T. H. COLDING AND W. P. MINICOZZI’S
"UNIQUENESS OF BLOWUPS AND LOJASIEWICZ INEQUALITIES"

CHRISTOS MANTOULIDIS

Abstract. In March 2014 I gave a thorough lecture series at Stanford on Toby Colding and Bill Minicozzi’s 2013 paper titled “Uniqueness of Blowups and Lojasiewicz Inequalities.” These notes served as a guideline for my presentation. I was asked if I could upload them, so here they are. They go through the entire paper in depth, and contain proofs that were as explicit as I could make them be for my presentation. There are some minor expository differences between these notes and the original paper. These changes were made solely for my own convenience and better understanding. I do not claim originality over the contents of the proofs, except perhaps for any mistakes I may have accidentally introduced here and there. I would like to take this opportunity to thank Toby and Bill for having been very responsive and very kind in our email correspondence over the past few months.

Contents

1. Introduction 1
   1.1. Mean Curvature Flow 1
   1.2. Sign and notation conventions 1
   1.3. $F$ functional, entropy 1
   1.4. Parabolic rescaling 2
   1.5. Monotonicity 2
   1.6. Mean convex flows 2
   1.7. Cylindrical flows are generic 2
   1.8. Question of uniqueness 2
   1.9. Rescaled Mean Curvature Flow 2
2. Synopsis 3
   2.1. Ultimate goal 3
   2.2. Graphical and Entropy scales 3
   2.3. Lojasiewicz inequalities 4
   2.4. Everything put together 5
   2.5. A note on $F$ and its variations 6
3. First Lojasiewicz Inequality 6
   3.1. L1 on spheres 6
   3.2. Dimension reduction 7
   3.3. Vertical translation 7
   3.4. Cylindrical estimates 10
   3.5. Proof of First Lojasiewicz inequality 10
4. Second Lojasiewicz Inequality 11
   4.1. Understanding the kernel of $L$ 11
   4.2. Preliminary gradient estimate 12
   4.3. Proof of Second Lojasiewicz inequality 16
5. Scale Comparison Theorem 17
   5.1. Improvement argument 17
   5.2. Extension argument 18
   5.3. Proof of Scale comparison theorem 22

Date: June 14, 2014.
1. Introduction

1.1. Mean Curvature Flow. MCF is the evolution equation

$$\partial_t x = H$$

for closed manifolds $M^n \subset \mathbb{R}^{n+1}$. We assume that the first singular time is $\tau = 0$, and thus that $\tau < 0$ everywhere below. We use $\tau$ for the time parameter because we will predominantly work with a modified flow, for which we reserve the $t$ parameter.

1.2. Sign and notation conventions.

- (1) $n$ is the outer unit normal.
- (2) $H$ is the mean curvature scalar ($H > 0$ on spheres), but also possibly the mean curvature vector depending on the context.

1.3. $F$ functional, entropy. For $M^n \subset \mathbb{R}^{n+1}$ we define

$$F(M) = (4\pi)^{-n/2} \int_M \exp \left( - \frac{|x|^2}{4} \right) dV(x)$$

Occasionally we write $F_{x_0, \lambda}(M)$ for $F(\lambda^{-1}(M - x_0))$. Under a normal perturbation $M \mapsto M + \psi n$, the first variation of $F$ is:

$$F_M(\psi) = (4\pi)^{-n/2} \int_M \psi \left( H - \frac{1}{2} x \cdot n \right) \exp \left( - \frac{|x|^2}{4} \right) dV(x)$$

The quantity $\frac{1}{2} x \cdot n - H$ is denoted by $\phi$. Notice that this is the negative of the Euler-Lagrange operator with respect to the Gaussian measure $d\mu(x) = e^{-|x|^2/4} dV(x)$.

We define the entropy of $M^n$ to be $\lambda(M) \triangleq \sup_{x_0, \lambda} F_{x_0, \lambda}(M)$. Bounded entropy is important because it gives a Hölder inequality on the Gaussian weighted $L^p$ spaces and also that

$$F(M \setminus B_R) \leq C(n, \gamma) \lambda_0 R^{\rho(n)} \exp \left( - \gamma \cdot \frac{R^2}{4} \right)$$

for any $\gamma \in (0, 1)$. This allows us to localize our estimates to compact sets.

1.4. Parabolic rescaling. Parabolic rescaling is bestphrased in terms of spacetime flows. Let $\mathcal{M}$ be a spacetime MCF for $t < 0$, and let $\lambda > 0$. Then

$$\lambda \# \mathcal{M} \triangleq \{ (\lambda y, \lambda^2 t) : (y, t) \in \mathcal{M} \}$$

is another MCF, said to be a parabolic dilation at $(0, 0)$. If $\lambda_i \uparrow \infty$ and $(0, 0) \in \overline{\mathcal{M}}$, then after possibly passing to a subsequence

$$\mathcal{M}' \triangleq \lim_i (\lambda_i) \# \mathcal{M}$$

is an ancient Brakke flow which is in fact backwards self-similar, i.e.

$$M'_{\lambda x^2} = \lambda M'_{-1}$$

for any $\lambda > 0$. The limit $\mathcal{M}'$, called a tangent flow, is likely to depend on the particular choice of $\lambda_i \uparrow \infty$.

1.5. Monotonicity. Huisken’s monotonicity can be expressed as follows: the parabolic density function

$$\Theta(\mathcal{M}, (x_0, t_0), \lambda) \triangleq (4\pi \lambda^2)^{-n/2} \int_{M_{t_0 - \lambda^2}} \exp \left( - \frac{|x - x_0|^2}{4\lambda^2} \right) dV(x)$$

is increasing in $\lambda > 0$. Notice that $\Theta(\mathcal{M}, (x_0, t_0), \lambda) = F(\lambda^{-1}(M_{t_0 - \lambda^2} - x_0))$. We also write $\Theta(\mathcal{M}, X) \triangleq \lim_{\lambda \downarrow 0} \Theta(\mathcal{M}, X, \lambda)$.

1.6. Mean convex flows. A flow $\mathcal{M}$ is called mean convex if the initial time slice satisfies $H \geq 0$. Brian White has shown that all tangent flows at the first singular time of a mean convex flow are elements of

$$C_k \triangleq \{ \text{all rotations of round cylinders } \mathbb{S}^k_{\sqrt{2k}} \times \mathbb{R}^{n-k} \subset \mathbb{R}^{n+1} \}$$

i.e., all tangent flows of a mean convex flow are cylindrical.
1.7. **Cylindrical flows are generic.** Colding, Ilmanen, and Minicozzi have shown that cylindrical flows are generic in some sense. In particular, if one tangent flow at a point is cylindrical, then they all are. Furthermore, there exists a neighborhood of the point in which any other singularities are also cylindrical.

1.8. **Question of uniqueness.** There remains the question of uniqueness for cylindrical flows: do we always get the same $\mathbb{R}^{n-k}$ axis independently of how we dilate the flow? In this paper, Colding and Minicozzi answer this in the affirmative.

1.9. **Rescaled Mean Curvature Flow.** Colding and Minicozzi work in the context of RMCF. RMCF is the evolution equation
\[ \partial_t x = H + \frac{1}{2} x^\perp = \phi n \]
for closed manifolds $\Sigma_t$. RMCF is the negative $L^2$-gradient flow of $F$, i.e. along $\nabla_{\Sigma}^\phi F = -\phi$. Furthermore $t \mapsto \lambda(\Sigma_t)$ decreases along a RMCF. The stationary points of RMCF are called shrinkers and are those for which $\phi = 0 \iff H = \frac{1}{2} x \cdot n$.

RMCF is obtained from MCF by setting $t = -\log(-\tau)$, $\Sigma_t = \frac{1}{\sqrt{-\tau}} M_t$ and in particular exists even for $t \to \infty$. Notice that picking $t_i \uparrow \infty$ in the RMCF corresponds to studying the time $t = -1$ slice of limit tangent flow. As such, the sequence $t \mapsto \Sigma_t$ may have many limit points. In other words, there is only one tangent flow if and only if $t \mapsto \Sigma_t$ has a unique limit. This is what we're going to show.

2. **Synopsis**

2.1. **Ultimate goal.** Uniqueness will follow from the following "package" theorem.

**Theorem (6.14).** There exist $K > 0$ (large), $R > 0$ (large), $\varepsilon_0 > 0$ (small), and $\tau > 0$ (small) depending on $n$, $\lambda_0$ such that if

1. $\Sigma_s$, $s \in [t-1, t+1]$ is a RMCF,
2. $\lambda(\Sigma_s) \leq \lambda_0$ for $s \in [t-1, t+1]$, and
3. there is a fixed cylinder $C \in C_k$ over which $\Sigma_s \cap B_R$ is a graph with $\|\cdot\|_{2, \alpha} \leq \varepsilon$ for all $s \in [t-1, t+1]$,

Then $|F(\Sigma_t) - F(C_k)|^{2-\tau} \leq K(F(\Sigma_{t-1}) - F(\Sigma_{t+1}))$.

**Remark.** The finite dimensional Lojasiewicz approach concerns the evolution equation $\dot{x} = -\nabla f(x)$, where $f$ is real analytic. One can prove that the curve $t \mapsto x(t)$ has finite total length, i.e.
\[ \int_0^\infty |\dot{x}(t)| \, dt < \infty \]
provided it has at least one limit point. Therefore that limit point is necessarily the unique limit point, i.e. $t \mapsto x(t)$ converges. The motivation is to extend this to infinitely many dimensions.

**Remark.** The observation that’s necessary to go from 6.14 to proving uniqueness is that for any $\varepsilon_0 > 0$ there exists $\delta > 0$ such that if $\Sigma_{t_1}$ is $C^{2, \alpha, \varepsilon_0/2}$ close to a cylinder on some ball (of arbitrary radius), and
\[ \int_{t_1}^{t_2} |\phi|_{L^1(\Sigma_t)} \, dt \leq \delta \]
then every $\Sigma_t$, $t \in [a, b]$, is a graph over the same cylinder inside the ball of the same radius. This is because we can control the $L^1$ distance between the graphs in terms of the integral of $\int |\nabla_x F| \, dt$, since the evolution is in terms of $\nabla_x F$. The $L^1$ bound improves to a Hölder estimate by parabolic Schauder theory. Once we have a result like this, then
\[ \int_0^\infty |\phi|_{L^1(\Sigma_t)} \, dt < \infty \]
will imply that (for large times) the surfaces are graphical over a unique cylinder.
Proof of: 6.14 ⇒ Uniqueness. By a contradiction argument (and White's theorem) we know that for \( t \) large enough, the \( (\Sigma_s)_{s \in [t-1, t + 1]} \) can all be written as a uniformly small cylindrical graph over a cylinder depending only on \( t \). Then by 6.14, \( |F(\Sigma_t) - F(\Sigma_{t+1})|^2 < K(\Sigma_{t-1} - F(\Sigma_{t+1})) \) for all \( t \) large enough, and \( \tau > 0 \) small and fixed. By a functional inequality (Lemma 6.22) that translates this estimate into a polynomial decay as in the standard Łojasiewicz inequalities, \( |F(\Sigma_t) - F(\Sigma_k)| \leq C \frac{1}{t^{1/(1-\tau)}} \), we get

\[
\sum_{j=1}^{\infty} \left[ F(\Sigma_j) - F(\Sigma_{j+1}) \right]^{1/2} < \infty
\]

At this point we can estimate

\[
\int_1^\infty |\nabla F|_{L^1} dt = \int_1^\infty |\phi|_{L^1(\Sigma_t)} dt \leq F(\Sigma_0)^{1/2} \sum_{j=1}^{\infty} \left[ F(\Sigma_j) - F(\Sigma_{j+1}) \right]^{1/2} < \infty
\]

and therefore the total length of the flow is finite, so there is a unique limit point. \( \square \)

2.2. Graphical and Entropy scales.

**Definition** (Graphical radius). Let \( \Sigma^n \subset \mathbb{R}^{n+1} \) be a given surface that is "approximately" cylindrical. Its graphical scale is the maximal radius up to which it is \( C^{\alpha} \)-close to a cylinder. For given \( \varepsilon_0 > 0 \) (small), \( \ell \in \mathbb{N} \) (large), \( C_{\ell} > 0 \) (large), we define \( R_{\mathrm{graph}, \ell, C_{\ell}}(\Sigma) \) to be the largest radius \( R \) for which:

1. \( |\nabla F| |A| \leq C_{\ell} \) on \( \Sigma \cap B_R \), and
2. \( \Sigma \cap B_R \) is \( C^{\alpha} \)-\( \varepsilon_0 \)-close to a cylinder in \( C_k \).

**Remark.** The constant \( \varepsilon_0 \) will be chosen universally by two lemmas that come from pure functional analysis and will depend only on \( n \). It is \( \ell, C_{\ell} \) that we have to be careful with.

**Definition** (Entropy radius). Let \( \Sigma_s \) be a RMCF for \( s \in [t - 1, t + 1] \). The entropy radius is an artificial radius that is large when we are close to a stationary point of \( F \), and therefore to a cylinder (Colding, Ilmanen, Minicozzi). It is defined indirectly at time \( t \) via the relation

\[
\exp \left( -\frac{R_{\mathrm{entr}}(\Sigma_t)^2}{2} \right) = F(\Sigma_{t-1}) - F(\Sigma_{t+1})
\]

**Remark.** The graphical scale is more convenient to work with analytically while the entropy radius is more natural to the problem. The key heuristic to have in mind is that:

\[
\exp \left( -\frac{R_{\mathrm{entr}}(\Sigma_t)^2}{2} \right) = F(\Sigma_{t-1}) - F(\Sigma_{t+1}) \approx |\delta F_t| \approx |\nabla F|_{L^2}^2 \equiv [\phi]_{L^2(\Sigma_t)}^2
\]

\[
\Rightarrow \exp \left( -\frac{R_{\mathrm{entr}}(\Sigma_t)^2}{4} \right) \approx [\phi]_{L^2}^2
\]

We often run into expressions of the form:

\[
[\phi]_{L^2(\Sigma \cap B_R)} + \exp \left( -\frac{R^2}{4} \right)
\]

or powers thereof. Note the following points:

1. The goal is to write everything in terms of \( \phi \). We want to have \( [\phi]_{L^2(B_R)} \) be the dominant term; the non-compact error term should be swallowed into the \( [\phi] \).
2. Since \( [\phi]_{L^2(\Sigma_t)}^2 \approx \exp(-R_{\mathrm{entr}}(\Sigma_t)^2/2) \), we have a problem when \( R \ll R_{\mathrm{entr}}(\Sigma_t) \): the noncompact error term dominates.
3. This is where the main scale comparison theorem (Theorem 5.3) comes in. It says that the graphical radius subsumes the entropy radius.

**Theorem** (5.3). There exist \( \mu > 0 \) (small), \( \ell_0 \geq 1 \) (large), \( C > 0 \) (large) depending on \( n, \lambda_0, \varepsilon_0 \), such that if:

1. \( \Sigma_t \) is a RMCF on \([-1, 1]\), and
2. \( t \geq \ell_0 \),
then there exists $C_\ell > 0$ (large) depending on $n$, $\lambda_0$, $\varepsilon_0$, $\ell$, such that
\[(1 + \mu)R_{\text{entr}}(\Sigma_0) \leq \min_{-1/2 \leq \ell \leq 1} R_{\text{graph}}^{c_\ell,C}(\Sigma_\ell) + C\]

### 2.3. Lojasiewicz inequalities.**

We will reach Theorem 6.14 through two Lojasiewicz inequalities. The first one is based on very general cylindrical estimates that have nothing to do with MCF/RMCF.

**Theorem (0.24 - LI1).** There exist $\rho > 0$ ($\rho = 5n + 10$) and $\ell_0 \geq 1$ (large) depending on $n$, and a sequence $c_{\ell,n} \in (0,1) \uparrow 1$ as $\ell \to \infty$, such that:

1. $\Sigma^n \subset \mathbb{R}^{n+1}$ has $\lambda(\Sigma) \leq \lambda_0$,
2. $\ell \geq \ell_0$, $C_\ell < \infty$, and
3. $R < R_{\text{graph}}^{c_{\ell,n},C_\ell}$,

then
\[
\inf_{C \in \mathcal{C}_k} \left[w_C - \sqrt{2(R_n^2)_{L^2(\Sigma_0 \cap B_R)}} \right] \leq C R^\rho \left\{ [\phi]_{L^1(\Sigma_0 \cap B_R)}^{c_{\ell,n}} + \exp \left( - c_{\ell,n} \cdot \frac{R^2}{4} \right) \right\}
\]
for $C = C(n,\lambda_0,\varepsilon_0,\ell,C_\ell)$, where $w_C$ denotes the distance function to the axis of the cylinder $C$.

**Remark.** One should interpret LI1 as saying that the (squared) $L^2$ closeness to an optimal cylinder can be controlled (up to a power slightly weaker than one) by the $L^1$ norm of $\phi$ alongside an exponentially decaying noncompact error term.

The second Lojasiewicz inequality requires more effort, but it will follow from the first Lojasiewicz inequality after careful analysis of the kernel $K$ of the linearization $L$ of $\nabla_\Sigma F$.

**Theorem (0.26 - LI2).** There exists $\rho > 0$ and $\ell_0 \geq 1$ (large) depending on $n$, and a sequence $c_{\ell,n} \in (0,1) \uparrow 1$ as $\ell \to \infty$ such that:

1. $\Sigma^n \subset \mathbb{R}^{n+1}$ has $\lambda(\Sigma) \leq \lambda_0$,
2. $\ell \geq \ell_0$, $C_\ell < \infty$,
3. $R < R_{\text{graph}}^{c_{\ell,n},C_\ell}$, and
4. $\beta \in [0,1]$,

then
\[
|F(\Sigma) - F(C_\ell)| \leq C R^\rho \left\{ [\phi]_{L^2(\Sigma_0 \cap B_R)}^{c_{\ell,n}} + \exp \left( - c_{\ell,n} \cdot \frac{3 + \beta}{2(1 + \beta)} \cdot \frac{R^2}{4} \right) + \exp \left( - \frac{3 + \beta}{4} \cdot \frac{R^2}{4} \right) \right\}
\]
for $C = C(n,\lambda_0,\varepsilon_0,\ell,C_\ell)$.

**Remark.** This corresponds to the “gradient Lojasiewicz” inequality. It will be crucial in the scale comparison theorem (§5 in the paper) as well as in the proof of the “package” theorem, 6.14. The first two terms should be understood as coming from LI1 when we are “close” to $\text{ker} L$ (i.e. the normal component to the kernel is reasonably small), and the last term is really a crude estimate that we can use when we are “very far” from the kernel of $L$ (i.e. the kernel component is much, much smaller than the normal component).

### 2.4. Everything put together.

In this subsection we show how LI2 and the scale comparison theorem imply the main theorem, 6.14.

**Theorem (6.14).** There exist $K > 0$ (large), $R > 0$ (large), $\varepsilon_0 > 0$ (small), and $\tau > 0$ (small) depending on $n$, $\lambda_0$ such that if

1. $\Sigma_s$, $s \in [t-1,t+1]$ is a RMCF,
2. $\lambda(\Sigma_s) \leq \lambda_0$ for $s \in [t-1,t+1]$, and
3. there is a fixed cylinder $C \in \mathcal{C}_k$ over which $\Sigma_s \cap B_R$ is a graph with $\|\cdot\|_{2,\alpha} \leq \varepsilon$ for all $s \in [t-1,t+1]$ Then $|F(\Sigma_t) - F(C_\ell)|^{2-\tau} \leq K (F(\Sigma_{t-1}) - F(\Sigma_{t+1}))$.

**Proof of: 0.26 + 5.3 ⇒ 6.14.** By LI2 it is true that
\[
|F(\Sigma_t) - F(C_\ell)| \leq C R^\rho \left\{ [\phi]_{L^2(\Sigma_0 \cap B_R)}^{c_{\ell,n}} + \exp \left( - c_{\ell,n} \cdot \frac{3 + \beta}{4} \cdot \frac{R^2}{4} \right) + \exp \left( - \frac{3 + \beta}{4} \cdot \frac{R^2}{4} \right) \right\}
\]
for the maximal choice of radius, \( R \triangleq R_{\text{graph}}^C \). It suffices to bound each term on the right by a power greater than \( 1/2 \) of \( \delta F_t \triangleq F(\Sigma_{t-1}) - F(\Sigma_{t+1}) \).

By the scale comparison theorem,
\[
R \geq (1 + \mu) R_{\text{entr}}(\Sigma_t) - C
\]
as long as \( C > \infty \) is large enough depending on \( \ell \). The point is that \( \mu, C \) are independent of \( \ell \), since we’ll pick \( \ell \) large in a little bit.

Then the last term can be bounded by
\[
\exp \left( -\frac{3 + \beta}{4} \cdot R^2 \right) \leq C \exp \left( -\frac{3 + \beta}{4} \cdot (1 + \mu) \cdot \frac{R_{\text{entr}}(\Sigma_t)^2}{4} \right) \leq (\delta F_t)^{\frac{3 + \beta}{4} \cdot \frac{1 + \mu}{2}}.
\]
Picking \( \beta \in (0,1) \) close enough to 1 makes the exponent larger than \( 1/2 \).

2.5. A note on \( F \) and its variations. We’ve already shown that the Euler-Lagrange functional of \( F \) with respect to Gaussian measure, \( d\mu \), is
\[
\nabla d\mu F = -\phi = H - \frac{1}{2} x \cdot n.
\]
If \( \Sigma^\alpha \subset \mathbb{R}^{n+1} \) is some hypersurface and \( u : \Sigma^\alpha \to \mathbb{R} \) is "small" then the graphical surface \( \Sigma_u \) has its own Euler-Lagrange functional of \( F \) with respect to Gaussian measure, \( \phi_u \). If we like, we can pull that back to \( \Sigma \). It will be convenient to pull back the Euler-Lagrange operator of \( F \) with respect to \( dV(x) \) instead. We denote that by \( \mathcal{M} u \). When \( \Sigma \) is a cylinder,
\[
\mathcal{M} u = \nabla u \left( H_u - \frac{1}{2} \eta_u \right) \exp \left( -\frac{2\sqrt{2k} u + u^2}{4} \right).
\]
where
(1) \( \nu_u \) is the relative volume element,
(2) \( H_u \) is the mean curvature scalar of the cylinder,
(3) \( \eta_u \) is the support function, \( \langle \cdot + u \cdot \rangle n(\cdot), n_u(\cdot) \), and
(4) \( w_u \) is the speed function \( \langle \nabla \text{dist}_\Sigma, n_u \rangle^{-1} \).

The linearization of \( \mathcal{M} u \) at \( u = 0 \) is the \( L \) operator,
\[
L = \mathcal{L} + |A|^2 + \frac{1}{2} = \Delta - \frac{1}{2} \nabla_x^2 + |A|^2 + \frac{1}{2}.
\]
When \( \Sigma \) is a cylinder,
\[
L = \mathcal{L} + 1 = \Delta - \frac{1}{2} \nabla_x^2 + 1.
\]

3. First Łojasiewicz Inequality

The first Łojasiewicz inequality states that \( \phi \) correctly captures the \( L^2 \)-closeness of \( \Sigma \) to a cylinder, up to a noncompact error term with exponential decay. The plan of action to prove this is:

(1) Show it first for the special case of almost spherical shrinkers. Those are compact. We get a stronger \( C^{2,\alpha} \) inequality, and no error term. This is Lemma 2.5.

(2) Present a dimension reduction technique that can let us reduce the case of noncompact cylinders to that of compact spheres at the waist of the cylinders. This is Lemma 2.11.

(3) Apply the strong \( C^{2,\alpha} \) spherical inequality on the waist of the cylinder and then extend in the flat directions under the assumption that we indeed have reasonably flat directions. This gives a weaker \( C^1 \) inequality, and it is Proposition 2.1.
(4) Formulate a criterion for the non-spherical directions to be sufficiently flat in terms of just $[\phi]_{L^1}$.
This is Corollary 1.27.
(5) Put it all together, obtaining LI1 - Theorem 0.24.

3.1. LI1 on spheres.

**Lemma** (2.5). There exist $\varepsilon_0 > 0$ (small) and $C > 0$ (large) depending on $k, \alpha \in (0, 1)$ such that if:

1. $\Sigma^k \subset \mathbb{R}^{k+1}$ is a graph over $S^k_{\sqrt{2k}}$ with $\|u\|_2 \leq \varepsilon_0$, then $\|u\|_{2,\alpha} \leq C\|\phi\|_{0,\alpha}$.

**Proof sketch.** The first three eigenvalues of $L = \Delta + 1$ on $S^k_{\sqrt{2k}}$ are $-1$, $-\frac{1}{2}$, $\frac{1}{2}$, so 0 is not an eigenvalue, so by Schauder theory

$$\|u\|_{2,\alpha} \leq C(k, \alpha) \|Lu\|_{0,\alpha}$$

On the other hand $L$ is the linearization of $\nabla_2 F = \phi$, so $\|\phi - Lu\|_{0,\alpha} \leq C(k, \alpha) \|u\|_2 \|u\|_{2,\alpha}$ provided we require $\|u\|_2 \leq \varepsilon_0$ to be uniformly bounded (so the constant $C$ is finite, and independent of $\varepsilon_0$). Therefore:

$$\|\phi - Lu\|_{0,\alpha} \leq C(k, \alpha) \varepsilon_0 \|u\|_{2,\alpha}$$

By taking $\varepsilon_0 \downarrow 0$ depending on $k, \alpha$ we can use the second inequality inside the first, and the result follows. $\square$

3.2. Dimension reduction. To reduce to the case of spheres we need a dimension reduction argument. The dimension reduction argument relies on the idea that if we slice an almost-cylindrical almost-shrinker almost-orthogonally to a plane, then we get an almost-cylindrical almost-shrinker of one flat dimension lower.

**Lemma** (2.11). Let $\Sigma^k \subset \mathbb{R}^{k+1}$, $\Sigma_0 = \Sigma \cap \{x_k = 0\}$, $x_0 \in \Sigma_0$ be a transverse intersection point, and suppose $\varepsilon \in (0, 1/2)$. If

1. $|\nabla^l x_{k+1}| \geq 1 - \varepsilon$ at $x_0$,
2. $|\nabla^l |\nabla x_{k+1}| | \leq \varepsilon$ at $x_0$, and
3. the tensor norms $|A(\cdot, \nabla^l x_{k+1})| + |(\nabla A)(\cdot, \nabla^l x_{k+1})| \leq \varepsilon$ on the indicated restricted ("almost-flat") directions at $x_0$,

then if $\phi_0$ is the $\phi$-function for $\Sigma_0$,

$$|\phi - \phi_0| + |\nabla^{\Sigma_0} (\phi - \phi_0)| \leq C \varepsilon \left(1 + |\phi| + |\nabla \phi|\right)$$

for a universal constant $C > 0$.

**Sketch of proof.** The proof is reasonably straightforward. We compute

$$\phi - \phi_0 = \frac{|v| - 1}{|v|} \phi + \frac{1}{|v|} A\left(\frac{v}{|v|}, \frac{v}{|v|}\right)$$

where $v = \nabla^l x_{k+1}$ near $x_0$. Immediately we see $|\phi - \phi_0| \leq C \varepsilon + C \varepsilon |\phi|$. By differentiating along $\Sigma_0$ and using Kato’s inequality we get $|\nabla^{\Sigma_0} (\phi - \phi_0)| \leq C \varepsilon + C \varepsilon |\phi| + C \varepsilon |\nabla \phi|$. $\square$

3.3. Vertical translation. The point now is to apply the spherical inequality on the waist of a cylinder and the translate all bounds vertically.

**Proposition** (2.1). There exist $\varepsilon_1, \varepsilon_2 > 0$ (small) and $C > 0$ (large) depending on $n, \delta > 0, M > 0$, so that if

1. $H \geq \delta > 0$, $|A| + |\nabla A| \leq M$ on $\Sigma \cap B_R$,
2. $\Sigma \cap B_{5\sqrt{n}} \subset C^2 \varepsilon_1$-close to a cylinder in $C_k$, $k \geq 1$, and
3. $5\sqrt{2n} \leq r < R$ is such that

$$r \|\phi\|_{1, B_{5\sqrt{n}}} + r^5 \|\nabla(A/H)\|_{1, B_r} \leq \varepsilon_2$$

then $\Sigma \cap B_{\sqrt{n} - 5\varepsilon_2}$ is a graph over a cylinder in $C_k$ with

$$\|\cdot\|_1 \leq C \left\{ r \|\phi\|_{1, B_{5\sqrt{n}}} + r^5 \|\nabla(A/H)\|_{1, B_r} \right\}$$
Remark. (1) The first assumption is a crude assumption on a large scale. (2) The second assumption is a small-scale fit that allows us to dimension-reduce to something reasonable. (3) The third assumption is twofold: the first term will get the waist to be $C^{2,\alpha}$-close to $S^k\sqrt{2k}$, and the second term says that $A/H$ is almost parallel and allows us to extend estimates in the flat directions.

Remark. The radius $5\sqrt{2n}$ is not arbitrary. We expect the waist to have radius $\sqrt{2k}$, so we're allowing for $k \leq n$ and simultaneously giving some extra wiggle room.

Remark. It's actually better to state a stronger result, seeing as to how it's what is actually proved and is what is going to be invoked later. For $y \in B_{\sqrt{r^2-3k}}$ and $u$ the graph function over the cylinder,

$$|u(y) + |\nabla u|(y)| \leq C \left\{ r^2 \|\phi\|_{1,B_{\sqrt{2n}}} + r^5 \|\nabla(A/H)\|_{1,B_{r\epsilon}} \right\}$$

Sketch of proof of 2.1. We write $\tau$ for $A/H$, $\varepsilon_{\tau}(r) = |\nabla\tau|_{1,B_{r\epsilon}}$, and $\varepsilon_{\phi}(r) = \|\phi\|_{1,B_{r\epsilon}}$.

**Step 1. Setup.** Pick $p \in \Sigma \cap B_{5\sqrt{2n}}$ arbitrary and fixed for the duration of the proof.

Since $\Sigma \cap B_{5\sqrt{2n}}$ is $C^2\varepsilon_1$-close to a cylinder, by possibly shrinking $\varepsilon_1$ depending on $n$ alone we can get an orthonormal basis of eigenvectors (flat:) $v_1(p), \ldots, v_{n-k}(p)$, (spherical:) $z_1(p), \ldots, z_k(p)$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_{n-k}, \sigma_1, \ldots, \sigma_k$ such that

$$|\lambda_i| \leq \frac{1}{\sqrt{100n}} \quad \text{and} \quad |\sigma_j| \geq \frac{1}{\sqrt{4n}}$$

Let's also define $u(p)$ to be the outer unit normal at $p$. By a general lemma of Colding and Minicozzi from a different paper, having $\lambda_i$ and $\sigma_j$ far apart forces a better yet bound on the flat eigenvalues $\lambda_i$ in terms of $|\nabla\tau|_1$:

$$|\lambda_i| \leq C(n, \gamma) \varepsilon_{\tau}(5\sqrt{2n})$$

where the $\delta$ came in because we have to multiply through by $H \geq \delta$ to get from $A$ to $\tau = A/H$.

Define the linear "coordinate" functions

$$f_i(x) \triangleq x \cdot v_i(p), \quad g_j(x) \triangleq x \cdot z_j(p), \quad g_{k+1}(x) \triangleq x \cdot n(p), \quad i \leq n-k, \quad j \leq k$$

and the tangent vectors

$$v_i(x) \triangleq \nabla f_i(x), \quad i \leq n-k$$

We will be slicing with the plane $P = \{f_1 = \ldots = f_{n-k} = 0\}$ and we will be interested in the cylinder normal to this plane. The radial distance to the axis of that cylinder is given by:

$$w(x)^2 \triangleq \sum_{j=1}^{k+1} g_j(x)^2$$

Without loss of generality we may translate our surface so that $p \in P$.

**Step 2. Bounds near $p$.** For $r > 5\sqrt{2n}$ let $\Omega_r = \Sigma \cap B_{5\sqrt{2n}}^\circ(p)$, the latter ball being a geodesic ball. The $v_i$ enjoy the following properties in $\Omega_r$:

1. $|v_i - v_i(p)| \leq C(n, \delta, M) r^2 \varepsilon_{\tau}(r)$,
2. $|\nabla v_i| \leq C(n, \delta, M) r^2 \varepsilon_{\tau}(r)$, and
3. $|D_{v_i} A| \leq C(n, \delta, M) r^2 \varepsilon_{\tau}(r)$.

Let $\gamma(t)$ be a curve of length $\leq 3r$ starting at $p$, and let $w(t)$ be the unit parallel translation of $v_i(p)$ along $\gamma(t)$. Note that

$$|w| = 1, \quad |\gamma| \leq 1, \quad |\nabla\tau| \leq \varepsilon_{\tau}(r) \text{ on } \Omega_r \Rightarrow |\nabla_{\gamma(t)}(\tau(w(t)))| \leq \varepsilon_{\tau}(r)$$

so

$$|\tau(w(t))| \leq (\text{length} \gamma) \varepsilon_{\tau}(r) + |\tau(v_i(p))| \leq 3r \varepsilon_{\tau}(r) + C(n, \delta) \varepsilon_{\tau}(5\sqrt{2n}) \leq C(n, \delta) r \varepsilon_{\tau}(r)$$

so

$$|A(w(t))| = |H| |\tau(w(t))| \leq C(n, \delta, M) r \varepsilon_{\tau}(r)$$

Since $\nabla_{\gamma} w = A(\gamma, w) n$, the FTC gives

$$|w(t) - v_i(p)| \leq (\text{length} \gamma) C(n, \delta, M) r \varepsilon_{\tau}(r) \leq C(n, \delta) r^2 \varepsilon_{\tau}(r)$$
Of course $w(t)$ is some tangent vector $x$ and $v_i(x)$ is the one closest to $v_i(p)$, so the inequality above also holds true for $|v_i - v_i(p)|$:

$$|v_i - v_i(p)| \leq (\text{length } \gamma) C(n, \delta, M) r \varepsilon_\gamma(r) \leq C(n, \delta) r^2 \varepsilon_\gamma(r) \Rightarrow (1)$$

Furthermore by the bound on $|\nabla \tau|$, the proof is the same for multiple simultaneous slices.

$$|\tau(v_i)| \leq |\tau(v_i) - \tau(v_i(p))| + |\tau(v_i(p))| \leq C(n, \delta, M) r^2 \varepsilon_\gamma(r) \Rightarrow (2)$$

Equation (3) follows from (2) by the Codazzi equation.

**Step 3. Dimension reduction.** We seek to slice $\Sigma$ with $P = \{f_1 = \ldots = f_{n-k} = 0\}$. We use (a modification of) 2.11 that allows multiple slices at once. The point is that slicing by $P$ gives a compact topological sphere that is $C^2$ close to $S^k_{\sqrt{2n}}$, whereas slicing with lower dimensional things gives noncompact objects. The proof is the same for multiple simultaneous slices.

The coordinates we are slicing with respect to are $f_1, \ldots, f_{n-k}$ so $\Sigma_0 = \Sigma \cap P$ is a compact topological sphere contained well inside $B_{5\sqrt{2n}}$. We claim that the assumptions of 2.11 are satisfied on $\Sigma_0 \approx S^k_{\sqrt{2n}} \subset \Omega_{5\sqrt{2n}}$. Indeed, at any point $x$ on the topological sphere:

(1) $|\nabla^t f_i| \geq |\nabla^t f_i(p)| - |\nabla^t f_i - \nabla^t f_i(p)| \geq 1 - C r_0^2 \varepsilon_\gamma(r_0)$,

(2) for $X \in T_x \Sigma$,

$$\nabla X \nabla^t f_i = \nabla X (\nabla f_i - (\nabla f_i \cdot n)n)$$

$$= -X(\nabla f_i \cdot n)n - (\nabla f_i \cdot n)\nabla X n$$

$$\Rightarrow \nabla X \nabla^t f_i = (\nabla f_i \cdot n) A(X)$$

$$= \left([\nabla f_i - \nabla^t f_i] \cdot n\right) A(X)$$

so $|\nabla^t \nabla^t f_i| \leq C r_0^2 \varepsilon_\gamma(r_0)$; finally,

(3) $|A(\cdot, \nabla^t f_i) + |(\nabla^t A)(\cdot, \nabla^t f_i)| \leq C r_0^2 \varepsilon_\gamma(r_0)$.

We’re writing $r_0 = 5\sqrt{2n}$ above. In view of the lower bound $r \geq 5\sqrt{2n}$, as long as we require $\varepsilon_2$ to be small enough depending on $n$, $\delta$, $M$ to make the constants above comparable to $1/2$, we can apply 2.11 and get that $\Sigma_0$ is $C^{2,\alpha}$ $(\varepsilon_{\phi} + \varepsilon_\gamma)$-close to $S^k_{\sqrt{2n}}$. Therefore the radial axis distance function satisfies:

$$\|w - \sqrt{2k}\|_{2,\alpha, \Sigma_0} \leq C(n, \delta, M) (\varepsilon_\phi(5\sqrt{2n}) + \varepsilon_\gamma(5\sqrt{2n}))$$

**Step 4. Extending vertically.** We extend the bound by constructing a vertical radial flow with speed $f(x)^2 = \sum_{i=1}^{n-k} f_i(x)^2$.

Notice that $f^2 + w^2 = |x|^2$. Notice that by defining

$$v = \frac{\nabla^t f}{|\nabla^t f|^2}$$

and flowing $\Sigma_0$ by $v$ moves $\Sigma_0$ through the level sets of $f$. We need to make sure the speed of $v$ is not too large. Since

$$\nabla^t f = \sum_i \frac{f_i}{f} v_i = \sum_i \frac{f_i}{f} v_i(p) + \sum_i \frac{f_i}{f} (v_i - v_i(p))$$

and since the $v_i(p)$ are orthonormal and $\sum_i f_i^2 = f$ we conclude $|1 - |\nabla^t f|| \leq \sum_i |v_i - v_i(p)| \leq C r^2 \varepsilon_\gamma(r)$ on $\Omega_e$. By possibly shrinking $\varepsilon_2$ depending on $n$, $\delta$, $M$, we can control the right hand side by $1/2$ and thus get $|v| \leq 2$ on $\Omega_e$. Note that

$$\nabla v_i \nabla g_j = 0 \Rightarrow \nabla v_i \nabla^t g_j = -\nabla v_i \nabla^t f \cdot n \nabla g_j + \nabla v_i \nabla g_j \cdot n$$

$$= \nabla v_i (\nabla g_j \cdot n)n + (\nabla g_j \cdot n) \nabla v_i n = (\nabla g_j \cdot \nabla v_i n)n$$

and therefore $|\nabla v_i \nabla^t g_j| \leq C(n) |A(v_j, \cdot)| \leq C r^2 \varepsilon_\gamma(r)$. Since $|v| \leq 2$, we conclude that

$$|\nabla v \nabla^t g_j| \leq C r^2 \varepsilon_\gamma(r)$$
in $\Omega_r$ and, consequently, that
\[
|\nabla_\nu \nabla^i w |^2 \leq 2C(n) \sum_{j=1}^{k+1} |\nabla_\nu (g_j \nabla^i g_j) | \leq 2C(n) \sum_{j=1}^{k+1} |\nabla_\nu g_j| |\nabla^i g_j| + |g_j| |\nabla_\nu \nabla^i g_j| \leq C r^3 \varepsilon_\tau(r)
\]
in $\Omega_r$, the extra $r$ coming from the linear $g_j$ on the right.

Let $\Omega_{r,f}$ be the region obtained by integrating $\Sigma_0$ along $v$ for time $t^2 \leq r^2 - 3k$, intersected with $\Omega_r$. By the fundamental theorem of calculus,
\[
|\nabla^i w|^2 \leq \sup_{\Sigma_0} |\nabla^i w|^2 + C r \sup_{\Omega_{r,f}} |\nabla_\nu \nabla^i w|^2 \leq C \varepsilon_\phi(r_0) + C r^4 \varepsilon_\tau(r)
\]
as well as
\[
|w|^2 - 2k \leq C r \varepsilon_\phi(r_0) + C r^5 \varepsilon_\tau(r)
\]
in $\Omega_{r,f}$. Since $w$ is bounded from below, these bounds can be turned into $C^1$ bounds for $w - \sqrt{2k}$. These in turn extend to $C^1$ bounds on the graph function $u$. \hfill $\square$

### 3.4. Cylindrical estimates.
In order to obtain the first Lojasiewicz inequality from Proposition 2.1 we need to absorb the $\|\nabla(A/H)\|$ term into a $\phi$ term. This is obtained by very generic estimates for surfaces that resemble cylinders in some sense.

The key operators are the drift Laplacian $L = \Delta - \frac{1}{2} \nabla_x^2$ and $L = L(A) + |A|^2 + \frac{1}{2}$. Recall that $L$ is the linearization of $\mathcal{M}$, the E-L functional of $F$. The second fundamental form and the function $\phi$ are related by a PDE:
\[
LA = A + \nabla^2 \phi + \phi A^2
\]
from which we get as a corollary that:
\[
\mathcal{L}_{s,E}(A) = \frac{\nabla^2 \phi + \phi A^2}{H} + \frac{A(\Delta \phi + \phi |A|^2)}{H^2}
\]
\[
\mathcal{L}_{s,E}[A/H]^2 = 2 \|\nabla(A/H)\|^2 + 2 \left\langle \nabla^2 \phi + \phi A^2, A \right\rangle + 2 |A|^2 \left\langle (\Delta \phi + \phi |A|^2) \right\rangle
\]
where $\mathcal{L}_f \sigma = \sigma + \langle \nabla \log f, \nabla \sigma \rangle$ for any tensor (or function) $\sigma$.

**Remark.** One point is that every term involving $A$ above is paired up with something linear in $\phi$ and we also have no derivatives of $A$ showing up. The exception to both is the term $|\nabla(A/H)|^2$ which thankfully has a favorable sign in the PDE.

After integrating the second relationship by parts with suitable use of the Gaussian measure we obtain the following weighted $L^2$ estimate:

**Corollary (1.24).** If $\Sigma \cap B_R$ is smooth, $H \geq \delta > 0$ and $|A| \leq M$, then for all $s \in (0, R)$ we can estimate
\[
|\nabla(A/H)|^2 \leq C \left\{ \frac{1}{\sqrt{s}} \text{vol}(\Sigma \cap B_R) \exp \left( - \frac{(R - s)^2}{4} \right) + [\phi]_{L^2(\Sigma \cap B_R)} + [\nabla^2 \phi]_{L^1(\Sigma \cap B_R)} \right\}
\]
for $C = C(n, \delta, M)$.

By interpolation we can get $L^\infty$ estimates from these $L^2$ estimates provided we control a high enough derivative.

**Corollary (1.27).** If $\Sigma \cap B_R$ is smooth, $H \geq \delta > 0$, $|A| + |\nabla^\ell A| \leq M$ for some $\ell \geq 2$, and $\lambda(\Sigma) \leq \lambda_0$, then for all $|y| < R - 2$:
\[
|\nabla(A/H)|^2 + |\nabla^2(A/H)|^2 \leq C R^{2n} \left\{ [\phi]_{L^2(\Sigma \cap B_R)}^{\ell} \exp \left( - \frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right\} \exp \left( \frac{1}{2} \frac{|y|^2}{4} \right)
\]
where $C = C(n, \lambda_0, \delta, \ell, M)$, and $c_{\ell,n} \in (0, 1) \uparrow 1$ as $\ell \to \infty$. 


3.5. Proof of First Łojasiewicz Inequality.

**Theorem (0.24 - LI1).** There exist \( \rho > 0 \) (\( \rho = 5n + 10 \)) and \( \ell_0 \geq 1 \) (large) depending on \( n \), and a sequence \( c_{\ell,n} \in (0,1) \uparrow 1 \) as \( \ell \to \infty \) such that if:

1. \( \Sigma^n \subset \mathbb{R}^{n+1} \) has \( \lambda(\Sigma) \leq \lambda_0 \),
2. \( \ell \geq \ell_0, \, C_\ell < \infty \), and
3. \( R < R^{c_{\ell_0,\ell,C_\ell}} \),

then

\[
\inf_{C \in C} \left[ w_C - \sqrt{2R} \right]_{L^2(\Sigma \cap B_R)} \leq C R^\rho \left\{ \left[ \phi \right]_{L^1(\Sigma \cap B_R)}^{c_{\ell,n}} + \exp \left( -c_{\ell,n} \cdot \frac{R^2}{4} \right) \right\}
\]

for \( C = C(n, \lambda_0, \varepsilon_0, \ell, C_\ell) \), where \( w_C \) denotes the distance function to the axis of the cylinder \( C \).

**Proof.** Suppose we are given \( n, \lambda_0, \varepsilon_0, \ell, C_\ell \) as above. Then by writing out the graphs and interpolation, \( |A| + |\nabla A| \leq M(\varepsilon_0, \ell, C_\ell) \) and \( H = \delta(\varepsilon_0) \) on \( \Sigma \cap B_R \). In fact

\[
\left\| \phi \right\|_{1, B_{\varepsilon_0 \sqrt{\pi}}} \leq C(n, \ell, C_\ell) \frac{\left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}}^{1/4} \| \nabla \phi \|_{L^2, B_{10 \sqrt{\pi}}}}{\left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}}}
\]

by interpolation. Choosing \( \ell_0 = \ell_0(n) \) so that \( a_{\ell,n} \geq 3/4 \) for all \( \ell \geq \ell_0 \) gives

\[
\left\| \phi \right\|_{1, B_{\varepsilon_0 \sqrt{\pi}}} \leq C(n, \ell, C_\ell) \frac{1}{\left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}}}
\]

seeing as to how we may suppose \( \left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}} \leq 1 \).

We may suppose that the quantity:

\[
R_1 \triangleq \sup \left\{ r \leq R : C(n, \lambda_0, \delta, \ell, M) \frac{R^{2n+5} \left[ \left[ \phi \right]_{L^1, B_{10 \sqrt{\pi}}}^{c_{\ell,n}} + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right]}{\left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}}} \leq \varepsilon_2 \}
\]

(\( \delta = \delta(\varepsilon_0) \), \( M = M(\varepsilon_0, \ell, C_\ell) \)) by writing out the graphs, and \( \varepsilon_2 = \varepsilon_2(n, \delta, M) \) as in 2.1) is larger than \( 5 \sqrt{2n} \), \( C \) here being the constant in the interpolation corollary 1.27 plus the constant \( C(n, \ell, C_\ell) \) from a few lines above. If this is not true, we are immediately done since the constant \( C \) in the conclusion is allowed to depend on precisely the same parameters as the ones we have: \( n, \lambda_0, \varepsilon_0, \ell, C_\ell \).

By 1.27 and the \( \phi \)-estimate above we can bound for every \( 5 \sqrt{2n} \leq r \leq R_1 \):

\[
r^2 \left\| \phi \right\|_{1, B_{\varepsilon_0 \sqrt{\pi}}} + r^5 \left\| \nabla (A/H) \right\|_{1, B_r} 
\leq C \left\{ r^2 \left\| \phi \right\|_{L^1, B_{10 \sqrt{\pi}}}^{3/4} + r^5 \left[ \left[ \phi \right]_{L^1, B_{10 \sqrt{\pi}}}^{c_{\ell,n}} + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right] \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) \right\}
\leq C R^{2n+5} \left[ \left[ \phi \right]_{L^1, B_{10 \sqrt{\pi}}}^{c_{\ell,n}} + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right] \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) = \varepsilon_1
\]

and therefore by 2.1, \( \Sigma \cap B_r \) is a graph over some cylinder \( C \), with \( \| u \|_1 \leq \) the expression above prior to \( \varepsilon_2 \). We want to take \( r = R_1 \).

**Remark.** In taking \( r = R_1 \) we get a crude term \( \exp(R_1^2/8) \) on the right hand side above that we cannot offset with the Gaussian measure. This is an expository issue in the way Proposition 2.1 is packaged. Instead one observes that the proof of 2.1 allows for us to have pointwise \( |u| + |\nabla u| \) bounds that depend on the distance to the origin. This way indeed we can pick \( r = R_1 \) and conclude that:

\[
|u(y)| + |\nabla u(y)| \leq C R^{2n+5} \left[ \left[ \phi \right]_{L^1, B_{10 \sqrt{\pi}}}^{c_{\ell,n}} + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right] \exp \left( \frac{1}{2} \cdot \frac{|y|^2}{4} \right)
\]

for \( y \in B_{R_1} \).
At this point we’re done because by the bound on $|u(y)|$ in the remark above
\[ |u|_{L^2, B_R}^2 = \int_{B_R} |u(y)|^2 \exp \left(-\frac{|y|^2}{4}\right) dV(y) \leq C \text{vol}(B_R) R^{4n+10} \left[ |\phi|_{L^1, B_R}^\infty + \exp \left(-c_{\ell,n} \cdot \frac{R^2}{4}\right) \right] \leq C R^\rho(n) \left[ |\phi|_{L^1, B_R}^\infty + \exp \left(-c_{\ell,n} \cdot \frac{R^2}{4}\right) \right] \Rightarrow |w - \sqrt{2k}|_{L^2, B_R}^2 \leq C R^\rho(n) \left[ |\phi|_{L^1, B_R}^\infty + \exp \left(-c_{\ell,n} \cdot \frac{R^2}{4}\right) \right]
\]
and on $B_R \setminus B_{R_1}$ we have the crude estimate $|w - \sqrt{2k}| \leq |x|$, so that
\[
|w - \sqrt{2k}|_{L^2, B_R \setminus B_{R_1}}^2 \leq C R^{n+2} \exp \left(-\frac{R^2}{4}\right) \leq C R^{5n+12} \left[ |\phi|_{L^1, B_R}^\infty + \exp \left(-c_{\ell,n} \cdot \frac{R^2}{4}\right) \right]
\]
the last inequality following from the definition of $R_1$. The result follows by adding these two inequalities. □

4. SECOND LOJASIEWICZ INEQUALITY

The second Lojasiewicz inequality captures the proximity of the $F$-value of a surface to that of a critical point of the $F$ functional in terms of the gradient $\phi = \nabla_S F$, and a noncompact error term with exponential decay. To prove this, we will:

1. Establish elliptic estimates for the operator $L$ and relevant Sobolev estimates on $\text{ker } L$ and $(\text{ker } L)^\perp$.
2. Using those estimates and by looking at what happens near $\text{ker } L$ versus near $(\text{ker } L)^\perp$ we will establish Proposition 4.1, which bounds the same quantity as $\text{LI}_2 \left( F(\Sigma) - F(C) \right)$ in terms of $\phi$, a noncompact error term with exponential decay, and also a measure of $L^2$-closeness to a cylinder.
3. We will then swallow the $L^2$-closeness term into the prior two terms in view of LI1.

4.1. Understanding the kernel of $L$. In this entire subsection $\Sigma$ refers to a round cylinder in $C_k$, in which case $L$ collapses to $L + 1$, seeing as to how $|A|^2 = 1/2$. Furthermore, all $L^p$ and Sobolev norms are weighted by the Gauss measure.

The following lemma summarizes all the tools we have available for us from functional analysis:

**Lemma (3.2).** The following are true for $\mathcal{L}$:

1. It is self-adjoint on $H^2$, with
   \[ \int_{\Sigma} u \mathcal{L} u \exp \left(-\frac{|x|^2}{4}\right) = -\int_{\Sigma} \langle \nabla u, \nabla v \rangle \exp \left(-\frac{|x|^2}{4}\right). \]
2. The embedding $H^1 \hookrightarrow L^2$ is compact.
3. The operator $\mathcal{L}$ has discrete spectrum with finite multiplicity in $H^2$.
4. $L^2$ admits a complete orthonormal basis of $C^\infty$ functions.

We can also show

**Lemma (3.4).** There exists $C$ depending on $n$ such that
\[
||x||_{L^2}^2 \leq C \left( [u]_{L^2}^2 + [\nabla R^{n+1} u]_{L^2}^2 \right) \leq C [u]_{H^1}^2,
\]
for all $u \in H^1$.

**Remark.** There is nothing surprising here. The inequality on the right is trivial, and the inequality on the left says that the $|x|$ weight can be swallowed into the gradient integral on the right, in fact even just the gradient in the flat directions.
\begin{proof}
We provide the sketch of this proof because really all weighted Sobolev norms proofs in this paper are more or less the same, so it’s good to outline a few of them.

Without loss of generality \( u \) has compact support. If we write \( y = x' \) for the \( \mathbb{R}^{n-k} \) factor coordinates, then \( |x|^2 = y^2 + 2k \) on any cylinder in \( C_0 \) it suffices to bound \( ||y|u|_{L^2}^2 \) instead of \( ||x|u|_{L^2}^2 \). We compute

\[
\begin{align*}
\text{div}_\Sigma \left( u^2 y \exp \left( -\frac{|x|^2}{4} \right) \right) &= u^2 (\nabla u, y) + (n-k) u^2 - 2 \left\| \frac{|y|^2}{4} \right\| \\
&\leq 4 \left\| \nabla_{\mathbb{R}^{n-k}} u \right\|^2 + (n-k) u^2 - 2 \left\| \frac{|y|^2}{4} \right\|
\end{align*}
\]

The result follows after integrating.
\end{proof}

The following lemma provides rudimentary elliptic estimates for the projections \( u_K, u^\perp \) on \( K = \ker L \) and \( (\ker L)^\perp \) (\( \perp \)'s taken in \( L^2 \)).

\textbf{Lemma (3.11).} There exist \( C > 0 \) (large), \( \mu > 0 \) (small) depending on \( n, k \) such that

\[
\mu \left[ u^\perp \right]_{H^2} \leq [Lu]_{L^2} \leq C [u]_{H^2} \leq C [u_K]_{L^2}
\]

\textbf{Proof.} The inequality \( [Lu]_{L^2} \leq [u]_{H^2} \) is a corollary of 3.4.

The inequality \( \mu \left[ u^\perp \right]_{H^2} \leq [Lu]_{L^2} \) is slightly stronger than the standard elliptic inequality (which would give an \( H^1 \) estimate, not \( H^2 \)), so let’s see how to prove that. We claim that

\[
[u]_{H^2} \leq C \left[ [u]_{L^2} + [Lu]_{L^2} \right].
\]

Notice that this is standard for \( H^1 \) on the right (integrate by parts and use self adjointness of \( L \)), so we really just need to show the same upper bound for \( \nabla^2 u \).

\[
\begin{align*}
\text{div}_\Sigma \left\{ \nabla^2 u (\nabla u) - (Lu) \nabla u \right\} &\leq \left\{ \frac{1}{2} \nabla^2 |u|^2 \right\} - \left\{ (\nabla L u, \nabla u) + Lu \Delta u \right\} - \left\{ \frac{1}{2} \nabla_x \nabla u, \nabla u \right\} - Lu \frac{1}{2} \nabla_x u,
\end{align*}
\]

where we’ve used (in the following order) the facts that:

(1) \( \nabla_x \nabla u = \nabla \nabla_x u \) on the cylinder,
(2) the Bochner formula \( \frac{1}{2} \Delta |\nabla u|^2 = |\nabla^2 u|^2 + \text{Ric}(\nabla u, \nabla u) + (\nabla \Delta u, \nabla u) \), and
(3) \( \text{Ric} \geq 0 \) on the cylinder

in the last four steps. Integrating, we get \( \left| \nabla^2 u \right|_{L^2}^2 \leq [Lu]_{L^2}^2 \) as claimed.

By applying this elliptic estimate to \( u^\perp \),

\[
\left[ u^\perp \right]_{H^2} \leq C \left[ [u^\perp]_{L^2} + [Lu^\perp]_{L^2} \right] = C \left[ [u^\perp]_{L^2} + [Lu - u^\perp]_{L^2} \right] \leq C \left[ [u^\perp]_{L^2} + [Lu]_{L^2} \right] \leq C [Lu]_{L^2},
\]
the last step holding true by ellipticity.

For the last claim observe that \( \nabla u_K = 0 \Rightarrow \nabla u_K = -u_K \), and that by the previous application of Bochner’s theorem this forces \( \|\nabla^2 u\|_{L^2} \leq \|\nabla u\|_{L^2} = |u_K|_{L^2} \). This bounds the second derivative. For the first derivative, integrating \( \nabla u_K = -u_K \) by parts gives \( \|\nabla u_K\|_{L^2} = |u_K|_{L^2} \). The result follows.

Next we seek to understand \( K = \text{ker} L \). In fact one can compute exactly what elements make up the kernel:

**Lemma** (3.25). Each \( v \in K \) can be written as \( v(y, \theta) = q(y) + \sum_i y_i f_i(\theta) + c \), where \( q \) is a homogeneous quadratic polynomial on the \( \mathbb{R}^{n-k} \) factor, \( f_i \) is an eigenfunction on \( \mathbb{S}^k_\sqrt{2} \) with eigenvalue \( 1/2 \), and \( c \in \mathbb{R} \).

**Proof.** This essentially follows from the fact that square-integrable (with respect to the weighted measure) \( L \)-harmonic functions are constant. Indeed suppose \( f \in C^\infty(\Sigma) \cap L^2(\Sigma) \) be such that \( \nabla f = 0 \). If \( \zeta \) is an arbitrary test function then \( \nabla f = 0 \Rightarrow \langle \zeta^2 f, \nabla f \rangle = 0 \Rightarrow \langle \nabla (\zeta^2 f), \nabla f \rangle = 0 \Rightarrow \langle \zeta \nabla f \rangle_{L^2}^2 = -\langle f \nabla \zeta^2, \nabla f \rangle \). For appropriate cutoff functions \( \zeta \), \( \|\nabla \zeta^2\| \leq C \zeta \), so

\[
\|\zeta \nabla f\|_{L^2}^2 \leq C \|\zeta \|_{L^1} \leq \frac{1}{2} \|\zeta \nabla f\|_{L^2}^2 + C \|f\|_{L^2}^2,
\]

i.e., \( \|\zeta \nabla f\|_{L^2} \leq C \|f\|_{L^2} \). Since \( \zeta \) were arbitrary, \( f \in H^1 \). Repeating this for the \( L^2 \)-harmonic function \( \nabla f \) we deduce that \( f \in H^2 \) so \( \nabla f \in L^2 \). Therefore \( \|\nabla f\|_{L^2}^2 = -\langle f, \nabla f \rangle = 0 \), i.e. \( f \) is constant.

The rest of the argument is clearly described in the paper. \( \square \)

**Corollary** (3.36). There exists \( C \) depending on \( n \) such that for any \( v \in K \), then

\[
|v| \leq C (1 + |y|^2) |v|_{L^2}, \quad |\nabla v| \leq C (1 + |y|) |v|_{L^2}, \quad |\nabla^2 v| \leq C |v|_{L^2}, \quad |\nabla^2 v| \leq C (1 + |y|) |v|_{L^2}.
\]

**Remark.** This result is important because it says that the only Jacobi fields with respect to \( e^{-|x|^2/4}dv(x) \) (i.e. with exponential growth) actually never grow faster than quadratically. By treating the quadratic growth as the product of two linear growths, in combination with the previous Poincare-type lemma we get:

**Lemma** (3.22). There exists \( C \) depending on \( n \) such that for any \( u \in H^2 \),

\[
\{u_K\}_2^2 \leq C \|u_K\|_{L^2}^2 \quad \text{and} \quad \{u^+\}^2 \leq C \|u\|_2 \{u^+\}_{H^2}^2
\]

where the quantity \( \{u\}^2 \) is defined to be

\[
\{u\}^2 \triangleq \left[ |u|^2 + |\nabla u|^2 + |\nabla^2 u(\cdot, \mathbb{R}^{n-k})|^2 + (1 + |x|)^{-1} |\nabla^2 u|^2 \right]_{L^2(\Sigma)}
\]

**Remark.** Observe that \( \{uv\} = |u| \{v\} \). One should think of \( \{u\} \) as a weighted \( W^{2,4} \) norm. Therefore the lemma (using the corollary) controls the weighted \( W^{2,4} \) norm of elements of \( K \) by the \( L^2 \) norm.

### 4.2. Preliminary gradient estimate.

Observe that since \( L \) is the linearization of \( \nabla \Sigma F \), one can prove the following standard nonlinear estimate:

**Lemma** (4.3). There exist \( \tau > 0 \) (small) and \( C > 0 \) (large) depending on \( n \) such that if \( u : C \to \mathbb{R} \) has \( \|u\|_{2,\Sigma} \leq \tau \), then

\[
\|\nabla u - \nabla (C_k - \frac{1}{2} u)\|_{L^2(\Sigma)} \leq C \{u\}^2
\]

**Sketch of proof.** One can compute \( \nabla \) to be of the form

\[
\nabla (C_k - \frac{1}{2} u) = f(u, \nabla u) + (p, V(u, \nabla u)) + (\Phi(u, \nabla u), \nabla^2 u)
\]

where \( f, V, \Phi \) depend smoothly on \( (s, y) \) for \( |s| \) small. Since \( L \) is the linearization of \( \nabla \Sigma F \) and cylinders are a critical point for \( \nabla \), we have the standard nonlinear (pointwise) estimate

\[
\|\nabla u - \nabla \| \leq C (1 + |x|) (|u| + |\nabla u|)^2 + C (|u| + |\nabla u|) |\nabla^2 u|
\]

where \( C = C(n) \). The \((1 + |x|)\) term shows up because of the naked \( p \) in the expression of \( \nabla u \). Integrating in space and using Young’s inequality,

\[
\|\nabla u\|_{L^2} \leq C (|1 + |x||^2)_{L^2} + C (|1 + |x|| |\nabla u|)_{L^2} + C (|1 + |x||^{-1} |\nabla^2 u|)
\]
The first claim follows by Lemma 3.4. The second claim follows by integration. □

**Proposition (4.1).** There exists $C > 0$ (large) depending on $n$, $\lambda_0$, and $\varepsilon > 0$ (small) depending on $n$ such that if:

1. $\lambda(\Sigma) \leq \lambda_0$,
2. $\Sigma \cap B_R$ is the graph of some function $u$ over $C \cap B_R$, with $\|u\|_{L^2(\Sigma \cap B_R)} \leq \varepsilon$, and
3. $\beta \in [0, 1)$,

then

$$|F(\Sigma) - F(C)| \leq C [\phi \frac{3+\beta}{2} \frac{R^3}{L^2(\Sigma \cap B_R)}] + C R^{n-1} \exp \left(-\frac{3 + \beta}{4} \cdot \frac{R^2}{4}\right) + C [u]^{3+\beta}.$$  

**Remark.** The first two terms on the right are important when $u_K$ is tiny, and the last term is important when $u^+$ is somewhat small.

**Proof of 4.1.** First we may extend $u$ to a function defined on the entire cylinder by multiplying with an appropriate cutoff function. If $\Sigma_u$ denotes the graph of this entire function $u$ then we have

$$|F(\Sigma) - F(\Sigma_u)| \leq C(n, \gamma) \lambda_0 R^{n-1} \exp \left(-\gamma \cdot \frac{R^2}{4}\right)$$

for $\gamma \in (0, 1)$ which we can just take to be $\frac{3+\beta}{2} < \gamma < 1$. Therefore we might as focus on bounding $F_0(u) \equiv F(\Sigma_u) - F(C)$. The point is to use the nonlinear comparison inequalities we’ve come up with:

1. $|\langle \phi u \rangle_{L^2} - [Lu]_{L^2}| \leq C \{u\}^2$, and
2. $|F_0(u) - \frac{1}{2} [u^+, Lu^+]| \leq C [u]_{L^2} \{u\}^2$.

In the second item above we’ve implicitly used that $\langle u, Lu \rangle = \langle u^+, Lu^+ \rangle$. We also need to make the most that we can out of the crude estimate

$$[\mathcal{M} u] \leq C [\phi]_{L^2} + \exp \left(-\frac{R^2}{8}\right).$$

Indeed:

$$|F_0(u)| \leq \frac{1}{2} |\langle u^+, Lu^+ \rangle| + C [u]_{L^2} \{u\}^2$$

$$\leq \frac{1}{2} [u^+]_{L^2} [Lu^+]_{L^2} + C [u]_{L^2} \{u_K\}^2 + C [u]_{L^2} \{u^+\}^2$$

$$\leq C [u]_{L^2} [u^+]_{H^2} + C [u]_{L^2} \{u_K\}^2 + \varepsilon C [u]_{L^2} [u^+]_{H^2}$$

$$\leq C [u]_{L^2} [u^+]_{H^2} + C [u]_{L^2} \{u^+\}^2$$

where we’ve used $[Lu^+] \leq [u^+]_{H^2}$ and $\{u^+\}^2 \leq C \|u\|_{L^2} [u^+]_{H^2}$. Here $C = C(n)$.

Now there are two cases to consider:

Suppose that the kernel component of $u$ is tiny, i.e. $u$ is essentially normal to the kernel. Then we can essentially control everything with a power of $[u^+]_{H^2}$ that is sufficiently large to make up for our crude $[\mathcal{M} u]$ bound. In particular, suppose

$$\{u_K\}^2 \leq \eta [u^+]_{H^2}^{1+\beta}$$

Then

$$\{u\}^2 \leq 2 \{u_K\}^2 + 2 \{u^+\}^2 \leq 2 \eta [u^+]_{H^2} + 2 C [u]_{L^2} [u^+]_{H^2} \leq \frac{1}{2} \mu [u^+]_{H^2}$$

as long as we require $2 \eta + 2 C \varepsilon \leq \frac{1}{2} C^{-1} \mu$, a quantity depending on $n$ (and $C = C(n)$ being the constant below). Then

$$[\mathcal{M} u]_{L^2} \geq [Lu]_{L^2} - C \{u\}^2 \geq \frac{1}{2} \mu [u^+]_{H^2}$$

Since $[u]_{L^2} \leq \{u_K\} + [u^+]_{H^2} \leq C \{u_K\} + [u^+]_{H^2}$, we can continue our $F_0(u)$ estimate as follows:

$$|F_0(u)| \leq C [u]_{L^2} [u^+]_{H^2} + C [u]_{L^2} \{u_K\}^2$$

$$\leq C \{u_K\} + [u^+]_{H^2} [u^+]_{H^2} + C \{u_K\} + [u^+]_{H^2} \{u_K\}^2$$

$$\leq C \{u_K\}^2 + C \{u_K\} [u^+]_{H^2} + C \{u_K\}^2 [u^+]_{H^2} + C [u^+]_{H^2}$$

$$\leq C \eta [u^+]^{3+\beta} + C [u^+]_{H^2} + C [u^+]_{H^2} \leq C [u^+]^{3+\beta}$$
because the middle term dominates for small $u$. But now we’re done because

$$|F_0(u)| \leq C[u_1]^{3+\beta} \leq 2C \mu^{-1} |\mathcal{M} u|_{L^2}^{3+\beta} \leq C[\phi]_{L^2}^{3+\beta} + C \exp \left( -\frac{3 + \beta}{4} \cdot \frac{R^2}{4} \right).$$

The second case to consider is that of the component normal to the kernel being somewhat small. Then since we have excellent estimates near the kernel, we are in good shape. That is, if $R$ is large enough that $C$ for

$$4.3.$$  \text{the graphical cylinder into the quantitative cylinder.}$

particular cylinder that may or may not match that of the graphical radius. We’re going to have to transplant

associated with it, while quantitative results like Proposition 2.1 provide concrete graphical estimates on a

Remark. Why is this lemma important? Because the graphical radius does not have a fixed cylinder

throughout the paper.

This and LI1 on spheres are the two lemmas that determine the value of $\varepsilon_0$ that is to be used throughout the paper.

Remark. Why is this lemma important? Because the graphical radius does not have a fixed cylinder associated with it, while quantitative results like Proposition 2.1 provide concrete graphical estimates on a particular cylinder that may or may not match that of the graphical radius. We’re going to have to transplant the graphical cylinder into the quantitative cylinder.

4.3. Proof of Second Lojasiewicz inequality. Recall the statement of LI2:

Theorem (0.26 - LI2). There exists $\varepsilon_0 > 0$ (small) depending on $\varepsilon$, $n$, such that if a surface $\Sigma$ is

(1) an $R_1$-graph over a cylinder $C_1$ with $\| \cdot \|_{1,B_{R_1}} \leq \varepsilon_0$, and

(2) an $R_2$-graph over another cylinder $C_2$ with $\| \cdot \|_{2,\alpha,B_{R_2}}$

with $5\sqrt{2n} \leq R_1 \leq R_2$, then it is also an $R$-graph, $R = \min\{2R_1, R_2\}$, over $C_1$ with $\| \cdot \|_{2,B_{R}} \leq \varepsilon$.

Remark. This and LI1 on spheres are the two lemmas that determine the value of $\varepsilon_0$ that is to be used throughout the paper.

Theorem (0.26 - LI2). There exists $\rho > 0$ and $\ell_0 \geq 1$ (large) depending on $n$, and a sequence $c_{\ell,n} \in (0,1) \uparrow 1$ as $\ell \to \infty$ such that if:

(1) $\Sigma^n \subset \mathbb{R}^{n+1}$ has $\lambda(\Sigma) \leq \lambda_0$,

(2) $\ell \geq \ell_0$, $C_\ell < \infty$,

(3) $R < R_{\text{graph}}^{C_\ell}$, and

(4) $\beta \in [0,1),$

then

$$|F(\Sigma) - F(C_{\ell})| \leq C R^2 \left\{ [\phi]_{L^2(\Sigma \cap B_R)}^{3+\beta} \right\} + \exp \left( -c_{\ell,n} \frac{1}{2(1+\beta)} \cdot \frac{R^2}{4} \right) + \exp \left( -\frac{3 + \beta}{4} \cdot \frac{R^2}{4} \right)$$

for $C = C(n, \lambda_0, \varepsilon_0, \ell, C_\ell)$.

Proof. Arguing as in LI1, we may assume that

$$R_1 \triangleq \sup \left\{ r \leq R : R^{2n+5} \left[ [\phi]_{L^2(\Sigma \cap B_R)}^{3+\beta} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right\} \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) \leq \eta \right\}$$

is large enough that $R_1 > 5\sqrt{2n}$, for a choice of $\eta$ that allows us to employ Proposition 2.1. According to 2.1, there exists a cylinder $C \in C_k$ over which $\Sigma \cap B_{R_1}$ is a graph of a function $u$ with $\|u\|_{1,C \cap B_{R_1}} \leq C R^2 \left[ [\phi]_{L^2,B_{R_1}}^{3+\beta} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right\} \exp \left( \frac{1}{2} \cdot \frac{R_1^2}{4} \right) \leq \varepsilon_0$
Notice that the LI1 bound can be extended from the estimate for MCF. We will use an improvement and extension argument to get good curvature estimates in comparison to Theorem 5.3. The key tool in proving this is going to be Brian White’s local curvature comparison theorem. The latter by the very definition of $R_1$. That is, $\|u\|_{2, \Sigma \cap B_{R_2}} \leq \varepsilon$ and

$$\|u\|_{L^2(\Sigma \cap B_{R_2 \setminus B_{R_1}})} \leq C \varepsilon^2 R^\alpha \exp\left(\frac{-R_1^2}{4}\right)$$

where $C = C(n, \lambda_0, \varepsilon_0, \ell, C_t)$.

By 4.1 we can estimate

$$|F(\Sigma) - F(C)| \leq C |\phi|_{L^2(\Sigma \cap B_R)}^{\frac{4+\beta}{2}} + C R_2^{n-1} \exp\left(\frac{-3+\beta}{4} \cdot \frac{R_2^2}{4}\right) + C |u|_{L^2(\Sigma \cap B_{R_2})}^{\frac{3+\beta}{2}}. $$

The $|\phi|$ term is going to be dominated by a smaller power of $|\phi|$ soon, so we won’t worry about it. The last term is bounded by LI1 in a favorable way.

As for the middle term, if $R_2 = R$ then we are done because the bound is favorable. If $R_2 < R$ then $R_2 = 2R_1$, so

$$\exp\left(\frac{-3+\beta}{4} \cdot \frac{R_2^2}{4}\right) = \exp\left(\frac{-3+\beta}{4} \cdot \frac{R_1^2}{4}\right) = C R^\alpha \left[|\phi|_{L^2(\Sigma \cap B_R)}^{\frac{3+\beta}{2}} + \exp\left(-c_{\ell,n} \left(3 + \beta \cdot \frac{R^2}{4}\right)\right)\right]$$

all of which are favorable. Combining all the dominant terms in these estimates we get LI2.

5. Scale Comparison Theorem

Now we discuss what is essentially the final and most important building block of the proof: the scale comparison theorem (Theorem 5.3). The key tool in proving this is going to be Brian White’s local curvature estimate for MCF. We will use an improvement and extension argument to get good curvature estimates forward in time, and then drag those backwards.

The plan of action is an improvement/extension scheme:

1. As long as we are below the shrinker scale and we have crude bounds on some large ball $B_R$, we can find strong $C^{2,\alpha}$ estimates on a slightly smaller ball $B_{(1-\tau)R}$. This is called the improvement scheme.
2. As long as we have strong $C^{2,\alpha}$ bounds on a ball $B_R$, we can find crude estimates on a larger ball $B_{(1+\theta)R}$. This is called the extension scheme.

Remark. The statements in this section of the companion (particularly in the extension section) differ somewhat from those of the paper. The reason is that I decided to keep very explicit track of the constants, and that required some rephrasing here and there. The statements here make more sense to “me”. We will assume throughout that we are working on the RMCF time interval $[-1,1]$.

5.1. Improvement argument. The first step in the improvement argument is an extension of LI1 which improves the $C^1$ estimate obtained in Proposition 2.1 into a $C^{2,\alpha}$ estimate using interpolation.

**Theorem (2.54).** There exists $\ell_0 > 0$ (large) depending only on $n$, and there exist $R_0 > 0$ (large) and $\sigma_0 > 0$ (small) depending on $n$, $\lambda_0$, $\varepsilon_0$, $\mu$, $\ell \geq \ell_0$, $M$ such that if:

1. $\Sigma^n \subset \mathbb{R}^{n+1}$ has $\lambda(\Sigma) \leq \lambda_0$,
2. $\Sigma \cap B_R$ has $|\mathcal{A}| + |\nabla^2 \mathcal{A}| \leq M$ and $H \geq \mu > 0$ for some $R > R_0$, $\ell \geq \ell_0$, $M$.
3. $\Sigma \cap B_{R_0}$ is $C^2$ $\sigma_0$-close to a cylinder in $C_k$. 

17
then $\Sigma \cap B_{R_1}$ is $C^{2,\alpha}$ $\delta_0$-close to a possibly different cylinder in $C_k$, where

$$ R_1 \triangleq \sup \left\{ r \leq R : R^{2n+5} \left[ \left| \phi \right|_{L^\infty(\Sigma \cap B_R)} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) \leq \gamma \right\} $$

and where $\gamma = \gamma(n, \lambda_0, \mu, \ell, M)$.

**Proof.** We mimic the proof of Li1 for the most part. As before we have

$$ \| \phi \|_{1, B_{R_0}} \leq C(n, \ell) \left[ \| \phi \|_{L^1, B_{100 \sqrt{\phi}}} + \| \phi \|_{L^{1, \infty, B_{100 \sqrt{\phi}}} \| \nabla \phi \|_{L^{1, \infty, B_{100 \sqrt{\phi}}}}} \right] $$

and so we continue to choose $\ell_0 = \ell_0(n) \geq 1$ large enough so that $a_{\ell,n} \geq 3/4$ for all $\ell \geq \ell_0$, because then $\| \phi \|_{1, B_{\sqrt{\phi}}} \leq C(n, \ell, M) \| \phi \|_{L^{1, \infty, B_{100 \sqrt{\phi}}} \| \nabla \phi \|_{L^{1, \infty, B_{100 \sqrt{\phi}}}}}$. We may suppose that

$$ R_1 \triangleq \sup \left\{ r \leq R : C(n, \lambda_0, \mu, \ell, M) R^{2n+5} \left[ \left| \phi \right|_{L^\infty(\Sigma \cap B_R)} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) \leq \varepsilon_2 \right\} $$

is $\geq 5 \sqrt{2n}$. Here $\varepsilon_2$ is actually $\leq \varepsilon_2 = \varepsilon_2(n, \mu, M)$ as in 2.1, and also squeezed further depending on $n, \lambda_0, \mu, \ell, M, \delta_0$ in a way that will be determined at the end of the proof. The $C$ is the constant of 1.27 plus the constant $C(n, \ell, M)$ above.

This assumption is valid because we may squeeze the negative $R$ exponential to be small by making $R_0$ large and $\| \phi \|_{L^1}$ to be small by making $\sigma_0$ small depending on $n, \lambda_0, \mu, \ell, M$. Of course we also need to have $\sigma_0 \leq \varepsilon_1(n, \mu, M)$ of 2.1.

At this point we may apply 2.1 like we did before to get that $\Sigma \cap B_{R_1}$ is $C^1$ close to a cylinder, with

$$ \| \cdot \|_{1, B_{R_1}} \leq C(n, \lambda_0, \mu, \ell, M) R^{2n+5} \left[ \left| \phi \right|_{L^\infty(\Sigma \cap B_R)} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \exp \left( \frac{1}{2} \cdot \frac{r^2}{4} \right) $$

In other words we can control the $C^1$ norm of the graph function on $B_{R_1}$, but also the $(\ell + 2)$-nd norm in view of $|\nabla^{\ell} A| \leq M$ on $B_{R_2}, R \geq R_1$. By interpolation,

$$ \| \cdot \|_{2, a, B_{R_1}} \leq \eta \| \cdot \|_{\ell+2, B_{R_1}} + C(n, \ell) \| \cdot \|_{1, B_{R_1}} $$

$$ \leq \eta C(n, M, \ell) \varepsilon_2 $$

Pick $\eta > 0$ small enough depending on $n, M$ such that the first term is $\leq \delta_0/2$. Now the choice of $\varepsilon_2$ is clear: pick it depending on $\eta, \ell, \delta$ so that the second term is $\leq \delta_0/2$.

**Claim** (Improvement step). There exist $\ell_0 \geq 1$ (large) depending on $n, \tau$, and $R_0 > 0$ (large) depending on $n, \lambda_0, \varepsilon_0, \delta_0, K, \tau, \ell, C_k$ such that if:

1. $\Sigma_s$ is RMCF with $\lambda(\Sigma_s) \leq \lambda_0$ for $s \in [-1, 1]$,
2. $\ell \geq \ell_0, C_k < \infty$,
3. $s_0 \in [-1, 1], \tau \in (0, 1/2]$,
4. $R \geq R_0, R \leq R_{\text{entr}}(\Sigma_0), R \leq R_{\text{graph}}^0, C_k(\Sigma_0)$, and
5. $|\nabla_{\Sigma_0} F|_{B_R}^2 \leq K (F(\Sigma_{-1}) - F(\Sigma_1))$,

then $\Sigma_0 \cap B_{(1 - \tau) R}$ is $C^{2,\alpha}$ $\delta$-close to a cylinder.

**Proof.** Observe that

$$ \left| \phi \right|_{L^2(\Sigma_0 \cap B_R)} \leq \| \nabla \nabla_{\Sigma_0} F \|_{B_R} \leq K \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) $$

The key quantity in Theorem 2.54 at $r = (1 - \tau) R$ is (bounded from above by)

$$ R^p \left[ \left| \phi \right|_{L^2(\Sigma_0 \cap B_R)} \right] + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \exp \left( (1 - \tau)^2 \cdot \frac{R^2}{4} \right) $$

$$ \leq K R^p \left[ \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) + \exp \left( -\frac{1}{2} c_{\ell,n} \cdot \frac{R^2}{4} \right) \right] \exp \left( (1 - \tau)^2 \cdot \frac{R^2}{4} \right) $$

$$ \leq 2 K R^p \exp \left( ((1 - \tau)^2 - c_{\ell,n}) \cdot \frac{R^2}{4} \right) \leq 2 K \exp \left( -\tau \cdot \frac{R^2}{4} \right) $$
since \( R \leq R_{\text{entr}}(\Sigma_0) \), and since we may require \( \ell_0 \) to be larger yet depending on \( n, \tau \) so that \( c_{\ell,n} \geq 1 - \tau \) for all \( \ell \geq \ell_0 \). Additionally require \( \ell_0 \) to exceed the \( \ell_0 \) in Theorem 2.54.

Then, requiring in addition that \( R_0 \) be large depending on all the parameters listed in the statement and the fixed choice of \( \ell_0 \), Theorem 2.54 kicks in to give the required \( C^{2,\alpha} \) estimate.

### 5.2. Extension argument.

Key in the extension argument is the following result due to Brian White:

**Claim.** There exist \( \sigma = \sigma(n) > 0 \) and \( C = C(n) > 0 \) such that if:

\[
(4\pi \tau)^{-n/2} \int_{M_{-\tau}} \exp \left( -\frac{|x - x_0|^2}{4\tau} \right) dV(x) \leq 1 + \sigma
\]

then for all \( s \in [-\tau/4, 0] \) we can estimate:

\[
|A|^2 \leq C \tau^{-1} \text{ on } M_s \cap B_{\sqrt{\tau/2}}(x_0).
\]

Here \( s \mapsto M_s \) is a MCF, not RMCF.

**Remark.** This is not the way the theorem was stated in Brian White’s paper, but rather Toby Colding and Bill Minicozzi’s statement. It was not immediately clear to me that Brian White’s theorem implies this, but Brian thinks it’s fine. Note, by the way, that the symbol \( \tau \) plays an entirely different role here in the extension section than it did in the improvement section.

**Lemma (5.39).** There exists \( \delta_0 > 0 \) (small), \( \theta > 0 \) (small) depending on \( n, \varepsilon_{0}, M \), such that if:

1. \( R > \sqrt{2n} \),
2. \( M_{\tau} \) is MCF such that \( ||A||_{3, M_{\tau} \cap B_{R+2}} \leq M \) for \( \tau \in [-1 - 1/M, -1 + 1/M] \), and
3. \( M_{-\tau} \) is \( C^{2,\alpha} \) \( \delta_0 \)-close to a cylinder \( C \in \mathcal{C}_k \),

then \( M_{\tau} \cap B_R \) is \( C^{2,\alpha} \) \( \varepsilon_0 \)-close to \( \sqrt{-\tau} C \) for \( \tau \in [-1 - \theta, -1 + \theta] \).

**Remark.** The importance of this lemma is clearer when it is reformulated in terms of RMCF. If

1. \( \Sigma_t \) is RMCF with curvature bounds on \([t - 1/M, t + 1/M] \), within \( B_R \), that depend on \( M \), and
2. \( \Sigma_t \cap B_R \) is \( C^{2,\alpha} \) \( \delta_0 \)-close to \( C \in \mathcal{C}_k \),
3. \( R \geq R_0 \),

then \( \Sigma_{t+\eta} \cap B_{(1+\kappa)R} \) is \( C^{2,\alpha} \) \( \varepsilon_0 \)-close to the same cylinder, for \( \eta, \kappa, \delta_0, R_0 \) depending on \( \varepsilon, \varepsilon_0, M \). It is certainly not surprising that the radius gets larger as time moves forward, since our reference frame is a shrinking cylinder of the original flow.

Another important ingredient is the following mean-value type estimate:

**Lemma (5.32).** There exists a constant \( C > 0 \) (large) depending on \( M \) such that if:

1. \( \Sigma_t \) is RMCF on \([t_1, t_2] \),
2. \( \beta \in (0, t_2 - t_1) \), \( R > 0 \), and
3. \( |A| \leq M \) on \( \Sigma_s \cap B_{R+1} \), for all \( s \in [t_1, t_2] \),

then

\[
\max_{s \in [t_1, t_2]} [ |\phi|^2 ]_{L^2(\Sigma_s \cap B_R)} \leq (C + \beta^{-1}) (F(\Sigma_{t_1}) - F(\Sigma_{t_2})).
\]

**Proof.** One may compute that

\[
(\partial_t - \mathcal{L}) \phi^2 = \phi^2 (2 |A|^2 + 1) - 2 |\nabla \phi|^2.
\]

If \( \zeta \) is a smooth cutoff function between \( B_R \subset B_{R+1} \), with \( 0 \leq \zeta \leq 1 \), \( |\nabla \zeta| \leq 2 \), then by taking the manifold movement into consideration, we see:

\[
\partial_t (\phi^2 \zeta^2) = \phi^3 (\nabla \zeta^2, n) + \zeta^2 \left( \mathcal{L} \phi^2 + \phi^2 (2 |A|^2 + 1) - 2 |\nabla \phi|^2 \right).
\]
Then by direct computation we can estimate
\[
\partial_t \int_{\Sigma_t} \phi^2 \zeta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) = - \int_{\Sigma_t} \phi^4 \zeta \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
+ \int_{\Sigma_t} (2 |A|^2 + 1) \phi^2 \zeta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
- 2 \int_{\Sigma_t} |\nabla \phi|^2 \zeta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
+ \int_{\Sigma_t} \phi^3 \langle \nabla \zeta^2, n \rangle \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
- \int_{\Sigma_t} \langle \nabla \phi^2, \nabla \zeta^2 \rangle \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
\leq \int_{\Sigma_t} \left( (2 |A|^2 + 1) \zeta^2 + 5 |\nabla \zeta|^2 \right) \phi^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
- \int_{\Sigma_t} |\nabla \phi|^2 \zeta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x)
\]
The last term may be dropped (we need to keep it to get a spacetime $\nabla \phi L^2$-estimate). In any case, for any $s \in [t_1 + \beta, t_2]$:  
\[
\int_{\Sigma_s} \phi^2 \zeta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
\leq \min_{t \in [t_1, t_1 + \beta]} \int_{\Sigma_t} \phi^2 \eta^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) + C \int_{t_1}^{t_2} \int_{\Sigma_t} \phi^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x) \\
\leq (C + \beta^{-1}) \int_{t_1}^{t_2} \int_{\Sigma_t} \phi^2 \exp \left( - \frac{|x|^2}{4} \right) dV(x)
\]
as claimed. \[\square\]

**Remark.** This is not quite the way the lemma is stated in the paper. It was more convenient for me to state it this way. Also, I have not stated one of the estimates of the lemma that I’m not going to use anywhere.

Another important ingredient is the following result:

**Corollary** (5.15). There exists $R_0 = R_0(n, \lambda_0, \sigma, \tau) > 0$ such that if:

1. $\Sigma_s$ is RMCF with $\lambda(\Sigma_s) \leq \lambda_0$, $s \in [-1, 1]$,
2. $R \geq R_0 + 2$, $R \leq R_{\text{entr}}(\Sigma_0)$, $x_0 \in B_{R - R_0}$, $\tau \in (0, 1/2]$, and
3. we can control
\[
(4\pi\tau)^{-n/2} \int_{\Sigma_1} \exp \left( - \frac{|x - x_0|^2}{4\tau} \right) dV(x) \leq 1 + \frac{\sigma}{2}
\]
then we can also control
\[
(4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp \left( - \frac{|x - x_0|^2}{4\tau} \right) dV(x) \leq 1 + \sigma
\]
for all $s \in [-1, 1]$.

**Proof.** In view of the entropy bound there exists $r_0 = r_0(n, \lambda_0, \sigma) > 0$ uniform over $s \in [-1, 1]$, $y \in \mathbb{R}^{n+1}$ such that
\[
(4\pi\tau)^{-n/2} \int_{\Sigma_s \setminus B_{r_0}\tau(y)} \exp \left( - \frac{|x - y|^2}{4\tau} \right) dV(x) \leq \frac{\sigma}{4}
\]
By use of a cutoff function, it’s relatively simple to show that
\[
\int_{\Sigma_s \cap B_R} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) \\
\leq \int_{\Sigma_s \cap B_{R+2}} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) + \frac{R+2}{\tau} \int_{-1}^1 \int_{\Sigma_s \cap B_{R+2}} |\phi| \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) + |\phi|^2 dV(x) dt
\]
The second term on the last line can be estimated by Cauchy-Schwarz and our entropy bounds:
\[
\frac{R+2}{\tau} \int_{-1}^1 \int_{\Sigma_s \cap B_{R+2}} |\phi| \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) \\
\leq \frac{R+2}{\tau} \sqrt{(4\pi\tau)^{n/2} \lambda_0} \exp \left( \frac{(R+2)^2}{8} \right) \left[ \int_{-1}^1 |\phi|_{L^2(\Sigma_s)}^2 dt \right]^{1/2} \\
\leq \frac{R+2}{\tau} \sqrt{(4\pi\tau)^{n/2} \lambda_0} \exp \left( \frac{(R+2)^2}{8} \right) \exp \left( - \frac{R_{\text{entr}}(\Sigma_0)^2}{4} \right) \leq \mu(n, \lambda_0, R_0) \tau^{-1}
\]
where \( \mu(n, \lambda_0, R_0) \) can be made arbitrarily small by requiring that \( R_0 \) be large (depending on \( n, \lambda_0 \), and the smallness factor). Along the same lines,
\[
\int_{-1}^1 \int_{\Sigma_s \cap B_{R+2}} |\phi|^2 dV(x) \\
\leq \exp \left( \frac{(R+2)^2}{4} \right) \int_{-1}^1 |\phi|_{L^2(\Sigma_s)}^2 dt \leq \exp \left( \frac{(R+2)^2}{4} \right) \exp \left( - \frac{R_{\text{entr}}(\Sigma_0)^2}{2} \right) \leq \mu'(R_0)
\]
which can once again be made arbitrarily small by choosing \( R_0 \) large. Putting it altogether, for \( s \in [-1,1] \) and \( x_0 \in B_{R-R_0} \) arbitrary we can estimate:
\[
(4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) \\
= (4\pi\tau)^{-n/2} \int_{\Sigma_s \cap B_R} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) + (4\pi\tau)^{-n/2} \int_{\Sigma_s \setminus B_R} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) \\
\leq (4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp \left( - \frac{|x-x_0|^2}{4\tau} \right) dV(x) + (4\pi\tau)^{-n/2} \mu(n, \lambda_0, R_0) \tau^{-1} + (4\pi\tau)^{-n/2} \mu'(R_0) + \frac{\sigma}{4} \\
\leq 1 + \frac{3\sigma}{4} + (4\pi\tau)^{-n/2} \mu(n, \lambda_0, R_0) \tau^{-1} + (4\pi\tau)^{-n/2} \mu'(R_0)
\]
This can be made \( \leq 1 + \sigma \) by choosing \( R_0 = R_0(n, \lambda_0, \sigma, \tau) \) sufficiently large.

**Remark.** Once again, this is not the way the result is stated in the paper.

The second important lemma makes use of Brian White’s local regularity theorem.

**Proposition (5.6).** There exists \( \delta_0 > 0 \) (small) depending on \( n \), and there exists \( R_0 > 0 \) (large) depending on \( n, \lambda_0, \tau \), and also \( C_L > 0 \) (large) depending on \( n, \ell \geq 2 \) such that if:

1. \( \Sigma_\ell \) is RMCF with \( \lambda(\Sigma_s) \leq \lambda_0, \ s \in [-1,1], \)
2. \( R \geq R_0 + 2, \ R \leq R_{\text{entr}}(\Sigma_0), \ x_0 \in B_{R-R_0}, \ \tau \in (0,1/2], \)
3. \( \ell \) can estimate \( |A|^2 \leq \delta_0 \tau^{-1} \) on \( \Sigma_\ell \cap B_{R_0 + \sqrt{\tau}(x_0)} \)

then for all \( t \in [-1,1], \ s \in [t - \log(1 - 7\tau/8), t - \log(1 - \tau)] \cap [-1,1], \) we can further estimate:
\[
|A|^2 + \tau^4 |\nabla^\ell A|^2 \leq C_L \tau \text{ on } \Sigma_s \cap B_{\sqrt{\tau}}(e^{\frac{4}{3}(s-t)}x_0).
\]

**Sketch of proof.** In view of the entropy bound there exists \( R_0 > 0 \) uniform over \( s \in [-1,1], \ y \in \mathbb{R}^{n+1} \) such that
\[
(4\pi\tau)^{-n/2} \int_{\Sigma_s \setminus B_{R_0 + \sqrt{\tau}}(y)} \exp \left( - \frac{|x-y|^2}{4\tau} \right) dV(x) \leq \frac{\sigma}{4}
\]
where \( \sigma \) is as in Brian White’s local regularity theorem (referred to as \( \varepsilon \) there).

If one picks \( x_0 \in B_{R-R_0} \) and dilates by \( \tau^{-1/2} \) around that center point then the curvature bound becomes:

\[
|A|^2 \leq \delta_0 \text{ on } \tau^{-1/2}(\Sigma_1 - x_0) \cap B_{R_0}
\]

Provided \( \delta_0 \) is small enough depending on \( \sigma \), the region \( \tau^{-1/2}(\Sigma_1 - x_0) \cap B_{R_0} = \tau^{-1/2}(\Sigma_1 \cap B_{R_0 \sqrt{\tau}}(x_0) - x_0) \)

is going to look essentially planar centered at the origin and the non-compact contribution of the exponential integral is small; i.e.,

\[
(4\pi\tau)^{-n/2} \int_{\Sigma_1} \exp\left(-\frac{|x-x_0|^2}{4\tau}\right) dV(x) \leq 1 + \frac{\sigma}{2}
\]

Since we’re now in the setup of Corollary 5.15, it is also true that

\[
(4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp\left(-\frac{|x-x_0|^2}{4\tau}\right) dV(x) \leq 1 + \sigma
\]

holds true for \( x_0 \in B_{R-R_0}, s \in [-1,1] \), arbitrary since \( R \geq R_0 + 2 \).

Since the density bound is true at time \( s \in [-1,1] \), by looking at the corresponding MCF Brian White’s local regularity theorem gives a crude curvature bound forward in MCF-time on each ball \( B_{\sqrt{\tau}/2}(x_0) \), \( x_0 \) as above. Translating that back to RMCF, we need to keep track of the fact that an MCF-static \( x_0 \) moves under RMCF, so instead the bound we got was the one claimed. The \( \sqrt{\tau}/2 \) radius shrinks to a \( \sqrt{\tau}/3 \) radius once we use Shi-like estimates to bound \(|\nabla^\ell A|\) given the \(|A|\) bound.

**Remark.** Observe that in view of the latter estimate being over the moving point \( e^{\frac{1}{2}(s-t)} x_0 \) in \([t - \log(1 - 7\tau/8), t - \log(1 - \tau)] \cap [-1,1] \), Proposition 5.6 gives us curvature bounds on a multiplicatively larger ball.

**Claim (Extension step).** There exist \( R_0 > 0 \) (large), \( \delta_0 > 0 \) (small), \( \theta > 0 \) (small), \( K > 0 \) (large) depending on \( n, \lambda_0, \varepsilon_0 \), and there exists \( C_\ell > 0 \) (large) depending on \( n, \lambda_0, \varepsilon_0, \ell \) such that if:

1. \( R \geq R_0 + 2, R \leq R_{\text{entr}}(\Sigma_i) \)
2. \( \Sigma_s \cap B_R \) is \( C^{2,\alpha} \delta_0 \)-close to a cylinder (depending on \( s \)) for each \( s \in [-1/2,1] \),

then for every \( s \in [-1/2,1] \):

\[
R_{\text{graph}}^{C_\ell}(\Sigma_s) \geq (1 + \theta) R \quad \text{and} \quad |\nabla \Sigma_s F|^2_{L^2(B_{(1+\theta)R})} \leq K (F(\Sigma_{-1}) - F(\Sigma_1))
\]

**Proof.** Let \( \delta_0 = \delta_0(n) \) be as in Proposition 5.6. Let \( \tau = \tau(n, \delta_0) > 0 \) be such that \(|A|^2 \leq \delta_0/\tau \) for each point in \( \Sigma_s \cap B_{R}, s \in [-1/2,1] \). This can be done seeing as to how our surfaces are \( C^{2,\alpha} \delta_0 \)-close to a cylinder at each time. By 5.6 we may estimate

\[
|A|^2 + \tau^\ell |\nabla^\ell A|^2 \leq \frac{C_\ell}{\tau} \text{ on } \Sigma_s \cap B_{\sqrt{\tau}/3}(e^{\frac{1}{2}(s-t)} x_0)
\]

for all \( x_0 \in B_{R-R_0}, R_0 = R_0(n, \lambda_0, \tau) \), and all \( s \in [t - \log(1 - 7\tau/8), t - \log(1 - \tau)] \cap [-1,1] \), and all \( t \in [-1,1] \). Said otherwise, we have curvature (and curvature \( \ell \)-th order derivative) bounds on a ball \( B_{(1+\theta)R} \) on a time interval \( s \in [-1 - \log(1 - 7\tau/8), 1] \).

The crude curvature bounds immediately give the required estimate:

\[
\max_{s \in [-3/4,1]} \|\phi\|^2_{L^2(\Sigma_s \cap B_{(1+\theta)R})} \leq K \left( F(\Sigma_{-1}) - F(\Sigma_1) \right)
\]

which implies the second inequality claimed in the statement.[1] Also,

\[
\max_{s \in [-3/4,1]} \|\phi\|^2_{L^2(\Sigma_s \cap B_{(1+\theta)R})} \leq K \exp \left( -\frac{R_{\text{entr}}(\Sigma_0)^2}{2} \right) \leq K \exp \left( -\frac{R^2}{2} \right)
\]

By interpolation,

\[
\|\phi\|_{2,\alpha, \Sigma_s \cap B_{(1+\theta)R}} \leq C \|\phi\|_{L^2(\Sigma_s \cap B_{(1+\theta)R})} + C \|\phi\|_{L^2(\Sigma_s \cap B_{(1+\theta)R})}
\]

\[
\leq C \exp \left( \frac{R^2}{4} \right) \|\phi\|_{L^2(\Sigma_s \cap B_{(1+\theta)R})} + C(C_3, \tau)
\]

\[
\leq C(K, R_0, C_3, \tau) = C(n, \lambda_0, \varepsilon_0, \delta_0, C_3)
\]

---

[1] The remainder of the proof can be simplified after a clarification kindly offered by Bill Minicozzi and Toby Colding.
In view of these crude bounds on the speed of evolution in the RMCF there exists \( \xi = \xi(n, \lambda_0, \varepsilon_0, \delta_0, C_3) \) such that we can go back in time by from \(-1/2\) to \(-1/2 - \xi\) and have our \( C^{2,\alpha} \) \( \delta_0 \)-bounds on \( B_R \) become \( C^{2,\alpha} \) \( 2\delta_0 \)-bounds at worse, on \( B_R \) still except earlier in time. If \( \delta_0 = \delta_0(n, \varepsilon_0, C_3) > 0 \) is small enough, by the uniform stability of cylinder lemma, 5.39, these extend to \( C^{2,\alpha} \) \( \varepsilon_0 \)-bounds on a ball \( B_{(1+\theta)R} \) slightly forward in time from the instant \( t = -1/2 - \xi \), say starting at \(-1/2\), where \( \theta = \theta(n, \varepsilon_0, C_3, \xi) \) can be chosen small enough that \( \theta \leq \sigma \).

In summary, there exists \( \theta = \theta(n, \lambda_0, \varepsilon_0) > 0 \) such that we have \( C^{2,\alpha} \) \( \varepsilon_0 \)-tight bounds on \( B_{(1+\theta)R} \) on the original time interval \([-1/2, 1]\), and curvature bounds \( |A|^2 + \tau^t |\nabla \tau|^2 \leq \frac{C}{\tau} \). By replacing \( C_\ell \) by \((\tau^{t-1} C_\ell)^{1/2}\), this implies that \( R_{\text{entr}}^{0,\ell,C_\ell} \geq (1 + \theta) R \) as claimed. \( \square \)

5.3. **Proof of Scale comparison theorem.** We remind the reader that the theorem read:

**Theorem (5.3).** There exist \( \mu > 0 \) (small), \( \ell_0 \geq 1 \) (large), \( C > 0 \) (large) depending on \( n, \lambda_0, \varepsilon_0 \), such that:

1. \( \Sigma_t \) is a RMCF on \([-1, 1]\), and
2. \( \ell \geq \ell_0 \),

then there exists \( C_\ell > 0 \) (large) depending on \( n, \lambda_0, \varepsilon_0, \ell \), such that

\[
(1 + \mu) R_{\text{entr}}(\Sigma_0) \leq \min_{-1/2 \leq t \leq 1} R_{\text{graph}}(\Sigma_t) + C
\]

**Proof.** Let \( R_0, \delta, \theta, K \) be as in the extension step, depending on \( n, \lambda_0, \varepsilon_0 \). Let \( \tau = \theta/2 \) be the parameter that is to be used in the improvement step. Let \( \ell_0 \geq 1 \) be as in the improvement step, depending on \( n, \tau \). Let \( C_{\ell_0} > 0 \) be as in the extension step, depending on \( n, \lambda_0, \varepsilon_0, \lambda_0 \). Finally, let \( R_0 > 0 \) be as in the improvement step depending on \( n, \lambda_0, \varepsilon_0, \delta, K, \ell_0, C_{\ell_0} \), quantities that have already been determined. The dependence of everything so far traces back down to \( n, \lambda_0, \varepsilon_0 \).

Notice that if the minimum graphical radius

\[
\rho \triangleq \min_{s \in [t-1/2, t+1]} R_{\text{graph}}^{0,\ell_0,C_{\ell_0}}(\Sigma_s)
\]

is smaller than \((1 - \tau)^{-1} R_0\) there is nothing to show: we may take \( C = R_0 \). So let’s suppose that \( \rho \geq (1 - \tau)^{-1} R_0 \), and of course that it’s \( \leq R_{\text{entr}}(\Sigma_t) \).

Then by the improvement step, \( \Sigma_s \cap B_{(1-\tau)\rho} \) is \( C^{2,\alpha} \) close to a cylinder, for every \( s \in [t-1/2, t+1] \). By the extension step,

\[
\min_{s \in [t-1/2, t+1]} R_{\text{graph}}^{0,\ell_0,C_{\ell_0}}(\Sigma_s) \geq (1 + \theta)(1 - \tau)\rho
\]

We may iterate this until

\[
\min_{s \in [t-1/2, t+1]} R_{\text{graph}}^{0,\ell_0,C_{\ell_0}}(\Sigma_s) \geq (1 + \theta)(1 - \tau)\rho
\]

At which time we may apply the extension step once more to get a multiplicatively better factor. The result follows for \( \ell = \ell_0 \), and in particular we have curvature bounds all the way up to the entropy radius. This lets us apply the same argument for any \( \ell \geq \ell_0 \). \( \square \)