Uniqueness of blowups and Łojasiewicz inequalities

UNIQUENESS OF BLOWUPS AND LOJASIEWICZ INEQUALITIES

TOBIAS HOLCK COLDING AND WILLIAM P. MINICOZZI II

We dedicate this article to Leon Simon in recognition of his fundamental contributions to analysis and geometry.

Abstract. Once one knows that singularities occur, one naturally wonders what the singularities are like. For minimal varieties the first answer, already known to Federer-Fleming in 1959, is that they weakly resemble cones[1]. For mean curvature flow, by the combined work of Huisken, Ilmanen, and White, singularities weakly resemble shrinkers. Unfortunately, the simple proofs leave open the possibility that a minimal variety or a mean curvature flow looked at under a microscope will resemble one blowup, but under higher magnification, it might (as far as anyone knows) resemble a completely different blowup. Whether this ever happens is perhaps the most fundamental question about singularities. It is this long standing open question that we settle here for mean curvature flow at all generic singularities and for mean convex mean curvature flow at all singularities.

0. Introduction

We show that at each generic singularity of a mean curvature flow the blowup is unique; that is independent of the sequence of rescalings. This settles a major open problem that was open even in the case of mean convex hypersurfaces where it was known that all singularities are generic. Moreover, it is the first general uniqueness theorem for blowups to a Geometric PDE at a non-compact singularity.

Uniqueness of blowups is perhaps the most fundamental question that one can ask about singularities and implies regularity of the singular set; see [CM5].

To prove our uniqueness result, we prove two completely new infinite dimensional Lojasiewicz type inequalities. Infinite dimensional Lojasiewicz inequalities were pioneered thirty years ago by Leon Simon. However, unlike all other infinite dimensional Lojasiewicz inequalities we know of, ours do not follow from a reduction to the classical finite-dimensional Lojasiewicz inequalities from the 1960s from algebraic geometry, rather we prove our inequalities directly and do not rely on Lojasiewicz's arguments or results.

It is well-known that to deal with non-compact singularities requires entirely new ideas and techniques as one cannot argue as in Simon's work, and all the later work that uses his ideas. Partly because of this we expect that the techniques and ideas developed here have applications to other flows. Our results hold in all dimensions.

The authors were partially supported by NSF Grants DMS 11040934, DMS 1206827, and NSF FRG grants DMS 0854774 and DMS 0853501.

[1]See Brian White [W4] section "Uniqueness of tangent cone" from which part of this discussion is taken and where one can find more discussion of uniqueness for minimal varieties.

This paper focuses on mean curvature flow (or MCF) of hypersurfaces. This is a non-linear parabolic evolution equation where a hypersurface evolves over time by locally moving in the direction of steepest descent for the volume element. It has been used and studied in material science for almost a century[2] to model things like cell, grain, and bubble growth[3]. Unlike some of the other earlier papers in material science both von Neumann's 1952 paper and Mullins 1956 paper had explicit equations. In his paper von Neumann discussed soap foams whose interface tend to have constant mean curvature whereas Mullins is describing coarsening in metals, in which interfaces are not generally of constant mean curvature. Partly as a consequence, Mullins may have been the first to write down the MCF equation in general. Mullins also found some of the basic self-similar solutions like the translating solution now known as the Grim Reaper. To be precise, suppose that $M_t \subset \mathbf{R}^{n+1}$ is a one-parameter family of smooth hypersurfaces, then we say that M_t flows by the MCF if

$$x_t = -H\,\mathbf{n}\,, \tag{0.1}$$

where H and $\mathbf{n}$ are the mean curvature and unit normal, respectively, of M_t at the point x.

To understand singularities past the first singular time, we need weak solutions of MCF. The weak solutions that we will use are the Brakke flows considered by White in [W3][4]. By theorem 7.4 in [W3], this includes flows starting from any closed embedded hypersurface.

0.1. **Tangent flows.** By definition, a tangent flow is the limit of a sequence of rescalings at a singularity, where the convergence is on compact subsets.[5] For instance, a tangent flow to M_t at the origin in space-time is the limit of a sequence of rescaled flows $\frac{1}{\delta_i}\,M_{\delta_i^2\,t}$ where $\delta_i \to 0$. A priori, different sequences δ_i could give different tangent flows and the question of the uniqueness of the blowup - independent of the sequence - is a major question in many geometric problems. By a monotonicity formula of Huisken, [H1], and an argument of Ilmanen and White, [I1], [W3], tangent flows are shrinkers, i.e., self-similar solutions of MCF that evolve by rescaling. The only generic shrinkers are round cylinders by [CM1].

We will say that a singular point is *cylindrical* if at least one tangent flow is a multiplicity one cylinder $\mathbf{S}^k \times \mathbf{R}^{n-k}$. Our main application of our analytical inequalities is the following theorem that shows that tangent flows at generic singularities are unique:

Theorem 0.2. Let M_t be a MCF in $\mathbf{R}^{n+1}$. At each cylindrical singular point the tangent flow is unique. That is, any other tangent flow is also a cylinder with the same $\mathbf{R}^k$ factor that points in the same direction.

---

[2]See, e.g., the early work in material science from the 1920s, 1940s, and 1950s of T. Sutoki, [Su], D. Harker and E. Parker, [HaP], J. Burke, [Bu], P.A. Beck, [Be], J. von Neumann, [N], and W.W. Mullins, [M].

[3]For instance, annealing, in metallurgy and materials science, is a heat treatment that alters a material to increase its ductility and to make it more workable. It involves heating material above its critical temperature, maintaining a suitable temperature, and then cooling. Annealing can induce ductility, soften material, relieve internal stresses, refine the structure by making it homogeneous, and improve cold working properties. The three stages of the annealing process that proceed as the temperature of the material is increased are: recovery, recrystallization, and grain growth. Grain growth is the increase in size of grains (crystallites) in a material at high temperature. This occurs when recovery and recrystallisation are complete and further reduction in the internal energy can only be achieved by reducing the total area of grain boundary [by mean curvature flow].

[4]That is, Brakke flows in the class $S(\lambda_0, n, n+1)$ defined in section 7 of [W3] for some $\lambda_0 > 1$.

[5]This is analogous to a tangent cone at a singularity of a minimal variety, cf. [FFl].

This theorem solves a major open problem; see, e.g., page 534 of [W2]. Even in the case of the evolution of mean convex hypersurfaces where all singularities are cylindrical, uniqueness of the axis was unknown; see [HS1], [HS2], [W1], [SS], [An] and [HaK].[6]

In recent joint work with Tom Ilmanen, [CIM], we showed that if one tangent flow at a singular point of a MCF is a multiplicity one cylinder, then all are. However, [CIM] left open the possibility that the direction of the axis (the $\mathbf{R}^k$ factor) depended on the sequence of rescalings. Our proof of Theorem 0.2 and, in particular, our first Lojasiewicz type inequality, has its roots in some ideas and inequalities from [CIM] and in fact implicitly use that cylinders are isolated among shrinkers by [CIM].

Uniqueness is a key question for the regularity of Geometric PDE's. Two of the most prominent early works on uniqueness of tangent cones are Leon Simon's hugely influential paper [Si1] from 1983, where he proves uniqueness for tangent cones of minimal varieties with smooth cross-section. The other is Allard-Almgren's 1981 paper [AA], where uniqueness of tangent cones with smooth cross-section is proven under an additional integrability assumption on the cross-section; see also [Si2], [Hr], [CM4] for additional references.

Our results are the first general uniqueness theorems for tangent flows to a geometric flow at a non-compact singularity. (In fact, not only are the singularities that we deal with here non-compact but they are also non-integrable; see Section 3.) Some special cases of uniqueness of tangent flows for MCF were previously analyzed assuming either some sort of convexity or that the hypersurface is a surface of rotation; see [H1], [H2], [HS1], [HS2], [W1], [SS], [AAG], section 3.2 in the book [GGS], and [GK], [GKS], [GS]. In contrast, uniqueness for blowups at compact singularities is better understood; cf. [AA], [Si1], [H3], [Sc], and [Se].

In fact, using the results of this paper we showed in [CM5] that, for a MCF of closed embedded hypersurfaces in $\mathbf{R}^{n+1}$ with only cylindrical singularities, the space-time singular set is contained in finitely many compact embedded $(n-1)$-dimensional Lipschitz submanifolds together with a set of dimension at most $n-2$. In particular, if the initial hypersurface is mean convex, then all singularities are generic and the results apply. In fact, in [CM5] we showed that the entire stratification of the space-time singular set is rectifiable in a very strong sense; cf., e.g., [Si3], [Si4], [Si5], [BrCoL] and [HrLi].

One of the significant difficulties that we overcome in this paper, and sets it apart from all other work we know of, is that our singularities are noncompact. This causes major analytical difficulties and to address them requires entirely new techniques and ideas. This is not so much because of the subtleties of analysis on noncompact domains, though this is an issue, but crucially because the evolving hypersurface cannot be written as an entire graph over the singularity no matter how close we get to the singularity. Rather, the geometry of the situation dictates that only part of the evolving hypersurface can be written as a graph over a compact piece of the singularity.[7]

0.2. **Lojasiewicz inequalities.** The main technical tools that we prove are two Lojasiewicz–type inequalities.

[6]Our results not only give uniqueness of tangent flows but also a definite rate where the rescaled MCF converges to the relevant cylinder. The distance to the cylinder is decaying to zero at a definite rate over balls whose radii are increasing at a definite rate to infinity.

[7]In the end, what comes out of our analysis is that the domain the evolving hypersurface is a graph over is expanding in time and at a definite rate, but this is not all all clear from the outset; see also footnote 3.

In real algebraic geometry, the Lojasiewicz inequality, [L], named after Stanislaw Lojasiewicz, gives an upper bound for the distance from a point to the nearest zero of a given real analytic function. Specifically, let $f : U \to \mathbf{R}$ be a real-analytic function on an open set U in $\mathbf{R}^n$, and let Z be the zero locus of f. Assume that Z is not empty. Then for any compact set K in U, there exist $\alpha \geq 2$ and a positive constant C such that, for all $x \in K$

$$\inf_{z \in Z} |x - z|^\alpha \leq C\, |f(x)| \, . \tag{0.3}$$

Here α can be large.

Lojasiewicz, [L], also proved the following inequality[8]: With the same assumptions on f, for every $p \in U$, there is a possibly smaller neighborhood W of p and constants $\beta \in (0,1)$ and $C > 0$ such that for all $x \in W$

$$|f(x) - f(p)|^\beta \leq C\, |\nabla_x f| \, . \tag{0.4}$$

Note that this inequality is trivial unless p is a critical point for f.

An immediate consequence of (0.4) is that every critical point of f has a neighborhood where every other critical point has the same value.[9]

0.3. **Lojasiewicz inequalities for non-compact hypersurfaces and MCF.** The infinite dimensional Lojasiewicz-type inequalities that we prove are for the F functional on the space of hypersurfaces.

The F-functional is given by integrating the Gaussian over a hypersurface $\Sigma \subset \mathbf{R}^{n+1}$. This is also often referred to as the Gaussian surface area and is defined by

$$F(\Sigma) = (4\pi)^{-n/2} \int_\Sigma \mathrm{e}^{-\frac{|x|^2}{4}} \, d\mu \, . \tag{0.5}$$

The entropy $\lambda(\Sigma)$ is the supremum of the Gaussian surface areas over all centers and scales.

It follows from the first variation formula that the gradient of F is

$$\nabla_\Sigma F(\psi) = \int_\Sigma \left(H - \frac{\langle x, \mathbf{n} \rangle}{2} \right) \psi \, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{0.6}$$

Thus, the critical points of F are shrinkers, i.e., hypersurfaces with $H = \frac{\langle x, \mathbf{n}\rangle}{2}$. The most important shrinkers are the generalized cylinders $\mathcal{C}$; these are the generic ones by [CM1]. The space $\mathcal{C}$ is the union of $\mathcal{C}_k$ for $k \geq 1$, where $\mathcal{C}_k$ is the space of cylinders $\mathbf{S}^k \times \mathbf{R}^{n-k}$, where the $\mathbf{S}^k$ is centered at 0 and has radius $\sqrt{2k}$ and we allow all possible rotations by $SO(n+1)$.

A family of hypersurfaces Σ_s evolves by the negative gradient flow for the F-functional if it satisfies the equation

$$(\partial_s x)^\perp = -H\, \mathbf{n} + x^\perp / 2 \, . \tag{0.7}$$

This flow is called the rescaled MCF since Σ_s is obtained from a MCF M_t by setting $\Sigma_s = \frac{1}{\sqrt{-t}} M_t$, $s = -\log(-t)$, $t < 0$. By (0.6), critical points for the F-functional or, equivalently, stationary points for the rescaled MCF, are the shrinkers for the MCF that become extinct at the origin in space-time. A rescaled MCF has a unique asymptotic limit if and only if the corresponding MCF has a unique tangent flow at that singularity.

[8] Lojasiewicz called this inequality the gradient inequality.

[9] This consequence of (0.4) for the F functional near a cylinder is implied by the rigidity result of [CIM].

We will prove versions of the two Lojasiewicz inequalities for the F functional on a general hypersurface Σ. Roughly speaking, we will show that

$$\text{dist}(\Sigma,\mathcal{C})^2 \leq C\,|\nabla_\Sigma F|\,, \tag{0.8}$$

$$(F(\Sigma)-F(\mathcal{C}))^{\frac{2}{3}} \leq C\,|\nabla_\Sigma F|\,. \tag{0.9}$$

Equation (0.8) will correspond to Lojasiewicz's first inequality whereas (0.9) will correspond to his second inequality. The precise statements of these inequalities will be much more complicated than this, but they will be of the same flavor.

0.4. **First Lojasiewicz with $\alpha = 2$ implies the second with $\beta = \frac{2}{3}$.** In this subsection we will explain how the second Lojasiewicz inequality for a function f in a neighborhood of an isolated critical point follows from the first when the first holds for ∇f and with $\alpha = 2$. (We will later extend this argument to infinite dimensions.)

Suppose that $f : \mathbf{R}^n \to \mathbf{R}$ is smooth function with $f(0) = 0$ and $\nabla f(0) = 0$; without loss of generality we may assume that at 0 the Hessian is in diagonal form and we will write the coordinates as $x = (y, z)$ where y are the coordinates where the Hessian is nondegenerate. By Taylor's formula in a small neighborhood of 0, we have that

$$f(x) = \frac{a_i}{2}\,y_i^2 + O(|x|^3)\,. \tag{0.10}$$

$$f_{y_i}(x) = a_i\,y_i + O(|x|^2)\,. \tag{0.11}$$

$$f_{z_i}(x) = O(|x|^2)\,. \tag{0.12}$$

It follows from this that the second of the two Lojasiewicz inequalities holds for f and $\beta = \frac{2}{3}$ provided that $|z|^2 \leq \epsilon\,|y|$ for some sufficiently small $\epsilon > 0$. Namely, if $|z|^2 \leq \epsilon\,|y|$, then

$$C\,|y| \leq |\nabla_x f| \text{ and } |f(x)| \leq C^{-1}\,|y|^{\frac{3}{2}} \tag{0.13}$$

for some positive constant C and, hence,

$$|f(x)|^{\frac{2}{3}} \leq C\,|\nabla_x f|\,. \tag{0.14}$$

Therefore, we only need to prove the second Lojasiewicz inequality for f in the region $|z|^2 \geq \epsilon\,|y|$. We will do this using the first Lojasiewicz inequality for ∇f. Since 0 is an isolated critical point for f, the first Lojasiewicz inequality for ∇f gives that

$$|\nabla_x f| \geq C\,|x|^2\,. \tag{0.15}$$

By assumption on the region and the Taylor expansion for f, we get that in this region

$$|f(x)| \leq C\,|y|^2 + C\,|z|^3 \leq C\,|z|^3 \leq C\,|x|^3\,. \tag{0.16}$$

Combining these two inequalities gives

$$|f(x)|^{\frac{2}{3}} \leq C\,|x|^2 \leq |\nabla_x f|\,. \tag{0.17}$$

This proves the second Lojasiewicz inequality for f with $\beta = \frac{2}{3}$.

In Section 4, we extend the above argument to general Banach spaces.

Lojasiewicz used his second inequality to show the "Lojasiewicz theorem": If $f : \mathbf{R}^n \to \mathbf{R}$ is an analytic function, $x = x(t) : [0,\infty) \to \mathbf{R}^n$ is a curve with $x'(t) = -\nabla f$ and $x(t)$ has a limit point x_∞, then the length of the curve is finite and $\lim_{t\to\infty} x(t) = x_\infty$. Moreover, x_∞ is a critical point for f.

Even in $\mathbf{R}^2$, it is easy to construct smooth functions where the Lojasiewicz theorem does not hold, but instead there are negative gradient flow lines with multiple limits.

We will discuss the Lojasiewicz theorem in a slightly more general setting at the end of the next subsection after briefly discussing infinite dimensional Lojasiewicz inequalities.

0.5. **Infinite dimensional Lojasiewicz inequalities and applications.** Infinite dimensional versions of Lojasiewicz inequalities were proven in a celebrated work of Leon Simon, [Si1], for the area and related functionals and used, in particular, to prove a fundamental result about uniqueness of tangent cones with smooth cross section of minimal surfaces. Simon's proof of the Lojasiewicz inequality is done by reducing the infinite dimensional version to the classical Lojasiewicz inequality by a Lyapunov-Schmidt reduction argument. Infinite dimensional Lojasiewicz inequalities proven using Lyapunov-Schmidt reduction, as in the work of Simon, have had a profound impact on various areas of analysis and geometry and are usually referred to as Lojasiewicz-Simon inequalities.

As already mentioned, we will also prove two infinite dimensional Lojasiewicz inequalities and use them to prove uniqueness of blowups for MCF (or, equivalently, convergence of the rescaled flow). However, unlike all other infinite dimensional Lojasiewicz inequalities we know of, ours do not follow from a reduction to the classical Lojasiewicz inequalities; rather we prove our inequalities directly and do not rely on Lojasiewicz's arguments or results. In fact, we prove our infinite dimensional analog of the first Lojasiewicz inequality directly and use this together with an infinite dimensional analog of the argument in the previous subsection to show our second Lojasiewicz inequality. The reason why we cannot argue as in Simon's work, and all the later work that make use his ideas, comes from that our singularities are noncompact. In particular, even near the singularities, the evolving hypersurface cannot be written as an entire graph over the singularity. Rather, only part of the evolving hypersurface can be written as a graph over a compact piece of the singularity.

Next we will explain how the second Lojasiewicz inequality is typically used to show uniqueness. Before we do that, observe first that in the second inequality we always work in a small neighborhood of p so that, in particular, $|f(x) - f(p)| \leq 1$ and hence smaller powers on the left hand side of the inequality imply the inequality for higher powers. As it turns out, we will see that any positive power strictly less than 1 would do for uniqueness.

Suppose now that X is a Banach space and $f : X \to \mathbf{R}$ is a Frechet differentiable function. Let $x = x(t)$ be a curve on X parametrized on $[0, \infty)$ whose velocity $x' = -\nabla f$. We would like to show that if the second inequality of Lojasiewicz holds for f with a power $1 > \beta > 1/2$, then the Lojasiewicz theorem mentioned above holds. That is, if $x(t)$ has a limit point x_∞, then the length of the curve is finite and $\lim_{t\to\infty} x(t) = x_\infty$. Since x_∞ is a limit point of $x(t)$ and f is non-increasing along the curve, x_∞ must be a critical point for f.

To see that $x(t)$ converges to x_∞, assume that $f(x_\infty) = 0$ and note that if we set $f(t) = f(x(t))$, then $f' = -|\nabla f|^2$. Moreover, by the second Lojasiewicz inequality, we get that $f' \leq -f^{2\beta}$ if $x(t)$ is sufficiently close to x_∞. (Assume for simplicity below that $x(t)$ stays in a small neighborhood x_∞ for t sufficiently large so that this inequality holds; the general case follows with trivial changes.) Then this inequality can be rewritten as $(f^{1-2\beta})' \geq (2\beta - 1)$

which integrates to

$$f(t) \leq C\, t^{\frac{-1}{2\beta - 1}} \, . \tag{0.18}$$

We need to show that (0.18) implies that $\int_0^\infty |\nabla f|\, ds$ is finite. This shows that $x(t)$ converges to x_∞ as $t \to \infty$. To see that $\int_0^\infty |\nabla f|\, ds$ is finite, observe by the Cauchy-Schwarz inequality that

$$\int_0^\infty |\nabla f|\, ds = \int_0^\infty \sqrt{-f'}\, ds \leq \left(- \int_0^\infty f'\, s^{1+\epsilon}\, ds \right)^{\frac{1}{2}} \left(\int_0^\infty s^{-1-\epsilon}\, ds \right)^{\frac{1}{2}} \, . \tag{0.19}$$

It suffices therefore to show that

$$- \int_0^T f'\, s^{1+\epsilon}\, ds \tag{0.20}$$

is uniformly bounded. Integrating by parts gives

$$\int_0^T f'\, s^{1+\epsilon}\, ds = |f\, s^{1+\epsilon}|_0^T - (1+\epsilon) \int_0^T f\, s^{\epsilon}\, ds \, . \tag{0.21}$$

If we choose $\epsilon > 0$ sufficiently small depending on β, then we see that this is bounded independent of T and hence $\int_0^\infty |\nabla f|\, ds$ is finite.

We will use an extension of this argument where the assumption $f^{2\beta}(t) \leq -f'(t)$ is replaced by the assumption that $f^{2\beta}(t) \leq f(t-1) - f(t+1)$; see Lemma 6.22. This assumption is exactly what comes out of our analog for the rescaled MCF of the gradient Lojasiewicz inequality, i.e., out of Theorem 0.26.

0.6. **The two Lojasiewicz inequalities.** We will now state the two Lojasiewicz-type inequalities for the F functional on the space of hypersurfaces.

Suppose that $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface and fix some small $\epsilon_0 > 0$ (this will be chosen small enough to satisfy Lemmas 2.5 and 4.3). Given an integer ℓ and constant C_ℓ, we let $\mathbf{r}_\ell(\Sigma)$ be the maximal radius so that

- $B_{\mathbf{r}_\ell(\Sigma)} \cap \Sigma$ is the graph over a cylinder in $\mathcal{C}_k$ of a function u with $\|u\|_{C^{2,\alpha}} \leq \epsilon_0$ and $|\nabla^\ell A| \leq C_\ell$.

The parameters ℓ and C_ℓ will be left free until the proof of the main theorem (Theorem 0.2) and will then be chosen large.

In the next theorem, we will use a Gaussian L^2 distance $d_{\mathcal{C}}(R)$ to the space $\mathcal{C}_k$ in the ball of radius R. To define this, given $\Sigma_k \in \mathcal{C}_k$, let $w_{\Sigma_k} : \mathbf{R}^{n+1} \to \mathbf{R}$ denote the distance to the axis of Σ_k (i.e., to the space of translations that leave Σ_k invariant). Then we define

$$d_{\mathcal{C}}^2(R) = \inf_{\Sigma_k \in \mathcal{C}_k} \|w_{\Sigma_k} - \sqrt{2k}\|_{L^2(B_R)}^2 \equiv \inf_{\Sigma_k \in \mathcal{C}_k} \int_{B_R \cap \Sigma_k} (w_{\Sigma_k} - \sqrt{2k})^2\, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{0.22}$$

The Gaussian L^p norm on the ball B_R is $\|u\|_{L^p(B_R)}^p = \int_{B_R} |u|^p\, \mathrm{e}^{-\frac{|x|^2}{4}}$.

Given a general hypersurface Σ, it is also convenient to define the function ϕ by

$$\phi = \frac{\langle x, \mathbf{n} \rangle}{2} - H \, , \tag{0.23}$$

so that ϕ is minus the gradient of the functional F.

The main tools that we develop here are the following two analogs for non-compact hypersurfaces of the well-known Lojasiewicz's inequalities for analytic functions on $\mathbf{R}^n$.

Theorem 0.24. (A Lojasiewicz inequality for non-compact hypersurfaces). If $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface with $\lambda(\Sigma) \leq \lambda_0$ and $R \in [1, \mathbf{r}_\ell(\Sigma) - 1]$, then

$$d^2_{\mathcal{C}}(R) \leq C\, R^\rho \left\{ \|\phi\|^{b_{\ell,n}}_{L^1(B_R)} + \mathrm{e}^{-\frac{b_{\ell,n}\, R^2}{4}} \right\} , \tag{0.25}$$

where $C = C(n, \ell, C_\ell, \lambda_0)$, $\rho = \rho(n)$ and $b_{\ell,n} \in (0,1)$ satisfies $\lim_{\ell\to\infty} b_{\ell,n} = 1$.

The theorem bounds the L^2 distance to $\mathcal{C}_k$ by a power of $\|\phi\|_{L^1}$, with an error term that comes from a cutoff argument since Σ is non-compact and is not globally a graph of the cylinder.[10] This theorem is essentially sharp. Namely, the estimate (0.25) does not hold for any exponent $b_{\ell,n}$ larger than one, but Theorem 0.24 lets us take $b_{\ell,n}$ arbitrarily close to one.

We will also see that the above inequality implies the following gradient type Lojasiewicz inequality. This inequality bounds the difference of the F functional near a critical point by two terms. The first is essentially a power of ∇F, while the second (exponentially decaying) term comes from that Σ is not a graph over the entire cylinder.

Theorem 0.26. (A gradient Lojasiewicz inequality for non-compact hypersurfaces). If $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface with $\lambda(\Sigma) \leq \lambda_0$, $\beta \in [0,1)$, and $R \in [1, \mathbf{r}_\ell(\Sigma) - 1]$, then

$$|F(\Sigma) - F(\mathcal{C}_k)| \leq C\, R^\rho \left\{ \|\phi\|^{c_{\ell,n}\frac{3+\beta}{2+2\beta}}_{L^2(B_R)} + \mathrm{e}^{-\frac{c_{\ell,n}\,(3+\beta)R^2}{8(1+\beta)}} + \mathrm{e}^{-\frac{(3+\beta)(R-1)^2}{16}} \right\} , \tag{0.27}$$

where $C = C(n, \ell, C_\ell, \lambda_0)$, $\rho = \rho(n)$ and $c_{\ell,n} \in (0,1)$ satisfies $\lim_{\ell\to\infty} c_{\ell,n} = 1$.

When we apply the theorem, the parameters β and ℓ will be chosen to make the exponent greater than one on the ∇F term, essentially giving that $|F(\Sigma) - F(\mathcal{C}_k)|$ is bounded by a power greater than one of $|\nabla F|$. A separate argument will be needed to handle the exponentially decaying error terms.

Throughout the paper C will denote a constant that can change from line to line.

We will show that when Σ_t are flowing by the rescaled MCF, then both terms on the right-hand side of (0.27) are bounded by a power greater than one of $\|\phi\|_{L^2}$ (the corresponding statement holds for Theorem 0.24). Thus, we will essentially get the inequalities

$$d^2_{\mathcal{C}} \leq C\, |\nabla_{\Sigma_t} F| , \tag{0.28}$$

$$(F(\Sigma_t) - F(\mathcal{C}))^{\frac{2}{3}} \leq C\, |\nabla_{\Sigma_t} F| . \tag{0.29}$$

These two inequalities can be thought of as analogs for the rescaled MCF of Lojasiewicz inequalities from real algebraic geometry; cf. (0.8) and (0.9).

See [CM6] for a survey on Lojasiewicz inequalities and their applications.

[10] This is a Lojasiewicz inequality for the gradient of the F functional (ϕ is the gradient of F). This follows since, by [CIM], cylinders are isolated critical points for F and, thus, $d_{\mathcal{C}}$ locally measures the distance to the nearest critical point.

1. Cylindrical estimates for a general hypersurface

In this section, we will prove estimates for a general hypersurface $\Sigma \subset \mathbf{R}^{n+1}$. The main results are bounds for $\nabla \frac{A}{H}$ when the mean curvature H is positive on a large set.

1.1. **A general Simons equation.** In this subsection, we will show that the second fundamental form A of Σ satisfies an elliptic differential equation similar to Simons' equation for minimal surfaces. The elliptic operator will be the L operator from [CM1] given by

$$L \equiv \mathcal{L} + |A|^2 + \frac{1}{2} \equiv \Delta - \frac{1}{2} \nabla_{x^T} + |A|^2 + \frac{1}{2} \, . \tag{1.1}$$

Namely, we will prove the following proposition:

Proposition 1.2. If $\phi = \frac{1}{2}\langle x, \mathbf{n} \rangle - H$, then

$$L\, A = A + \text{Hess}_{\phi} + \phi\, A^2 \, , \tag{1.3}$$

where the tensor A^2 is given in orthonormal frame by $(A^2)_{ij} = A_{ik}\, A_{kj}$.

Note that ϕ vanishes precisely when Σ is a shrinker and, in this case, we recover the Simons' equation for A for shrinkers from [CM1].

We will use the following general version of Simons' equation for the second fundamental form of a hypersurface:

Lemma 1.4. The second fundamental form A satisfies

$$\left(\Delta + |A|^2 \right) A = -H\, A^2 - \text{Hess}_H \, . \tag{1.5}$$

See, e.g., [CM3] for a proof.

The next lemma computes the Hessian of the support function $\langle x, \mathbf{n} \rangle$.

Lemma 1.6. The Hessian of $\langle x, \mathbf{n} \rangle$ is given by

$$\text{Hess}_{\langle x, \mathbf{n} \rangle} = -\nabla_{x^T} A - A - A^2 \, \langle x, \mathbf{n} \rangle \, . \tag{1.7}$$

Proof. Fix a point $p \in \Sigma$. Let e_i be a local orthonormal frame for Σ with $\nabla^T_{e_i} e_j = 0$ at p for every i and j. Thus, at p, we have

$$\nabla_{e_i} e_j = A_{ij}\, \mathbf{n} \, . \tag{1.8}$$

Finally, using this and $\nabla_{e_i} \mathbf{n} = -A_{ik}\, e_k$ (which holds at all points), we compute at p

$$\begin{aligned} \text{Hess}_{\langle x, \mathbf{n} \rangle}(e_i, e_j) &= \langle x, \mathbf{n} \rangle_{ij} = \langle x, \nabla_{e_i} \mathbf{n} \rangle_j = - \left(A_{ik} \, \langle x, e_k \rangle \right)_j \\ &= -A_{ikj} \, \langle x, e_k \rangle - A_{ik}\, \delta_{jk} - A_{ik} \langle x, A_{jk}\, \mathbf{n} \rangle \\ &= - \left(\nabla_{x^T} A \right) (e_i, e_j) - A(e_i, e_j) - \langle x, \mathbf{n} \rangle A^2(e_i, e_j) \, , \end{aligned} \tag{1.9}$$

where the last equality used the Codazzi equation $A_{ikj} = A_{ijk}$. □

Proof of Proposition 1.2. Since $L = \mathcal{L} + |A|^2 + \frac{1}{2}$ and $\mathcal{L} = \Delta - \frac{1}{2} \nabla_{x^T}$, Lemma 1.4 gives

$$LA = \left(\Delta + |A|^2 \right) A + \frac{1}{2} A - \frac{1}{2} \nabla_{x^T} A = -H\, A^2 - \text{Hess}_H + \frac{1}{2} A - \frac{1}{2} \nabla_{x^T} A \, . \tag{1.10}$$

On the other hand, Lemma 1.6 gives

$$\text{Hess}_\phi = \frac{1}{2}\text{Hess}_{\langle x,\mathbf{n}\rangle} - \text{Hess}_H = -\text{Hess}_H - \frac{1}{2}\nabla_{x^T} A - \frac{1}{2}A - \frac{1}{2}A^2 \langle x, \mathbf{n}\rangle \, , \tag{1.11}$$

so we have $L\,A - \text{Hess}_\phi = A + \phi\, A^2$. □

1.2. **An integral bound when the mean curvature is positive.** We will show that the tensor $\tau = A/H$ is almost parallel when H is positive and ϕ is small. This generalizes an estimate from [CIM] in the case where Σ is a shrinker (i.e., $\phi \equiv 0$) with $H > 0$.

Given $f > 0$, define a weighted divergence operator div_f and drift Laplacian $\mathcal{L}_f$ by

$$\text{div}_f(V) = \frac{1}{f}\, \mathrm{e}^{|x|^2/4}\, \text{div}_\Sigma \, \left(f\, \mathrm{e}^{-|x|^2/4}\, V\right) \, , \tag{1.12}$$

$$\mathcal{L}_f\, u \equiv \text{div}_f(\nabla u) = \mathcal{L}\, u + \langle \nabla \log f, \nabla u\rangle \, . \tag{1.13}$$

Here u may also be a tensor; in this case the divergence traces only with ∇. Note that $\mathcal{L} = \mathcal{L}_1$. We recall the quotient rule (see lemma 4.3 in [CIM]):

Lemma 1.14. Given a tensor τ and a function g with $g \neq 0$, then

$$\mathcal{L}_{g^2}\, \frac{\tau}{g} = \frac{g\, \mathcal{L}\, \tau - \tau\, \mathcal{L}\, g}{g^2} = \frac{g\, L\, \tau - \tau\, L\, g}{g^2} \, . \tag{1.15}$$

Proposition 1.16. On the set where $H > 0$, we have

$$\mathcal{L}_{H^2}\, \frac{A}{H} = \frac{\text{Hess}_\phi + \phi\, A^2}{H} + \frac{A\, (\Delta \phi + \phi\, |A|^2)}{H^2} \, , \tag{1.17}$$

$$\mathcal{L}_{H^2}\, \frac{|A|^2}{H^2} = 2\, \left| \nabla \frac{A}{H}\right|^2 + 2\, \frac{\langle \text{Hess}_\phi + \phi\, A^2 , A\rangle}{H^2} + 2\, \frac{|A|^2\, (\Delta \phi + \phi\, |A|^2)}{H^3} \, . \tag{1.18}$$

Proof. The trace of Proposition 1.2 (H is minus the trace of A by convention) gives

$$L\, H = H - \Delta \phi - \phi\, |A|^2 \, , \tag{1.19}$$

where we also used that the trace of A^2 is $|A|^2$ since A is symmetric. Using the quotient rule (Lemma 1.14) and the equations for LH and LA (from Proposition 1.2) gives

$$\begin{aligned} \mathcal{L}_{H^2}\, \frac{A}{H} &= \frac{H\, LA - A\, LH}{H^2} = \frac{H\, (A + \text{Hess}_\phi + \phi\, A^2) - A\, (H - \Delta\phi - \phi\, |A|^2)}{H^2} \\ &= \frac{\text{Hess}_\phi + \phi\, A^2}{H} + \frac{A\, (\Delta \phi + \phi\, |A|^2)}{H^2} \, , \end{aligned} \tag{1.20}$$

giving the first claim. The second claim follows from the first since $\frac{|A|^2}{H^2} = \langle \frac{A}{H}, \frac{A}{H}\rangle$ and

$$\frac{1}{2}\, \mathcal{L}_{H^2} \langle \frac{A}{H}, \frac{A}{H}\rangle = \langle \mathcal{L}_{H^2}\, \frac{A}{H}, \frac{A}{H}\rangle + \left| \nabla \frac{A}{H}\right|^2 \, . \tag{1.21}$$

□

The next proposition gives exponentially decaying integral bounds for $\nabla(A/H)$ when H is positive on a large ball. It will be important that these bounds decay rapidly.

Proposition 1.22. If $B_R \cap \Sigma$ is smooth with $H > 0$, then for $s \in (0, R)$ we have

$$\int_{B_{R-s}\cap\Sigma} \left|\nabla \frac{A}{H}\right|^2 H^2 \,\mathrm{e}^{-\frac{|x|^2}{4}} \leq \frac{4}{s^2} \sup_{B_R\cap\Sigma} |A|^2 \operatorname{Vol}(B_R\cap\Sigma)\,\mathrm{e}^{-\frac{(R-s)^2}{4}}$$

$$+ 2\int_{B_R\cap\Sigma} \left\{ \left| \langle \mathrm{Hess}_\phi, A\rangle + \frac{|A|^2}{H}\Delta\phi\right| + \left|\langle A^2, A\rangle + \frac{|A|^4}{H}\right| |\phi| \right\} \mathrm{e}^{-\frac{|x|^2}{4}} . \tag{1.23}$$

Proof. Set $\tau = A/H$ and $u = |\tau|^2 = |A|^2/H^2$. It will be convenient within this proof to use square brackets $[\cdot]$ to denote Gaussian integrals over $B_R \cap \Sigma$, i.e. $[f] = \int_{B_R\cap\Sigma} f\,\mathrm{e}^{-|x|^2/4}$.

Let ψ be a function with support in B_R. Using the divergence theorem, the formula from Proposition 1.16 for $\mathcal{L}_{H^2}u$, and the absorbing inequality $4ab \leq a^2 + 4b^2$, we get

$$\begin{aligned}
0 &= \left[\operatorname{div}_{H^2}\left(\psi^2\,\nabla u\right) H^2\right] = \left[\left(\psi^2\,\mathcal{L}_{H^2}u + 2\psi\langle\nabla\psi,\nabla u\rangle\right) H^2\right] \\
&= \left[\left\{ 2\,\psi^2\,|\nabla\tau|^2 + 2\,\psi^2\left(\frac{\langle \mathrm{Hess}_\phi + \phi\,A^2, A\rangle}{H^2} + \frac{|A|^2\,(\Delta\phi+\phi\,|A|^2)}{H^3}\right) + 4\psi\langle\nabla\psi, \tau\cdot\nabla\tau\rangle\right\} H^2\right] \\
&\geq \left[\left(\psi^2\,|\nabla\tau|^2 - 4\,|\tau|^2\,|\nabla\psi|^2\right) H^2\right] + 2\,\left[\psi^2\langle \mathrm{Hess}_\phi + \phi\,A^2, A\rangle\right] + 2\,\left[\psi^2\frac{|A|^2\,(\Delta\phi + \phi\,|A|^2)}{H}\right],
\end{aligned}$$

from which we obtain

$$\left[\psi^2\,|\nabla\tau|^2\,H^2\right] \leq 4\,\left[|\nabla\psi|^2\,|A|^2\right] - 2\,\left[\psi^2\,\langle \mathrm{Hess}_\phi + \phi A^2, A\rangle\right] - 2\,\left[\psi^2\,\Delta\phi\,\frac{|A|^2}{H} + \psi^2\,\phi\frac{|A|^4}{H}\right] .$$

The proposition follows by choosing $\psi \equiv 1$ on B_{R-s} and going to zero linearly on ∂B_R. $\square$

We record the following corollary:

Corollary 1.24. If $B_R \cap \Sigma$ is smooth with $H > \delta > 0$ and $|A| \leq C_1$, then there exists $C_2 = C_2(n, \delta, C_1)$ so that for $s \in (0, R)$ we have

$$\int_{B_{R-s}\cap\Sigma} \left|\nabla\frac{A}{H}\right|^2 \mathrm{e}^{-\frac{|x|^2}{4}} \leq \frac{C_2}{s^2} \operatorname{Vol}(B_R\cap\Sigma)\,\mathrm{e}^{-\frac{(R-s)^2}{4}} + C_2\int_{B_R\cap\Sigma}\{|\mathrm{Hess}_\phi| + |\phi|\}\,\mathrm{e}^{-\frac{|x|^2}{4}} . \tag{1.25}$$

Remark 1.26. Corollary 1.24 essentially bounds the distance *squared* to the space of cylinders by $\|\phi\|_{L^1}$. This is sharp: it is not possible to get the sharper bound where the powers are the same. This is a general fact when there is a non-integrable kernel. Namely, if we perturb in the direction of the kernel, then ϕ vanishes quadratically in the distance.

The next corollary combines the Gaussian L^2 bound on $\nabla\tau$ from Corollary 1.24 with standard interpolation inequalities to get pointwise bounds on $\nabla\tau$ and $\nabla^2\tau$.

Corollary 1.27. If $B_R \cap \Sigma$ is smooth with $H > \delta > 0$, $|A| + \left|\nabla^{\ell+1}A\right| \leq C_1$, and $\lambda(\Sigma) \leq \lambda_0$, then there exists $C_3 = C_3(n, \lambda_0, \delta, \ell, C_1)$ so that for $|y| + \frac{1}{1+|y|} < R - 1$, we have

$$\left|\nabla\frac{A}{H}\right|(y) + \left|\nabla^2\frac{A}{H}\right|(y) \leq C_3\,R^{2n}\left\{\mathrm{e}^{-d_{\ell,n}\frac{(R-1)^2}{8}} + \|\phi\|_{L^1(B_R)}^{\frac{d_{\ell,n}}{2}}\right\}\mathrm{e}^{\frac{|y|^2}{8}} , \tag{1.28}$$

where the exponent $d_{\ell,n} \in (0, 1)$ has $\lim_{\ell\to\infty} d_{\ell,n} = 1$.

Proof. Set $\tau = A/H$ and note that $\left|\nabla^{\ell+1}\tau\right|$ is bounded by a constant depending on δ, ℓ and C_1. Define the ball B^y and constant δ_y by

$$B^y = B_{\frac{1}{1+|y|}}(y) \text{ and } \delta_y = \int_{B^y \cap \Sigma} |\nabla \tau| \, . \tag{1.29}$$

Applying Lemma B.1 on B^y gives

$$|\nabla \tau|(y) \leq C' \left\{ R^n \, \delta_y + \delta_y^{a_{\ell,n}} \, \|\nabla^{\ell+1}\tau\|_{L^\infty(B^y)}^{1-a_{\ell,n}} \right\} \leq C \, \left\{ R^n \, \delta_y + \delta_y^{a_{\ell,n}} \right\} \, ,$$

$$|\nabla^2 \tau|(y) \leq C' \left\{ R^{n+1} \, \delta_y + \delta_y^{b_{\ell,n}} \, \|\nabla^{\ell+1}\tau\|_{L^\infty(B^y)}^{1-b_{\ell,n}} \right\} \leq C \, \left\{ R^{n+1} \, \delta_y + \delta_y^{b_{\ell,n}} \right\} \, ,$$

where the powers are given by $a_{\ell,n} = \frac{2\ell}{2\ell+n}$ and $b_{\ell,n} = \frac{2\ell-2}{2\ell+n}$, and $C = C(n, \delta, \ell, C_1)$.

To get the bound on δ_y, observe that

$$\inf_{B^y} \mathrm{e}^{-\frac{|x|^2}{4}} \geq \mathrm{e}^{-\frac{|y|^2}{4} - 1} \, , \tag{1.30}$$

so that Cauchy-Schwarz gives

$$(1+|y|)^n \, \mathrm{e}^{\frac{-|y|^2}{4} - 1} \, \delta_y^2 \leq C \, \mathrm{e}^{\frac{-|y|^2}{4} - 1} \int_{B^y \cap \Sigma} |\nabla \tau|^2 \leq C \int_{B^y \cap \Sigma} |\nabla \tau|^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq C_2 \, \gamma \, , \tag{1.31}$$

where the last inequality is Corollary 1.24, $C_2 = C_2(n, \lambda_0, \delta, C_1)$ and γ is

$$\gamma = R^n \, \mathrm{e}^{-\frac{(R-1)^2}{4}} + \int_{B_{R-1/2} \cap \Sigma} \left\{ |\mathrm{Hess}_\phi| + |\phi| \right\} \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{1.32}$$

To bound the Hessian term, first choose balls $B^i = B_{\frac{1}{1+|z_i|}}(z_i)$ so that

- $B_{R-1/2} \cap \Sigma$ is contained in the union of the half-balls $\frac{1}{2} \, B^i$.
- Each point is in at most $c = c(n)$ of the balls.

To simplify notation, set $r_i = \frac{1}{1+|z_i|}$. Applying Lemma B.1 on B^i gives

$$\sup_{\frac{1}{2} B^i} |\mathrm{Hess}_\phi| \leq C \left\{ r_i^{-n-2} \int_{B^i} |\phi| + \left(\int_{B^i} |\phi| \right)^{c_{\ell,n}} \right\} \, , \tag{1.33}$$

where $c_{\ell,n} \in (0,1)$ goes to one as $\ell \to \infty$. Note that the Gaussian weight has bounded oscillation on B^i (this is why the radius r_i was chosen). It follows that

$$\begin{aligned} \int_{B_{R-1/2} \cap \Sigma} |\mathrm{Hess}_\phi| \, \mathrm{e}^{-\frac{|x|^2}{4}} &\leq C \sum \left\{ r_i^{-2} \int_{B^i} |\phi| + r_i^n \left(\int_{B^i} |\phi| \right)^{c_{\ell,n}} \right\} \mathrm{e}^{-\frac{|z_i|^2}{4}} \\ &\leq C \, R^2 \, \|\phi\|_{L^1(B_R)} + C \sum \left(\int_{B^i} |\phi| \right)^{c_{\ell,n}} \mathrm{e}^{-\frac{|z_i|^2}{4}} \\ &\leq C \, R^2 \, \|\phi\|_{L^1(B_R)} + C \, \|\phi\|_{L^1(B_R)}^{c_{\ell,n}} \, , \end{aligned} \tag{1.34}$$

where the last inequality uses the Hölder inequality for sums and the bound for $F(\Sigma)$. Since $\|\phi\|_{L^1}$ is bounded (we are interested in the case where it is much less than one), the lower power is dominant and we conclude that

$$\mathrm{e}^{\frac{-|y|^2}{4} - 1} \, \delta_y^2 \leq C_2 \, \gamma \leq C \, R^n \, \mathrm{e}^{-\frac{(R-1)^2}{4}} + C \, R^2 \, \|\phi\|_{L^1(B_R)}^{c_{\ell,n}} \, . \tag{1.35}$$

Arguing similarly and using this in the bounds for $\nabla\tau$ gives

$$|\nabla \tau|(y) \leq C\, R^n\, \delta_y^{a_{\ell,n}} \leq C\, R^{\frac{3n}{2}} \left\{ \mathrm{e}^{\frac{|y|^2-(R-1)^2}{8}} + \mathrm{e}^{\frac{|y|^2}{8}}\, \|\phi\|_{L^1(B_R)}^{\frac{c_{\ell,n}}{2}} \right\}^{a_{\ell,n}} , \tag{1.36}$$

$$\left|\nabla^2 \tau\right|(y) \leq C\, R^{n+1}\, \delta_y^{b_{\ell,n}} \leq C\, R^{\frac{3n+2}{2}} \left\{ \mathrm{e}^{\frac{|y|^2-(R-1)^2}{8}} + \mathrm{e}^{\frac{|y|^2}{8}}\, \|\phi\|_{L^1(B_R)}^{\frac{c_{\ell,n}}{2}} \right\}^{b_{\ell,n}} . \tag{1.37}$$

□

2. Distance to cylinders and the first Lojasiewicz inequality

In this section, we will prove the first Lojasiewicz inequality that bounds the distance squared to the space $\mathcal{C}_k$ of all rotations of the cylinder $\mathbf{S}^k_{\sqrt{2k}} \times \mathbf{R}^{n-k}$ by a power close to one of the gradient of the F functional. This will follow from the bounds on the tensor $\tau = \frac{A}{H}$ in the previous section together with the following proposition:

Proposition 2.1. Given n, $\delta > 0$ and C_1, there exist $\epsilon_0 > 0$, $\epsilon_1 > 0$ and C so that if $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface (possibly with boundary) that satisfies:

(1) $H \geq \delta > 0$ and $|A| + |\nabla A| \leq C_1$ on $B_R \cap \Sigma$.
(2) $B_{5\sqrt{2n}} \cap \Sigma$ is ϵ_0 C^2-close to a cylinder in $\mathcal{C}_k$ for some $k \geq 1$,

then, for any $r \in (5\sqrt{2n}, R)$ with

$$r^2 \sup_{B_{5\sqrt{2n}}} \left(|\phi| + |\nabla \phi|\right) + r^5 \sup_{B_r} \left(|\nabla \tau| + |\nabla^2 \tau|\right) \leq \epsilon_1 , \tag{2.2}$$

we have that $B_{\sqrt{r^2-3k}} \cap \Sigma$ is the graph over (a subset of) a cylinder in $\mathcal{C}_k$ of u with

$$|u| + |\nabla u| \leq C \left\{ r^2 \sup_{B_{5\sqrt{2n}}} \left(|\phi| + |\nabla \phi|\right) + r^5 \sup_{B_r} \left(|\nabla \tau| + |\nabla^2 \tau|\right) \right\} . \tag{2.3}$$

This proposition shows that Σ must be close to a cylinder as long as H is positive, ϕ is small, τ is almost parallel and Σ is close to a cylinder on a fixed small ball. Together with Tom Ilmanen, we proved a similar result in proposition 2.2 in [CIM] in the special case where Σ is a shrinker (i.e., when $\phi \equiv 0$) and this proposition is inspired by that one.

We will prove the proposition over the next two subsections and then turn to the proof of the first Lojasiewicz inequality.

2.1. Ingredients in the proof of Proposition 2.1.

2.1. Ingredients in the proof of Proposition 2.1. This subsection contains the ingredients for the proof of Proposition 2.1. The first is the following result from [CIM] (see corollary 4.22 in [CIM]):

Corollary 2.4 ([CIM]). If $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface (possibly with boundary) with

- $0 < \delta \leq H$ on Σ,
- the tensor $\tau \equiv A/H$ satisfies $|\nabla \tau| + |\nabla^2 \tau| \leq \epsilon \leq 1$,
- At the point $p \in \Sigma$, τ_p has at least two distinct eigenvalues $\kappa_1 \neq \kappa_2$,

then

$$|\kappa_1 \kappa_2| \leq \frac{2\,\epsilon}{\delta^2} \left(\frac{1}{|\kappa_1 - \kappa_2|} + \frac{1}{|\kappa_1 - \kappa_2|^2} \right) .$$

We will use two additional lemmas in the proof of Proposition 2.1. The next lemma shows that ϕ controls the distance to the shrinking sphere in a neighborhood of the sphere. This, of course, implies that the shrinking sphere is isolated in the space of shrinkers. The proof uses that the linearized operator is invertible.

Lemma 2.5. Given k and $\alpha > 0$, there exist $\epsilon_0 > 0$ and C so that if $\Sigma_0 \subset \mathbf{R}^{k+1}$ is the graph of a $C^{2,\alpha}$ function u over $\mathbf{S}^k_{\sqrt{2k}}$ with $\|u\|_{C^2} \leq \epsilon_0$, then

$$\|u\|_{C^{2,\alpha}} \leq C \, \|\phi\|_{C^\alpha} \, . \tag{2.6}$$

Proof. On the sphere, the linearized operator L for ϕ is given by $L = \Delta + 1$ since $|A|^2 = 1/2$ and the drift term vanishes. The eigenvalues for Δ on the sphere of radius one occur in clusters with the m-th cluster at $m^2 + (k-1)\, m$. Scaling this to the sphere of radius $\sqrt{2k}$, the m-th cluster is now at

$$\frac{m^2 + (k-1)\, m}{2k} \, , \tag{2.7}$$

and, thus, the first three eigenvalues for $L = \Delta + 1$ occur at -1, $-\frac{1}{2}$ and $\frac{1}{k}$. In particular, 0 is not an eigenvalue and, thus, L is invertible and, by the Schauder estimates, we have

$$\|u\|_{C^{2,\alpha}} \leq C \, \|L\, u\|_{C^\alpha} \, , \tag{2.8}$$

where C depends only on k and α. The lemma follows from this and the fact that the linearization of ϕ is L and the error is quadratic (cf. Lemma 4.10 below) so we have

$$\|\phi - Lu\|_{C^\alpha} \leq C \, \|u\|_{C^2} \, \|u\|_{C^{2,\alpha}} \, , \tag{2.9}$$

where C again depends only on k and α. Combining the last two inequalities gives

$$\|u\|_{C^{2,\alpha}} \leq C \, \|\phi\|_{C^\alpha} + C \, \|u\|_{C^2} \, \|u\|_{C^{2,\alpha}} \leq C \, \|\phi\|_{C^\alpha} + C \, \epsilon_0 \, \|u\|_{C^{2,\alpha}} \, , \tag{2.10}$$

which gives the claim after choosing $\epsilon_0 > 0$ so that $C\, \epsilon_0 = \frac{1}{2}$. □

The next lemma shows if Σ has an approximate translation and is almost a shrinker, then slicing Σ orthogonally to the translation gives a submanifold Σ_0 of one dimension less that is also almost a shrinker. We will use this to repeatedly slice an almost cylinder to get down to the almost sphere. We let ϕ_0 be the ϕ of Σ_0 (so $\Sigma_0 \subset \mathbf{R}^k$ is a shrinker when $\phi_0 \equiv 0$).

Lemma 2.11. Let $\Sigma \subset \mathbf{R}^{k+1}$ be a hypersurface, $\Sigma_0 = \{x_{k+1} = 0\} \cap \Sigma$, and $x \in \Sigma_0$ a point where Σ intersects the hyperplane $\{x_{k+1} = 0\}$ transversely. If we have

- $\left| \nabla^T x_{k+1} \right| \geq 1 - \epsilon > 1/2$,
- $\left| \nabla^T \, \nabla^T x_{k+1} \right| \leq \epsilon$,
- $\left| A(\cdot, \nabla^T x_{k+1}) \right| + \left| (\nabla A) \, (\cdot, \nabla^T x_{k+1}) \right| \leq \epsilon$.

Then at x

$$|\phi - \phi_0| + |\nabla_{\Sigma_0} (\phi - \phi_0)| \leq 24 \, \epsilon \, \{1 + |\phi| + |\nabla \phi|\} \, . \tag{2.12}$$

Proof. Set $v = \nabla^T x_{k+1} = \partial^T_{k+1}$. Let $e_1, \dots, e_{k-1}$ be an orthonormal frame for Σ_0, so that

$$e_1, \dots, e_{k-1}, \frac{v}{|v|} \tag{2.13}$$

gives an orthonormal frame for Σ. If $\mathbf{n} \in \mathbf{R}^{k+1}$ and $\mathbf{n}_0 \in \mathbf{R}^k$ denote the normals to Σ and Σ_0, respectively, then

$$\mathbf{n} = |v|\,\mathbf{n}_0 + \langle \partial_{k+1}, \mathbf{n}\rangle\, \partial_{k+1}\,. \tag{2.14}$$

(To see this, check that this unit vector is orthogonal to the frame (2.13).) Since $\langle \nabla_{e_i} e_j, \partial_{k+1}\rangle = 0$, the expression for $\mathbf{n}$ gives $\langle \nabla_{e_i} e_j, \mathbf{n}\rangle = |v|\,\langle \nabla_{e_i} e_j, \mathbf{n}_0\rangle$. It follows that

$$\begin{aligned} H - H_0 &= -\left\{ A(e_i,e_i) + A\left(\frac{v}{|v|},\frac{v}{|v|}\right)\right\} + \langle \nabla_{e_i} e_i, \mathbf{n}_0\rangle \\ &= \frac{1-|v|}{|v|}\,A(e_i,e_i) - A\left(\frac{v}{|v|},\frac{v}{|v|}\right) = \frac{|v|-1}{|v|}\,H - \frac{1}{|v|}A\left(\frac{v}{|v|},\frac{v}{|v|}\right)\,. \end{aligned} \tag{2.15}$$

Similarly, given $x \in \Sigma_0$, we have $x_{k+1} = 0$ and, thus,

$$\langle x, \mathbf{n}\rangle - \langle x_0, \mathbf{n}_0\rangle = \langle x, \mathbf{n}\rangle - \langle x, \mathbf{n}_0\rangle = \frac{|v|-1}{|v|}\,\langle x, \mathbf{n}\rangle\,. \tag{2.16}$$

Combining the last two equations gives for $x \in \Sigma_0$ that

$$\begin{aligned} \phi - \phi_0 &= \frac{1}{2}\left(\langle x, \mathbf{n}\rangle - \langle x_0, \mathbf{n}_0\rangle\right) - (H - H_0) = \frac{|v|-1}{|v|}\,\left\{\frac{1}{2}\langle x, \mathbf{n}\rangle - H\right\} + \frac{1}{|v|}A\left(\frac{v}{|v|},\frac{v}{|v|}\right) \\ &= \frac{|v|-1}{|v|}\,\phi + \frac{1}{|v|}A\left(\frac{v}{|v|},\frac{v}{|v|}\right)\,. \end{aligned} \tag{2.17}$$

Since $|v| \geq 1/2$ and $1 - |v| \leq \epsilon$, it follows that

$$|\phi - \phi_0| \leq 2\epsilon\,|\phi| + 8\,\,|A(v,v)| \leq 2\epsilon\,|\phi| + 8\,\epsilon\,. \tag{2.18}$$

Similarly, we bound the derivative by

$$\begin{aligned} |\nabla(\phi - \phi_0)| &\leq 2(1-|v|)\,|\nabla\phi| + 2\,|\nabla v|\,|\phi| + 4\,(1-|v|)|\nabla v|\,|\phi| \\ &\quad + 16\,|\nabla v|\,|A(v,v)| + 8\,\,|\nabla A(v,v)| + 16\,\,|A(v,\nabla v)| \\ &\leq 2\,\epsilon\,|\nabla\phi| + 4\,\epsilon\,|\phi| + 16\,\epsilon\,. \end{aligned} \tag{2.19}$$

□

2.2. **The proof of Proposition 2.1.**

Proof of Proposition 2.1. Within the proof, it will be convenient to set

$$\epsilon_\tau(r) = \sup_{B_r}\,\left(|\nabla\tau| + |\nabla^2\tau|\right) \text{ and } \epsilon_\phi(r) = \sup_{B_r}\,\left(|\phi| + |\nabla\phi|\right)\,. \tag{2.20}$$

Step 1: The approximate translations. Using the C^2-closeness in (2), at every p in $\Sigma \cap B_{5\sqrt{2n}}$ there are $n-k$ orthonormal eigenvectors

$$v_1(p), \ldots, v_{n-k}(p),$$

of A with eigenvalues $\kappa_1, \ldots \kappa_{n-k}$ with absolute value less than $1/\sqrt{100n}$, plus $k \geq 1$ eigenvectors with eigenvalues $\sigma_1, \ldots \sigma_k$ with absolute value at least $1/\sqrt{4n}$. By (1), we can apply Corollary 2.4 to obtain

$$|\kappa_j(p)| \leq C\,\epsilon_\tau(5\sqrt{2n})\,, \qquad j = 1, \ldots, n-k, \tag{2.21}$$

where C depends only on n and δ.

Now fix some p in $\Sigma \cap B_{2\sqrt{2n}}$ and define $n-k$ linear functions f_i on $\mathbf{R}^{n+1}$ and tangential vector fields v_i on Σ by

$$f_i(x) = \langle v_i(p), x \rangle \text{ and } v_i = \nabla^T f_i = v_i(p) - \langle v_i(p), \mathbf{n} \rangle \, \mathbf{n} \, .$$

Step 2: Extending the bounds away from p. For each $r > 5\sqrt{2n}$, let Ω_r denote the set of points in $B_r \cap \Sigma$ that can be reached from p by a path in $B_r \cap \Sigma$ of length at most $3r$. The v_i's have the following three properties on Ω_r:

$$|v_i - v_i(p)| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.22}$$

$$|\tau(v_i)| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.23}$$

$$|\nabla_{v_i} A| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.24}$$

where C depends only on n, δ and C_1.

To prove (2.22) and (2.23), suppose that $\gamma : [0, 3r] \to \Sigma$ is a curve with $\gamma(0) = p$ and $|\gamma'| \leq 1$ and that w is a parallel unit vector field along γ with $w(0) = v_i(p)$. Therefore, the bound on $\nabla \tau$ gives $|\nabla_{\gamma'} \tau(w)| \leq \epsilon_\tau(r)$ and, thus,

$$|\tau(w)| \leq 3 \, r \, \epsilon_\tau(r) + |\tau_p(v_i(p))| \leq (C + 3r) \, \epsilon_\tau(r) \leq C \, r \, \epsilon_\tau(r) \, . \tag{2.25}$$

In particular, we also have

$$|A(w)| = |H| \; |\tau(w)| \leq C \, r \, \epsilon_\tau(r) \, . \tag{2.26}$$

Therefore, since $\nabla_{\gamma'}^{\mathbf{R}^{n+1}} w = A(\gamma', w) \, \mathbf{n}$, the fundamental theorem of calculus gives

$$|w(t) - v_i(p)| = |w(t) - w(0)| \leq \int_0^{3r} |A(w(s))| \, ds \leq C \, r^2 \, \epsilon_\tau(r) \, . \tag{2.27}$$

Since $w(t)$ is tangential, we see that $|\langle v_i(p), \mathbf{n} \rangle| \leq C \, r^2 \, \epsilon_\tau(r)$, giving (2.22). Similarly, (2.27) gives that

$$|w(t) - v_i| = \left| (w(t) - v_i(p))^T \right| \leq |w(t) - v_i(p)| \leq C \, r^2 \, \epsilon_\tau(r) \, . \tag{2.28}$$

If we combine this (and the boundedness of τ) with (2.25), the triangle inequality gives

$$|\tau(v_i)| \leq |\tau(w)| + |\tau \, (w - v_i)| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.29}$$

where we used the lower bound on r to bound r by r^2. This gives (2.23).

We will see that (2.23) implies (2.24). Namely, given unit vector fields x and y, the Codazzi equation gives

$$\begin{aligned} |(\nabla_{v_i} A)(x, y)| &= |(\nabla_x A)(v_i, y)| = |(\nabla_x (H \, \tau))(v_i, y)| \\ &= |H \, (\nabla_x \tau)(v_i, y)| + |(\nabla_x H) \; \tau(v_i, y)| \leq C \, \epsilon_\tau(r) + C \, r^2 \, \epsilon_\tau \; , \end{aligned} \tag{2.30}$$

where the last inequality used that $|H|$ and $|\nabla H|$ are bounded by (1). This gives (2.24).

Step 3: The sphere. From the ϵ_0 closeness to $\mathcal{C}_k$ in $B_{5\sqrt{2n}}$ in (2), we know that

$$\Sigma_0 \equiv B_{5\sqrt{2n}} \cap \Sigma \cap \{ f_1 = \cdots = f_{n-k} = 0 \}$$

is a compact topological $\mathbf{S}^k$ of radius fixed close to $\sqrt{2k}$. Using (2.22)–(2.24), we can apply Lemma 2.11 $(n-k)$ times to get that Σ_0 has

$$\| \phi_0 \|_{C^1} \leq C \; (\epsilon_\tau + \epsilon_\phi) \; , \tag{2.31}$$

where ϵ_τ and ϵ_ϕ are evaluated at $r = 5\sqrt{2n}$. We can now apply Lemma 2.5 to get that Σ_0 is a graph over $\mathbf{S}^k_{\sqrt{2k}}$ of a function u_0 with

$$\|u_0\|_{C^{2,\alpha}} \leq C \, (\epsilon_\tau + \epsilon_\phi) \, . \tag{2.32}$$

Step 4: The translations and extending the bound. Let $y_1, \dots, y_{k+1}$ be an orthonormal basis of linear functions orthogonal to the f_i's. Define the function w by

$$w^2 \equiv \sum_{i=1}^{k+1} y_i^2 \, , \tag{2.33}$$

so that w would be identically equal to $\sqrt{2k}$ if Σ was in $\mathcal{C}_k$. In our case, it follows from (2.32) that the restriction w_0 of w to Σ_0 satisfies

$$\|w_0 - \sqrt{2k}\|_{C^{2,\alpha}(\Sigma_0)} \leq C \, (\epsilon_\tau + \epsilon_\phi) \, . \tag{2.34}$$

We will use the v_j's to extend the bounds away from Σ_0 inside Ω_r. Namely, for each y_i and v_j and any point in Ω_r, we have

$$\left|\nabla_{v_j} \nabla^T y_i\right| = \left|\nabla_{v_j} \nabla^\perp y_i\right| \leq |A(v_j, \cdot)| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.35}$$

where the last inequality used (2.23) and the positive lower bound for H.

We will extend the bounds by constructing a "radial flow". First, define a function f by

$$f^2 = \sum_{i=1}^{n-k} f_i^2 \, ,$$

and then define the vector field v by

$$v = \frac{\nabla^T f}{|\nabla^T f|^2} \, .$$

Thus, the flow by v preserves the level sets of f. Note that

$$\nabla^T f = \frac{\sum f_i \nabla^T f_i}{f} = \sum \frac{f_i}{f} \, v_i = \sum \frac{f_i}{f} \, v_i(p) + \sum \frac{f_i}{f} \, (v_i - v_i(p)) \, . \tag{2.36}$$

Since the $v_i(p)$'s are orthonormal and $\sum \left(\frac{f_i}{f}\right)^2 = 1$, it follows that

$$\left| \sum \frac{f_i}{f} \, v_i(p) \right| = 1 \, .$$

Combining this with the triangle inequality and (2.22) gives that

$$\sup_{\Omega_r} \left|1 - \left|\nabla^T f\right|\right| \leq \sum |v_i(p) - v_i| \leq C \, r^2 \, \epsilon_\tau(r) \, , \tag{2.37}$$

where C depends only on n, δ and C_1. We will assume from now on that r satisfies

$$C \, r^2 \, \epsilon_\tau(r) \leq \frac{1}{2} \, , \tag{2.38}$$

so that $\left|1 - \left|\nabla^T f\right|\right| \leq \frac{1}{2}$ and, thus, that $\sup_{\Omega_r} |v| \leq 2$. Since v is in the span of the v_i's and $|v| \leq 2$, it follows from (2.35) that

$$\sup_{\Omega_r} \left|\nabla_v \nabla^T y_i\right| \leq C\, r^2 \, \epsilon_\tau(r)\, . \tag{2.39}$$

Since $\langle \nabla y_i, v_j\rangle = 0$ at p and $|v_i - v_i(p)| \leq C\, r^2\, \epsilon_\tau(r)$ on Ω_r by (2.22), we know that

$$\left|\nabla^T_{v_j} y_i\right| \leq C\, r^2\, \epsilon_\tau(r) \text{ on } \Omega_r\, .$$

Hence, since v is in the span of the v_j's and $|v| \leq 2$, $|\nabla_v y_i| \leq C\, r^2\, \epsilon_\tau(r)$ on Ω_r. Combining this and (2.39) gives

$$\sup_{\Omega_r} \left|\nabla_v \nabla^T w^2\right| = 2 \sup_{\Omega_r} \left|\nabla_v (y_i \nabla^T y_i)\right| \leq 2(k+1) \sup_{\Omega_r} \left\{ |\nabla_v y_i|\, |\nabla^T y_i| + |y_i|\, \left|\nabla_v \nabla^T y_i\right| \right\}$$

$$\leq C\, r^3\, \epsilon_\tau(r)\, . \tag{2.40}$$

We will now define a subset $\Omega_{r,f}$ of Ω_r given by flowing Σ_0 outwards along the vector field v. To do this, let $\Phi(q,t)$ to be the flow by v at time t starting from q and set

$$\Omega_{r,f} = \left\{ \Phi(q,t) \,|\, q \in \Sigma_0,\ t^2 \leq r^2 - 3k \text{ and } \Phi(q,s) \in \Omega_r \text{ for all } s \leq t \right\}\, . \tag{2.41}$$

By integrating (2.40) up from Σ_0, we conclude that

$$\sup_{\Omega_{r,f}} \left|\nabla^T w^2\right| \leq \sup_{\Sigma_0} \left|\nabla^T w^2\right| + 6\, r \sup_{\Omega_r} \left|\nabla_v \nabla^T w^2\right| \leq C\, \epsilon_\phi + C\, r^4\, \epsilon_\tau(r)\, . \tag{2.42}$$

Integrating (2.42) from Σ_0 gives that

$$\sup_{\Omega_{r,f}} \left|w^2 - 2k\right| \leq C\, r\, \epsilon_\phi + C\, r^5\, \epsilon_\tau(r)\, . \tag{2.43}$$

Observe next that as long as

$$C\, r\, \epsilon_\phi + C\, r^5\, \epsilon_\tau(r) \leq k\, , \tag{2.44}$$

then we can conclude that

$$\Omega_{r,f} = \{f^2 \leq r^2 - 3k\} \cap \Sigma\, . \tag{2.45}$$

This gives a positive lower bound for w on $\Omega_{r,f}$ so the bound on $\nabla^T w^2$ then gives

$$\sup_{\Omega_{r,f}} \left|\nabla^T w\right| \leq C\, \epsilon_\phi + C\, r^4\, \epsilon_\tau(r)\, , \tag{2.46}$$

so the C^1 bound on w, and thus also on u, hold as claimed.

□

2.3. **Proving the first Lojasiewicz inequality.** In this subsection, we will prove Theorem 0.24. The proof not only gives the L^2 closeness to a cylinder, but also gives pointwise closeness on a scale that depends on ϕ and the initial graphical scale of Σ.

Proof of Theorem 0.24. We have that $B_R \cap \Sigma$ is a smooth graph over a cylinder of a function $\bar{u}$ with $\|\bar{u}\|_{C^{2,\alpha}} \leq \epsilon$ and $|\nabla^\ell \bar{u}| \leq C_\ell$ and that Σ satisfies:

1. $H \geq \delta > 0$ and $|A| + |\nabla A| \leq C_1$ on $B_R \cap \Sigma$.
2. $B_{5\sqrt{2n}} \cap \Sigma$ is ϵ_0 C^2-close to a cylinder in $\mathcal{C}_k$ for some $k \geq 1$.

The starting point is Proposition 2.1 which gives, for any $r \in (5\sqrt{2n}, R)$ with

$$r^2 \sup_{B_{5\sqrt{2n}}} \left(|\phi| + |\nabla \phi| \right) + r^5 \sup_{B_r} \left(|\nabla \tau| + |\nabla^2 \tau| \right) \leq \epsilon_1 \, , \tag{2.47}$$

we have that $B_{\sqrt{r^2-3k}} \cap \Sigma$ is the graph over (a subset of) a cylinder in $\mathcal{C}_k$ of u with

$$|u| + |\nabla u| \leq C \left\{ r^2 \sup_{B_{5\sqrt{2n}}} \left(|\phi| + |\nabla \phi| \right) + r^5 \sup_{B_r} \left(|\nabla \tau| + |\nabla^2 \tau| \right) \right\} . \tag{2.48}$$

Using the a priori bounds and assuming that ℓ is large enough,[11] we can use the interpolation inequalities of Lemma B.1 to get that

$$\sup_{B_{5\sqrt{2n}}} \left(|\phi| + |\nabla \phi| \right) \leq C_4 \, \|\phi\|_{L^1(B_R)}^{\frac{3}{4}} \, , \tag{2.49}$$

where $C_4 = C_4(n)$ and $L^1(B_R)$ denotes the Gaussian L^1 norm on B_R.

To get bounds on $\nabla \tau$ and $\nabla^2 \tau$, we apply Corollary 1.27 to get $C_3 = C_3(n, \lambda_0, \ell, C_\ell)$ so that for $r + \frac{1}{1+r} < R - 1$, we have

$$\sup_{B_r} \left(|\nabla \tau| + |\nabla^2 \tau| \right) \leq C_3 \, R^{2n} \left\{ \mathrm{e}^{-d_{\ell,n} \, \frac{(R-1)^2}{8}} + \|\phi\|_{L^1(B_R)}^{\frac{d_{\ell,n}}{2}} \right\} \mathrm{e}^{\frac{r^2}{8}} \, , \tag{2.50}$$

where the exponent $d_{\ell,n} \in (0,1)$ has $\lim_{\ell \to \infty} d_{\ell,n} = 1$.

Thus, we see that $B_{\sqrt{r^2-3k}} \cap \Sigma$ is the graph over (a subset of) a cylinder $\Sigma_k \in \mathcal{C}_k$ of u with

$$\begin{aligned} |u| + |\nabla u| &\leq C \left\{ r^2 \, \|\phi\|_{L^1}^{\frac{3}{4}} + r^5 \, R^{2n} \left\{ \mathrm{e}^{-d_{\ell,n} \, \frac{(R-1)^2}{8}} + \|\phi\|_{L^1(B_R)}^{\frac{d_{\ell,n}}{2}} \right\} \mathrm{e}^{\frac{r^2}{8}} \right\} \\ &\leq C \, R^{2n+5} \left\{ \mathrm{e}^{-d_{\ell,n} \, \frac{(R-1)^2}{8}} + \|\phi\|_{L^1(B_R)}^{\frac{d_{\ell,n}}{2}} \right\} \mathrm{e}^{\frac{r^2}{8}} \, , \end{aligned} \tag{2.51}$$

where $C = C(n, \lambda_0, \ell, C_\ell)$ and this holds so long as the right hand side is at most $\epsilon_1 > 0$. Define the radius $R_1 \leq R - 1$ to be the maximal radius where this holds.

To get the L^2 bound, we first use (2.51) on B_{R_1} to get

$$\int_{B_{R_1}} \left| w_{\Sigma_k} - \sqrt{2k} \right|^2 \mathrm{e}^{-\frac{|x|^2}{4}} \leq C \, R^{5n+10} \left\{ \mathrm{e}^{-d_{\ell,n} \, \frac{(R-1)^2}{4}} + \|\phi\|_{L^1(B_R)}^{d_{\ell,n}} \right\} \, , \tag{2.52}$$

and then use that $\left| w_{\Sigma_k} - \sqrt{2k} \right|^2 (x) \leq |x|^2$ to get that

$$\int_{B_R \setminus B_{R_1}} \left| w_{\Sigma_k} - \sqrt{2k} \right|^2 \mathrm{e}^{-\frac{|x|^2}{4}} \leq C \, R^{n+2} \, \mathrm{e}^{-\frac{R_1^2}{4}} \leq C \, R^{5n+12} \left\{ \mathrm{e}^{-d_{\ell,n} \, \frac{(R-1)^2}{4}} + \|\phi\|_{L^1(B_R)}^{d_{\ell,n}} \right\} \, , \tag{2.53}$$

where the last inequality is the definition of R_1. Combining these completes the proof.

□

We will later also need a variation on this, where we assume bounds on A and H on a large scale and conclude that Σ is a graph over a cylinder on a large set.

Theorem 2.54. There exist R_0, ℓ_0 and $\delta > 0$ so that if $\Sigma \subset \mathbf{R}^{n+1}$ has $\lambda(\Sigma) \leq \lambda_0$ and

[11] We will take ℓ large later; we could replace 3/4 by any constant less than one by taking ℓ larger.

(1) for some $R > R_0$, we have on $B_R \cap \Sigma$ that $|A| + |\nabla^{\ell_0} A| \leq C_0$ and $H \geq \delta_0 > 0$,
(2) $B_{R_0} \cap \Sigma$ is a C^2 graph over some cylinder in $\mathcal{C}_k$ with norm at most δ.

Then there is a cylinder $\tilde{\Sigma} \in \mathcal{C}_k$ so that

(3) $B_{R_1-2} \cap \Sigma$ is the graph of u over $\tilde{\Sigma}$ with $\|u\|_{C^{2,\alpha}} \leq \epsilon_0$,

where R_1 is given by

$$R_1 = \max \left\{ r \leq R - 1 \,\Big|\, R^{2n+5} \left(\mathrm{e}^{-b_{\ell_0,n} \frac{(R-1)^2}{8}} + \|\phi\|_{L^1(B_R)}^{\frac{b_{\ell_0,n}}{2}} \right) \mathrm{e}^{\frac{r^2}{8}} \leq \tilde{C} \right\} , \tag{2.55}$$

the exponent $b_{\ell_0,n} \in (0,1)$ satisfies $\lim_{\ell_0 \to \infty} b_{\ell_0,n} = 1$ and $\tilde{C} = \tilde{C}(n, \lambda_0, \delta_0, C_0)$.

Proof. We follow the proof of Theorem 0.24 up through (2.51) to get $\tilde{\Sigma} \in \mathcal{C}_k$ and a function u so that $B_{R_1-1} \cap \Sigma$ is the graph of u over $\tilde{\Sigma}$, R_1 is defined by (2.55), and

$$|u| + |\nabla u| \leq 2\,\delta\,. \tag{2.56}$$

Finally, we use interpolation and the $\nabla^{\ell_0} A$ bound to get the desired $C^{2,\alpha}$ bound when $\delta > 0$ is sufficiently small.

□

3. Analysis on the cylinder

In this section, we will prove estimates for the $\mathcal{L}$ and L operators on a cylinder $\Sigma \in \mathcal{C}_k$ with $k \in \{1, \dots, n-1\}$. These estimates will be used in the next section to prove our second Lojasiewicz inequality. Note that $L = \mathcal{L} + 1$ on Σ since $|A|^2 \equiv \frac{1}{2}$.

We will use the Gaussian L^2-norm $\|u\|_{L^2}^2 = \int u^2 \, \mathrm{e}^{-\frac{|x|^2}{4}}$, as well as the associated Gaussian $W^{1,2}$ and $W^{2,2}$ norms

$$\|u\|_{W^{1,2}}^2 = \int \left(u^2 + |\nabla u|^2 \right) \mathrm{e}^{-\frac{|x|^2}{4}} \text{ and } \|u\|_{W^{2,2}}^2 = \int \left(u^2 + |\nabla u|^2 + |\mathrm{Hess}_u|^2 \right) \mathrm{e}^{-\frac{|x|^2}{4}}\,. \tag{3.1}$$

3.1. Symmetry, the spectrum of $\mathcal{L}$ and a Poincaré inequality.

The starting point is the following elementary lemma that summarizes the key properties of the $\mathcal{L}$ operator on $\Sigma \in \mathcal{C}_k$:

Lemma 3.2. The operator $\mathcal{L}$ on Σ is symmetric on $W^{2,2}$ with

$$\int_\Sigma u \, \mathcal{L} v \, \mathrm{e}^{-\frac{|x|^2}{4}} = - \int_\Sigma \langle \nabla u, \nabla v \rangle \, \mathrm{e}^{-\frac{|x|^2}{4}}\,. \tag{3.3}$$

The space $W^{1,2}$ embeds compactly into L^2 and $\mathcal{L}$ has discrete spectrum with finite multiplicity on $W^{2,2}$ with a complete basis of smooth L^2-orthonormal eigenfunctions.

Proof. The first claim follows from integration by parts. The second follows from [BE] since Σ has positive Bakry-Émery Ricci curvature and finite weighted volume. Finally, the last claim is a consequence of the first two (cf. theorem 10.20 in [Gr]). □

We will also use the following Gaussian Poincaré inequality on $\Sigma = \mathbf{S}^k_{\sqrt{2k}} \times \mathbf{R}^{n-k}$. The middle term does not use the full gradient, but only the gradient in the translation directions.

Lemma 3.4. There exists $C = C(k,n)$ so that if $\Sigma \in \mathcal{C}_k$ and $u \in W^{1,2}$, then

$$\| |x| \, u \|_{L^2}^2 \leq C \, \left(\|u\|_{L^2}^2 + \|\nabla_{\mathbf{R}^{n-k}} u\|_{L^2}^2 \right) \leq C \, \|u\|_{W^{1,2}}^2\,. \tag{3.5}$$

Proof. Let y be coordinates on the $\mathbf{R}^{n-k}$ factor, so that

$$x^T = y \text{ and } |x|^2 = |y|^2 + 2k \,. \tag{3.6}$$

We compute

$$\begin{aligned} \mathrm{e}^{\frac{|x|^2}{4}} \, \mathrm{div}_\Sigma \left(u^2 \, y \, \mathrm{e}^{-\frac{|x|^2}{4}} \right) &= 2 \, u \langle \nabla u, y \rangle + (n-k) \, u^2 - u^2 \, \frac{|y|^2}{2} \\ &\leq 4 \, |\nabla_{\mathbf{R}^{n-k}} u|^2 + (n-k) \, u^2 - u^2 \, \frac{|y|^2}{4} \,, \end{aligned} \tag{3.7}$$

where the inequality used the absorbing inequality $2ab \leq \frac{a^2}{4} + 4 \, b^2$.

By approximation, we can assume that u has compact support on Σ and, thus, Stokes' theorem gives

$$\frac{1}{4} \int_\Sigma u^2 \, |y|^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq \int_\Sigma \left\{ (n-k) \, u^2 + 4 \, |\nabla_{\mathbf{R}^{n-k}} u|^2 \right\} \, \mathrm{e}^{-\frac{|x|^2}{4}} \,. \tag{3.8}$$

The lemma follows since $u^2 \, |x|^2 = u^2 \, (|y|^2 + 2k)$.

□

3.2. **Estimates for the projection onto the kernel of L.** Let $\mathcal{K}$ be the kernel of L

$$\mathcal{K} = \{ v \in W^{2,2} \, | \, Lv = 0 \} \,. \tag{3.9}$$

Given any $u \in W^{2,2}$, we let $u_{\mathcal{K}}$ denote the L^2-orthogonal projection of u onto $\mathcal{K}$ and

$$u^\perp = u - u_{\mathcal{K}} \tag{3.10}$$

the projection onto the L^2-orthogonal complement of $\mathcal{K}$.

The next lemma shows that L is bounded from $W^{2,2}$ to L^2, L is uniformly invertible on $\mathcal{K}^\perp$ and the projection onto $\mathcal{K}$ is bounded from L^2 to $W^{2,2}$.

Lemma 3.11. Given n, there exist C and $\mu > 0$ so that on $\mathcal{C}_k$

$$\|Lu\|_{L^2} \leq C \, \|u\|_{W^{2,2}} \,, \tag{3.12}$$

$$\mu \, \|u^\perp\|_{W^{2,2}} \leq \|Lu\|_{L^2} \,, \tag{3.13}$$

$$\|u_{\mathcal{K}}\|_{W^{2,2}} \leq C \, \|u\|_{L^2} \,. \tag{3.14}$$

Proof. Since $L = \Delta + \frac{1}{2} \nabla_{x^T} + 1$ on the cylinder, we have

$$\|Lu\|_{L^2} \leq \|\Delta u\|_{L^2} + \|u\|_{L^2} + \frac{1}{2} \, \| |x| \, |\nabla u| \|_{L^2} \,. \tag{3.15}$$

The first claim follows from this and using Lemma 3.4 to get the bound

$$\| |x| \, |\nabla u| \|_{L^2} \leq C \, \| |\nabla u| \|_{W^{1,2}} \leq C \, \left\{ \| |\nabla u| \|_{L^2} + \| \mathrm{Hess}_u \|_{L^2} \right\} \,. \tag{3.16}$$

To get the second claim, we will need the "Gaussian elliptic estimate"

$$\|v\|_{W^{2,2}} \leq C \, \left(\|v\|_{L^2} + \|\mathcal{L} v\|_{L^2} \right) \,, \tag{3.17}$$

where C depends on n and the estimate holds for any $v \in W^{2,2}$. To prove (3.17), we first integrate by parts to get

$$\|\nabla v\|_{L^2}^2 = |\langle v, \mathcal{L} v \rangle_{L^2}| \leq \|v\|_{L^2} \, \|\mathcal{L} v\|_{L^2} \leq \frac{1}{2} \|v\|_{L^2}^2 + \frac{1}{2} \|\mathcal{L} v\|_{L^2}^2 \,. \tag{3.18}$$

Thus, we see that $\|v\|_{W^{1,2}}$ is bounded by the right hand side of (3.17). It remains to bound the L^2 norm of the Hessian of v. This will follow from what we've done and the divergence theorem since

$$\begin{aligned} \mathrm{e}^{\frac{|x|^2}{4}} \operatorname{div}_\Sigma \left(\{v_{ij} v_i - (\mathcal{L} v)\, v_j\} \, \mathrm{e}^{-\frac{|x|^2}{4}} \right) &= \frac{1}{2} \mathcal{L}\, |\nabla v|^2 - (\mathcal{L}\, v)^2 - \langle \nabla \mathcal{L}\, v, \nabla v \rangle \\ &\geq |\mathrm{Hess}_v|^2 - (\mathcal{L} v)^2 \,, \end{aligned} \tag{3.19}$$

where the last inequality used the Bochner formula for the drift Laplacian on the cylinder.[12]

The second claim now follows by first applying Lemma 3.2 to get $\mu_0 > 0$ so that

$$\mu_0 \, \|u^{\perp}\|_{L^2} \leq \|L u^{\perp}\|_{L^2} = \|L u\|_{L^2} \tag{3.20}$$

and then using (3.17) to bound the $W^{2,2}$ norm.

The final claim follows from the trivial projection bound $\|u_{\mathcal{K}}\|_{L^2} \leq \|u\|_{L^2}$ and the bound

$$\|u_{\mathcal{K}}\|_{W^{2,2}} \leq C \, \|u_{\mathcal{K}}\|_{L^2} \,. \tag{3.21}$$

To see (3.21), first use the equation $\mathcal{L} u_{\mathcal{K}} = -u_{\mathcal{K}}$ to get $\|\nabla u_{\mathcal{K}}\|_{L^2} = \|u_{\mathcal{K}}\|_{L^2}$, and then use the Bochner formula as in (3.19) to bound the Hessian of $u_{\mathcal{K}}$ in terms of $\|u_{\mathcal{K}}\|_{W^{1,2}}$. □

We will also need the next lemma that bounds the Gaussian L^2 norm of a quadratic expression in $u, \nabla u, \mathrm{Hess}_u$ that bounds the error term in the linear approximation of the gradient of the F functional. When $u \in \mathcal{K}$, the bound is the square of the Gaussian L^2 norm[13] while we obtain a weaker bound when u is orthogonal to $\mathcal{K}$.

Lemma 3.22. There exist $C_K = C_K(n)$ and $C_0 = C_0(n)$ so that if $u \in W^{2,2}$, then

$$\left\| u_{\mathcal{K}}^2 + |\nabla u_{\mathcal{K}}|^2 + \left| \mathrm{Hess}_{u_{\mathcal{K}}}(\cdot, \mathbf{R}^{n-k}) \right|^2 + (1+|x|)^{-1} \, |\mathrm{Hess}_{u_{\mathcal{K}}}|^2 \right\|_{L^2} \leq C_K \, \|u_{\mathcal{K}}\|_{L^2}^2 \,, \tag{3.23}$$

$$\left\| (u^{\perp})^2 + |\nabla u^{\perp}|^2 + \left| \mathrm{Hess}_{u^{\perp}}(\cdot, \mathbf{R}^{n-k}) \right|^2 + (1+|x|)^{-1} \, |\mathrm{Hess}_{u^{\perp}}|^2 \right\|_{L^2} \leq C_0 \, \|u\|_{C^2} \, \|u^{\perp}\|_{W^{2,2}} \,. \tag{3.24}$$

The key for proving both claims is an explicit description of $\mathcal{K}$. Namely, $\mathcal{K}$ is generated by multiplying a polynomial eigenfunction of $\mathcal{L}_{\mathbf{R}^{n-k}}$ times a spherical eigenfunction of $\Delta_{\mathbf{S}^k_{\sqrt{2k}}}$. To state this, let y_i be coordinates on the $\mathbf{R}^{n-k}$ factor and let θ be in the $\mathbf{S}^k$ factor.

Lemma 3.25. Each $v \in \mathcal{K}$ can be written as

$$v(y, \theta) = q(y) + \sum_i y_i f_i(\theta) + c \,, \tag{3.26}$$

where q is a homogeneous quadratic polynomial on $\mathbf{R}^{n-k}$, each f_i is an eigenfunction on $\mathbf{S}^k_{\sqrt{2k}}$ with eigenvalue $\frac{1}{2}$, and c is a constant.

Proof. The operator L splits as

$$L = \mathcal{L} + 1 = \Delta_\theta + \mathcal{L}_y + 1 \,, \tag{3.27}$$

where Δ_θ is the Laplacian on $\mathbf{S}^k_{\sqrt{2k}}$ and $\mathcal{L}_y$ is the drift operator on $\mathbf{R}^{n-k}$.

[12] See, for instance, (6.7) below.

[13] This would be obvious if the C^2 norm of v were bounded by the L^2 norm, but this is not the case.

The first obervation is that differentiating with respect to y_i lowers the eigenvalue by $\frac{1}{2}$. Thus, if we set $v_i = \frac{\partial v}{\partial y_i}$ and $v_{ij} = \frac{\partial^2 v}{\partial y_i \partial y_j}$, then

$$\mathcal{L} v_i = -\frac{1}{2} v_i \,, \tag{3.28}$$

$$\mathcal{L} v_{ij} = 0 \,. \tag{3.29}$$

Since every L^2 $\mathcal{L}$-harmonic function must be constant, we conclude that v_{ij} is constant. As a consequence, the function v can be written as

$$v = \sum_{i,j} a_{ij}\, y_i\, y_j + \sum_i f_i(\theta)\, y_i + g(\theta) \,, \tag{3.30}$$

where $a_{ij} \in \mathbf{R}$, each f_i is a function on $\mathbf{S}^k_{\sqrt{2k}}$ and g is a function on $\mathbf{S}^k_{\sqrt{2k}}$.

Note that

$$\mathcal{L}_y y_i = -\frac{1}{2}\, y_i \,, \tag{3.31}$$

$$\mathcal{L}_y \left(y_i y_j \right) = 2 \delta_{ij} - y_i y_j \,. \tag{3.32}$$

Using this and the decomposition of L from (3.27), we get that

$$0 = L v = \sum_{i,j} a_{ij}\, \left(2 \delta_{ij} \right) + \sum_i \left[y_i\, \Delta_\theta f_i(\theta) + \frac{1}{2} f_i(\theta)\, y_i \right] + \left(\Delta_\theta + 1 \right) g(\theta) \,. \tag{3.33}$$

Observe first that only the middle terms depend on y. Setting these equal to zero, we conclude that each f_i satisfies

$$\Delta_\theta f_i = -\frac{1}{2} f_i \,. \tag{3.34}$$

It follows that $g + 2 \sum a_{ii}$ is a $\mathbf{S}^k_{\sqrt{2k}}$ eigenfunction with eigenvalue one, i.e.,

$$\Delta_\theta \left(g + 2 \sum a_{ii} \right) = - \left(g + 2 \sum a_{ii} \right) \,. \tag{3.35}$$

However, one is not an eigenvalue of Δ_θ (the eigenvalues jump from $1/2$ to $(k+1)/k$; see (2.7)), so we have $g \equiv -2 \sum a_{ii}$.

□

It is interesting to note that the $y_i\, f_i$ part of $\mathcal{K}$ corresponds to rotations. However, by [CIM], the quadratic polynomials in the kernel are not generated by one-parameter families of shrinkers. In particular, the kernel $\mathcal{K}$ contains non-integrable functions.

As a corollary of Lemma 3.25, we get C^2 pointwise estimates for functions in the kernel of L that grow at most quadratically in $|y|$:

Corollary 3.36. There exists C depending on n so that if $v \in \mathcal{K}$, then

$$\sup |v| \leq C \,(1 + |y|^2)\, \|v\|_{L^2} \,, \tag{3.37}$$

$$\sup |\nabla v| \leq C \,(1 + |y|)\, \|v\|_{L^2} \,, \tag{3.38}$$

$$\sup \left| \mathrm{Hess}_v (\cdot, \mathbf{R}^{n-k}) \right| \leq C\, \|v\|_{L^2} \,, \tag{3.39}$$

$$\sup |\mathrm{Hess}_v| \leq C \,(1 + |y|)\, \|v\|_{L^2} \,. \tag{3.40}$$

Remark 3.41. The point of (3.39) is that, as opposed to (3.40), we get a better bound, that does not grow in y, if we restrict to the Hessian in the Euclidean factor. This is useful later.

Proof of Corollary 3.36. Since $\mathcal{K}$ is finite dimensional, the estimates (3.37)–(3.40) will follow for all of $\mathcal{K}$ from the squared triangle inequality once we show that there is an orthogonal basis for $\mathcal{K}$ where each element in the basis satisfies (3.37)–(3.40).

The key for this is Lemma 3.25 which shows that $\mathcal{K}$ can be written as

$$\mathcal{K} = \mathcal{K}_1 \oplus \mathcal{K}_2 \, , \text{ where} \tag{3.42}$$

- Each $v_1 \in \mathcal{K}_1$ is given by $\sum_i y_i \, f_i$ where f_i is a $\mathbf{S}^k_{\sqrt{2k}}$ eigenfunction with eigenvalue $\frac{1}{2}$.
- Each $v_2 \in \mathcal{K}_2$ is a constant plus a homogeneous quadratic polynomial in y.

In particular, (3.42) is a L^2-orthogonal decomposition.

Case 1. If f_i, f_j are $\mathbf{S}^k_{\sqrt{2k}}$ eigenfunctions with eigenvalue $\frac{1}{2}$, then

$$\langle y_i \, f_i, y_j \, f_j \rangle_{L^2} = 0 \text{ if } i \neq j \, , \tag{3.43}$$

so we get an orthogonal basis for $\mathcal{K}_1$ consisting of a single y_i times an f. Suppose that

$$v_1 = y_i \, f \, , \tag{3.44}$$

where f is a $\mathbf{S}^k_{\sqrt{2k}}$ eigenfunction with eigenvalue $\frac{1}{2}$. Note that

$$\|v_1\|^2_{L^2} = \mathrm{e}^{-\frac{k}{2}} \, \|f\|^2_{L^2_\theta} \int_{\mathbf{R}^{n-k}} y_i^2 \, \mathrm{e}^{-\frac{|y|^2}{4}} \, dy \equiv C_k \, \|f\|^2_{L^2_\theta} \, , \tag{3.45}$$

where the constant $C_k > 0$ depends only on k and the sub θ denotes the norms on $\mathbf{S}^k_{\sqrt{2k}}$.

Using elliptic estimates for the compact manifold $\mathbf{S}^k_{\sqrt{2k}}$, we have $c_0 = c_0(k)$ so that

$$\|f\|_{C^2_\theta} \leq c_0 \, \|f\|_{L^2_\theta} \, . \tag{3.46}$$

Therefore, at each point, we have that

$$|v_1|^2 = y_i^2 \, f^2 \leq c_0^2 \, y_i^2 \, \|f\|^2_{L^2_\theta} \, , \tag{3.47}$$

$$|\nabla v_1|^2 = y_i^2 \, |\nabla_\theta f|^2 + f^2 \leq c_0^2 \, \left(1 + y_i^2\right) \, \|f\|^2_{L^2_\theta} \, , \tag{3.48}$$

$$|\mathrm{Hess}_{v_1}|^2 = y_i^2 \, |\mathrm{Hess}_f|^2 \leq c_0^2 \, y_i^2 \, \|f\|^2_{L^2_\theta} \, , \tag{3.49}$$

$$\left|\mathrm{Hess}_{v_1}(\cdot, \mathbf{R}^{n-k})\right|^2 \leq |\nabla_\theta f|^2 \leq c_0^2 \, , \tag{3.50}$$

giving the desired bounds in this case (the first bound is even better than needed).

Case 2. It is easy to see that an orthogonal basis for $\mathcal{K}_2$ is given by

$$\{y_i y_j - 2\delta_{ij} \, | \, i \leq j\} \, . \tag{3.51}$$

Therefore, it suffices to show (3.37)–(3.40) when

$$v_2 = y_i y_j - 2\delta_{ij} \, . \tag{3.52}$$

However, this follows immediately since the L^2 norms are nonzero and v_2 is a quadratic polynomial in y (in this case, the Hessian bound is even better than needed). □

We will now use the estimates from the corollary to prove Lemma 3.22.

Proof of Lemma 3.22. To simplify notation, set

$$\|v\|_2 \equiv \left\| v^2 + |\nabla v|^2 + \left|\text{Hess}_v(\cdot, \mathbf{R}^{n-k})\right|^2 + (1+|x|)^{-1} \left|\text{Hess}_v\right|^2 \right\|_{L^2} . \tag{3.53}$$

Given $a \in \mathbf{R}$, note that $\|a\, v\|_2 = a^2 \, \|v\|_2$.

We will show that there is a constant C_K so that

$$C_K \equiv \sup \left\{ \|w\|_2 \, \middle| \, w \in \mathcal{K} \text{ and } \|w\|_{L^2} = 1 \right\} < \infty . \tag{3.54}$$

Once we have this, then for a general $v \in \mathcal{K}$, we set $w = \frac{v}{\|v\|_{L^2}}$ so that

$$\|v\|_2 = \| \, \|v\|_{L^2} \, w\|_2 = \|v\|_{L^2}^2 \, \|w\|_2 \leq C_K \, \|v\|_{L^2}^2 , \tag{3.55}$$

giving the first claim (3.23).

To establish (3.54), apply Corollary 3.36 to get $C = C(n)$ so that

$$|w|^4 + |\nabla w|^4 + |\text{Hess}_w|^4 \leq C \, (1+|y|^2)^4 . \tag{3.56}$$

Integrating this polynomially growing bound against the exponential decaying Gaussian weight gives the desired uniform bound on $\|w\|_2^2$.

To prove (3.24), we will show that

$$\left\|(u^\perp)^2\right\|_{L^2} , \; \left\||\nabla u^\perp|^2\right\|_{L^2} , \; \left\|\left|\text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k})\right|^2\right\|_{L^2} \text{ and } \left\|(1+|x|)^{-1} \left|\text{Hess}_{u^\perp}\right|^2\right\|_{L^2} \tag{3.57}$$

are each bounded by $C_0 \, \|u\|_{C^2} \, \|u^\perp\|_{W^{2,2}}$ for a constant C_0 depending only on the dimension n. The key point will be the bounds (3.37)–(3.40) on $u_{\mathcal{K}}$ from Corollary 3.36.

For the first term, we use (3.37) to get

$$\begin{aligned} (u^\perp)^2 = (u - u_{\mathcal{K}}) \, u^\perp &\leq \left(\|u\|_{C^0} + C \, (1+|x|^2) \|u_{\mathcal{K}}\|_{L^2} \right) \, |u^\perp| \\ &\leq C \, \|u\|_{C^0} \, (1+|x|^2) \, |u^\perp| , \end{aligned} \tag{3.58}$$

where the last inequality used the projection inequality $\|u_{\mathcal{K}}\|_{L^2} \leq \|u\|_{L^2}$ and the trivial inequality $\|u\|_{L^2} \leq C \, \|u\|_{C^0}$ that follows since Σ has finite Gaussian area. Integrating and applying Lemma 3.4 twice gives

$$\|(u^\perp)^2\|_{L^2} \leq C \, \|u\|_{C^0} \, \|(1+|x|^2) \, u^\perp\|_{L^2} \leq C \, \|u\|_{C^0} \, \|u^\perp\|_{W^{2,2}} . \tag{3.59}$$

For the second term, we use the triangle inequality and (3.38) to get

$$\begin{aligned} |\nabla u^\perp|^2 &\leq (|\nabla u| + |\nabla u_{\mathcal{K}}|) \, |\nabla u^\perp| \leq (\|u\|_{C^1} + C(1+|x|) \, \|u_{\mathcal{K}}\|_{L^2}) \, |\nabla u^\perp| \\ &\leq C \, \|u\|_{C^1} (1+|x|) \, |\nabla u^\perp| , \end{aligned} \tag{3.60}$$

where the last inequality follows as above. Integrating and applying Lemma 3.4 gives

$$\||\nabla u^\perp|^2\|_{L^2} \leq C \, \|u\|_{C^1} \, \|(1+|x|) \, |\nabla u^\perp|\|_{L^2} \leq C \, \|u\|_{C^1} \, \|u^\perp\|_{W^{2,2}} . \tag{3.61}$$

For the third term, we use the triangle inequality and (3.39) to get

$$\begin{aligned} \left|\text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k})\right|^2 &\leq \left\{ \left|\text{Hess}_u(\cdot, \mathbf{R}^{n-k})\right| + \left|\text{Hess}_{u_{\mathcal{K}}}(\cdot, \mathbf{R}^{n-k})\right| \right\} \, \left|\text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k})\right| \\ &\leq \left\{ \|u\|_{C^2} + C \, \|u_{\mathcal{K}}\|_{L^2} \right\} \, \left|\text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k})\right| \\ &\leq C \, \|u\|_{C^2} \, \left|\text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k})\right| . \end{aligned} \tag{3.62}$$

Integrating this gives

$$\left\| \left| \text{Hess}_{u^\perp}(\cdot, \mathbf{R}^{n-k}) \right|^2 \right\|_{L^2} \leq C \, \|u\|_{C^2} \, \|\text{Hess}_{u^\perp}\|_{L^2} \, . \tag{3.63}$$

Finally, for the fourth (last) term, we use the triangle inequality and (3.40) to get

$$\begin{aligned} (1+|x|)^{-1} \, |\text{Hess}_{u^\perp}|^2 &\leq (1+|x|)^{-1} \left(|\text{Hess}_u| + |\text{Hess}_{u_{\mathcal{K}}}| \right) |\text{Hess}_{u^\perp}| \\ &\leq (1+|x|)^{-1} \left\{ \|u\|_{C^2} + C\,(1+|x|) \, \|u_{\mathcal{K}}\|_{L^2} \right\} |\text{Hess}_{u^\perp}| \\ &\leq C \, \|u\|_{C^2} \, |\text{Hess}_{u^\perp}| \, . \end{aligned} \tag{3.64}$$

To bound the last term and complete the proof of (3.24), we integrate this to get

$$\left\| (1+|x|)^{-1} \, |\text{Hess}_{u^\perp}|^2 \right\|_{L^2} \leq C \, \|u\|_{C^2} \, \|\text{Hess}_{u^\perp}\|_{L^2} \, .$$

□

4. The gradient Lojasiewicz inequality for F

In this section, we will prove a gradient Lojasiewicz inequality for F in a neighborhood of a cylinder $\Sigma \in \mathcal{C}_k$. The inequality will hold for graphs over part of Σ with small C^2 norm. The key technical ingredient is the next proposition which shows that our first Lojasiewicz implies our gradient Lojasiewicz inequality.

Proposition 4.1. There exist $C = C(n, \lambda_0)$ and $\bar{\epsilon} = \bar{\epsilon}(n) > 0$ so that if $\lambda(\Sigma) \leq \lambda_0$ and $B_{\tilde{R}} \cap \Sigma$ is the graph of $\tilde{u}$ over a cylinder in $\mathcal{C}_k$ with $\|\tilde{u}\|_{C^2} \leq \bar{\epsilon}$, then for any $\beta \in [0, 1)$

$$|F(\Sigma) - F(\mathcal{C}_k)| \leq C \|\phi\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{2}} + C\,(1 + \tilde{R}^{n-1}) \, \mathrm{e}^{-\frac{(3+\beta)(\tilde{R}-1)^2}{16}} + C \, \|\tilde{u}\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{1+\beta}} \, . \tag{4.2}$$

The proof of Proposition 4.1 is an infinite dimensional version of the model argument using Taylor expansion given in Subsection 0.4. The simple model was done with $\beta = 0$, but would have worked with any $\beta \in [0, 1)$. However, the simple model did not include a cutoff and, to bound the exponential term in (4.2), we will need to choose β close to one.[14]

4.1. **The linearization of the gradient of the** F **functional.** Given a graph Σ_u of a function u over a cylinder $\Sigma \in \mathcal{C}_k$, we let $F(u) \equiv F(\Sigma_u)$ and then let $\mathcal{M}(u)$ be the gradient of F. The next lemma gives linear and quadratic approximations for $\mathcal{M}$ and F, respectively.

Lemma 4.3. There exists C_1 so that if the C^2 norm of u is sufficiently small and u is defined on the entire cylinder, then

$$\|\mathcal{M}(u) - Lu\|_{L^2} \leq C_1 \, \left\| u^2 + |\nabla u|^2 + |\nabla_{\mathbf{R}^{n-k}} |\nabla u||^2 + (1+|x|)^{-1} \, |\text{Hess}_u|^2 \right\|_{L^2} \, , \tag{4.4}$$

$$\left| F(u) - F(\mathcal{C}_k) - \frac{1}{2} \langle u, Lu \rangle_{L^2} \right| \leq C_1 \, \|u\|_{L^2} \, \left\| u^2 + |\nabla u|^2 + |\nabla_{\mathbf{R}^{n-k}} |\nabla u||^2 + (1+|x|)^{-1} \, |\text{Hess}_u|^2 \right\|_{L^2} \, . \tag{4.5}$$

[14]The $\|\phi\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{2}}$ term is fine for any $\beta > 0$ and the $\|\tilde{u}\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{1+\beta}}$ term is fine if $\beta < 1$.

The bound in the first inequality in Lemma 4.3 is essentially quadratic in u. For example, it is bounded by $C \, \|u\|_{C^2} \, \|u\|_{W^{2,2}}$. Ideally, we would have liked the bound to be quadratic in $\|u\|_{W^{2,2}}$, but the exponential decay in the Gaussian norm makes this impossible and, thus, leads to technical complications.

We will prove Lemma 4.3 in this subsection. The starting point is the next lemma computing $\mathcal{M}(u)$ in terms of $u, \nabla u$ and Hess_u.

Lemma 4.6. If Σ is a cylinder in $\mathcal{C}_k$ and $p \in \Sigma$, then

$$\mathcal{M}(u)(p) = f(u(p), \nabla u(p)) + \langle p, V(u(p), \nabla u(p)) \rangle + \Phi^{\alpha\beta}(u(p), \nabla u(p)) \, u_{\alpha\beta}(p) \, , \tag{4.7}$$

where f, V and $\Phi^{\alpha\beta}$ depend smoothly on (s, y) for $|s|$ small.

Lemma 4.6 is proven in Appendix A.

The next lemma shows, for a general map $u \to \mathcal{M}(u)$ of the form (4.7), that the linearization gives a good approximation up to quadratic error. To state this precisely, consider a general map $\mathcal{N}(u)$ of the form

$$\mathcal{N}(u)(p) = f(p, u(p), \nabla u(p)) + \Phi^{\alpha\beta}(p, u(p), \nabla u(p)) \, u_{\alpha\beta}(p) \, , \tag{4.8}$$

where f and $\Phi^{\alpha\beta}$ are smooth functions of (p, s, y) where p is the point, $s \in \mathbf{R}$, and y is a tangent vector at p. The linearization of $\mathcal{N}$ at u is defined to be

$$L_u \, v = \frac{d}{dt}\Big|_{t=0} \mathcal{N}(u + tv) = f_s \, v + f_{y_\alpha} v_\alpha + \Phi^{\alpha\beta} \, v_{\alpha\beta} + u_{\alpha\beta} \, \left(\Phi^{\alpha\beta}_s \, v + \Phi^{\alpha\beta}_{y_\gamma} \, v_\gamma \right) \, , \tag{4.9}$$

where all functions are evaluated at the same point p and we have left out the obvious dependence of f and Φ on $(p, u(p), \nabla u(p))$.

Lemma 4.10. If $\mathcal{N}(u)$ is given by (4.8), then we get at each point p that

$$|\mathcal{N}(u+v) - \mathcal{N}(u) - L_u v| \leq C_1 \, \left(|v| + |\nabla v|\right)^2 + C_2 \, \left(|v| + |\nabla v|\right) \, |\text{Hess}_v| \, , \tag{4.11}$$

where the constants $C_1 = C_1(p)$ and $C_2 = C_2(p)$ are given by

$$C_1 = \text{Lip}_p(f_s) + |u_{\alpha\beta}| \, \text{Lip}_p \left(\Phi^{\alpha\beta}_s\right) + \text{Lip}_p(f_{y_\gamma}) + |u_{\alpha\beta}| \, \text{Lip}_p \left(\Phi^{\alpha\beta}_{y_\gamma}\right) \, , \tag{4.12}$$

$$C_2 = \left|\Phi^{\alpha\beta}_s\right| + \left|\Phi^{\alpha\beta}_{y_\gamma}\right| + \text{Lip}_p \left(\Phi^{\alpha\beta}\right) \, . \tag{4.13}$$

Here Lip_p denotes the Lipschitz norm at p with respect to the s and y variables.

Proof. Using (4.9), we get at p that for any w that

$$\begin{aligned} |L_{u+w}v - L_u v| \leq \; & |f_s(p, u+w, \nabla u + \nabla w) - f_s(p, u, \nabla u)| \, |v| \\ & + \left|(u_{\alpha\beta} + w_{\alpha\beta}) \, \Phi^{\alpha\beta}_s(p, u+w, \nabla u + \nabla w) - u_{\alpha\beta} \Phi^{\alpha\beta}_s(p, u, \nabla u)\right| \, |v| \\ & + |f_{y_\alpha}(p, u+w, \nabla u + \nabla w) - f_{y_\alpha}(p, u, \nabla u)| \, |v_\alpha| \\ & + \left|(u_{\alpha\beta} + w_{\alpha\beta}) \, \Phi^{\alpha\beta}_{y_\gamma}(p, u+w, \nabla u + \nabla w) - u_{\alpha\beta} \Phi^{\alpha\beta}_{y_\gamma}(p, u, \nabla u)\right| \, |v_\gamma| \\ & + \left|\Phi^{\alpha\beta}(p, u+w, \nabla u + \nabla w) - \Phi^{\alpha\beta}(p, u, \nabla u)\right| \, |v_{\alpha\beta}| \, . \end{aligned} \tag{4.14}$$

Bounding these terms gives

$$
\begin{aligned}
|L_{u+w}v - L_u v| \leq & \left\{ \left[\mathrm{Lip}_p(f_s) + |u_{\alpha\beta}| \ \mathrm{Lip}_p\left(\Phi_s^{\alpha\beta}\right)\right] \ (|w| + |\nabla w|) + |w_{\alpha\beta}| \ \left|\Phi_s^{\alpha\beta}\right| \right\} \ |v| \\
& + \left\{ \left[\mathrm{Lip}_p(f_{y_\gamma}) + |u_{\alpha\beta}| \ \mathrm{Lip}_p\left(\Phi_{y_\gamma}^{\alpha\beta}\right)\right] \ (|w| + |\nabla w|) + |w_{\alpha\beta}| \ \left|\Phi_{y_\gamma}^{\alpha\beta}\right| \right\} \ |v_\gamma| \\
& + \mathrm{Lip}_p\left(\Phi^{\alpha\beta}\right) \ (|w| + |\nabla w|) \ |v_{\alpha\beta}| \ . \qquad (4.15)
\end{aligned}
$$

The fundamental theorem of calculus in one variable gives

$$
\mathcal{N}(u+v) - \mathcal{N}(u) = \int_0^1 \left(\frac{d}{dt}\Big|_{t=0} \mathcal{N}(u+tv) \right) \, dt = \int_0^1 \, L_{u+tv} \, v \, dt \, . \qquad (4.16)
$$

Finally, combining this with (4.15) gives that (again at p)

$$
\begin{aligned}
|\mathcal{N}(u+v) - \mathcal{N}(u) - L_u v| & \leq \sup_{t \in [0,1]} |L_{u+tv} \, v - L_u \, v| \\
& \leq \left\{ \left|\Phi_s^{\alpha\beta}\right| + \left|\Phi_{y_\gamma}^{\alpha\beta}\right| + \mathrm{Lip}_p\left(\Phi^{\alpha\beta}\right) \right\} \ (|v| + |\nabla v|) \ |\mathrm{Hess}_v| \\
& + \left\{ \mathrm{Lip}_p(f_s) + |u_{\alpha\beta}| \ \mathrm{Lip}_p\left(\Phi_s^{\alpha\beta}\right) + \mathrm{Lip}_p(f_{y_\gamma}) + |u_{\alpha\beta}| \ \mathrm{Lip}_p\left(\Phi_{y_\gamma}^{\alpha\beta}\right) \right\} \ (|v| + |\nabla v|)^2 \ . \qquad (4.17)
\end{aligned}
$$

□

Proof of Lemma 4.3. By Lemma 4.6, $\mathcal{M}(u)$ is of the form (4.8). Since 0 is a critical point for F, we have $\mathcal{M}(0) = 0$. Therefore, Lemma 4.10 gives

$$
|\mathcal{M}(u) - Lu| \leq C_1 \ (|u| + |\nabla u|)^2 + C_2 \ (|u| + |\nabla u|) \, |\mathrm{Hess}_u| \ , \qquad (4.18)
$$

where the constant $C_1 = C_1(x)$ is bounded by $C\,(1+|x|)$ and the constant C_2 is uniformly bounded independent of $x \in \Sigma$ (these bounds follow from Lemma 4.6).

Integrating in space (against the Gaussian weight) gives

$$
\begin{aligned}
\|\mathcal{M}(u) - Lu\|_{L^2} & \leq C \ \left(\|(1+|x|)\, u^2\|_{L^2} + \|(1+|x|) \ |\nabla u|^2\|_{L^2}\right) + C_2 \, \| \, (|u| + |\nabla u|) \, |\mathrm{Hess}_u| \|_{L^2} \\
& \leq C \ \left(\|(1+|x|)\, u^2\|_{L^2} + \|(1+|x|) \ |\nabla u|^2\|_{L^2}\right) + C \, \|(1+|x|)^{-1} \, |\mathrm{Hess}_u|^2\|_{L^2} \ , \qquad (4.19)
\end{aligned}
$$

where the last inequality used the absorbing inequality

$$
2 \ (|u| + |\nabla u|) \, |\mathrm{Hess}_u| \leq (1+|x|) \ (|u| + |\nabla u|)^2 + (1+|x|)^{-1} \, |\mathrm{Hess}_u|^2 \, . \qquad (4.20)
$$

To get rid of the $|x|$'s in the first two terms, we use Lemma 3.4 to get

$$
\| |x| \, u^2\|_{L^2} \leq C \, \|u^2\|_{W^{1,2}} \leq C \, \|u^2 + |\nabla u|^2\|_{L^2} \ , \qquad (4.21)
$$

$$
\| |x| \, |\nabla u|^2\|_{L^2} \leq C \, \| |\nabla u|^2 + |\nabla_{\mathbf{R}^{n-k}} |\nabla u||^2 \, \|_{L^2} \, . \qquad (4.22)
$$

To get the first claim, we substitute these bounds back into (4.19)

$$
\|\mathcal{M}(u) - Lu\|_{L^2} \leq C \, \left\| u^2 + |\nabla u|^2 + |\nabla_{\mathbf{R}^{n-k}} |\nabla u||^2 + (1+|x|)^{-1} \, |\mathrm{Hess}_u|^2 \right\|_{L^2} \ . \qquad (4.23)
$$

To get the second claim, we first use the fundamental theorem of calculus and the definition of the gradient to get

$$F(u) - F(\mathcal{C}_k) - \frac{1}{2} \langle u, Lu \rangle_{L^2} = \int_0^1 \frac{d}{dt} \left[F(tu) - \frac{t^2}{2} \langle u, Lu \rangle_{L^2} \right] dt$$

$$= \int_0^1 \langle u, \mathcal{M}(tu) - t\, Lu \rangle_{L^2} \, dt \, . \tag{4.24}$$

Since the Cauchy-Schwarz inequality bounds the integrand by $\|u\|_{L^2} \, \|\mathcal{M}(tu) - t\, Lu\|_{L^2}$, the second claim now follows from the first. □

4.2. **The gradient Lojasiewicz inequality.** We will prove Proposition 4.1 and then use it to prove our gradient Lojasiewicz inequality using our first Lojasiewicz inequality.

Proof of Proposition 4.1. **Step 1**: *Cutting off to get a compactly supported perturbation of the cylinder.* Unlike this proposition, both Lemma 4.3 and the results of the previous section are for entire graphs over a cylinder. Thus, we fix a cutoff function ψ with $0 \leq \psi \leq 1$ that is one on $B_{\tilde{R}-1}$ and zero outside of $B_{\tilde{R}}$ and set

$$u = \psi \, \tilde{u} \, . \tag{4.25}$$

Observe that u has $\|u\|_{C^2} \leq C_n \, \|\tilde{u}\|_{C^2} \leq C_n \, \epsilon_0$, where C_n depends on the C^2 norm of ψ and, thus, depends only on n. Since ψ is supported in $B_{\tilde{R}}$ and $|\psi| \leq 1$, we have

$$\|u\|_{L^2}^2 \leq \|\tilde{u}\|_{L^2(B_{\tilde{R}})}^2 \, . \tag{4.26}$$

Finally, using the exponential decay of the Gaussian, we see that

$$|F(\Sigma) - F(\text{Graph}_u)| \leq C \, \lambda_0 \, \tilde{R}^{n-1} \, \mathrm{e}^{-\frac{(\tilde{R}-1)^2}{4}} \, , \tag{4.27}$$

$$\|\mathcal{M}(u)\|_{L^2} \leq C \, \|\phi\|_{L^2(B_{\tilde{R}})} + C_n \, \mathrm{e}^{-\frac{(\tilde{R}-1)^2}{8}} \, , \tag{4.28}$$

where ϕ here is the ϕ for Σ and C, C_n depend only on n.

Step 2: *The gradient Lojasiewicz inequality for the compact perturbation.* To simplify notation, define $F_0(u)$ by

$$F_0(u) = F(\text{Graph}_u) - F(\mathcal{C}_k) \, , \tag{4.29}$$

and, given a function v, set

$$\|v\|_2 \equiv \left\| v^2 + |\nabla v|^2 + \left|\text{Hess}_v(\cdot, \mathbf{R}^{n-k})\right|^2 + (1+|x|)^{-1} \, |\text{Hess}_v|^2 \right\|_{L^2} \, . \tag{4.30}$$

Assuming that $\|u\|_{C^2}$ is sufficiently small, then Lemma 4.3 gives C_1 so that

(L1) $\left| \|\mathcal{M}(u)\|_{L^2} - \|Lu\|_{L^2} \right| \leq C_1 \, \|u\|_2$.

(L2) $\left| F_0(u) - \frac{1}{2} \langle u, Lu \rangle_{L^2} \right| \leq C_1 \, \|u\|_{L^2} \, \|u\|_2$.

Here we also used the Kato inequality

$$\left|\nabla_{\mathbf{R}^{n-k}} |\nabla v|\right|^2 \leq \left|\text{Hess}_v(\cdot, \mathbf{R}^{n-k})\right|^2 \, . \tag{4.31}$$

We will divide into cases depending on the projection of u to the kernel $\mathcal{K}$ of L. Let C_1 be the constant from (L1) and (L2).

Case 1: Suppose first that u satisfies

$$\|u_{\mathcal{K}}\|_2 \leq \epsilon \, \|u^{\perp}\|_{W^{2,2}}^{1+\beta} \, , \tag{4.32}$$

where $\epsilon > 0$ will be chosen below and $\beta \in [0,1)$. Using the squared triangle inequality and then (4.32) plus[15] Lemma 3.22 gives

$$\|u\|_2 \leq 2\,\|u_{\mathcal{K}}\|_2 + 2\,\|u^{\perp}\|_2 \leq 2\,\left(\epsilon + C_0 \|u\|_{C^2}\right)\,\|u^{\perp}\|_{W^{2,2}} \, . \tag{4.33}$$

Using (L1) and (3.13) and then using (4.33) gives

$$\begin{aligned} \|\mathcal{M}(u)\|_{L^2} &\geq \|Lu\|_{L^2} - C_1\,\|u\|_2 \geq \mu\,\|u^{\perp}\|_{W^{2,2}} - C_1\,\|u\|_2 \\ &\geq \left(\mu - 2\,C_1\,[\epsilon + C_0\|u\|_{C^2}]\right)\,\|u^{\perp}\|_{W^{2,2}} \, . \end{aligned} \tag{4.34}$$

We now choose $\epsilon > 0$ and a bound for $\|u\|_{C^2}$ so that $2\,C_1\,[\epsilon + C_0\|u\|_{C^2}] \leq \frac{\mu}{2}$ and, thus,

$$\|\mathcal{M}(u)\|_{L^2} \geq \frac{\mu}{2}\,\|u^{\perp}\|_{W^{2,2}} \, . \tag{4.35}$$

We will show that $F_0(u)$ is higher order in $\|u^{\perp}\|_{W^{2,2}}$. Since L is symmetric and $Lu_{\mathcal{K}} = 0$, Cauchy-Schwarz and the bound on L from $W^{2,2}$ to L^2 by (3.12) give

$$|\langle u, Lu\rangle_{L^2}| = \left|\langle u^{\perp}, Lu^{\perp}\rangle_{L^2}\right| \leq C\,\|u^{\perp}\|_{L^2}\,\|u^{\perp}\|_{W^{2,2}} \, . \tag{4.36}$$

Substituting this into (L2), using that $\|u\|_2 \leq C\,\|u^{\perp}\|_{W^{2,2}}$ (by (4.33)), and applying the triangle inequality $\|u\|_{L^2} \leq \|u_{\mathcal{K}}\|_{L^2} + \|u^{\perp}\|_{L^2}$ gives

$$\begin{aligned} |F_0(u)| &\leq C\,\|u^{\perp}\|_{L^2}\,\|u^{\perp}\|_{W^{2,2}} + C\,\|u\|_{L^2}\,\|u\|_2 \\ &\leq C\,\|u^{\perp}\|_{L^2}\,\|u^{\perp}\|_{W^{2,2}} + C\,\|u_{\mathcal{K}}\|_{L^2}\,\|u^{\perp}\|_{W^{2,2}} \, . \end{aligned} \tag{4.37}$$

The first term on the right side is trivially bounded by $C\,\|u^{\perp}\|_{W^{2,2}}^2$. To bound the last term, we use that the cylinder has finite Gaussian area so that

$$\|u_{\mathcal{K}}\|_{L^2}^2 \leq C\,\|u_{\mathcal{K}}^2\|_{L^2} \leq C\,\|u_{\mathcal{K}}\|_2 \, , \tag{4.38}$$

to get that

$$\|u_{\mathcal{K}}\|_{L^2}\,\|u^{\perp}\|_{W^{2,2}} \leq C\,\|u_{\mathcal{K}}\|_2^{\frac{1}{2}}\,\|u^{\perp}\|_{W^{2,2}} \leq C\,\|u^{\perp}\|_{W^{2,2}}^{\frac{3+\beta}{2}} \, , \tag{4.39}$$

where the last inequality used (4.32). Putting all of this together (and noting that $\|u^{\perp}\|_{W^{2,2}}$ is bounded) gives

$$|F_0(u)| \leq C\,\|u^{\perp}\|_{W^{2,2}}^{\frac{3+\beta}{2}} \leq C\,\|\mathcal{M}(u)\|_{L^2}^{\frac{3+\beta}{2}} \, , \tag{4.40}$$

where the last inequality is (4.35). Combining this with the bound on $\|\mathcal{M}(u)\|_{L^2}$ from (4.28) gives

$$|F_0(u)| \leq C\|\phi\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{2}} + C\,\mathrm{e}^{-\frac{(3+\beta)(\tilde{R}-1)^2}{16}} \, . \tag{4.41}$$

Case 2: Suppose now that u satisfies

$$\|u_{\mathcal{K}}\|_2 > \epsilon\,\|u^{\perp}\|_{W^{2,2}}^{1+\beta} \, . \tag{4.42}$$

[15]Note that $\|u^{\perp}\|_{W^{2,2}}$ is small so $\|u^{\perp}\|_{W^{2,2}}^{1+\beta} \leq \|u^{\perp}\|_{W^{2,2}}$.

Lemma 3.22 gives C_0 so that

$$\|u^\perp\|_2 \leq C_0 \|u\|_{C^2} \|u^\perp\|_{W^{2,2}} \leq C \|u_{\mathcal{K}}\|_2^{\frac{1}{1+\beta}} , \tag{4.43}$$

where the last inequality is (4.42). Using the squared triangle inequality and (4.43) gives

$$\|u\|_2 \leq 2 \|u_{\mathcal{K}}\|_2 + 2 \|u^\perp\|_2 \leq C \|u_{\mathcal{K}}\|_2^{\frac{1}{1+\beta}} , \tag{4.44}$$

where the last inequality uses that $\|u_{\mathcal{K}}\|_2$ is bounded. Using (L2) and (3.12), then (4.42) and (4.44) (and the projection inequality $\|u^\perp\|_{L^2} \leq \|u\|_{L^2}$), we get

$$|F_0(u)| \leq C \|u^\perp\|_{L^2} \|u^\perp\|_{W^{2,2}} + C_1 \|u\|_{L^2} \|u\|_2 \leq 2\, C \|u\|_{L^2} \|u_{\mathcal{K}}\|_2^{\frac{1}{1+\beta}} . \tag{4.45}$$

However, since Lemma 3.22 and the projection inequality $\|u_{\mathcal{K}}\|_{L^2} \leq \|u\|_{L^2}$ give that

$$\|u_{\mathcal{K}}\|_2 \leq C_K \|u_{\mathcal{K}}\|_{L^2}^2 \leq C_K \|u\|_{L^2}^2 , \tag{4.46}$$

we conclude that

$$|F_0(u)| \leq C \|u\|_{L^2}^{\frac{3+\beta}{1+\beta}} \leq C \|\tilde{u}\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{1+\beta}} , \tag{4.47}$$

where the last inequality is the Gaussian L^2 bound on u from (4.26).

If we now combine the bounds from the two cases, then we see that

$$|F_0(u)| \leq C\|\phi\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{2}} + C\, \mathrm{e}^{-\frac{(3+\beta)(\tilde{R}-1)^2}{16}} + C \|\tilde{u}\|_{L^2(B_{\tilde{R}})}^{\frac{3+\beta}{1+\beta}} . \tag{4.48}$$

Finally, we use the triangle inequality to combine this with the bound (4.27) on the F functional from Step 2 and Step 3 to complete the proof.

□

We will use the following elementary lemma to control graphical bounds when we write a surface as a graph over two nearby cylinders.

Lemma 4.49. There exists $\epsilon_0 = \epsilon_0(n) > 0$ so that if $\Sigma_1, \Sigma_2 \in \mathcal{C}_k$, $5\sqrt{2n} \leq R_1 < R_2$ and

- $B_{R_1} \cap \Sigma$ is the graph of u_1 over Σ_1 with $|u_1| + |\nabla u_1| \leq \epsilon_0$,
- $B_{R_2} \cap \Sigma$ is the graph of u_2 over Σ_2 with $\|u_2\|_{C^{2,\alpha}} \leq \epsilon_0$,

then we get for $R = \min\{2\, R_1, R_2\}$ that

- $B_R \cap \Sigma$ is the graph of u_1 over Σ_1 with $\|u_1\|_{C^2} \leq \bar{\epsilon}$.

Proof. Since $B_{R_1} \cap \Sigma$ is ϵ_0 C^1-close to Σ_1 and ϵ_0 close to Σ_2, we get that the distance between Σ_1 and Σ_2 in B_{R_1} is at most $2\epsilon_0$. Since the distance between cylinders grows linearly in the radius, we conclude that the distance between Σ_1 and Σ_2 in B_R is at most $4\epsilon_0$. The lemma follows easily from this. □

Proof of Theorem 0.26. The result will follow by combining the L^2 closeness to a cylinder given by the first Lojasiewicz inequality and Proposition 4.1. Note that we can assume that R is large and $\|\phi\|_{L^2(B_R)}$ is small since the inequality is otherwise trivially true.

Step 1: *Fixing the nearby cylinder.* The Lojasiewicz inequality of Theorem 0.24 gives a cylinder $\Sigma_k \in \mathcal{C}_k$ so that $B_{\tilde{R}} \cap \Sigma$ is the graph of $\tilde{u}$ over Σ_k with $\|\tilde{u}\|_{C^1} \leq \epsilon_0$, where $b_{\ell,n} \in (0,1)$

satisfies $\lim_{\ell\to\infty} b_{\ell,n} = 1$, and[16]

$$\tilde{R} = \max \left\{ r \leq R \,\Big|\, R^{2n+5} \left\{ \mathrm{e}^{-b_{\ell,n}\frac{R^2}{8}} + \|\phi\|_{L^2(B_R)}^{\frac{b_{\ell,n}}{2}} \right\} \mathrm{e}^{\frac{r^2}{8}} \leq \tilde{C} \right\} , \tag{4.50}$$

where $\tilde{C}$ depends on $n, \lambda_0, \ell, C_\ell$. Combining this with Lemma 4.49, we extend $\tilde{u}$ out to $\bar{R} = \min\{2\tilde{R}, R\}$ so that

($\star_1$) $B_{\bar{R}} \cap \Sigma$ is the graph of $\tilde{u}$ over Σ_k with $\|\tilde{u}\|_{C^2} \leq \bar{\epsilon}$,

($\star_2$) $\|\tilde{u}\|^2_{L^2(B_{\bar{R}})} \leq C\, R^\rho \left\{ \|\phi\|^{b_{\ell,n}}_{L^2(B_R)} + \mathrm{e}^{-\frac{b_{\ell,n} R^2}{4}} \right\}$,

where $C = C(n, \ell, C_\ell, \lambda_0)$ and $\rho = \rho(n)$.

Step 2: *Using the first Lojasiewicz to get the second*. Proposition 4.1 gives

$$|F(\Sigma) - F(\mathcal{C}_k)| \leq C \left\{ \|\phi\|^{\frac{3+\beta}{2}}_{L^2(B_{\bar{R}})} + (1 + \bar{R}^{n-1})\, \mathrm{e}^{-\frac{(3+\beta)(\bar{R}-1)^2}{16}} + \|\tilde{u}\|^{\frac{3+\beta}{1+\beta}}_{L^2(B_{\bar{R}})} \right\} , \tag{4.51}$$

where $C = C(n, \lambda_0)$. To bound the last term in (4.51), we use ($\star_2$) to get

$$\|\tilde{u}\|^{\frac{3+\beta}{1+\beta}}_{L^2(B_{\bar{R}})} \leq C\, R^{\frac{3+\beta}{2+2\beta}\rho} \left\{ \|\phi\|^{b_{\ell,n}\frac{3+\beta}{2+2\beta}}_{L^2(B_R)} + \mathrm{e}^{-\frac{b_{\ell,n}(3+\beta)R^2}{8(1+\beta)}} \right\} . \tag{4.52}$$

To deal with the exponential term in (4.51), we consider two cases. Suppose first that $\bar{R} < R$, so that $\bar{R} = 2\tilde{R}$ and the definition of $\tilde{R}$ gives

$$R^{2n+5} \left\{ \mathrm{e}^{-b_{\ell,n}\frac{R^2}{8}} + \|\phi\|^{\frac{b_{\ell,n}}{2}}_{L^2(B_R)} \right\} \mathrm{e}^{\frac{\tilde{R}^2}{8}} = \tilde{C} . \tag{4.53}$$

Since $\bar{R} = 2\tilde{R}$ in this case, we have

$$\mathrm{e}^{-\frac{\bar{R}^2}{8}} = \left[\mathrm{e}^{-\frac{\tilde{R}^2}{8}} \right]^4 \leq C\, R^{8n+20} \left\{ \mathrm{e}^{-b_{\ell,n}\frac{R^2}{2}} + \|\phi\|^{2b_{\ell,n}}_{L^2(B_R)} \right\} . \tag{4.54}$$

We can assume that $\bar{R} > 4$ so that $\left(\frac{\bar{R}-1}{\bar{R}}\right)^2 > 1/2$. Raising (4.54) to the $\frac{3+\beta}{2}\left(\frac{\bar{R}-1}{\bar{R}}\right)^2 > \frac{3+\beta}{4}$ power, we bound the exponential term in (4.51) by a constant times a power of R times

$$\mathrm{e}^{-b_{\ell,n}\frac{(3+\beta)R^2}{8}} + \|\phi\|^{b_{\ell,n}\frac{(3+\beta)}{2}}_{L^2(B_R)} \leq \mathrm{e}^{-\frac{(3+\beta)(R-1)^2}{16}} + \|\phi\|^{b_{\ell,n}\frac{(3+\beta)}{2}}_{L^2(B_R)} , \tag{4.55}$$

where the last inequality used also that $b_{\ell,n}$ is close to one, and in particular at least 1/2. We proved this inequality in the case where $\bar{R} < R$, but it also obviously holds in the case when when $\bar{R} = R$ (and the ϕ term is unnecessary).

Putting it all together, $|F(\Sigma) - F(\mathcal{C}_k)|$ is bounded by $C\, R^{\rho'}$ times

$$\|\phi\|^{\frac{3+\beta}{2}}_{L^2(B_{\bar{R}})} + \left\{ \|\phi\|^{b_{\ell,n}\frac{3+\beta}{2+2\beta}}_{L^2(B_R)} + \mathrm{e}^{-\frac{b_{\ell,n}(3+\beta)R^2}{8(1+\beta)}} \right\} + \left\{ \mathrm{e}^{-\frac{(3+\beta)(R-1)^2}{16}} + \|\phi\|^{b_{\ell,n}\frac{(3+\beta)}{2}}_{L^2(B_R)} \right\} ,$$

where we have grouped terms together based on where they came from in (4.51). Finally, the first and fifth terms can be absorbed in the second term.

□

[16]We choose $\tilde{R}$ to make $\|\tilde{u}\|_{C^1}$ small by (2.51).

5. Compatibility of the shrinker and cylindrical scales

One of the main difficulties in this paper is that the singularities are not compact and, thus, surfaces cannot generally be written as entire graphs over a cylinder. As a result, our estimates include "error terms" coming from cut off functions. Thus, a surface is close to the cylinder if a large part of it can be written as a small graph over the cylinder.

Given a hypersurface $\Sigma \subset \mathbf{R}^{n+1}$, we will prove a lower bound for the scale on which it is "roughly cylindrical" in Theorem 5.3 below. This essentially bounds the error terms in our Lojasiewicz inequalities by a power greater than one of $|\nabla_\Sigma F|$, which is crucial in the next section when we prove uniqueness of tangent flows. It will also imply that the size of the graphical region is growing at a definite rate under the rescaled MCF.

5.1. **The cylindrical scale and the shrinker scale.** Recall that the cylindrical scale $\mathbf{r}_\ell(\Sigma)$ is the largest radius where Σ can be written as a small $C^{2,\alpha}$ graph over a cylinder with a uniform bound on $\nabla^\ell A$. Namely, given a fixed $\epsilon_0 > 0$, an integer ℓ and a constant C_ℓ, $\mathbf{r}_\ell(\Sigma)$ is the maximal radius where

- $B_{\mathbf{r}_\ell(\Sigma)} \cap \Sigma$ is the graph over a cylinder in $\mathcal{C}_k$ of a function u with $\|u\|_{C^{2,\alpha}} \leq \epsilon_0$ and $|\nabla^\ell A| \leq C_\ell$.

The constant ϵ_0 is fixed, but we have yet to choose ℓ and C_ℓ. (The constant ℓ will be chosen large to get good bounds on lower derivatives by interpolation and then C_ℓ will be chosen.)

The point of this section is to prove that these cylindrical scales are large enough that the error terms in our Lojasiewicz inequalities can be absorbed. The scale R that we have to beat[17] is roughly given by $\mathrm{e}^{-\frac{R^2}{4}} = |\nabla_\Sigma F|$. Thus, we define a "shrinker scale" $R(\Sigma)$ by

$$\mathrm{e}^{-\frac{R^2(\Sigma)}{2}} = |\nabla_\Sigma F|^2 \,, \tag{5.1}$$

with the convention that $R(\Sigma)$ is infinite when Σ is a complete shrinker. When Σ_t flows by the rescaled MCF, we define the shrinker scale (also denoted by $R(\Sigma_t)$) to be

$$\mathrm{e}^{-\frac{R^2(\Sigma_t)}{2}} = \int_{t-1}^{t+1} |\nabla_{\Sigma_s} F|^2 \, ds = F(\Sigma_{t-1}) - F(\Sigma_{t+1}) \,. \tag{5.2}$$

The main result of this section is the following theorem that shows that the cylindrical scale is a fixed factor larger than the shrinker scale :

Theorem 5.3. There exist $\mu > 0$ and C so that if Σ_t flows by the rescaled MCF and $\lambda(\Sigma_t) \leq \lambda_0$, then, given any ℓ, there exists C_ℓ (depending on ℓ) so that

$$(1+\mu)\, R(\Sigma_t) \leq \min_{t-1/2 \leq s \leq t+1} \mathbf{r}_\ell(\Sigma_s) + C \,. \tag{5.4}$$

To understand this, observe that Theorem 2.54 gives uniform graphical estimates on any scale less than $R(\Sigma)$. To apply Theorem 2.54, we need uniform curvature bounds and a lower bound for H on this larger scale. We will establish these uniform bounds on the larger scale by an extension and improvement argument, where Theorem 2.54 gives uniform bounds on larger scales (this is the improvement part). Roughly speaking, the extension argument will

[17] To see this, note that the larger exponentially decaying term on the right in Theorem 0.26 is essentially $\mathrm{e}^{-\frac{R^2}{4}}$. We need to bound this by a power greater than one of $|\nabla_\Sigma F| = \|\phi\|_{L^2}$.

use curvature estimates for MCF to get bounds on a larger scale forward in time, then use the bounds on ϕ to pull these bounds backwards in time under the rescaled MCF. Repeating this gets us as close as we want to the scale $R(\Sigma)$ and gets us uniform curvature bounds on a larger scale than $R(\Sigma)$. The final step is to get graphical estimates on a larger scale too; for this, we cannot use Theorem 2.54. Rather, we get these graphical estimates from estimates for MCF and a scaling argument to relate MCF and rescaled MCF.[18]

5.2. **Backward curvature estimates.** Recall that, when $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface, ϕ is defined to be $\phi = \frac{\langle x, \mathbf{n} \rangle}{2} - H$, so that Σ_t flows by the rescaled MCF if $\partial_t x = \phi \, \mathbf{n}$ and

$$|\nabla_\Sigma F|^2 = \|\phi\|_{L^2}^2 \equiv \int_\Sigma \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{5.5}$$

In this subsection, we will show the following curvature estimate for the rescaled MCF:

Proposition 5.6. Given n, λ_0, there exists $s \geq 2$ and $\delta > 0$ so that the following holds: Given $1/2 \geq \tau > 0$, there exists $\mu > 0$, such that if Σ_t flows by the rescaled MCF, $\lambda_0 \geq \lambda(\Sigma_t)$, $t_2 \geq t_1 + \tau$, $x_0 \in B_{R-s}$ and

$$\int_{t_1}^{t_2} \int_{B_{R+2} \cap \Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq \frac{\mu^2 \, \mathrm{e}^{-\frac{(R+2)^2}{4}}}{R^2 \, (t_2 - t_1 + 1)} \, , \tag{5.7}$$

$$\sup_{B_{s\sqrt{\tau}}(x_0) \cap \Sigma_{t_2}} |A|^2 \leq \delta \, \tau^{-1} \, , \tag{5.8}$$

then for all $t \in [t_1 - \log(1 - 7\tau/8), t_1 - \log(1 - \tau)]$ and any ℓ

$$\sup_{B_{\frac{\sqrt{\tau}}{3}}\left(\mathrm{e}^{\frac{1}{2}(t-t_1)} \, x_0\right) \cap \Sigma_t} \left\{ |A|^2 + \tau^\ell \, \left|\nabla^\ell A\right|^2 \right\} \leq \frac{C_\ell}{\tau} \, , \tag{5.9}$$

where C_ℓ depends on n and ℓ.

We will need the following elementary lemma:

Lemma 5.10. If Σ_t flows by the rescaled MCF and $0 \leq \eta \leq 1$ is a smooth compactly-supported function on $\mathbf{R}^{n+1}$, then for $t_1 \leq t_2$ and $\tau > 0$

$$\begin{aligned} &\int_{\Sigma_{t_2}} \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} - \int_{\Sigma_{t_1}} \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \\ &\quad = \int_{t_1}^{t_2} \int_{\Sigma_t} \langle \nabla \eta, \mathbf{n} \rangle \, \phi \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2} \int_{\Sigma_t} \frac{\langle x_0, \mathbf{n} \rangle}{2\tau} \, \phi \, \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \\ &\quad\quad + \left(1 - \frac{1}{\tau}\right) \int_{t_1}^{t_2} \int_{\Sigma_t} \frac{\langle x, \mathbf{n} \rangle}{2} \, \phi \, \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} - \int_{t_1}^{t_2} \int_{\Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \, . \end{aligned} \tag{5.11}$$

[18] This final step can not be iterated since it has a loss in the estimates and we can no longer apply Theorem 2.54 to get rid of the loss.

Proof. If $f \in \mathbf{R}^{n+1} \to \mathbf{R}$ is a smooth function with compact support, then

$$\begin{aligned} \frac{d}{dt} \int_{\Sigma_t} f \, \mathrm{e}^{-\frac{|x|^2}{4}} &= \int_{\Sigma_t} \langle \nabla f, \mathbf{n} \rangle \, \phi \, \mathrm{e}^{-\frac{|x|^2}{4}} - \int_{\Sigma_t} f \, \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \\ &= \int_{\Sigma_t} \langle \nabla \log f, \mathbf{n} \rangle \, \phi \, f \, \mathrm{e}^{-\frac{|x|^2}{4}} - \int_{\Sigma_t} \phi^2 \, f \, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \end{aligned} \tag{5.12}$$

If we set $f(x) = \eta \, \mathrm{e}^{\frac{|x|^2}{4}} \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}}$, then

$$\nabla \log f = \nabla \log \eta + \frac{x}{2} - \frac{x - x_0}{2\tau} = \nabla \log \eta + \frac{x_0}{2\tau} + \frac{1}{2} \left(1 - \frac{1}{\tau} \right) x \, . \tag{5.13}$$

Therefore

$$\begin{aligned} \frac{d}{dt} \int_{\Sigma_t} \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} &= \int_{\Sigma_t} \langle \nabla \eta, \mathbf{n} \rangle \, \phi \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} + \int_{\Sigma_t} \frac{\langle x_0, \mathbf{n} \rangle}{2\tau} \, \phi \, \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \\ &+ \left(1 - \frac{1}{\tau} \right) \int_{\Sigma_t} \frac{\langle x, \mathbf{n} \rangle}{2} \, \phi \, \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} - \int_{\Sigma_t} \phi^2 \, \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \, . \end{aligned} \tag{5.14}$$

The lemma now follows by integrating from t_1 to t_2. □

Corollary 5.15. Given $\epsilon > 0$, $1 \geq \tau > 0$, and λ_0, there exists $\mu = \mu(\epsilon, \tau, \lambda_0) > 0$, $s = s(\epsilon, \lambda_0) \geq 2$ such that if Σ_t flows by the rescaled MCF, $\lambda_0 \geq \lambda(\Sigma_t)$, $x_0 \in B_{R-s}$, and $t_2 > t_1$

$$\int_{t_1}^{t_2} \int_{B_{R+2} \cap \Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq \frac{\mu^2 \, \mathrm{e}^{-\frac{(R+2)^2}{4}}}{R^2 \, (t_2 - t_1 + 1)} \, , \tag{5.16}$$

$$(4\pi\tau)^{-\frac{n}{2}} \int_{\Sigma_{t_2}} \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \leq 1 + \frac{\epsilon}{2} \, , \tag{5.17}$$

then

$$(4\pi\tau)^{-\frac{n}{2}} \int_{\Sigma_{t_1}} \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \leq 1 + \epsilon \, . \tag{5.18}$$

Proof. Observe first that by the entropy bound $\lambda(\Sigma_t) \leq \lambda_0$, there exists $s > 0$ such that for all $y \in \mathbf{R}^{n+1}$ and all t

$$(4\pi\tau)^{-\frac{n}{2}} \int_{\Sigma_t \setminus B_{s\sqrt{\tau}}(y)} \mathrm{e}^{-\frac{|x-y|^2}{4\tau}} \leq \frac{\epsilon}{4} \, . \tag{5.19}$$

If we choose a non-negative function η with $\eta \leq 1$, $|\nabla \eta| \leq 1$, $\eta = 1$ on B_R, and $\eta = 0$ outside B_{R+2}, then Lemma 5.10 gives

$$\begin{aligned} \int_{B_R \cap \Sigma_{t_1}} \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} &\leq \int_{\Sigma_{t_1}} \eta \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \leq \int_{B_{R+2} \cap \Sigma_{t_2}} \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2} \int_{(B_{R+2} \setminus B_R) \cap \Sigma_t} |\phi| \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \\ &+ \left(\frac{1}{\tau} - 1 \right) \int_{t_1}^{t_2} \int_{B_{R+2} \cap \Sigma_t} \frac{|\langle x, \mathbf{n} \rangle|}{2} \, |\phi| \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2} \int_{B_{R+2} \cap \Sigma_t} \frac{|\langle x_0, \mathbf{n} \rangle|}{2\tau} \, |\phi| \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \\ &+ \int_{t_1}^{t_2} \int_{B_{R+2} \cap \Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \, . \end{aligned} \tag{5.20}$$

Combining the terms that are linear in ϕ gives

$$\int_{B_R\cap\Sigma_{t_1}} e^{-\frac{|x-x_0|^2}{4\tau}} \leq \int_{B_{R+2}\cap\Sigma_{t_2}} e^{-\frac{|x-x_0|^2}{4\tau}} + \left(1+\frac{\frac{|x_0|}{\tau}+\left(\frac{1}{\tau}-1\right)(R+2)}{2}\right)\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}|\phi|\, e^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2$$

$$\leq \int_{B_{R+2}\cap\Sigma_{t_2}} e^{-\frac{|x-x_0|^2}{4\tau}} + \frac{R+2}{\tau}\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}|\phi|\, e^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\,. \tag{5.21}$$

By the entropy bound $\lambda(\Sigma_t)\leq\lambda_0$ and the Cauchy-Schwarz inequality, we have

$$\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}|\phi|\, e^{-\frac{|x-x_0|^2}{4\tau}} \leq \left(\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\, e^{-\frac{|x-x_0|^2}{4\tau}}\right)^{\frac{1}{2}}\sqrt{(4\pi\tau)^{\frac{n}{2}}(t_2-t_1)\,\lambda_0}$$

$$\leq \sqrt{(4\pi\tau)^{\frac{n}{2}}(t_2-t_1)\,\lambda_0}\left(\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\right)^{\frac{1}{2}} \tag{5.22}$$

$$\leq \sqrt{(4\pi\tau)^{\frac{n}{2}}(t_2-t_1)\,\lambda_0}\, e^{\frac{(R+2)^2}{8}}\left(\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\, e^{-\frac{|x|^2}{4}}\right)^{\frac{1}{2}}.$$

We can therefore bound the two last terms in (5.21) as follows

$$\frac{R+2}{\tau}\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}|\phi|\, e^{-\frac{|x-x_0|^2}{4\tau}} + \int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2$$

$$\leq C\,\frac{R+2}{\tau}\sqrt{(4\pi\tau)^{\frac{n}{2}}(t_2-t_1)}\, e^{\frac{(R+2)^2}{8}}\left(\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\, e^{-\frac{|x|^2}{4}}\right)^{\frac{1}{2}} + e^{\frac{(R+2)^2}{4}}\int_{t_1}^{t_2}\int_{B_{R+2}\cap\Sigma_t}\phi^2\, e^{-\frac{|x|^2}{4}}$$

$$\leq C\,(\mu/\tau+\mu^2)\,. \tag{5.23}$$

Using (5.19), (5.21), and (5.23) we get that

$$(4\pi\tau)^{-\frac{n}{2}}\int_{\Sigma_{t_1}} e^{-\frac{|x-x_0|^2}{4\tau}} = (4\pi\tau)^{-\frac{n}{2}}\int_{B_R\cap\Sigma_{t_1}} e^{-\frac{|x-x_0|^2}{4\tau}} + (4\pi\tau)^{-\frac{n}{2}}\int_{\Sigma_{t_1}\setminus B_R} e^{-\frac{|x-x_0|^2}{4\tau}}$$

$$\leq (4\pi\tau)^{-\frac{n}{2}}\int_{\Sigma_{t_2}} e^{-\frac{|x-x_0|^2}{4\tau}} + C\,(\mu/\tau+\mu^2)+\frac{\epsilon}{4}\,. \tag{5.24}$$

Choosing μ sufficiently small gives the corollary. □

We will apply this corollary in combination with Brakke's regularity result to get curvature estimates at an earlier time-slice in terms of curvature estimates at a later time-slice. By White's [W3] version of Brakke's regularity result [B], there exist constants ϵ and C_B depending on n and λ_0 such that if $M_s\subset\mathbf{R}^{n+1}$ flow ($s<0$) by the MCF, $\lambda(M_s)\leq\lambda_0$, and for some $s_0<0$

$$(-4\pi s_0)^{-\frac{n}{2}}\int_{M_{s_0}} e^{\frac{|x-x_0|^2}{4s_0}} \leq 1+\epsilon\,, \tag{5.25}$$

then for all $s \in [-\frac{s_0}{4}, 0]$

$$\sup_{M_s \cap B_{\frac{1}{2}\sqrt{-s_0}}(x_0)} |A|^2 \leq \frac{C_B}{-s_0}\,. \tag{5.26}$$

We can use the correspondence between MCF and rescaled MCF to translate this into a similar curvature estimate for rescaled MCF. Namely, if Σ_t is a rescaled MCF with entropy at most λ_0 and there is some $\tau \in (0, 1/2)$ so that

$$(4\pi\tau)^{-\frac{n}{2}} \int_{\Sigma_{t_0}} \mathrm{e}^{-\frac{|x-x_0|^2}{4\tau}} \leq 1 + \epsilon\,, \tag{5.27}$$

then for all $t \in [t_0 - \log(1 - 3\tau/4), t_0 - \log(1 - \tau)]$ we have

$$\sup_{\Sigma_t \cap B_{\frac{\sqrt{\tau}}{2}}\left(\mathrm{e}^{\frac{1}{2}(t-t_0)} x_0\right)} |A|^2 \leq \frac{C_B}{\tau}\,. \tag{5.28}$$

This is proven by writing the rescaled flow Σ_t as $\mathrm{e}^{\frac{1}{2}(t-t_0)} M_s$ where where $s = 1 - \mathrm{e}^{t_0 - t} - \tau$ and M_s is the MCF with $M_{-\tau} = \Sigma_{t_0}$. (Here we have used that the result of Brakke/White is uniform in Σ or more precise uniform in the point x_0 where it is centered as for the rescaled MCF when the point x_0 is fixed this mean that the original "fixed" point x_0 for the MCF evolves by $\mathrm{e}^{\frac{1}{2}(t-t_0)} x_0$.)

Proof of Proposition 5.6. Combining the above consequence of Brakke's theorem with Corollary 5.15 gives the $|A|$ bound in Proposition 5.6 for t in the time interval

$$[t_1 - \log(1 - 3\tau/4), t_1 - \log(1 - \tau)]\,. \tag{5.29}$$

The bounds on higher derivatives of A then follow from this and the interior estimates of Ecker and Huisken, [EH]. □

5.3. **A mean value inequality.** In the next lemma, we will use that if Σ_t flow by the rescaled MCF, then (see section 2 of [CIMW])

$$(\partial_t - L)\,\phi = 0 \text{ where } L = \mathcal{L} + |A|^2 + \frac{1}{2}\,. \tag{5.30}$$

Hence,

$$(\partial_t - \mathcal{L})\,\phi^2 = 2\,\phi\,(\partial_t - \mathcal{L})\,\phi - 2\,|\nabla\phi|^2 = \phi^2\,(2\,|A|^2 + 1) - 2\,|\nabla\phi|^2\,. \tag{5.31}$$

Lemma 5.32. There exists a constant C so that if Σ_t flow by the rescaled MCF for $t \in [t_1, t_2]$, $r + 1 \leq \min_{t_1 \leq s \leq t_2} \mathbf{r}(\Sigma_s)$ and $0 < \beta < (t_2 - t_1)/2$, then

$$\max_{s \in [t_1+\beta, t_2]} |\nabla_{\Sigma_s} F|^2_{B_r} \leq (C + 1/\beta)\,(F(\Sigma_{t_1}) - F(\Sigma_{t_2}))\,. \tag{5.33}$$

$$\int_{t_1+\beta}^{t_2} \int_{B_r \cap \Sigma_s} |\nabla\phi|^2\,\mathrm{e}^{-\frac{|x|^2}{4}} \leq (C + 1/\beta)\,(F(\Sigma_{t_1}) - F(\Sigma_{t_2}))\,. \tag{5.34}$$

Proof. Fix a compactly supported function η on $\mathbf{R}^{n+1}$ with $1 \leq \eta \leq 0$, η identically one on B_r, η vanishes outside B_{r+1}, and $|\nabla\eta| \leq 2$. If we restrict η to Σ_t, then the flow equation and (5.31) give

$$\partial_t\,\left(\phi^2\,\eta^2\right) = (\eta^2)_t\,\phi^2 + \eta^2\,\partial_t\phi^2 = \phi^3\,\langle\nabla\eta^2, \mathbf{n}\rangle + \eta^2\,\left(\mathcal{L}\phi^2 + \phi^2\,(2\,|A|^2 + 1) - 2\,|\nabla\phi|^2\right)\,. \tag{5.35}$$

Using this and the equation for the derivative of the weighted measure, and integrating by parts to take $\mathcal{L}$ off of ϕ^2, we get

$$\partial_t \left(\int_{\Sigma_t} \phi^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \right) = - \int_{\Sigma_t} \phi^4 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} + \int_{\Sigma_t} \left(2 \, |A|^2 + 1 \right) \phi^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}}$$
$$- 2 \int_{\Sigma_t} |\nabla \phi|^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} + \int_{\Sigma_t} \phi^3 \, \langle \nabla \eta^2, \mathbf{n} \rangle \, \mathrm{e}^{-\frac{|x|^2}{4}} - \int_{\Sigma_t} \langle \nabla \phi^2, \nabla \eta^2 \rangle \, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{5.36}$$

Using the absorbing inequalities $2\phi^3 \eta |\nabla \eta| \leq \phi^4 \eta^2 + \phi^2 |\nabla \eta|^2$ and $4 \, \eta |\phi| |\nabla \eta| |\nabla \phi| \leq \eta^2 |\nabla \phi|^2 + 4\phi^2 |\nabla \eta|^2$, we get

$$\partial_t \left(\int_{\Sigma_t} \phi^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \right) \leq \int_{\Sigma_t} \left\{ \left(2 \, |A|^2 + 1 \right) \eta^2 + 5 |\nabla \eta|^2 \right\} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} - \int_{\Sigma_t} |\nabla \phi|^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}}$$
$$\leq C \int_{\Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} - \int_{\Sigma_t} |\nabla \phi|^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \, . \tag{5.37}$$

Suppose that $s \in [t_1 + \beta, t_2]$. To prove (5.33), we integrate (5.37) to get

$$\int_{\Sigma_s} \phi^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq \min_{[t_1, t_1 + \beta]} \int_{\Sigma_t} \phi^2 \, \eta^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} + C \int_{t_1}^{s} \int_{\Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq (C + 1/\beta) \int_{t_1}^{s} \int_{\Sigma_t} \phi^2 \, \mathrm{e}^{-\frac{|x|^2}{4}}$$
$$\leq (C + 1/\beta) \, \left(F(\Sigma_{t_1}) - F(\Sigma_{t_2}) \right) \, . \tag{5.38}$$

Finally, to get (5.34), we integrate (5.37) from $t_1 + \beta$ to t_2 and use (5.33) to bound the contributions at the end points. □

5.4. **Uniform short time stability of the cylinder.** The last result that we will need for proving Theorem 5.3 is the following elementary short time uniform stability of the cylinder under MCF with bounded curvature:

Lemma 5.39. Given $R > \sqrt{2n}$, $\epsilon > 0$ and C_0, there exist $\delta > 0$ and $\theta > 0$ so that if M_t is a MCF with

(1) $B_{R+2} \cap M_{-1}$ is a $C^{2,\alpha}$ graph over $\Sigma \in \mathcal{C}_k$ with norm at most δ.
(2) $|A| + |\nabla A| + |\nabla^2 A| + |\nabla^3 A| \leq C_0$ on $B_{R+2} \cap M_t$ for $t \in [-1 - 1/C_0, -1 + 1/C_0]$.

Then for each $t \in [-1, \theta - 1]$ we have that

- $B_R \cap M_t$ is a $C^{2,\alpha}$ graph over $\sqrt{-t}\,\Sigma$ with norm at most ϵ.

Proof. Since $|A|$ is bounded, the MCF equation implies that $|\partial_t x|$ is also bounded. Likewise, the bound on $|\nabla A|$ (and thus on $|\nabla H|$) and the evolution equation for the normal (see lemma 7.5 in [HP]) imply that $|\partial_t \mathbf{n}|$ is also uniformly bounded. Combining these two bounds, it follows that $B_{R+1} \cap M_t$ remains a graph over Σ of a function u with a uniform bound

$$|\partial_t u| + |\partial_t \nabla u| \leq C_1 \text{ for } t \in [-1 - \theta_0, -1 + \theta_0] \, , \tag{5.40}$$

where $\theta_0 > 0$ and C_1 depend on C_0, ϵ, n. Similarly, the higher derivative bounds on A then yield bounds on higher derivatives of u and the lemma follows immediately. □

5.5. **Proof of Theorem 5.3.** We are now prepared to prove Theorem 5.3 which shows that the cylindrical scale is a fixed factor larger than the shrinker scale.

Proof. (of Theorem 5.3). The theorem follows by an extension and improvement argument that is inspired by a similar argument for shrinkers in [CIM].

(1) Extending the scale. Given ℓ, we will show that there exist $\delta > 0$, $\bar{s} > 0$, $\theta > 0$, R_0, C_2 and C_ℓ so that if

(A1) $B_R \cap \Sigma_s$ is a graph of u_1 over some $\Sigma_1 \in \mathcal{C}_k$ with $\|u_1\|_{C^{2,\alpha}} \leq \delta$ for each $s \in [t_0 - \bar{s}, t_0 + \bar{s}]$ for some $R \in [R_0, R(\Sigma_t)]$ and $t_0 \in [t - 1/2, t + 1 - \bar{s}]$

then, for every $s \in [t_0 - \bar{s}, t_0 + \bar{s}]$, we have

(A2) $\mathbf{r}_\ell(\Sigma_t) \geq (1+\theta)R$ and $|\nabla_{\Sigma_s} F|^2_{B_{(1+\theta)R}} \leq C_2 \, (F(\Sigma_{t-1}) - F(\Sigma_{t+1}))$.

The key observation is that the cylindrical estimates and global entropy bound imply that the local Gaussian densities on some fixed scale are almost one. Thus, White's Brakke estimate [W3] gives a curvature bound on a larger region $B_{(1+\kappa)R}$ with $\kappa > 0$ but at the cost of moving forward in time. However, Proposition 5.6 pulls this curvature bound backwards in time *while only coming in by a fixed additive amount*. As long as R is sufficiently large, the multiplicative gain beats the additive loss and, thus, the bound on A extends to a larger scale with no loss in time. Ecker-Huisken [EH] then gives uniform higher derivative bounds on A. We can now apply Lemma 5.39 on a unit scale but centered at points out to the extended scale to get the cylindrical estimates on the larger scale. Finally, using the curvature bounds, Lemma 5.32 gives the constant C_2 so that $|\nabla_{\Sigma_s} F|^2_{B_{(1+\theta)R}} \leq C_2 \, (F(\Sigma_{t-1}) - F(\Sigma_{t+1}))$.

(2) The improvement below the shrinker scale. The Lojasiewicz inequality of Theorem 2.54 will give an improved bound on the larger scale if we are below the shrinker scale: Given $\tau > 0$, $\delta > 0$, C_2, ℓ and C_ℓ, there exist ℓ_1 and R_1 so that if $\ell \geq \ell_1$ and $R \in [R_1, R(\Sigma_t)]$ satisfies

$$R \leq r_\ell(\Sigma_s) \text{ and } |\nabla_{\Sigma_s} F|^2_{B_R} \leq C_2 \, (F(\Sigma_{t-1}) - F(\Sigma_{t+1})) \, , \tag{5.41}$$

then $B_{(1-\tau)R} \cap \Sigma_s$ is a graph of u_3 over some $\Sigma_3 \in \mathcal{C}_k$ with $\|u_3\|_{C^{2,\alpha}} \leq \delta$.

Putting it together: The point is to choose τ much smaller than θ, so that the gain in scale from extending in (1) beats the loss in scale from the improvement in (2). We can then apply the two steps iteratively to get a fixed factor greater than one beyond the shrinker scale, giving the theorem. □

6. The gradient Lojasiewicz inequality and uniqueness

In this section, we will use the gradient Lojasiewicz inequality of Theorem 0.26 and the compatibility of the shrinker and cylindrical scales of the previous section to prove a gradient Lojasiewicz inequality for rescaled MCF. We will show that this inequality implies uniqueness of the tangent flow at a cylindrical singularity, thus completing the proof of Theorem 0.2.

6.1. **Mean value inequalities.** In this subsection, we will prove a mean value inequality that is needed for the gradient Lojasiewicz inequality. The argument follows that of Lemma 5.32 in the previous section with the gradient of ϕ in place of ϕ essentially using the equation that one gets from taking the derivative of equation (5.30) to get the following:

Lemma 6.1. There exists a constant C so that if Σ_t flow by the rescaled MCF, and $r \leq \min_{t-1/2 \leq s \leq t+1} \mathbf{r}_\ell(\Sigma_s)$, then

$$\max_{s \in [t-\frac{1}{4}, t+1]} \int_{B_r \cap \Sigma_t} |\nabla \phi|^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq C \, (1+r) \, (F(\Sigma_{t-1}) - F(\Sigma_{t+1})) \, . \tag{6.2}$$

$$\int_{t-\frac{1}{4}}^{t+1} \int_{B_r \cap \Sigma_s} |\mathrm{Hess}_\phi|^2 \, \mathrm{e}^{-\frac{|x|^2}{4}} \leq C \, (1+r) \, (F(\Sigma_{t-1}) - F(\Sigma_{t+1})) \, . \tag{6.3}$$

Lemma 6.1 will follow from the same argument as in the proof of Lemma 5.32 (together with the result of Lemma 5.32) provided we have the following:

Lemma 6.4. If Σ_t flow by the rescaled MCF, then

$$(\partial_t - \mathcal{L}) \, |\nabla \phi|^2 \leq -2 \, |\mathrm{Hess}_\phi|^2 + C \, |\nabla \phi|^2 + \phi^2 \, , \tag{6.5}$$

where C depends only on n and the bounds for A and ∇A.

Proof. To prove this, note first that if $\Sigma \subset \mathbf{R}^{n+1}$ is a hypersurface, $f : \mathbf{R}^{n+1} \to \mathbf{R}$ is a smooth function, $X, Y \in T_x\Sigma$, then

$$\mathrm{Hess}_f^{\mathbf{R}^{n+1}}(X,Y) = \langle \nabla_X \left(\nabla^T f + \langle \nabla f, \mathbf{n} \rangle \, \mathbf{n} \right), Y \rangle = \mathrm{Hess}_f^{\Sigma}(X,Y) - \langle \nabla f, \mathbf{n} \rangle \, A(X,Y) \, . \tag{6.6}$$

Recall also that if Σ is a manifold (not necessarily embedded in Euclidean space), then the Bochner formula for the drift Laplacian $\Delta_f \, u = \Delta \, u - \langle \nabla f, \nabla u \rangle$ is

$$\frac{1}{2} \Delta_f \, |\nabla u|^2 = |\mathrm{Hess}_u|^2 + \langle \nabla \Delta_f u, \nabla u \rangle + \mathrm{Ric}_f(\nabla u, \nabla u) \, . \tag{6.7}$$

Here $\mathrm{Ric}_f = \mathrm{Ric} + \mathrm{Hess}_f$ is the Bakry-Émery Ricci curvature.

We will use that if $\Sigma_t \subset \mathbf{R}^{n+1}$ is a one-parameter family of hypersurfaces moving by the rescaled MCF and $u = u(x,t) : \Sigma_t \times \mathbf{R} \to \mathbf{R}$ is a smooth function, then

$$\partial_t \, |\nabla^T u|^2 = 2 \, \langle \nabla^T \partial_t u, \nabla^T u \rangle + 2 \, \left(\frac{\langle x, \mathbf{n} \rangle}{2} - H \right) \, A(\nabla^T u, \nabla^T u) \, . \tag{6.8}$$

Proof of (6.8): To see this, extend u to a function on $\mathbf{R}^{n+1} \times \mathbf{R}$ so that on Σ_t $\nabla u = \nabla^T u$ and $\partial_t u = \langle \nabla u, \partial_t x \rangle + u_t = u_t$, where u_t is the t derivative of u as a function on $\mathbf{R}^{n+1} \times \mathbf{R}$ and the rescaled MCF equation is

$$\partial_t x = \left(\frac{1}{2} \, \langle x, \mathbf{n} \rangle - H \right) \, \mathbf{n} \, . \tag{6.9}$$

Therefore, differentiating $\nabla u = \nabla u(x(t), t)$ on Σ_t, the chain rule gives

$$\partial_t \nabla u = \nabla_{\partial_t x} \nabla u + \nabla u_t = \left(\frac{1}{2} \, \langle x, \mathbf{n} \rangle - H \right) \, \nabla_{\mathbf{n}} \nabla u + \nabla \partial_t u \, . \tag{6.10}$$

Using this, the symmetry of the Hessian, and the definition of A gives

$$\frac{1}{2} \, \partial_t \, |\nabla^T u|^2 = \frac{1}{2} \, \partial_t \, |\nabla u|^2 = \langle \partial_t \nabla u, \nabla u \rangle = \left(\frac{1}{2} \, \langle x, \mathbf{n} \rangle - H \right) \, \langle \nabla_{\nabla u} \nabla u, \mathbf{n} \rangle + \langle \nabla \partial_t u, \nabla u \rangle$$

$$= \left(\frac{1}{2} \, \langle x, \mathbf{n} \rangle - H \right) \, A(\nabla u, \nabla u) + \langle \nabla \partial_t u, \nabla u \rangle \, , \tag{6.11}$$

completing the proof of (6.8).

Let $f = \frac{|x|^2}{4}$ so that (6.6) gives

$$\text{Ric}_f(\nabla u, \nabla u) = \text{Ric}(\nabla u, \nabla u) + \frac{|\nabla u|^2}{2} + \frac{\langle x, \mathbf{n}\rangle}{2} A(\nabla u, \nabla u)\,. \tag{6.12}$$

Therefore, combining (6.8) and the Bochner formula (6.7) gives

$$\begin{aligned}(\partial_t - \mathcal{L})\,|\nabla u|^2 &= -2\,|\text{Hess}_u|^2 + 2\,\langle \nabla (\partial_t - \mathcal{L})\,u, \nabla u\rangle \\ &\quad - 2\,\text{Ric}(\nabla u, \nabla u) - |\nabla u|^2 - 2\,H\,A(\nabla u, \nabla u)\,.\end{aligned} \tag{6.13}$$

Suppose now that $u = \phi$, so that $(\partial_t - \mathcal{L})\,\phi = (|A|^2 + \frac{1}{2})\,\phi$ by (5.30). Finally, (6.5) follows from (6.13) and the absorbing inequality $2\,|\phi\,\langle \nabla |A|^2, \nabla \phi\rangle| \leq \phi^2 + C\,|\nabla \phi|^2$. □

6.2. **A discrete gradient Lojasiewicz inequality for rescaled MCF.** The next theorem gives a discrete version of a gradient Lojasiewicz inequality for rescaled MCF.

Theorem 6.14. Given n and λ_0, there exist constants $K, \bar{R}, \epsilon$ and $\tau \in (1/3, 1)$ so that if Σ_s is a rescaled MCF for $s \in [t-1, t+1]$ satisfying

- $\lambda(\Sigma_s) \leq \lambda_0$.
- $B_{\bar{R}} \cap \Sigma_s$ is a $C^{2,\alpha}$ graph over some cylinder in $\mathcal{C}_k$ with norm at most ϵ for each s.

Then we have

$$(F(\Sigma_t) - F(\mathcal{C}))^{1+\tau} \leq K\,\left(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\right)\,. \tag{6.15}$$

Proof. Given any $\beta \in [0, 1)$ and $R \in [1, \mathbf{r}_\ell(\Sigma_t) - 2]$, Theorem 0.26 gives

$$|F(\Sigma_t) - F(\mathcal{C})| \leq C\,R^\rho\,\left\{ \|\phi\|_{L^2(B_R \cap \Sigma_t)}^{c_{\ell,n}\,\frac{3+\beta}{2+2\beta}} + \mathrm{e}^{-\frac{R^2}{4}\,c_{\ell,n}} + \mathrm{e}^{-\frac{R^2}{4}\,\left(\frac{3+\beta}{4}\right)} \right\}\,, \tag{6.16}$$

where $C = C(n, \ell, C_\ell, \lambda_0)$, $\rho = \rho(n)$ and $c_{\ell,n} \in (0,1)$ satisfies $\lim_{\ell \to \infty} c_{\ell,n} = 1$. We will bound each term by a power greater than 1/2 of $(F(\Sigma_{t-1}) - F(\Sigma_{t+1}))$.

We defined the shrinker scale $R(\Sigma_t)$ in (5.2) by

$$\mathrm{e}^{-\frac{R^2(\Sigma_t)}{2}} = \int_{t-1}^{t+1} |\nabla_{\Sigma_s} F|^2\,ds = F(\Sigma_{t-1}) - F(\Sigma_{t+1})\,. \tag{6.17}$$

If we set $R + 2 \equiv \min_{t-1/2 \leq s \leq t+1} \mathbf{r}_\ell(\Sigma_s)$, then Theorem 5.3 gives $\mu > 0$ and C so that

$$R \geq (1 + \mu)\,R(\Sigma_t) - C\,, \tag{6.18}$$

as long as we are willing to choose C_ℓ sufficiently large depending on ℓ. The crucial point is that μ does not change when we take ℓ larger, although C_ℓ does depend on ℓ.

Lemmas 5.32 gives a constant C so that

$$\|\phi\|^2_{L^2(B_R \cap \Sigma_t)} \leq C\,\left(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\right)\,. \tag{6.19}$$

We first choose $\beta \in [0, 1)$ so that

$$(1 + \mu)\,\left(\frac{3 + \beta}{4}\right) > 1\,. \tag{6.20}$$

This takes care of the third term in (6.16). Now we choose ℓ large so that

$$c_{\ell,n}\,\left(\frac{3 + \beta}{2 + 2\beta}\right) > 1 \text{ and } (1 + \mu)\,c_{\ell,n} > 1\,. \tag{6.21}$$

This takes care of the first two terms. Once we choose ℓ, then Theorem 5.3 gives C_ℓ and, thus, determines the multiplicative factor K.

□

6.3. **An extension of "Lojasiewicz theorem".** Lojasiewicz used the gradient Lojasiewicz inequality to prove convergence of flow lines for the negative gradient flow of an analytic function f. We will prove an analogous convergence result where the differential inequality $f^{2\beta}(t) \leq -f'(t)$ (that follows from the gradient Lojasiewicz) is replaced by the discrete inequality $f^{2\beta}(t) \leq f(t-1) - f(t+1)$. This assumption is exactly what comes out of our analog of the gradient Lojasiewicz inequality, i.e., out of Theorem 0.26.

The extension will rely on the following elementary lemma:

Lemma 6.22. If $f : [0,\infty) \to [0,\infty)$ is a non-increasing function, $\epsilon, K > 0$ and for $t \geq 1$

$$K\, f^{1+\epsilon}(t) \leq f(t-1) - f(t+1)\,, \tag{6.23}$$

then there exists a constant C such that

$$f(t) \leq C\, t^{-\frac{1}{\epsilon}}\,. \tag{6.24}$$

Moreover, if $\epsilon < 1$, then

$$\sum_{j=1}^{\infty} \left(f(j) - f(j+1)\right)^{\frac{1}{2}} < \infty\,. \tag{6.25}$$

Proof. After replacing f by f/C_0 for some positive constant C_0, we can assume without loss of generality that $0 < f(0) \leq 1$ and $K = 1$. Set $t_0 = 4\, 2^{\epsilon}\, f^{-\epsilon}(0)/\epsilon + 2$ and $C = f(0)\, t_0^{\frac{1}{\epsilon}}$, then $f(0) = C\, t_0^{-\frac{1}{\epsilon}}$ and hence (6.24) holds for all $t \leq t_0$. Next note that by assumption for all $t \geq 2$

$$f^{1+\epsilon}(t) \leq f^{1+\epsilon}(t-1) \leq f(t-2) - f(t)\,. \tag{6.26}$$

Or, equivalently, for all $t \geq 2$

$$f(t-2) \geq f(t)\,(1 + f^{\epsilon}(t))\,. \tag{6.27}$$

We would like to show that (6.24) holds; so suppose not and let t be a t where inequality (6.24) fails. After possibly replacing t by $t-2$ a finite number of times we may assume that (6.24) fails for t, but holds for $t-2$. From the choice of C it follows that $t > t_0 \geq 2$. Moreover,

$$f(t-2) \geq f(t)\,(1 + f^{\epsilon}(t)) > C\, t^{-\frac{1}{\epsilon}}\,(1 + C^{\epsilon}\, t^{-1})\,. \tag{6.28}$$

Combining this with the elementary inequality that $(1+h)^{-\epsilon} \leq 1 - 2^{-1-\epsilon}\,\epsilon\, h$ for all $h \leq 1$ and that both $C^{\epsilon}\, t_0^{-1} = f^{\epsilon}(0) \leq 1$ and $2^{-1-\epsilon}\,\epsilon\, C^{\epsilon} = 2^{-1-\epsilon}\,\epsilon\, f^{\epsilon}(0)\, t_0 \geq 2$ gives

$$f^{-\epsilon}(t-2) < C^{-\epsilon}\, t\,(1 + C^{\epsilon}\, t^{-1})^{-\epsilon} \leq C^{-\epsilon}\,(t - 2^{-1-\epsilon}\,\epsilon\, C^{\epsilon}) \leq C^{-\epsilon}\,(t-2)\,. \tag{6.29}$$

Contradicting that (6.24) holds for $t-2$ and, thus, completing the proof of the first claim.

Suppose now that $\epsilon < 1$ and fix some $p \in (1, 1/\epsilon)$. Cauchy-Schwarz gives that

$$\left[\sum_{j=1}^{\infty} \left(f(j) - f(j+1)\right)^{\frac{1}{2}}\right]^2 \leq \left[\sum_{j=1}^{\infty} \left(f(j) - f(j+1)\right) j^{p}\right] \left[\sum_{j=1}^{\infty} j^{-p}\right]\,. \tag{6.30}$$

The last term is finite since $p > 1$, so it suffices to prove that $(f(j)-f(j+1))\, j^p$ is summable. However, this follows from the summation by parts formula

$$\sum_{j=1}^{n} b_j\,(a_{j+1}-a_j) = [b_{n+1}a_{n+1}-b_1a_1] - \sum_{j=1}^{n-1} a_{j+1}\,(b_{j+1}-b_j) \tag{6.31}$$

with $a_j = f(j)$ and $b_j = j^p$ since the decay (6.24) implies that

$$(n+1)^p\, f(n+1) \le C\,(n+1)^{-\frac{1}{\epsilon}}\,(n+1)^p \to 0\,, \tag{6.32}$$

$$\sum_{j=1}^{\infty} f(j+1)\,[(j+1)^p - j^p] \le C\,p\,\sum_{j=1}^{\infty}(j+1)^{-\frac{1}{\epsilon}+p-1} < \infty\,. \tag{6.33}$$

The first inequality in the second line used that $[(j+1)^p - j^p] \le p\,(j+1)^{p-1}$.

□

6.4. **Uniqueness of tangent flows.** We are now prepared to prove the uniqueness of cylindrical tangent flows.

Proof of Theorem 0.2. Let Σ_t be the rescaled MCF associated to the cylindrical singularity. It follows from the uniqueness theorem of [CIM] that if a sequence $t_j \to \infty$, then there is a subsequence $t_j' \to \infty$ so that $\Sigma_{t_j'}$ converges with multiplicity one to a cylinder $\Sigma \in \mathcal{C}_k$. It follows from White's Brakke-type theorem, [W3], that this convergence is smooth on compact subsets. A priori, different sequences could lead to different cylinders (i.e., different rotations of the same cylinder); the point of this theorem is that this does not occur.

Given any fixed large ρ and small $\epsilon > 0$, it follows from the previous paragraph that there must be some T so that

- For each $t \ge T$, there is a cylinder in $\mathcal{C}_k$ so that, for each $s \in [t-1, t+1]$, $B_\rho \cap \Sigma_s$ is a $C^{2,\alpha}$ graph over this cylinder with norm at most ϵ.

Therefore, we can apply Theorem 6.14 to Σ_t for $t \ge T$ to get K and $\mu \in (1/3, 1)$ so that

$$(F(\Sigma_t) - F(\mathcal{C}))^{1+\mu} \le K\,\left(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\right)\,. \tag{6.34}$$

This "discrete differential inequality" allows us to apply Lemma 6.22 to conclude that

$$\sum_{j=1}^{\infty} (F(\Sigma_j) - F(\Sigma_{j+1})^{\frac{1}{2}} < \infty\,. \tag{6.35}$$

Using Cauchy-Schwarz and that rescaled MCF is the negative gradient flow for F, we have

$$\begin{aligned} \int_1^{\infty} \|\phi\|_{L^1(\Sigma_t)}\,dt &\le \sum_{j=1}^{\infty}\left(F(\Sigma_j)\,\int_j^{j+1} \|\phi\|^2_{L^2(\Sigma_t)}\,dt\right)^{\frac{1}{2}} \\ &\le \sqrt{F(\Sigma_0)}\,\sum_{j=1}^{\infty}(F(\Sigma_j) - F(\Sigma_{j+1})^{\frac{1}{2}} < \infty\,, \end{aligned} \tag{6.36}$$

where the last inequality is (6.35) and the L^1 and L^2 norms are all weighted Gaussian norms. The uniqueness now follows immediately from Lemma A.48. □

Appendix A. Geometric quantities on a graph

In this appendix, we will prove some technical results for the geometry of normal exponential graphs over a hypersurface. As one consequence, we will prove Lemma 4.6 which computes the gradient of the F functional on graphs over cylinders.

Throughout this appendix, Σ_u will denote the graph of a function u over a fixed hypersurface Σ (in most applications Σ will be a cylinder), where Σ_u is given by

$$\Sigma_u = \{x + u(x)\, \mathbf{n}(x) \,|\, x \in \Sigma\} . \tag{A.1}$$

We will assume that $|u|$ is small so Σ_u is contained in a tubular neighborhood of Σ where the normal exponential map is invertible. Let e_{n+1} be the gradient of the (signed) distance function to Σ; note that e_{n+1} equals $\mathbf{n}$ on Σ.

The geometric quantities that we need to compute on Σ_u are:

- The relative area element $\nu_u(p) = \sqrt{\det g^u_{ij}(p)}/\sqrt{\det g_{ij}(p)}$, where $g_{ij}(p)$ is the metric for Σ at p and $g^u_{ij}(p)$ is the pull-back metric from the graph of u at $(p + u(p)\, \mathbf{n}(p))$.
- The mean curvature $H_u(p)$ of Σ_u at $(p + u(p)\, \mathbf{n}(p))$.
- The support function $\eta_u(p) = \langle p + u(p)\, \mathbf{n}(p), \mathbf{n}_u\rangle$, where $\mathbf{n}_u$ is the normal to Σ_u.
- The speed function $w_u(p) = \langle e_{n+1}, \mathbf{n}_u\rangle^{-1}$ evaluated at $(p + u(p)\, \mathbf{n}(p))$.

The mean curvature and the support function directly appear in the shrinker equation. The speed function enters indirectly when we rewrite the equation in graphical form; the speed function adjusts for that the normal direction and vertical directions may not be the same. The relative area element will be used to compute the mean curvature and to relate the gradient of F to $\phi = \frac{1}{2}\langle x, \mathbf{n}\rangle - H$.

A.1. **Calculations.** The next lemma gives the expressions for the ν_u, η_u and w_u on a graph Σ_u over a general hypersurface Σ. The statement is rather technical and it is helpful to keep in mind the special case where Σ is the hyperplane $\mathbf{R}^n$ and the quantities are given by

$$\nu_u = \sqrt{1 + |\nabla u|^2} = w_u \text{ and } \eta_u = \frac{u - \langle p, \nabla u\rangle}{\sqrt{1 + |\nabla u|^2}} . \tag{A.2}$$

The first part of the lemma gives similar formulas for a general Σ. The second part uses the formulas to compute Taylor expansions of the quantities. Some of these computations are used to compute linear approximations here, while others are not used in this paper but are recorded for future reference and will be used elsewhere.

Lemma A.3. There are functions w, ν, η depending on $(p, s, y) \in \Sigma \times \mathbf{R} \times T_p\Sigma$ that are smooth for $|s|$ less than the normal injectivity radius of Σ so that:

$$w_u(p) = w(p, s, y) = \sqrt{1 + |B^{-1}(p, s)(y)|^2} , \tag{A.4}$$

$$\nu_u(p) = \nu(p, s, y) = w(p, s, y) \det \left(B(p, s)\right) , \tag{A.5}$$

$$\eta_u(p) = \eta(p, s, y) = \frac{\langle p, \mathbf{n}(p)\rangle + s - \langle p, B^{-1}(p, s)(y)\rangle}{w(p, s, y)} , \tag{A.6}$$

where the linear operator $B(p, s) \equiv \text{Id} - s\, A(p)$. Finally, the functions w, ν, and η satisfy:

- $w(p, s, 0) \equiv 1$, $\partial_s w(p, s, 0) = 0$, $\partial_{y_\alpha} w(p, s, 0) = 0$, and $\partial_{y_\alpha}\partial_{y_\beta} w(p, 0, 0) = \delta_{\alpha\beta}$.

- $\nu(p,0,0) = 1$; the non-zero first and second order terms are $\partial_s \nu(p,0,0) = H(p)$, $\partial_s^2 \nu(p,0,0) = H^2(p) - |A|^2(p)$, $\partial_{p_j}\partial_s \nu(p,0,0) = H_j(p)$, and $\partial_{y_\alpha}\partial_{y_\beta}\nu(p,0,0) = \delta_{\alpha\beta}$.
- $\eta(p,0,0) = \langle p, \mathbf{n}\rangle$, $\partial_s \eta(p,0,0) = 1$, and $\partial_{y_\alpha}\eta(p,0,0) = -p_\alpha$.

Proof. Let (p,s) be Fermi coordinates on the normal tubular neighborhood of Σ, so that s measures the signed distance to Σ. If we fix an s and a path $\gamma(t)$ in Σ, then applying the normal exponential map for time s sends $\gamma(t)$ to $\gamma(t) + s\,\mathbf{n}(\gamma(t))$. It follows that the differential is given by the symmetric linear operator

$$B(p,s) \equiv (\mathrm{Id} - s\,A(p)) : T_p\Sigma \to T_p\Sigma\,, \tag{A.7}$$

where we used that $-A$ is the differential of the Gauss map to differentiate $\mathbf{n}$ and the Gauss lemma to identify $T_p\Sigma$ with the tangent space to the level set of the distance to Σ.

We will use this to compute the relative area element for the graph Σ_u. Pushing forward an orthonormal frame e_i for Σ at p gives a frame E_i for Σ_u at $(p,u(p))$

$$E_i \equiv B(p,u)(e_i) + u_i(p)\,\partial_s\,. \tag{A.8}$$

Thus, the metric on the graph is given in this frame by

$$g^u_{ij}(p) \equiv \langle E_i, E_j\rangle = \langle B(p,u)(e_i), B(p,u)(e_j)\rangle + u_i\,u_j\,. \tag{A.9}$$

Since the e_i's are orthonormal on Σ, we get

$$\nu_u^2(p) = \det\,\left(B^2(p,u(p)) + \nabla u \otimes \nabla u(p)\right)\,. \tag{A.10}$$

Similarly, using the frame (A.8), we see that the vector field

$$\partial_s - B^{-1}(p,u(p))(\nabla u)(p) = e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p) \tag{A.11}$$

is normal to Σ_u. It follows that the speed function is given by

$$\begin{aligned} w_u(p) = \langle e_{n+1}, \mathbf{n}_u\rangle^{-1} &= \frac{|e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)|}{\langle e_{n+1}, e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)\rangle} \\ &= \sqrt{1 + |B^{-1}(p,u(p))(\nabla u(p))|^2}\,. \end{aligned} \tag{A.12}$$

To rewrite the relative area element, we will need two elementary facts. The first is that for $n \times n$ matrices M_1 and M_2, we have $\det(M_1\,M_2) = \det(M_1)\,\det(M_2)$. The second is that for a vector $v \in \mathbf{R}^n$, we have

$$\det(\mathrm{Id} + v \otimes v) = 1 + |v|^2\,. \tag{A.13}$$

Using these two facts, we now rewrite (A.10) as

$$\begin{aligned} \nu_u^2(p) &= \det\,\left\{B(p,u(p))\,\left(\mathrm{Id} + B^{-1}(p,u(p))(\nabla u(p)) \otimes B^{-1}(p,u(p))(\nabla u(p))\right)\,B(p,u(p))\right\} \\ &= \left[\det\,(B(p,u(p)))\,w_u(p)\right]^2\,. \end{aligned} \tag{A.14}$$

To compute the support function η_u, first use the formula (A.11) to get

$$\mathbf{n}_u = \frac{e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)}{|e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)|} = \frac{e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)}{w_u(p)}\,, \tag{A.15}$$

where $\mathbf{n}_u$ is evaluated at $p + u(p)\,\mathbf{n}(p)$. Thus, the support function is given by

$$\begin{aligned} w_u(p)\,\eta_u(p) &= \langle p + u(p)\,\mathbf{n}(p), e_{n+1} - B^{-1}(p,u(p))(\nabla u)(p)\rangle \\ &= \langle p, \mathbf{n}(p)\rangle + u(p) - \langle p, B^{-1}(p,u(p))(\nabla u(p))\rangle\,, \end{aligned} \tag{A.16}$$

where the last equality used that $\mathbf{n}(p)$ is equal to e_{n+1} at the point $p + s\,\mathbf{n}(p)$ for any s.

We have now established the formulas (A.4), (A.5) and (A.6) for the functions w, ν, and η. It is clear from the expressions for w, ν and η that they are smooth in the three variables provided that s is sufficiently small.

The next thing is to establish the second set of three claims that give the second order Taylor expansions for w, ν, and η. The function w appears in all three expressions, so it is convenient to start there. It follows immediately that $w(p,s,0) = 1$. To compute the partials involving y_α's, we get

$$\partial_{y_\alpha} w(p,s,y) = \frac{\sum_\beta \left(B^{-2}\right)_{\alpha\beta}(p,s)\, y_\beta}{w(p,s,y)} \, . \tag{A.17}$$

It follows that $\partial_{y_\alpha} w(p,s,0) = 0$. To get the Hessian, we differentiate (A.17) again

$$\partial_{y_\alpha}\partial_{y_\beta} w(p,0,0) = \frac{\left(B^{-2}\right)_{\alpha\beta}(p,0)}{w(p,0,0)} = \delta_{\alpha\beta} \, , \tag{A.18}$$

where the last equality used that $B(p,0) = \mathrm{Id}$.

Using (A.5), we have $\nu(p,s,y) = w(p,s,y)\,\mathcal{B}(p,s)$ where

$$\mathcal{B}(p,s) = \det\left(B(p,s)\right) = \det\left(\mathrm{Id} - s\,A(p)\right) . \tag{A.19}$$

We have $\mathcal{B}(p,0) \equiv 1$ and $\partial_s \mathcal{B}(p,0) = -\mathrm{Tr}(A(p)) = H(p)$. This also gives $\partial_s \partial_{p_j} \mathcal{B}(p,0) = H_j(p)$. To get the second derivative in s, observe that

$$\partial_s \log \mathcal{B}(p,s) = \mathrm{Tr}\left[B^{-1}(p,s)\, \partial_s B(p,s)\right] = -\mathrm{Tr}\left[\left(\mathrm{Id} - s\,A(p)\right)^{-1} A(p)\right] . \tag{A.20}$$

Thus, we see that

$$\partial_s^2\, \mathcal{B}(p,0) = \left(\partial_s\, \mathcal{B}(p,0)\right)\, H(p) - \mathcal{B}(p,0)\, |A|^2(p) = H^2(p) - |A|^2(p) \, . \tag{A.21}$$

Combining the calculations for $\mathcal{B}$ with the earlier ones for w, we can compute the first three Taylor series terms for ν. The constant term is $\nu(p,0,0) = 1$. The first order terms are

$$\partial_{p_j}\, \nu(p,0,0) = 0 \, , \tag{A.22}$$

$$\partial_s\, \nu(p,0,0) = \left(\partial_s \mathcal{B}(p,0)\right)\, w(p,0,0) + \left(\partial_s w(p,0,0)\right)\, \mathcal{B}(p,0) = H(p) \, , \tag{A.23}$$

$$\partial_{y_\alpha}\, \nu(p,0,0) = \left(\partial_{y_\alpha} w(p,0,0)\right)\, \mathcal{B}(p,0) = 0 \, . \tag{A.24}$$

The second order terms involving just s and p derivatives are simplified greatly since $w(p,s,0) \equiv 1$. These are $\partial_{p_j}\partial_{p_k}\, \nu(p,0,0) = 0$ and

$$\begin{aligned} \partial_s^2\, \nu(p,0,0) &= \left\{\left(\partial_s^2 \mathcal{B}\right)\, w + 2\,\left(\partial_s \mathcal{B}\right)\, \partial_s w + \left(\partial_s^2 w\right)\, \mathcal{B}\right\}(p,0,0) = \partial_s^2 \mathcal{B}(p,0) \\ &= H^2(p) - |A|^2(p) \, , \end{aligned} \tag{A.25}$$

$$\begin{aligned} \partial_{p_j}\partial_s\, \nu(p,0,0) &= \left\{\left(\partial_s \mathcal{B}\right)\, \partial_{p_j} w + \left(\partial_{p_j}\partial_s w\right)\, \mathcal{B} + \left(\partial_{p_j}\partial_s \mathcal{B}\right)\, w + \left(\partial_s w\right)\, \partial_{p_j} \mathcal{B}\right\}(p,0,0) \\ &= \partial_{p_j}\partial_s \mathcal{B}(p,0) = H_j(p) \, . \end{aligned} \tag{A.26}$$

To compute the terms involving y derivatives, it is useful to keep in mind that $\mathcal{B}$ does not depend on y. We get

$$\partial_{p_j}\partial_{y_\alpha}\,\nu(p,0,0) = \left\{\left(\partial_{p_j}\partial_{y_\alpha} w\right)\,\mathcal{B} + \left(\partial_{y_\alpha} w\right)\,\partial_{p_j}\mathcal{B}\right\}(p,0,0) = 0\,, \tag{A.27}$$

$$\partial_s\partial_{y_\alpha}\,\nu(p,0,0) = \left\{\left(\partial_s\partial_{y_\alpha} w\right)\,\mathcal{B} + \left(\partial_{y_\alpha} w\right)\,\partial_s\mathcal{B}\right\}(p,0,0) = 0\,, \tag{A.28}$$

$$\partial_{y_\beta}\partial_{y_\alpha}\,\nu(p,0,0) = \left(\partial_{y_\beta}\partial_{y_\alpha} w(p,0,0)\right)\,\mathcal{B}(p,0) = \delta_{\alpha\beta}\,. \tag{A.29}$$

Finally, using (A.6) and the fact that the first derivatives of w vanish at $(p,0,0)$, we get the first order expansion for η. □

A.2. **The mean curvature and its linearization via the first variation.** We will compute the mean curvature H_u using the first variation of the area of Σ_u. This gives a divergence form equation in u.

Corollary A.30. The mean curvature H_u of Σ_u is given by

$$\begin{aligned} H_u(p) &= \frac{w}{\nu}\left[\partial_s\nu - \operatorname{div}_\Sigma\left(\partial_{y_\alpha}\nu\right)\right] \\ &= \frac{w}{\nu}\,\left(\partial_s\nu - \partial_{p_\alpha}\partial_{y_\alpha}\nu - \left(\partial_s\partial_{y_\alpha}\nu\right)u_\alpha(p) - \left(\partial_{y_\beta}\partial_{y_\alpha}\nu\right)u_{\alpha\beta}(p)\right)\,, \end{aligned} \tag{A.31}$$

where w, ν and their derivatives are all evaluated at $(p,u(p),\nabla u(p))$.

Proof. By Lemma A.3, the area of the graph Σ_u is

$$\operatorname{Area}(\Sigma_u) = \int_\Sigma \nu_u\;dp_\Sigma = \int_\Sigma \nu(p,u(p),\nabla u(p))\,dp_\Sigma\,. \tag{A.32}$$

Given a one-parameter family of graphs Σ_{u+tv} with v compactly supported, differentiating the area gives

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0}\operatorname{Area}(\Sigma_{u+tv}) &= \int_\Sigma \left\{\partial_s\nu(p,u(p),\nabla u(p))\,v(p) + \partial_{y_\alpha}\nu(p,u(p),\nabla u(p))\,v_\alpha(p)\right\}\,dp_\Sigma \\ &= \int_\Sigma \left\{\partial_s\nu(p,u(p),\nabla u(p)) - \operatorname{div}_\Sigma\left(\partial_{y_\alpha}\nu(p,u(p),\nabla u(p))\right)\right\}\,v(p)\,dp_\Sigma\,. \end{aligned} \tag{A.33}$$

On the other hand, the variation vector field on Σ_u is given by $v\,e_{n+1}$ so the first variation formula (see, e.g., (1.45) in [CM3]) gives

$$\frac{d}{dt}\Big|_{t=0}\operatorname{Area}(\Sigma_{u+tv}) = \int_{\Sigma_u} H_u\,\langle v\,e_{n+1},\mathbf{n}_u\rangle = \int_\Sigma H_u(p)\,\frac{v(p)\,\nu_u(p)}{w_u(p)}\,dp_\Sigma\,, \tag{A.34}$$

where the second equality used the definition of the speed function $w_u = \langle e_{n+1},\mathbf{n}_u\rangle^{-1}$.

Equating these two expressions for the derivative of area, we conclude that

$$H_u(p)\,\frac{\nu(p,u(p),\nabla u(p))}{w(p,u(p),\nabla u(p))} = \partial_s\nu(p,u(p),\nabla u(p)) - \operatorname{div}_\Sigma\left(\partial_{y_\alpha}\nu(p,u,\nabla u)\right)\,. \tag{A.35}$$

This gives the first equality in (A.31); the second equality follows from the chain rule. □

A.3. **The F functional near a cylinder.** We now specialize to where Σ is a cylinder in $\mathcal{C}_k$ and $F(u)$ is the F functional of the graph Σ_u.

Lemma A.36. If $\Sigma \in \mathcal{C}_k$, then the gradient $\mathcal{M}(u)$ of the F functional is given by

$$\mathcal{M}(u) = \frac{\nu}{w} \, \left(H_u - \frac{1}{2} \eta \right) \, \mathrm{e}^{-\frac{2\sqrt{2k}\, u + u^2}{4}} \, , \tag{A.37}$$

where H_u is the mean curvature of Σ_u and ν, w, η are all evaluated at $(p, u(p), \nabla u(p))$.

Proof. Since we are using the Gaussian L^2 inner product, $\mathcal{M}(u)$ is defined by

$$\frac{d}{dt}\Big|_{t=0} F(u+tv) = (4\pi)^{-\frac{n}{2}} \int_{\Sigma} v \, \mathcal{M}(u) \, \mathrm{e}^{-\frac{|p|^2}{4}} \, d\mu_{\Sigma} \, . \tag{A.38}$$

On the other hand, the first variation formula for the F functional from [CM1] gives

$$\frac{d}{dt}\Big|_{t=0} F(u+tv) = (4\pi)^{-\frac{n}{2}} \int_{\Sigma_u} \langle v \, e_{n+1}, \mathbf{n}_u \rangle \, \left(H_u - \frac{1}{2} \langle \mathbf{n}_u , x \rangle \right) \, \mathrm{e}^{-\frac{|x|^2}{4}} \, d\mu_{\Sigma_u} \, , \tag{A.39}$$

where each quantity is evaluated on Σ_u. Given $p \in \Sigma$, we have

$$|p + u(p) \, \mathbf{n}(p)|^2 = |p|^2 + u^2 + 2 \, u \, \langle p , \mathbf{n} \rangle = |p|^2 + u^2 + 2 \sqrt{2k} \, u \, , \tag{A.40}$$

where the last equality used that $\Sigma \in \mathcal{C}_k$. Writing (A.39) as an integral over Σ gives

$$\frac{d}{dt}\Big|_{t=0} F(u+tv) = (4\pi)^{-\frac{n}{2}} \int_{\Sigma} \frac{v}{w} \, \left(H_u - \frac{1}{2} \eta \right) \, \mathrm{e}^{-\frac{2\sqrt{2k}\, u + u^2}{4}} \, \nu \, \mathrm{e}^{-\frac{|p|^2}{4}} \, d\mu_{\Sigma} \, . \tag{A.41}$$

The lemma follows by equating (A.38) and (A.41) □

Proof of Lemma 4.6. By Lemma A.36 and Corollary A.30, $\mathcal{M}(u)$ can be written as

$$\mathcal{M}(u) \, \mathrm{e}^{\frac{2\sqrt{2k}\, u + u^2}{4}} = \partial_s \nu - \partial_{p_\alpha} \partial_{y_\alpha} \nu - \left(\partial_s \partial_{y_\alpha} \nu \right) u_\alpha(p) - \left(\partial_{y_\beta} \partial_{y_\alpha} \nu \right) u_{\alpha \beta}(p) - \frac{\nu}{2 \, w} \, \eta \, . \tag{A.42}$$

Since the exponential term depends only on u, we have to show that each of the five terms on the right side can be expressed as either:

$$\text{(i) } f(u, \nabla u), \ \text{(ii) } \langle p, V(u, \nabla u) \rangle \text{ or (iii) } \Phi^{\alpha \beta}(u, \nabla u) \, u_{\alpha \beta} \, .$$

The proof will repeatedly use the calculations from Lemma A.3.

The key point is that A is parallel on cylinders and, thus, the linear operator $B(p,s)$ depends only on s (and not p). In particular, the function ν depends only on s and y (and not p). Thus, the first three terms on the right side of (A.42) are type (i) and the fourth term is type (iii). Similarly, w depends only on s and y, it suffices to show that $w \, \eta$ is a sum of terms of the three allowed types. Lemma A.3 gives

$$w \, \eta = \langle p, \mathbf{n}(p) \rangle + s - \langle p , B^{-1}(p,s)(y) \rangle \, . \tag{A.43}$$

The first term is constant (so trivially type (i)) and the second is also type (i). Finally, since B depends only on s, the third term is type (ii). □

A.4. **Rescaled MCF near a shrinker.** Let $\Sigma \subset \mathbf{R}^{n+1}$ be an embedded shrinker and $u(p,t)$ a smooth function on $\Sigma \times (-\epsilon, \epsilon)$, giving a one-parameter family of hypersurfaces Σ_u. We next derive the graphical rescaled MCF equation.

Lemma A.44. The graphs Σ_u flow by rescaled MCF if and only if u satisfies

$$\partial_t u(p,t) = w(p,u(p,t),\nabla u(p,t)) \left(\frac{1}{2}\, \eta(p,u(p,t),\nabla u(p,t)) - H_u \right) . \tag{A.45}$$

Proof. As in [EH], the rescaled MCF equation $x_t = \left(\frac{1}{2} \langle x, \mathbf{n} \rangle - H \right) \mathbf{n}$ is equivalent (up to tangential diffeomorphisms) to the equation

$$(x_t)^{\perp} = \frac{1}{2} \langle x, \mathbf{n} \rangle - H . \tag{A.46}$$

The variation vector field and unit normal for Σ_u are $\partial_t u(p,t)\, \mathbf{n}(p)$ and $\mathbf{n}_u$, respectively, at the point $p + u(p,t)\, \mathbf{n}(p)$, so we get the equation

$$\langle \mathbf{n}(p), \mathbf{n}_u \rangle \, \partial_t u(p,t) = \langle (\partial_t u(p,t))\, \mathbf{n}(p), \mathbf{n}_u \rangle = \frac{1}{2}\, \eta_u - H_u . \tag{A.47}$$

Finally, multiplying through by $w_u = \langle \mathbf{n}(p), \mathbf{n}_u \rangle^{-1}$ gives the lemma. □

We will use the following lemma bounding the distance between time slices of a rescaled MCF by the L^1 norm of the gradient of the F functional.

Lemma A.48. Given n, there exist C and $\delta > 0$ so that if $\Sigma \in \mathcal{C}_k$ and Σ_u is a graphical solution of rescaled MCF on $[t_1, t_2]$ with $\|u(\cdot, t)\|_{C^1} \leq \delta$, then

$$\int_{\Sigma} |u(p,t_2) - u(p,t_1)| \, \mathrm{e}^{-\frac{|p|^2}{4}} \leq C \int_{t_1}^{t_2} \int_{\Sigma_u(t=r)} \left| \frac{\langle x, \mathbf{n} \rangle}{2} - H \right| \mathrm{e}^{-\frac{|x|^2}{4}} \, dr . \tag{A.49}$$

Proof. By Lemma A.44, u satisfies

$$\partial_t u(p,t) = w(p,u(p,t),\nabla u(p,t)) \left(\frac{1}{2}\, \eta(p,u(p,t),\nabla u(p,t)) - H_u \right) . \tag{A.50}$$

Since $|u|$ and $|\nabla u|$ are small, Lemma A.3 gives that both w and the relative area element ν_u are uniformly bounded and (A.40) relates the Gaussians on Σ and Σ_u, so we get

$$\begin{aligned} \int_{\Sigma} |\partial_t u(p,t)| \, \mathrm{e}^{-\frac{|p|^2}{4}} &\leq C \int_{\Sigma} \left| \frac{1}{2}\, \eta(p,u(p,t),\nabla u(p,t)) - H_u \right| \mathrm{e}^{-\frac{|p|^2}{4}} \\ &\leq C' \int_{\Sigma} \left| \frac{1}{2}\, \eta(p,u(p,t),\nabla u(p,t)) - H_u \right| \nu_u \, \mathrm{e}^{-\frac{|p+u(p,t)\,\mathbf{n}|^2}{4}} \\ &= C' \int_{\Sigma_u} \left| \frac{\langle x, \mathbf{n} \rangle}{2} - H \right| \mathrm{e}^{-\frac{|x|^2}{4}} . \end{aligned} \tag{A.51}$$

The lemma follows from integrating this with respect to t, using the fundamental theorem of calculus and Fubini's theorem.

□

Appendix B. An interpolation inequality

We will use the following interpolation inequality which is well-known, but we are including the short proof since we do not have an exact reference. Unlike the rest of this paper, the L^1 norms below are unweighted.

Lemma B.1. There exists $C = C(k,n)$ so that if u is a C^k function on $B_{2r} \subset \mathbf{R}^n$, then

$$\|u\|_{L^\infty(B_r)} \leq C \left\{ r^{-n} \|u\|_{L^1(B_{2r})} + \|u\|^{a_{k,n}}_{L^1(B_{2r})} \|\nabla^k u\|^{1-a_{k,n}}_{L^\infty(B_{2r})} \right\} , \tag{B.2}$$

$$r \|\nabla u\|_{L^\infty(B_r)} \leq C \left\{ r^{-n} \|u\|_{L^1(B_{2r})} + r \|u\|^{b_{k,n}}_{L^1(B_{2r})} \|\nabla^k u\|^{1-b_{k,n}}_{L^\infty(B_{2r})} \right\} , \tag{B.3}$$

$$r^2 \|\nabla^2 u\|_{L^\infty(B_r)} \leq C \left\{ r^{-n} \|u\|_{L^1(B_{2r})} + r^2 \|u\|^{c_{k,n}}_{L^1(B_{2r})} \|\nabla^k u\|^{1-c_{k,n}}_{L^\infty(B_{2r})} \right\} , \tag{B.4}$$

where $a_{k,n} = \frac{k}{k+n}$, $b_{k,n} = \frac{k-1}{k+n}$ and $c_{k,n} = \frac{k-2}{k+n}$.

Proof. By scaling, it suffices to prove the case $r = 1$.

The starting point is the following standard consequence of the Bernstein/Kellogg inequality for polynomials, [K]:

(K) Given n and d, there exists $C_{d,n}$ so that if p is a polynomial of degree at most d on a ball $B_\delta \subset \mathbf{R}^n$ for some $\delta > 0$, then

$$\|p\|_{L^\infty(B_\delta)} + \delta \|\nabla p\|_{L^\infty(B_\delta)} + \delta^2 \|\nabla^2 p\|_{L^\infty(B_\delta)} \leq C_{d,n}\, \delta^{-n} \int_{B_\delta} |p| \, . \tag{B.5}$$

Set $m = \|\nabla^k u\|_{L^\infty(B_2)}$. Choose $x \in \overline{B_1}$ where $|u|$ achieves its maximum and let p be the degree $(k-1)$ polynomial giving the first $(k-1)$ terms of the Taylor series of u at x. In particular, given any $\delta \in (0,1]$, Taylor expansion gives

$$\int_{B_\delta(x)} |u - p| \leq C\, m\, \delta^{n+k} \, , \tag{B.6}$$

where C depends on n and k. Using this in (K) gives

$$\|u\|_{L^\infty(B_1)} = |p|(x) \leq C\, \delta^{-n} \int_{B_\delta(x)} |p| \leq C\, \delta^{-n} \left\{ \int_{B_\delta(x)} |u| + \int_{B_\delta(x)} |u - p| \right\}$$

$$\leq C\, \delta^{-n} \left\{ \|u\|_{L^1(B_2)} + C\, m\, \delta^{n+k} \right\} . \tag{B.7}$$

We now consider two cases. First, if $m \leq \|u\|_{L^1(B_2)}$, then (B.7) with $\delta = 1$ gives

$$\|u\|_{L^\infty(B_1)} \leq C \|u\|_{L^1(B_2)} \, . \tag{B.8}$$

Next, if $m > \|u\|_{L^1(B_2)}$, then we set $\delta^{n+k} = \frac{\|u\|_{L^1(B_2)}}{m}$ (which is less than one) and (B.7) gives

$$\|u\|_{L^\infty(B_1)} \leq C \|u\|^{\frac{k}{n+k}}_{L^1(B_2)} m^{\frac{n}{n+k}} \, . \tag{B.9}$$

Thus, we see that (B.2) holds in either case.

We will argue similarly to get the ∇u bound. This time, let $x \in \overline{B_1}$ be a point where $|\nabla u|$ achieves its maximum. Given $\delta \in (0,1]$, using (K) gives

$$|\nabla u|(x) = |\nabla p|(x) \leq C\, \delta^{-n-1} \left\{ \|u\|_{L^1(B_2)} + C\, m\, \delta^{n+k} \right\} . \tag{B.10}$$

In the case where $m \leq \|u\|_{L^1(B_2)}$, we get (B.3) by setting $\delta = 1$. On the other hand, when $m > \|u\|_{L^1(B_2)}$, then we set $\delta^{n+k} = \frac{\|u\|_{L^1(B_2)}}{m}$ (which is less than one) and (B.10) gives

$$|\nabla u|(x) \leq C \, \|u\|_{L^1(B_2)}^{\frac{k-1}{n+k}} \, m^{\frac{n+1}{n+k}} \,, \tag{B.11}$$

completing the proof of (B.3). The last bound (B.4) follows similarly.

□

References

[Al] W.K Allard, *On the first variation of a varifold*. Ann. of Math. (2) 95 (1972), 417–491.

[AA] W.K. Allard and F.J. Almgren, Jr, *On the radial behavior of minimal surfaces and the uniqueness of their tangent cones*. Ann. of Math. (2) 113 (1981), no. 2, 215–265.

[AAG] S. Altschuler, S.B. Angenent, and Y. Giga, *Mean curvature flow through singularities for surfaces of rotation*, Journal of geometric analysis, Volume 5, Number 3, (1995) 293–358.

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

[Be] P.A. Beck, Metal Interfaces, p. 208. Cleveland, Ohio: American Society for Testing Materials, 1952.

[B] K. Brakke, *The motion of a surface by its mean curvature*. Mathematical Notes, 20. Princeton University Press, Princeton, N.J., 1978.

[BE] D. Bakry and M. Émery, *Diffusions hypercontractives*. Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., 1123, Springer, Berlin, 1985.

[Br] S. Brendle, *An inscribed radius estimate for mean curvature flow in Riemannian manifolds*, http://arXiv:1310.3439.

[BrCoL] H. Brezis, J.-M. Coron and E. Lieb, *Harmonic maps with defects*, Comm. Math. Phys. 107 (1986), no. 4, 649–705.

[Bu] J. Burke, *Some factors affecting the rate of grain growth in metals* , AIME Transactions, (1949) vol. 180, pp. 73–91.

[CIM] T.H. Colding, T. Ilmanen and W.P. Minicozzi II, *Rigidity of generic singularities of mean curvature flow*, preprint, arXiv:1304.6356.

[CIMW] T.H. Colding, T. Ilmanen, W.P. Minicozzi II and B. White, *The round sphere minimizes entropy among closed self-shrinkers*, J. Differential Geom., vol. 95, no. 1 (2013) 53–69.

[CM1] T.H. Colding and W.P. Minicozzi II, *Generic mean curvature flow I; generic singularities*, Annals of Math., Volume 175 (2012), Issue 2, 755–833.

[CM2] ______, *Smooth compactness of self-shrinkers*, Comm. Math. Helv., 87 (2012) 463–475.

[CM3] ______, *A course in minimal surfaces*, Graduate Studies in Mathematics, Vol. 121, AMS (2011).

[CM4] ______, *On uniqueness of tangent cones for Einstein manifolds*, Invent. Math., to appear, http://arXiv:1206.4929.

[CM5] ______, *The singular set of mean curvature flow with generic singularities*, preprint.

[CM6] ______, *Lojasiewicz inequalities and applications*, preprint, http://arXiv:1402.5087.

[CMP] T.H. Colding, W.P. Minicozzi II and E.K. Pedersen, *Mean curvature flow as a tool to study topology of 4-manifolds*, preprint, http://arXiv:1208.5988.

[EH] K. Ecker and G. Huisken, *Interior estimates for hypersurfaces moving by mean curvature*. Invent. Math. 105 (1991), no. 3, 547–569.

[FFl] H. Federer and W. Fleming, *Normal and integral currents*, Ann. of Math. 72 (1960), 458–520.

[GK] Z. Gang and D. Knopf, *Universality in mean curvature flow neckpinches*, http://arxiv.org/abs/1308.5600.

[GKS] Z. Gang, D. Knopf, and I.M. Sigal, *Neckpinch dynamics for asymmetric surfaces evolving by mean curvature flow*, preprint, http://arXiv:1109.0939v1.

[GS] Z. Gang and I.M. Sigal, *Neck pinching dynamics under mean curvature flow*, preprint, http://arxiv.org/abs/0708.2938.

[GGS] M. Giga, Y. Giga, and J. Saal, *Nonlinear partial differential equations. Asymptotic behavior of solutions and self-similar solutions*. Progress in Nonlinear Differential Equations and their Applications, 79. Birkhäuser Boston, Inc., Boston, MA, 2010.

[Gr] A. Grigor'yan, *Heat kernel and analysis on manifolds*. AMS/IP Studies in Advanced Mathematics, 47. American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.

[Hr] R.M. Hardt, *Singularities of harmonic maps*. Bull. Amer. Math. Soc. (N.S.) 34 (1997), no. 1, 15–34.

[HrLi] R.M. Hardt and F.-H. Lin, *Stability of singularities of minimizing harmonic maps. J. Differential Geom.* 29 (1989), no. 1, 113–123.

[HaP] D. Harker and E. Parker, *Grain shape and grain growth* . Trans. Am. Soc. Met. (1945) 34:156–201.

[HaK] R. Haslhofer and B. Kleiner, *Mean curvature flow of mean convex hypersurfaces*, preprint, http://arXiv:1304.0926.

[H1] G. Huisken, *Asymptotic behavior for singularities of the mean curvature flow*. J. Differential Geom. 31 (1990), no. 1, 285–299.

[H2] ———, *Local and global behaviour of hypersurfaces moving by mean curvature*. Differential geometry: partial differential equations on manifolds (Los Angeles, CA, 1990), 175–191, Proc. Sympos. Pure Math., 54, Part 1, Amer. Math. Soc., Providence, RI, 1993.

[H3] ———, *Flow by mean curvature of convex surfaces into spheres*, J. Differential Geometry 20 (1984) 237-266.

[HP] G. Huisken and A. Polden, *Geometric evolution equations for hypersurfaces*. Calculus of variations and geometric evolution problems, 45–84, Lecture Notes in Math., 1713, Springer, Berlin, 1999.

[HS1] G. Huisken and C. Sinestrari, *Convexity estimates for mean curvature flow and singularities of mean convex surfaces*, Acta Math. 183 (1999) no. 1, 45–70.

[HS2] G. Huisken and C. Sinestrari, *Mean curvature flow singularities for mean convex surfaces*, Calc. Var. Partial Differ. Equ. 8 (1999), 1–14.

[I1] T. Ilmanen, *Singularities of Mean Curvature Flow of Surfaces*, preprint, 1995, http://www.math.ethz.ch/~/papers/pub.html.

[I2] T. Ilmanen, *Elliptic regularization and partial regularity for motion by mean curvature*, Memoirs Amer. Math. Soc. 520 (1994).

[K] O. Kellogg, *On bounded polynomials in several variables*, Math. Z. 27 (1928), no. 1, 55–64.

[L] S. Lojasiewicz, *Ensembles semi-analytiques*, IHES notes (1965).

[M] W.W. Mullins, Two-dimensional motion of idealized grain boundaries, J. Appl. Phys. 27, 900–904 (1956).

[N] J. von Neumann, in Metal Interfaces (ed. Herring, C.) 108110 (American Society for Metals, Cleveland, 1952).

[Sc] F. Schulze, *Uniqueness of compact tangent flows in mean curvature flow*, Crelle, to appear, arXiv:1107.4643.

[Se] N. Sesum, *Rate of convergence of the mean curvature flow*. Comm. Pure Appl. Math. 61 (2008), no. 4, 464–485.

[Si1] L. Simon, *Asymptotics for a class of evolution equations, with applications to geometric problems*, Annals of Math. 118 (1983), 525–571.

[Si2] ———, *A general asymptotic decay lemma for elliptic problems*. Handbook of geometric analysis. No. 1, 381–411, Adv. Lect. Math. (ALM), 7, Int. Press, Somerville, MA, 2008.

[Si3] ———, *Rectifiability of the singular set of energy minimizing maps*, Calc. Var. Partial. Differential Equations 3 (1995), no. 1, 1–65.

[Si4] ———, *Theorems on regularity and singularity of energy minimizing maps*, Lectures on geometric variational problems (Sendai, 1993), 115–150, Springer, Tokyo, 1996.

[Si5] ———, *Rectifiability of the singular sets of multiplicity 1 minimal surfaces and energy minimizing maps*, Surveys in differential geometry, Vol. II (Cambridge, MA, 1993), 246–305, Int. Press, Cambridge, MA, 1995.

[SS] H. Soner and P. Souganidis, *Singularities and uniqueness of cylindrically symmetric surfaces moving by mean curvature*. Comm. Partial Differential Equations 18 (1993), no. 5-6, 859–894.

[Su] T. Sutoki, *On the mechanism of crystal growth by annealing*, Scientific Reports of Tohoku. Imperial University, (1928) vol. 17, no. 2, pp. 857–876.

[T] J. Taylor, *Regularity of the singular sets of two-dimensional area-minimizing flat chains modulo 3 in* $\mathbf{R}^3$, Invent. Math. 22 (1973), 119–159.

[W1] B. White, *The nature of singularities in mean curvature flow of mean-convex sets*. J. Amer. Math. Soc. 16 (2003), no. 1, 123–138.

[W2] ———, *Evolution of curves and surfaces by mean curvature*. Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002), 525–538.

[W3] ———, *A local regularity theorem for mean curvature flow*. Ann. of Math. 161 (2005), 1487–1519.

[W4] ———, *The mathematics of F. J. Almgren, Jr.* J. Geom. Analysis 8 (1998), no. 5, 681–702.

[W5] ———, *Partial regularity of mean-convex hypersurfaces flowing by mean curvature*, Int. Math. Res. Notices (1994) 185–192.

MIT, Dept. of Math., 77 Massachusetts Avenue, Cambridge, MA 02139-4307.

E-mail address: `colding@math.mit.edu` and `minicozz@math.mit.edu`