^{n-1-\varepsilon} d\mu \leq C(M_0, \varepsilon) < \infty,$$

for any  $\varepsilon > 0$ . Motivated by this result, it is natural to conjecture:<sup>2</sup>

**Conjecture 1.1** ( $L^{n-1}$ -curvature conjecture). *If  $\{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$  is a mean curvature flow of two-convex closed embedded hypersurfaces, then*

$$(1.4) \quad \int_{M_t} |A|^{n-1} d\mu \leq C,$$

*for some constant  $C < \infty$  depending only on the geometric parameters of the initial hypersurface.*

---

<sup>1</sup>One of course cannot hope to get  $L^{n-1}$ -bounds assuming only mean convexity. Indeed, consider the case that  $M_0 \approx S_r^{n-2} \times S_R^2$  is a very thin rotationally symmetric torus, i.e.  $r$  is very small. Under the flow these small  $(n-2)$ -spheres will degenerate to points, and it is easy to see that  $\lim_{t \nearrow T} \int_{M_t} H^{n-1} d\mu = \infty$ .

<sup>2</sup>R.H. thanks John Head for introducing him to these conjectures during a visit to the Courant Institute in 2011. While we unfortunately don't know the precise history, the conjectures have certainly been discussed among experts well before 2011, c.f. Perelman's bounded diameter conjecture for 3d Ricci flow [23, Sec. 13.2].

**Conjecture 1.2** (bounded diameter conjecture). *If  $\{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$  is a mean curvature flow of two-convex closed embedded hypersurfaces, then*

$$(1.5) \quad \text{diam}(M_t, d_t) \leq C,$$

*for some constant  $C < \infty$  depending only on the geometric parameters of the initial hypersurface.*

Note that by the result of Topping [24, Thm. 1.1] an affirmative answer to Conjecture 1.1 would imply an affirmative answer to Conjecture 1.2.

Conjecture 1.1 states that one can get rid of the  $\varepsilon$  in the curvature estimate (1.3). The question of whether or not one can actually get rid of the  $\varepsilon$  in estimates like (1.3) is a very delicate question, that depends on the fine structure of singularities and regions of high curvature.

For comparison, it is useful to look at related elliptic questions. In recent impressive work [22], Naber-Valtorta improved the known  $L^{3-\varepsilon}$  estimates for the gradient of minimizing harmonic maps to sharp  $L^3_{\text{weak}}$ -estimates. The simple example  $f(x) = \frac{x}{|x|}$  in dimension three, shows that for minimizing harmonic maps the  $L^3_{\text{weak}}$  estimate actually *cannot* be replaced by an  $L^3$  estimate; cf. also the related work [20, 21, 10, 25].

Having discussed the subtleties of sharp integral estimates, let us now emphasize that our approach for the solution of Conjecture 1.1 and Conjecture 1.2 actually proceeds in the opposite order. Namely, we first prove that we can control the intrinsic diameter:

**Theorem 1.1** (Intrinsic diameter control). *If  $\{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$  is a mean curvature flow of two-convex closed embedded hypersurfaces, then*

$$(1.6) \quad \text{diam}(M_t, d_t) \leq C,$$

*for a constant  $C = C(\alpha, \beta, \gamma, \mathcal{A}) < \infty$ , which only depends on certain geometric parameters of the initial hypersurface  $M_0$  (see Section 2.1).*

The diameter bound from Theorem 1.1 is new even in the classical case of mean convex surfaces evolving in  $\mathbb{R}^3$  (note that for  $n = 2$  mean convex surfaces are automatically  $\beta$ -uniformly two-convex with  $\beta = 1$ ).

We then use Theorem 1.1 to derive sharp integral estimates for the second fundamental form. More precisely, we actually obtain sharp integral estimates for the regularity scale of the level set flow with two-convex initial data. Recall from [5], if  $\mathcal{M} = \{M_t\}_{t \in [0, T]}$  is a weak solution of the mean curvature flow (here we only consider the mean

convex case, where Brakke solutions and level set solutions are known to be equivalent) then the regularity scale  $r_{\mathcal{M}}(x, t)$  at a point  $(x, t) \in \mathcal{M}$  is defined as the supremum of  $0 \leq r \leq 1$  such that  $M_{t'} \cap B_r(x)$  is a smooth graph for all  $t - r^2 < t' < t + r^2$  and such that

$$(1.7) \quad \sup_{|x'-x|<r, |t'-t|<r^2} r|A|(x', t') \leq 1.$$

Obviously  $|A|(x, t) \leq r_{\mathcal{M}}(x, t)^{-1}$ , but of course a bound for  $r_{\mathcal{M}}(x, t)^{-1}$  captures geometric control on a whole parabolic neighborhood of definite size as opposed to just information at the single point  $(x, t)$ .

**Theorem 1.2** (Sharp regularity estimate). *If  $\mathcal{M} = \{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$  is a level set flow with two-convex initial data, then we have the sharp estimate*

$$(1.8) \quad \int_{M_t} r_{\mathcal{M}}(x, t)^{-(n-1)} d\mu_t(x) \leq C,$$

for a constant  $C = C(\alpha, \beta, \gamma, \mathcal{A}) < \infty$ , which only depends on certain geometric parameters of the initial hypersurface  $M_0$  (see Section 2.1).

Since  $|A|(x, t) \leq r_{\mathcal{M}}(x, t)^{-1}$ , Theorem 1.2 immediately implies a sharp integral estimate for the second fundamental form:

**Corollary 1.1** (Sharp curvature control). *If  $\mathcal{M} = \{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$  is a level set flow with two-convex initial data, then we have the sharp curvature estimate<sup>3</sup>*

$$(1.9) \quad \int_{M_t} |A|^{n-1} d\mu_t \leq C,$$

for a constant  $C = C(\alpha, \beta, \gamma, \mathcal{A}) < \infty$ , which only depends on certain geometric parameters of the initial hypersurface  $M_0$  (see Section 2.1).

In particular, Theorem 1.2 and Corollary 1.1 crucially improve the regularity estimates from Head [13] and Cheeger-Haslhofer-Naber [5].

Let us now sketch the main ideas for our proof of Theorem 1.1.

To get some intuition, imagine first the potential scenario of the so-called fractal tube, as illustrated in Figure 1. The worry is that necks might slowly but steadily change their axes as one moves over an uncontrolled number of scales, and arrange themselves in a Koch snowflake like way, to make the intrinsic diameter arbitrarily large.

---

<sup>3</sup>Note that  $M_t$  is  $n$ -rectifiable with multiplicity one for all  $t$ , and that the second fundamental form  $A$  is defined classically  $\mu_t$ -almost everywhere.

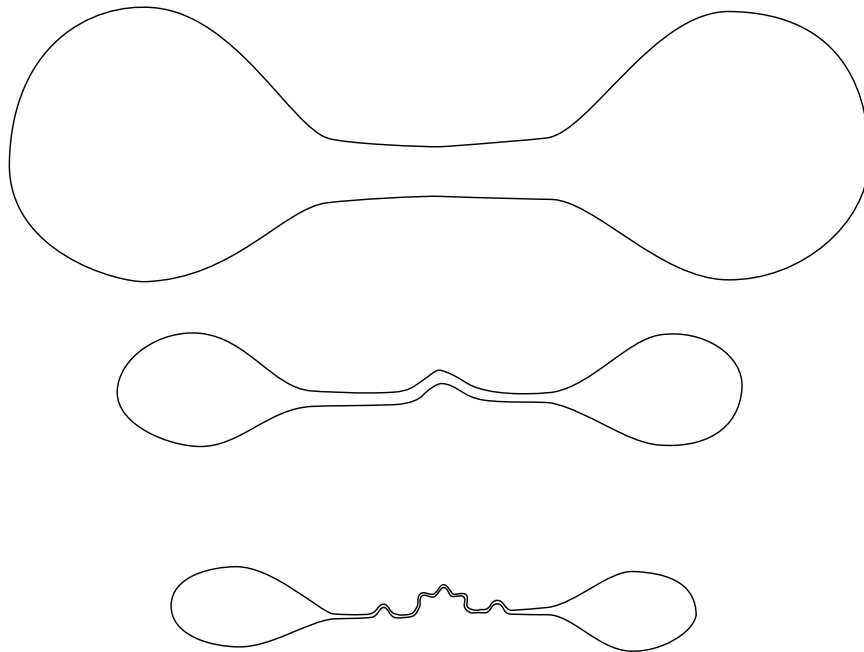


FIGURE 1. Fractal tube

In a recent breakthrough [6], Colding-Minicozzi proved a Łojasiewicz inequality for the mean curvature flow. Their Łojasiewicz inequality implies, roughly speaking, that cylindrical regions can actually only tilt by a controlled amount, *provided* one can ensure a priori that one is very close to a cylinder at all scales in consideration. Colding-Minicozzi applied the Łojasiewicz inequality to prove uniqueness of cylindrical tangent flows [6], to derive sharp results about the singular set of mean curvature flow with generic singularities [9], and also to derive related sharp results for the arrival time function of a mean convex flow [7, 8]. In particular, if the flow is two-convex then its space-time singular set is contained in finitely many compact embedded Lipschitz curves together with a countable set of points. This of course immediately rules out the example of an “exact fractal tube”, i.e. a fractal tube as above that becomes singular exactly on a fractal curve.

While the Łojasiewicz inequality from [6] has been applied many times to derive results about the *singular set*, it hasn’t been applied yet in a more quantitative way to derive results about *high curvature regions*. Note that a bound for the size of the singular set by itself, doesn’t yield any control for the diameter. For example, one could imagine the scenario of an “approximate fractal tube” that looks Koch-like shortly

before the first singular time, but then only becomes singular with a neck pinching off at one single point.

The main challenge in proving Theorem 1.1 is thus to ensure a priori that we can locate enough cylindrical regions propagating all the way from a microscopic scale to a macroscopic scale of definite size, to which we can apply the Łojasiewicz inequality. We achieve this as follows:

Given a two-convex mean curvature flow  $\mathcal{M} = \{M_t \subset \mathbb{R}^{n+1}\}_{t \in [0, T]}$ , and a time  $\bar{t} < T$ , we want to derive a bound for  $\text{diam}(M_{\bar{t}}, d_{\bar{t}})$  depending only on certain geometric parameters of the initial hypersurface  $M_0$ . Using in particular the canonical neighborhood theorem from [12], we first argue (see Proposition 3.1 and Proposition 3.2) that it is enough to control the quantity

$$(1.10) \quad L(M_{\bar{t}}) := \sup\{\text{diam}(N, d_{\bar{t}}) \mid N \subset M_{\bar{t}} \text{ is an } \varepsilon\text{-tube} \\ \text{with } H > \bar{H} \text{ on } N\}.$$

Here,  $\bar{H}$  denotes a very large curvature scale, that in particular is much larger than the curvature scale  $H_{\text{can}}$  from the canonical neighborhood theorem. An  $\varepsilon$ -tube is a subset  $N \subset M_{\bar{t}}$  diffeomorphic to  $S^{n-1} \times \mathbb{R}$  such that each  $x \in N$  lies on the central sphere of a very strong  $\varepsilon$ -neck at time  $\bar{t}$ . Roughly speaking, having a very strong  $\varepsilon$ -neck means that after rescaling to unit curvature one sees an almost round shrinking cylinder in a very large space-time neighborhood. The precision parameter  $\varepsilon$  enters in the rigorous definitions, replacing the informal words “almost” and “very large” (see Section 2.2 for the precise definitions). Moreover, we consider a central curve  $\gamma$  parametrized by arc-length such that  $N$  is covered by  $\varepsilon$ -necks centered at  $p \in \gamma$  and axis parallel to  $\partial_s \gamma(p)$ .

The central task is to understand how the flow arrived at the tube  $N$  at time  $t = \bar{t}$ . We show that there is a uniform  $\tau > 0$  such that around any  $p \in \gamma$  and at any time  $t \in [\bar{t} - \tau, \bar{t}]$  the flow has a strong  $\varepsilon_1$ -neck (here  $\varepsilon \ll \varepsilon_1 \ll 1$ ) centered at  $p$  of radius  $\sqrt{2(n-1)(t_p - t)}$ , where  $t_p$  essentially determines the scale of the neck and only depends on  $p$ .

To this end, first recall that by the canonical neighborhood theorem all points with  $H \geq H_{\text{can}}$  have a precise geometric description. The canonical neighborhoods are modeled on so-called  $(\alpha, \beta)$ -solutions (see Section 2.3), which in turn look neck-like away from cap-like pieces of controlled size. Inspired by [18], we prove a backwards-stability result for necks in  $(\alpha, \beta)$ -solutions (Proposition 4.2). Roughly speaking, the result says that if one sees a neck in an  $(\alpha, \beta)$ -solution then going far

enough back in time one sees an even more precise neck, and in particular not a cap. Combining the neck-stability result with a continuity argument and the assumption that we started off with a very strong  $\varepsilon$ -tube, we arrive at the conclusion that around any  $p \in \gamma$  one sees  $\varepsilon_1$ -necks of the right radius for a uniform amount  $\tau$  backwards in time.

Having done the above, we can finally apply the Łojasiewicz inequality from [6] to conclude that the total tilt of every neck in  $N$ , when followed via normal motion from time  $\bar{t}$  back to time  $\bar{t} - \tau$ , is in fact small (Proposition 4.1). Since at time  $\bar{t} - \tau$  all necks have macroscopic size, and since overlapping necks must have aligned axes, this gives the desired diameter control.

To prove Theorem 1.2, we first generalize Theorem 1.1 to the setting of mean curvature flow with surgery in the framework of [12]. The important point is that the diameter bound is independent of the surgery parameters, see (5.3). We then combine this with the canonical neighborhood theorem [12, Thm. 1.22] to establish uniform  $L^{n-1}$ -bounds for the mean curvature under mean curvature flow with surgery, see (5.4).

Given a fixed two-convex initial condition  $M_0$ , we then consider a sequence  $\{M_t^i\}$  of flows with surgery where the surgery parameters degenerate suitably. It follows from the work of Head [13] and Lauer [19], that for  $i \rightarrow \infty$  the sequence of flows with surgery converges to the level set flow. The convergence is in the space-time Hausdorff sense, and also in the sense of varifolds for every time. By lower-semicontinuity, this implies  $L^{n-1}$ -control for the mean curvature of level set flow with two-convex initial data. Finally, by the local curvature estimate [11, Thm. 1.8], this can be upgraded to  $L^{n-1}$ -control for the regularity scale.

This article is organized as follows: In Section 2, we introduce our notation and summarize the needed preliminaries. In Section 3, we carry out the reduction of the problem to the case of high curvature tubes. In Section 4, we prove the key propositions that have been outlined above and combine them to prove Theorem 1.1. Finally, in Section 5, we generalize the diameter bound to the setting of mean curvature flow with surgery and prove Theorem 1.2.

**Acknowledgments.** P.G. has been supported by a Fields-Ontario Postdoctoral Fellowship. R.H. has been supported by NSERC grant RGPIN-2016-04331, NSF grant DMS-1406394 and a Connaught New Researcher Award. We thank the Fields Institute and the University of Toronto for providing an excellent research environment. We thank

Nick Edelen, John Head, Mohammad Ivaki, Bruce Kleiner, and Felix Schulze for useful discussions.

## 2. NOTATION AND PRELIMINARIES

**2.1. Mean curvature flow.** Let  $M_0 \subset \mathbb{R}^{n+1}$  be a closed embedded hypersurface. Then there exist a unique smooth evolution by mean curvature flow  $\mathcal{M} = \{M_t\}_{t \in [0, T)}$  defined on a maximal time interval  $[0, T)$ . The maximal existence time  $T < \infty$  is characterized by

$$(2.1) \quad \lim_{t \nearrow T} \max_{M_t} |A| = \infty.$$

A closed embedded hypersurface  $M \subset \mathbb{R}^{n+1}$  is called *2-convex*, if

$$(2.2) \quad \lambda_1 + \lambda_2 > 0$$

for all  $p \in M$ , where  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  are the principal curvatures, i.e. the eigenvalues of the second fundamental form  $A$ . In particular, every 2-convex domain has positive mean curvature

$$(2.3) \quad H > 0.$$

We also need more quantitative notions of 2-convexity and embeddedness: A 2-convex hypersurface  $M$  is called  *$\beta$ -uniformly 2-convex*, if

$$(2.4) \quad \lambda_1 + \lambda_2 \geq \beta H$$

for all  $p \in M$ . An embedded hypersurface  $M \subset \mathbb{R}^{n+1}$  is called  *$\alpha$ -noncollapsed* if it has positive mean curvature and each  $p \in M$  admits interior and exterior balls of radius at least  $\alpha/H(p)$ . By [1, 17], all the above properties are preserved under the flow.

We also assume that  $H \leq \gamma$  initially. To keep track of the constants, we put  $\alpha = (\alpha, \beta, \gamma)$  and say that  $M_0$  is  $\alpha$ -controlled. By compactness, every 2-convex embedded hypersurfaces  $M_0$  is  $\alpha$ -controlled for some parameters  $\alpha, \beta > 0, \gamma < \infty$ . Some estimates will also depend on a bound  $\mathcal{A}$  for the initial area (recall that area is decreasing under the flow). While clearly most constants also depend on the dimension  $n$ , we usually don't explicitly indicate this in our notation.

**2.2. Necks and tubes.** In this section, we summarize our basic terminology about necks and tubes.

We say that an embedded hypersurface  $M^n \subset \mathbb{R}^{n+1}$  has an  $\varepsilon$ -neck with center  $p$  and radius  $r$  if the rescaled surface  $r^{-1}(M - p)$  is  $\varepsilon$ -close in



$C^{\lfloor 1/\varepsilon \rfloor}$  in  $B_{1/\varepsilon}(0)$  to a round cylinder  $S^{n-1} \times \mathbb{R}$  (up to rotation) with center 0 and radius 1.

As in [12, Def. 2.3], we say that a mean curvature flow  $\mathcal{M}$  has a *strong  $\varepsilon$ -neck* with center  $p$  and radius  $r$  at time  $t_0$ , if the parabolically rescaled flow  $\{r^{-1}(M_{t_0+r^2t} - p)\}_{t \in (-1, 0]}$  is  $\varepsilon$ -close in  $C^{\lfloor 1/\varepsilon \rfloor}$  in  $B_{1/\varepsilon}(0) \times (-1, 0]$  to the evolution of a round cylinder  $S^{n-1} \times \mathbb{R}$  (up to rotation) with center 0 and radius 1 at  $t = 0$ .

We say that  $\mathcal{M}$  has a *very strong  $\varepsilon$ -neck* with center  $p$  and radius  $r$  at time  $t_0$ , if in the above definition  $(-1, 0]$  can be replaced by  $(-2\mathcal{T}, 0]$ , where  $\mathcal{T} = \mathcal{T}(2\varepsilon_1, \frac{1}{2}\varepsilon_1, \alpha, \beta)$  denotes the constant from Proposition 4.2.<sup>4</sup>

A subset  $N \subset M_{\bar{t}}$  is called an  $\varepsilon$ -tube, if it is diffeomorphic to a cylinder and each  $x \in N$  lies on the central sphere of a very strong  $\varepsilon$ -neck at time  $\bar{t}$  (more precisely, the slice of a strong  $\varepsilon$ -neck at its maximal time).

Finally, as in [4], for each  $\varepsilon$ -tube we can find an  $\varepsilon$ -approximate central curve  $\gamma$ . In particular, for each  $p \in \gamma$  the vector  $\partial_s \gamma(p)$  determines the axis of the  $\varepsilon$ -neck centered at  $p$  (recall that the axis of an  $\varepsilon$ -neck is well defined up to errors of order  $\varepsilon$ ).

**2.3. Canonical neighborhoods.** We recall the canonical neighborhood theorem from Haslhofer-Kleiner [12] in the special case of smooth flows without surgeries.<sup>5</sup> The general case of the canonical neighborhood theorem for mean curvature flow with surgery will only be needed at the end, and its discussion will thus be deferred until Section 5.

**Theorem 2.1** (Special case of [12, Thm 1.22]). *For every  $\varepsilon_{can} > 0$ , there exist a constant  $H_{can}(\varepsilon_{can}) = H_{can}(\alpha, \varepsilon_{can}) < \infty$  with the following significance. If  $\mathcal{M}$  is a smooth mean curvature flow with  $\alpha$ -controlled initial data, then any  $(x, t) \in \mathcal{M}$  with  $H(x, t) \geq H_{can}(\varepsilon_{can})$  is  $\varepsilon_{can}$ -close to a  $\beta$ -uniformly 2-convex ancient  $\alpha$ -noncollapsed mean curvature flow.*

For brevity we refer to “ $\beta$ -uniformly 2-convex ancient  $\alpha$ -noncollapsed mean curvature flows” simply as  $(\alpha, \beta)$ -solutions. We recall that being  $\varepsilon_{can}$ -close is understood in a scale invariant sense. The conclusion in Theorem 2.1 thus means that the flow  $\mathcal{M}'$  which is obtained from  $\mathcal{M}$

<sup>4</sup>Here,  $\varepsilon_1$  is a certain quality parameter, that will be fixed in Section 4.3.

<sup>5</sup>In this special of smooth flows without surgeries the proof of the canonical neighborhood theorem simplifies quite a bit. In fact, Theorem 2.1 follows directly from the global convergence theorem for  $\alpha$ -noncollapsed flows [11, Thm. 1.10].

by shifting  $(x, t)$  to the origin and parabolically rescaling by  $H^{-1}(x, t)$  is  $\varepsilon_{\text{can}}$ -close in  $C^{\lfloor 1/\varepsilon_{\text{can}} \rfloor}$  in  $B_{1/\varepsilon_{\text{can}}}(0) \times (-\varepsilon_{\text{can}}^{-2}, 0]$  to an  $(\alpha, \beta)$ -solution.

The structure of  $(\alpha, \beta)$ -solutions has been analyzed in [12, Sec. 3.1]. In particular,  $(\alpha, \beta)$ -solutions are always convex and look neck-like away from caps of controlled size (in a scale invariant sense, as always).

**2.4. Lojasiewicz-Simon inequality.** For a hypersurface  $M \subset \mathbb{R}^{n+1}$  the Gaussian surface area is defined by

$$(2.5) \quad F(M) = (4\pi)^{-n/2} \int_M e^{-|x|^2/4} d\mathcal{H}^n.$$

The entropy  $\lambda$  is defined as supremum of the Gaussian area over all centers and scales:

$$(2.6) \quad \lambda(M) = \sup_{a>0, b \in \mathbb{R}^{n+1}} F(aM - b).$$

By Huisken's monotonicity formula [15] the entropy is nonincreasing under mean curvature flow. Thus, the entropy of all hypersurfaces under consideration will be bounded above by some constant  $\Lambda = \Lambda(\alpha, \gamma, \mathcal{A})$ . Also recall that if  $\{M_t \subset \mathbb{R}^{n+1}\}_{t<0}$  moves by mean curvature flow, then  $\Sigma_s = \frac{1}{\sqrt{-t}}M_t$ ,  $s = -\log(-t)$ , moves by rescaled mean curvature flow

$$(2.7) \quad \partial_s x = H + \frac{1}{2}x^\perp.$$

The rescaled mean curvature flow is the negative gradient flow of the  $F$ -functional.

**Theorem 2.2** (Colding-Minicozzi [6, Thm. 6.1]). *There exist constants  $K, \bar{R} < \infty, \varepsilon_L > 0, \mu < 1$  (depending only on the dimension and an upper bound for the entropy) such that if  $\{\Sigma_s\}_{s \in [t-1, t+1]}$  is a rescaled mean curvature flow such that  $B_{\bar{R}} \cap M_s$  is for each  $s$  a  $C^{2,\alpha}$  graph with norm at most  $\varepsilon_L$  over the cylinder  $Z = S^{\frac{n-1}{\sqrt{2(n-1)}}} \times \mathbb{R}$ , then*

$$(2.8) \quad |F(\Sigma_t) - F(Z)|^{1+\mu} \leq K (F(\Sigma_{t-1}) - F(\Sigma_{t+1})).$$

To see that (2.8) is indeed a discrete Lojasiewicz-Simon gradient inequality, it helps to rewrite the right hand side using

$$(2.9) \quad F(\Sigma_{t-1}) - F(\Sigma_{t+1}) = \int_{t-1}^{t+1} |\nabla_{\Sigma_s} F|^2 ds.$$

## 3. REDUCTION TO THE CYLINDRICAL CASE

Let  $\mathcal{M} = \{M_t\}_{t \in [0, T]}$  be a 2-convex flow with  $\alpha$ -controlled initial condition, and area bounded by  $\mathcal{A}$ .

Let  $\bar{t} < T$ . We want to show that  $\text{diam}(M_{\bar{t}}) \leq C$ , for some  $C < \infty$  depending only on  $\alpha$  and  $\mathcal{A}$ .

**3.1. Reduction to the high curvature case.** We will argue first that it is enough to estimate the length of geodesics that stay entirely in regions of high curvature. To this end, assume that  $\bar{H}$  is a constant with  $\bar{H} \geq H_{\text{can}}(\varepsilon_{\text{can}})$ , where the precision parameters  $\varepsilon > 0$  and  $\varepsilon_{\text{can}} = \varepsilon_{\text{can}}(\varepsilon, \alpha) > 0$  are small enough as in [12], and consider the quantity

$$(3.1) \quad D(M_{\bar{t}}) := \sup\{\ell(\gamma) \mid \gamma \text{ is a minimizing geodesic in } (M_{\bar{t}}, d_{\bar{t}}), \\ \text{and } H > 2\bar{H} \text{ along } \gamma\}.$$

**Proposition 3.1.** *There exists a constant  $\bar{N} = \bar{N}(\alpha, \mathcal{A}, \bar{H}) < \infty$  such that*

$$(3.2) \quad \text{diam}(M_{\bar{t}}) \leq \bar{N} + (\bar{N} + 1)D(M_{\bar{t}}).$$

*Proof.* Let  $\gamma : [0, L] \rightarrow (M_{\bar{t}}, d_{\bar{t}})$  be a minimizing geodesic parametrized by arclength. Choose a maximal collection  $s_1, \dots, s_N \in [0, L]$  such that  $H(\gamma(s_i)) \leq 2\bar{H}$  and  $|s_i - s_j| \geq 1$ . Consider then the geodesic balls  $B_i$  with center  $\gamma(s_i)$  and radius  $1/2$ .

We claim that the curvature is uniformly bounded in small balls centered at  $\gamma(s_i)$  of definite size (depending on  $\bar{H}$  and  $\alpha$ ): To this end, first observe that the  $\alpha$ -noncollapsing and the bound  $H \leq \gamma$  at the initial time imply that  $|A| \leq \sqrt{n}\gamma/\alpha$  at  $t = 0$ . By standard doubling estimates there is a constant  $T_1 = T_1(\alpha) > 0$  such that if  $\bar{t} \leq T_1$  then  $|A| \leq 2\sqrt{n}\gamma/\alpha$  on the whole hypersurface. If  $\bar{t} > T_1$  using that  $H(\gamma(s_i)) \leq 2\bar{H}$  we can apply the local curvature estimate [11, Thm. 1.8] to get curvature control in a ball centered at  $\gamma(s_i)$  of definite size.

By the above and the  $\alpha$ -noncollapsing we get a lower bound

$$(3.3) \quad \mathcal{H}^n(B_i) \geq c(\alpha, \bar{H}) > 0.$$

Since the balls  $B_i$  are disjoint, this implies that

$$(3.4) \quad N \leq \mathcal{A}/c =: \bar{N}.$$

Now restricting  $\gamma$  to the connected components of  $[0, L] \setminus \bigcup_{i=1}^N (s_i - \frac{1}{2}, s_i + \frac{1}{2})$  gives at most  $\bar{N} + 1$  minimizing geodesics with the additional property that  $H > 2\bar{H}$  along them. The assertion follows.  $\square$

**3.2. Applying the canonical neighborhood theorem.** Consider the quantity

$$(3.5) \quad L(M_{\bar{t}}) := \sup\{\text{diam}(N, d_{\bar{t}}) \mid N \subset M_{\bar{t}} \text{ is an } \varepsilon\text{-tube} \\ \text{with } H > \bar{H} \text{ on } N\}.$$

**Proposition 3.2.** *There exists a constant  $C = C(\alpha, \varepsilon, H_{\text{can}}) < \infty$  such that*

$$(3.6) \quad D(M_{\bar{t}}) \leq C + CL(M_{\bar{t}}).$$

*Proof.* Let  $\gamma$  be a minimizing geodesic in  $(M, d_{\bar{t}})$  with  $H > 2\bar{H}$  along  $\gamma$ . We can apply the canonical neighborhood theorem (Theorem 2.1) at each point  $(\gamma(s), \bar{t})$ . If the canonical model at some point is a convex solution of controlled geometry, then  $\gamma$  has controlled length and we are done. In all other cases, arguing as in the proof of [12, Cor. 1.25] we see that  $\gamma$  must be contained in an  $\varepsilon$ -tube, possibly capped at one or both ends, or with its ends identified. The caps have diameter at most  $C(\varepsilon)H_{\text{can}}^{-1}$ , and thus contribute at most a controlled amount to the length of  $\gamma$ . Similarly, in the case that the ends of the  $\varepsilon$ -tube are identified, we can also simply ignore a piece of small size, to reduce it again to the case of an  $\varepsilon$ -tube. The assertion follows.  $\square$

## 4. ESTIMATING THE LENGTH OF CYLINDRICAL REGIONS

**4.1. Small axis tilting under a priori assumptions.** In this section, we consider mean curvature flows under the a priori assumption that they are close to cylinders along a large – possibly uncontrolled – number of scales. We assume that the space-time center point is fixed, but a priori the axis is allowed to be different at every scale. Using methods from [6], we show that the total tilt of the axis is in fact small.

**Proposition 4.1.** *For all  $\varepsilon_0 > 0$  there exists an  $\varepsilon_1 = \varepsilon_1(\varepsilon_0, \Lambda) > 0$  with the following significance. Let  $\mathcal{M}$  be a mean curvature flow with entropy bounded by  $\Lambda$ , and suppose there are  $p \in \mathbb{R}^{n+1}$  and  $t_0 < t_1 < t_*$ , such that for all  $t \in [t_0, t_1]$ ,  $\mathcal{M}$  has a strong  $\varepsilon_1$ -neck with center  $p$  and radius  $\sqrt{2(n-1)(t_* - t)}$  at time  $t$ . Then there exists a  $v \in \mathbb{R}^{n+1}$  such*

that for all  $t \in [t_0, t_1]$ ,  $\mathcal{M}$  has a strong  $\varepsilon_0$ -neck with center  $p$ , radius  $\sqrt{2(n-1)(t_* - t)}$  at time  $t$ , and axis in the fixed direction  $v$ .

*Proof.* Our argument is a variant of the one from Colding-Minicozzi [6, Sec. 6]. Consider the rescaled mean curvature flow  $\Sigma_s = \frac{1}{\sqrt{t_* - t}} M_t$ ,  $s = -\log(t_* - t)$ . If  $\varepsilon_1$  is small enough (depending on  $\varepsilon_L = \varepsilon_L(\Lambda)$ ,  $\bar{R} = \bar{R}(\Lambda)$ ) then the hypothesis of Theorem 2.2 are satisfied. Hence,

$$(4.1) \quad |F(\Sigma_s) - F(Z)|^{1+\mu} \leq K(F(\Sigma_{s-1}) - F(\Sigma_{s+1})),$$

for every  $s \in [s_0, s_1]$ , where  $s_0 = \lceil -\log(t_* - t_0) + 1 \rceil$ ,  $s_1 = \lfloor -\log(t_* - t_1) - 1 \rfloor$ .

Applying the discrete Łojasiewicz lemma (Lemma A.1) for the function  $f(s) = F(\Sigma_s) - F(Z)$  we infer that for every  $\varepsilon > 0$  there exists an  $\bar{\varepsilon} = \bar{\varepsilon}(\varepsilon, \Lambda) > 0$  such that

$$(4.2) \quad \sum_{j=s_0+1}^{s_1} (F(\Sigma_j) - F(\Sigma_{j-1}))^{\frac{1}{2}} < \varepsilon,$$

provided  $\varepsilon_1 < \bar{\varepsilon}$ . Using the Cauchy-Schwarz inequality and the fact that the rescaled mean curvature flow is the negative gradient flow of  $F$ , this implies

$$(4.3) \quad \int_{s_0}^{s_1} \int_{\Sigma_s} |\vec{H} + \frac{1}{2}x^\perp| \frac{e^{-|x|^2/4}}{(4\pi)^{n/2}} d\mathcal{H}^n ds$$

$$(4.4) \quad \leq \Lambda^{1/2} \sum_{j=s_0+1}^{s_1} \left( \int_{j-1}^j \int_{\Sigma_s} |\vec{H} + \frac{1}{2}x^\perp|^2 \frac{e^{-|x|^2/4}}{(4\pi)^{n/2}} d\mathcal{H}^n ds \right)^{1/2}$$

$$(4.5) \quad = \Lambda^{1/2} \sum_{j=s_0+1}^{s_1} (F(\Sigma_j) - F(\Sigma_{j+1}))^{\frac{1}{2}} < \Lambda^{1/2} \varepsilon.$$

Choosing  $\varepsilon_1 = \varepsilon_1(\varepsilon_0, \Lambda)$  small enough, we can make the time-integral of the weighted  $L^1$ -norm of  $\vec{H} + \frac{1}{2}x^\perp$  as small as we want, and the assertion follows.  $\square$

**4.2. Backwards stability of necks.** Inspired by [18, Thm 6.1], we prove the following neck-stability result for  $(\alpha, \beta)$ -solutions:

**Proposition 4.2.** *There is a  $\delta_{neck} > 0$  such that for all  $\delta_0, \delta_1 \leq \delta_{neck}$  there is a  $\mathcal{T} = \mathcal{T}(\delta_0, \delta_1, \alpha, \beta) < \infty$  with the following property.*

*Suppose that  $\mathcal{M}$  is an  $(\alpha, \beta)$ -solution that has a  $\delta_0$ -neck with center  $p$  and radius  $\sqrt{2(n-1)}$  at time  $-1$ . Then for all  $t \in (-\infty, -\mathcal{T}]$  the*

flow  $\hat{\mathcal{M}}$  which is obtained from  $\mathcal{M}$  by shifting  $(p, 0)$  to the origin and parabolically rescaling by  $|t|^{-1/2}$  is  $\delta_1$ -close to a round shrinking cylinder  $S^{n-1} \times \mathbb{R}$  (up to rotation) that becomes extinct at time 0. In particular, for all  $t \in (-\infty, -\mathcal{T}]$  the unrescaled flow  $\mathcal{M}$  has a strong  $\delta_1$ -neck with center  $p$  and radius  $\sqrt{2(n-1)|t|}$  at time  $t$ .

*Proof.* We first fix small enough constants  $\delta_{\text{neck}} > 0$  and  $c > 0$  (depending only on the dimension) such that the Gaussian density (2.5) satisfies

$$(4.6) \quad F(M_{-1}) \geq \lambda(S^n) + c,$$

whenever  $\mathcal{M}$  has a  $\delta_0$ -neck with center 0 and radius  $\sqrt{2(n-1)}$  at time  $-1$ .

If the conclusion of the theorem didn't hold, then we could find a sequence  $\mathcal{M}^i$  of  $(\alpha, \beta)$ -solutions with a  $\delta_0$ -neck with center 0 and radius  $\sqrt{2(n-1)}$  at time  $-1$ , but such that the flows  $\hat{\mathcal{M}}^i$  which are obtained from  $\mathcal{M}^i$  by parabolically rescaling by  $|t_i|^{-1/2}$  are not  $\delta_1$ -close to a round shrinking cylinder, for some sequence  $t_i \rightarrow -\infty$ .

Consider Huisken's monotone quantity based at  $(0, 0)$ , namely

$$(4.7) \quad \Phi(M_t^i) = \int_{M_t^i} \frac{1}{(4\pi|t|)^{n/2}} e^{-|x|^2/4|t|} d\mathcal{H}^n \quad (t < 0).$$

Since for each fixed  $i$ , the flow  $\hat{\mathcal{M}}^i$  has a blowdown-limit (ancient soliton) which must be either a plane, a round shrinking sphere, or a round shrinking cylinder with only one  $\mathbb{R}$ -factor (see [11, Thm. 1.14, and its refinement for uniformly 2-convex flows]), we get the upper bound

$$(4.8) \quad \Phi(M_t^i) \leq \lambda(S^{n-1}).$$

After passing to a subsequence, we can find  $\tilde{t}_i \in (t_i, -1)$  with  $\tilde{t}_i/t_i \rightarrow 0$  such that

$$(4.9) \quad |\Phi(M_t^i) - \Phi(M_{\tilde{t}_i}^i)| < 1/i$$

for  $t \in [A_i \tilde{t}_i, A_i^{-1} \tilde{t}_i]$ , where  $A_i \rightarrow \infty$ .

Let  $\tilde{\mathcal{M}}^i$  be the sequence of flows that is obtained from  $\mathcal{M}^i$  by parabolically rescaling by  $|\tilde{t}_i|^{-1/2}$ . After passing to a subsequence we can pass to a limit  $\tilde{\mathcal{M}}^i \rightarrow \tilde{\mathcal{M}}^\infty$  which must be either a flat plane, a round shrinking sphere, or a round shrinking cylinder with only one  $\mathbb{R}$ -factor (see again [11, Thm. 1.14, and its refinement for uniformly

2-convex flows]). The plane and the sphere are excluded by (4.6). In particular, we see that  $\Phi(M_{\tilde{t}_i}^i) \rightarrow \lambda(S^{n-1})$ .

Thus, for  $i$  large enough  $\Phi(M_t^i)$  is almost constant on the interval  $(-\infty, \tilde{t}_i)$ . Consequently, after passing to another subsequence we can pass to a limit  $\hat{\mathcal{M}}^i \rightarrow \hat{\mathcal{M}}^\infty$  which must be a round shrinking cylinder with one  $\mathbb{R}$ -factor (see again [11, Thm. 1.14, and its refinement for uniformly 2-convex flows]); this gives the desired contradiction.  $\square$

**4.3. Conclusion of the argument.** Let  $\mathcal{M} = \{M_t\}_{t \in [0, T]}$  be a 2-convex flow with  $\alpha$ -controlled initial condition, and area bounded by  $\mathcal{A}$ . Let  $\bar{t} < T$ . We want to show that  $\text{diam}(M_{\bar{t}}) \leq C$ , for some constant  $C = C(\alpha, \mathcal{A}) < \infty$ .

The following argument depends on various quality parameters for necks. The logical order for choosing these parameters is that one first fixes a small constant  $\varepsilon_0$  (depending only on the dimension), and a large factor  $Q < \infty$  (e.g.  $Q=10$ ), and then successively determines suitable quality parameters  $\varepsilon_1, \varepsilon$ , and  $\varepsilon_{\text{can}}$  (where  $\varepsilon_{\text{can}} \ll \varepsilon \ll \varepsilon_1 \ll \varepsilon_0$ ) by reading the argument backwards.

Let  $N \subset M_{\bar{t}}$  by an  $\varepsilon$ -tube with  $H > \bar{H} = QH_{\text{can}}$  on  $N$ . By the reduction from Section 3, namely by Proposition 3.1 and Proposition 3.2, it is enough to estimate the length of  $N$ .

Let  $\gamma : [0, L] \rightarrow \mathbb{R}^{n+1}$  be an  $\varepsilon$ -approximate central curve for  $N$  (see Section 2.2), parametrized by arclength. We want to establish an upper bound for  $L$ , depending only on  $\alpha$  and  $\mathcal{A}$ . Suppose  $L > 2$  (otherwise we are done), and consider the truncated curve  $\gamma|_{[1, L-1]}$  (which we still denote by  $\gamma$ ) and the corresponding truncated  $\varepsilon$ -tube around it (which we still denote by  $N$ ).

Given any  $p \in \gamma$ , let  $x \in N$  be a point on the central sphere of the  $\varepsilon$ -neck centered at  $p$  of radius  $r_p$  at time  $\bar{t}$ . Note that  $|x - p| \leq (1 + \varepsilon)r_p$ . Also note that, since  $H(x) > \bar{H}$ , for  $\varepsilon$  small enough we have

$$(4.10) \quad r_p \leq 2(n-1)(QH_{\text{can}})^{-1}.$$

Set  $t_p = \bar{t} + \frac{1}{2(n-1)}r_p^2$ , and let  $\tau := \frac{1}{16}H_{\text{can}}^{-2}$ . We claim that for every  $p \in \gamma$  and every  $t \in [\bar{t} - \tau, \bar{t}]$ , the flow  $\mathcal{M}$  has a strong  $\varepsilon_1$ -neck with center  $p$  and radius

$$(4.11) \quad r(t) = \sqrt{2(n-1)(t_p - t)},$$

at time  $t$ , provided  $\varepsilon$  and  $\varepsilon_{\text{can}}$  are small enough. Note that the axis is a priori allowed to change when going from scale to scale.

Suppose towards a contradiction that  $t_0 \in [\bar{t} - \tau, \bar{t}]$  is the largest time such that  $\mathcal{M}$  does not have a strong  $\varepsilon_1$ -neck with center  $p$  and radius  $r(t_0)$  at time  $t_0$ .

Since every point in an  $\varepsilon$ -tube belongs to the final time slice of a very strong  $\varepsilon$ -neck, we immediately see that  $t_0 < \bar{t}$ . More precisely, given  $\varepsilon_1$  let  $\mathcal{T} = \mathcal{T}(2\varepsilon_1, \frac{1}{2}\varepsilon_1, \alpha, \beta)$  be the constant from Proposition 4.2 (backwards stability). Since there is a very strong  $\varepsilon$ -neck centered at  $p$  of radius  $r(\bar{t})$  at time  $\bar{t}$ , then for every  $t$  satisfying

$$(4.12) \quad t - r(t)^2 \geq \bar{t} - 2\mathcal{T}r(\bar{t})^2,$$

there is a strong  $\varepsilon_1$ -neck centered at  $p$  of radius  $r(t)$  at time  $t$ . At time  $t = t_0$  the inequality (4.12) must be violated, in other words

$$(4.13) \quad \frac{t_p - t_0}{t_p - \bar{t}} > \frac{4(n-1)\mathcal{T} + 1}{2(n-1) + 1} \geq \frac{3}{2}\mathcal{T},$$

where we tacitly assume that  $\mathcal{T}$  is large enough (depending only on the dimension). Since obviously  $\frac{t_p - t_0}{t_p - t_0} = 1$ , by the intermediate value theorem we can find a  $t_1 \in (t_0, \bar{t}]$  such that

$$(4.14) \quad \frac{t_p - t_0}{t_p - t_1} = \frac{3}{2}\mathcal{T}.$$

Since  $t_1 \in (t_0, \bar{t}]$ , the flow  $\mathcal{M}$  has an  $\varepsilon_1$ -neck centered at  $p$  of radius  $r(t_1)$  at time  $t = t_1$ . Hence, there is a point  $x_{t_1} \in M_{t_1}$  on the central sphere of that neck, which satisfies in particular

$$(4.15) \quad \frac{n-1}{(1+\varepsilon_1)r(t_1)} < H(x_{t_1}) < \frac{(1+\varepsilon_1)(n-1)}{r(t_1)},$$

and

$$(4.16) \quad |x_{t_1} - p| < (1 + \varepsilon_1)r(t_1).$$

Using (4.10), (4.11), and our choice of  $Q$  and  $\tau$  we see that

$$(4.17) \quad r^2(\bar{t} - \tau) = r_p^2 + 2(n-1)\tau \leq \frac{(n-1)^2}{4}H_{\text{can}}^{-2},$$

which together with (4.15) and the obvious inequality  $r(t_1) \leq r(\bar{t} - \tau)$ , implies that

$$(4.18) \quad H(x_{t_1}) \geq H_{\text{can}}.$$



Thus, by Theorem 2.1 (canonical neighborhoods) and (4.16) the flow  $\hat{\mathcal{M}}$  that is obtained from  $\mathcal{M}$  by shifting  $(p, t_1)$  to the origin and parabolically rescaling by  $H^{-1}(x_{t_1})$  is  $\varepsilon_{\text{can}}$ -close in  $C^{\lfloor 1/\varepsilon_{\text{can}} \rfloor}$  in  $B_{1/\varepsilon_{\text{can}}-2n}(0) \times (-\varepsilon_{\text{can}}^2, 0]$  to an  $(\alpha, \beta)$ -solution  $\mathcal{N}$ .

For  $\varepsilon_{\text{can}}$  small enough, the  $(\alpha, \beta)$ -solution  $\mathcal{N}$  has a  $2\varepsilon_1$ -neck with radius  $(n-1)$  and center 0 at time 0. By Proposition 4.2 (backwards stability) it follows that  $\mathcal{N}$  has a strong  $\frac{1}{2}\varepsilon_1$ -neck centered at 0 of radius  $\sqrt{2(n-1)(s+(n-1)/2)}$  at time  $-s$ , provided  $s \geq (n-1)(\mathcal{T}-1)/2$ .

Putting things together, it follows that if  $\varepsilon_{\text{can}} \ll \varepsilon_1$  (depending only on  $\alpha$ ) then  $\mathcal{M}$  has a strong  $\frac{3}{4}\varepsilon_1$ -neck centered at  $p$  of radius  $r(t_0) = \sqrt{2(n-1)(t_p-t_0)}$  at time  $t_0$ ; a contradiction. This proves the claim.

Now we can apply Proposition 4.1 (small tilt) with  $\varepsilon_1 = \varepsilon_1(\varepsilon_0, \Lambda)$  to conclude that for every  $p \in \gamma$  there exists  $O_p \in \text{SO}_{n+1}$  such that for all  $t \in [\bar{t}-\tau, \bar{t}]$  we have that  $M_t$  is  $\varepsilon_0$ -close to the cylinder  $Z_p = p + O_p(S^{\frac{n-1}{2}} \times \mathbb{R})$ , with fixed axis, in  $B_{\varepsilon_0^{-1}\sqrt{2(n-1)(t_p-t)}}(p)$ .

Finally, if  $p_1, p_2 \in \gamma$  are points with  $|p_1 - p_2| < \frac{1}{4}\varepsilon_0^{-1}\sqrt{\tau}$ , then the associated cylinders  $Z_{p_1}, Z_{p_2}$  have substantial overlap at time  $t = \bar{t} - \tau$ , hence  $\|O_{p_1} - O_{p_2}\| = O(\varepsilon_0)$  and  $\min\left\{\frac{t_{p_1}}{t_{p_2}}, \frac{t_{p_2}}{t_{p_1}}\right\} = 1 + O(\varepsilon_0)$ . Thus

$$(4.19) \quad d_\gamma(p_1, p_2) \leq (1 + O(\varepsilon_0))|p_1 - p_2|,$$

i.e. the intrinsic distance along  $\gamma$  between any two points  $p_1, p_2 \in \gamma$  with  $|p_1 - p_2| < \frac{1}{4}\varepsilon_0^{-1}\sqrt{\tau}$  is controlled by  $(1 + O(\varepsilon_0))$  times their extrinsic distance. Thus, the intersection of  $\gamma$  with any ball of radius  $\frac{1}{8}\varepsilon_0^{-1}\sqrt{\tau}$  is  $O(\varepsilon_0)$ -close to a linear segment. Since  $M_0$  is contained in large ball of radius  $R = R(\alpha, \mathcal{A}) < \infty$ , we conclude that the length of  $\gamma$  is bounded by some constant depending only on  $\alpha$  and  $\mathcal{A}$ . This finishes the proof of Theorem 1.1.

## 5. PROOF OF THEOREM 1.2

In this final section, we will prove Theorem 1.2.

We use the framework of mean curvature flow with surgery from [12]. Thus, for us a mean curvature flow with surgery is a  $(\alpha, \delta, \mathbb{H})$ -flow as defined in [12, Def. 1.17]. Recall in particular that  $\alpha = (\alpha, \beta, \gamma)$  denotes the control parameters for the two-convex initial hypersurface  $M_0$ , that  $\delta$  specifies the quality of the surgeries, and that the three curvature

scales  $\mathbb{H} = (H_{\text{trig}}, H_{\text{neck}}, H_{\text{th}})$  are used to specify more precisely when and how surgeries (and/or discarding) are performed.

The main existence theorem [12, Thm. 1.21] and the canonical neighborhood theorem [12, Thm. 1.22], tell us that for any  $\alpha$ -controlled initial hypersurface  $M_0$ , there exists an  $(\alpha, \delta, \mathbb{H})$ -flow starting at  $M_0$  such that all points with  $H \geq H_{\text{can}}(\alpha, \varepsilon_{\text{can}})$  possess canonical neighborhoods, provided  $\delta$  is small enough, and  $H_{\text{th}}$  as well as the ratios  $H_{\text{trig}}/H_{\text{neck}}$ ,  $H_{\text{neck}}/H_{\text{th}}$  are large enough (depending only on  $\alpha$ ).

We make the following two minor adjustments compared to [12]. First, we work with ‘very strong’ necks instead of ‘strong’ necks. This is just a cosmetic change, that only slightly alters some constants. Second, and more importantly, in the line before [12, Claim 4.7] instead of selecting an arbitrary minimal collection of disjoint  $\delta$ -necks that separates the thick part and the trigger part, we select an ‘innermost’ collection of such  $\delta$ -necks, i.e. we impose the additional condition that

$$(5.1) \quad \sum_{p \in \mathcal{J}_j} \text{dist}(p, \{H = H_{\text{trig}}\})$$

is minimal. If the surgeries are performed this way, then on the discarded components we have

$$(5.2) \quad H \geq cH_{\text{neck}}$$

for some  $c = c(\alpha) > 0$ .

The argument from the previous sections (taking into account the one additional case that canonical neighborhoods can now also be modelled on the evolution of a standard cap preceded by the evolution of a round shrinking cylinder) shows that every connected component  $M'_t \subset M_{\bar{t}}$  satisfies

$$(5.3) \quad \text{diam}(M'_t, d_{\bar{t}}) \leq C(\alpha, \mathcal{A}) < \infty$$

for all  $\bar{t} \geq 0$ , where  $d_{\bar{t}}$  denotes the intrinsic distance on  $M'_{\bar{t}}$ .

Moreover, by the canonical neighborhood theorem, the nature of the surgeries, and the  $\alpha$ -noncollapsing, the number of connected components is uniformly bounded by some  $N = N(\alpha, \mathcal{A}, H_{\text{thick}}) < \infty$ .

We want to show that for any  $\bar{t} \geq 0$  we have the estimate

$$(5.4) \quad \int_{M'_{\bar{t}}} H^{n-1} d\mu_{\bar{t}} \leq C(\alpha, \mathcal{A}) < \infty.$$

To this end, write  $M'_t = M_t^{\text{low}} \cup M_t^{\text{high}}$ , with  $M_t^{\text{low}} = \{x \in M'_t : H(x) \leq \bar{H}\}$  and  $M_t^{\text{high}} = \{x \in M'_t : H(x) > \bar{H}\}$ , where  $\bar{H} \gg H_{\text{can}}$  is a large but fixed constant as in the previous section.

Using the canonical neighborhood theorem, we can decompose  $M_t^{\text{high}}$  into the union of  $\varepsilon$ -tubes, caps, standard caps, and compact solutions of controlled geometry. It is easy to see that the latter three only contribute a controlled amount to the integral in (5.4).

Hence, let  $M_t^{\text{tubes}} \subset M_t^{\text{high}}$  be the union of the remaining  $\varepsilon$ -tubes, with curvature larger than  $\bar{H}$ . Note that at this point we allow the  $\varepsilon$ -tubes to have identified ends and combine tubes with their ends attached.

We can then separate the connected components of  $M_t^{\text{tubes}}$  in two types:  $\varepsilon$ -tubes  $\{T_i^{\text{low}}\}$  with curvature less than  $2\bar{H}$  and  $\varepsilon$ -tubes  $\{T_i^{\text{high}}\}$  that intersect the set  $\{H > 2\bar{H}\}$ . Since we obviously have

$$(5.5) \quad \int_{M_t \cap \{H \leq 2\bar{H}\}} H^{n-1} d\mu_t \leq (2\bar{H})^{n-1} \mathcal{A},$$

it remains to estimate

$$(5.6) \quad \int_{\bigcup_i T_i^{\text{high}}} H^{n-1} d\mu_t.$$

Now, we can write any  $\varepsilon$ -tube  $T^{\text{high}}$  as the union of a maximal collection of  $\varepsilon$ -necks  $N_i$  centered at  $p_i$  of radius  $r_i$  at time  $t = \bar{t}$ , such that any two balls  $B_{r_i/5}(p_i), B_{r_j/5}(p_j)$  are disjoint. We can then estimate

$$(5.7) \quad \int_{T^{\text{high}}} H^{n-1} d\mu_t \leq \sum_i \int_{N_i} H^{n-1} d\mu_t \leq c \sum_i r_i \leq cL,$$

where  $L$  is the length of the  $\varepsilon$ -tube  $N$  and  $c = c(n) < \infty$  is a numerical constant. Since the intrinsic diameter of each connected component is uniformly bounded, the integral (5.7) is bounded by some constant  $C = C(\alpha, \mathcal{A}) < \infty$ .

To bound the integral (5.6), we need to control the number of tubes in the union. Assume without essential loss of generality that  $M_t^{\text{low}} \neq \emptyset$  (the case  $M_t^{\text{low}} = \emptyset$  can be handled easily). Note that each tube  $T^{\text{high}}$  has a point  $x$  with  $H(x) = \frac{3}{2}\bar{H}$ . This implies that  $x$  belongs to an  $\varepsilon$ -neck that is entirely contained in  $T^{\text{high}}$  since its curvature is approximately  $\frac{3}{2}\bar{H}$ , so each tube  $T_i^{\text{high}}$  contributes a definite amount of area. Since the total area of  $M_t$  is bounded by  $\mathcal{A}$ , this suffices to control the number of

such tubes by a constant that depends only on  $\alpha$  and  $\mathcal{A}$ , and concludes the proof of (5.4).

Now, in order to prove the regularity estimate for the level set flow, fix the initial hypersurface  $M_0$ , and consider a sequence  $(\alpha, \delta, \mathbb{H}^j)$ -flows starting at  $M_0$ , with  $H_{\text{neck}}^j \rightarrow \infty$ , but  $H_{\text{thick}}^j$  bounded. By the above, we have the uniform estimate

$$(5.8) \quad \int_{M_t^j} H^{n-1} d\mu_t \leq C(\alpha, \mathcal{A}).$$

By the work of Head [13] and Lauer [19] (see also [11, Prop. 1.27, Rem. 4.11]; in particular, note that equation (5.2) ensures that the curvature of the discarded components goes to infinity) for  $j \rightarrow \infty$  the flows with surgery  $\{M_t^j\}$  converge to the level set flow  $\{M_t\}$  with initial condition  $M_0$ . The convergence is both in the space-time Hausdorff sense and also in the varifold sense for every time. By lower-semicontinuity of  $L^p$ -norms of the mean curvature under varifold convergence, we thus infer that the level set flow satisfies

$$(5.9) \quad \int_{M_t} H^{n-1} d\mu_t \leq C(\alpha, \mathcal{A})$$

for every  $t \geq 0$ . Finally, by the local curvature estimate [11, Thm. 1.8] we have  $r_{\mathcal{M}}^{-1} \leq C(\alpha)H$ , and the assertion of Theorem 1.2 follows.

## APPENDIX A. DISCRETE ŁOJASIEWICZ LEMMA

**Lemma A.1** (Discrete Łojasiewicz lemma, c.f. [6, Lem. 6.9]). *For every  $\varepsilon > 0$ , there exists a  $\delta = \delta(\varepsilon, K, \mu) > 0$  with the following significance. Suppose that  $f : \{0, 1, \dots, T\} \rightarrow \mathbb{R}$  is a non-increasing function such that for some  $K < \infty$  and  $\mu < 1$  it holds that*

$$(A.1) \quad |f(t)|^{1+\mu} \leq K(f(t-1) - f(t+1))$$

for  $t = 1, \dots, T-1$ , and suppose that  $|f| \leq \delta$ . Then

$$(A.2) \quad \sum_{j=1}^T (f(j) - f(j-1))^{\frac{1}{2}} \leq \varepsilon.$$

*Proof.* For most parts of the proof we only assume  $|f| \leq 1$ . We will impose the condition that  $|f| \leq \delta$  is actually small towards the end.

Let  $t_0 \in \{0, 1, \dots, T\}$  be the smallest integer with the property  $f(j) < 0$  for every  $t_0 < j \leq T$ .

If  $t_0 > 0$ , then as in [6, Lemma 6.9], there is a  $C = C(K, \mu) < \infty$  such that

$$(A.3) \quad f(t) \leq Ct^{-\frac{1}{\mu}},$$

for every  $t \in [0, t_0]$ . Moreover, for  $p \in (1, \frac{1}{\mu})$  and any  $j_0 \in [1, t_0]$ , using the Cauchy–Schwarz inequality we obtain

$$(A.4) \quad \left( \sum_{j=j_0}^{t_0} (f(j) - f(j+1))^{\frac{1}{2}} \right)^2 \leq \left( \sum_{j=j_0}^{t_0} (f(j) - f(j+1))j^p \right) \sum_{j=j_0}^{t_0} j^{-p}.$$

Estimating the right-hand side of (A.4) as in [6, Lemma 6.9] we find  $b_1 = b_1(\delta, K, \mu) < \infty$  such that if  $b_1 \leq j_0 \leq t_0$  then

$$(A.5) \quad \sum_{j=j_0}^{t_0} (f(j) - f(j+1))^{\frac{1}{2}} < \frac{\varepsilon}{4}.$$

In case that  $t_0 < T$ , we also consider the function  $\tilde{f}(t) := -f(T-t)$ . This function  $\tilde{f} : [0, T-t_0] \rightarrow [0, \infty)$  is non-increasing and satisfies

$$(A.6) \quad \begin{aligned} \tilde{f}(t)^{1+\mu} &= |f(T-t)|^{1+\mu} \leq K (f(T-t-1) - f(T-t+1)) \\ &= K (\tilde{f}(t-1) - \tilde{f}(t+1)), \end{aligned}$$

and  $|\tilde{f}| \leq 1$ . As above, we can find  $b_2 = b_2(\delta, K, \mu) < \infty$  such that if  $b_2 \leq j_0 \leq T-t_0-2$  then

$$(A.7) \quad \sum_{j=t_0+1}^{T-j_0-1} (f(j) - f(j+1))^{\frac{1}{2}} = \sum_{k=j_0}^{T-t_0-2} (\tilde{f}(k) - \tilde{f}(k+1))^{\frac{1}{2}} < \frac{\varepsilon}{4}.$$

Together with (A.5) this implies that for  $b = \max\{b_1, b_2 + 1\}$

$$(A.8) \quad \sum_{j=b}^{T-b} (f(j) - f(j+1))^{\frac{1}{2}} < \frac{\varepsilon}{2},$$

tacitly assuming that  $T > 2b$  (otherwise there is not much to prove, see the next sentence). Since only  $2b - 2$  terms of (A.2) missing from (A.8), and each is bounded by  $(f(j) - f(j+1))^{\frac{1}{2}} < (2\delta)^{\frac{1}{2}}$ , we conclude that (A.2) holds, provided we choose  $\delta$  small enough.  $\square$

## REFERENCES

1. B. Andrews, *Noncollapsing in mean-convex mean curvature flow*, *Geom. Topol.* **16** (2012), no. 3, 1413–1418.

2. S. Brendle and G. Huisken, *A fully nonlinear flow for two-convex hypersurfaces*, arXiv:1507.04651 (2015).
3. S. Brendle and G. Huisken, *Mean curvature flow with surgery of mean convex surfaces in  $\mathbb{R}^3$* , *Invent. Math.* **203** (2016), no. 2, 615–654.
4. R. Buzano, R. Haslhofer, and O. Hershkovits, *The moduli space of two-convex embedded spheres*, arXiv:1607.05604 (2016).
5. J. Cheeger, R. Haslhofer, and A. Naber, *Quantitative stratification and the regularity of mean curvature flow*, *Geom. Funct. Anal.* **23** (2013), no. 3, 828–847.
6. T. Colding and W. Minicozzi, *Uniqueness of blowups and Lojasiewicz inequalities*, *Ann. of Math. (2)* **182** (2015), no. 1, 221–285.
7. ———, *Differentiability of the arrival time*, *Comm. Pure Appl. Math.* **69** (2016), no. 12, 2349–2363.
8. ———, *Regularity of the level set flow*, arXiv:1606.05185 (2016).
9. ———, *The singular set of mean curvature flow with generic singularities*, *Invent. Math.* **204** (2016), no. 2, 443–471.
10. N. Edelen and M. Engelstein, *Quantitative stratification for some free-boundary problems*, arXiv:1702.04325 (2017).
11. R. Haslhofer and B. Kleiner, *Mean curvature flow of mean convex hypersurfaces*, *Comm. Pure Appl. Math.* **70** (2017), no. 3, 511–546.
12. ———, *Mean curvature flow with surgery*, *Duke Math. J.* **166** (2017), no. 9, 1591–1626.
13. J. Head, *On the mean curvature evolution of two-convex hypersurfaces*, *J. Differential Geom.* **94** (2013), no. 2, 241–266.
14. G. Huisken, *Flow by mean curvature of convex surfaces into spheres*, *J. Differential Geom.* **20** (1984), no. 1, 237–266.
15. ———, *Asymptotic behavior for singularities of the mean curvature flow*, *J. Differential Geom.* **31** (1990), no. 1, 285–299.
16. G. Huisken and A. Polden, *Geometric evolution equations for hypersurfaces*, *Calculus of variations and geometric evolution problems* (Cetraro, 1996), *Lecture Notes in Math.*, vol. 1713, Springer, Berlin, 1999, pp. 45–84.
17. G. Huisken and C. Sinestrari, *Mean curvature flow with surgeries of two-convex hypersurfaces*, *Invent. Math.* **175** (2009), no. 1, 137–221.
18. B. Kleiner and J. Lott, *Singular Ricci flows I*, arXiv:1408.2271 (2014).
19. J. Lauer, *Convergence of mean curvature flows with surgery*, *Comm. Anal. Geom.* **21** (2013), no. 2, 355–363.
20. A. Naber and D. Valtorta, *The Singular Structure and Regularity of Stationary and Minimizing Varifolds*, arXiv:1505.03428 (2015).
21. ———, *Stratification for the singular set of approximate harmonic maps*, arXiv:1611.03008 (2016).
22. A. Naber and D. Valtorta, *Rectifiable-Reifenberg and the regularity of stationary and minimizing harmonic maps*, *Ann. of Math. (2)* **185** (2017), no. 1, 131–227.
23. G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, arXiv:math/0211159, 2002.
24. P. Topping, *Relating diameter and mean curvature for submanifolds of Euclidean space*, *Comment. Math. Helv.* **83** (2008), no. 3, 539–546.

25. Y. Wang, *Quantitative stratification of stationary Yang-Mills fields*, arXiv:1610.00351 (2016).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, 40 ST  
GEORGE STREET, TORONTO, ON M5S 2E4, CANADA

*E-mail:* p.gianniotis@utoronto.ca, roberth@math.toronto.edu