T. H. COLDING AND W. P. MINICOZZI'S "UNIQUENESS OF BLOWUPS AND ŁOJASIEWICZ 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 Łojasiewicz 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 monthts.

Contents

Date: June 14, 2014.

1. Introduction

1.1. Mean Curvature Flow. MCF is the evolution equation

$$\partial_\tau x = H$$

for closed manifolds $M_\tau^n \subset \mathbb{R}^{n+1}$. We assume that the first singular time is $\tau = 0$, and thus that $\tau < 0$ everywhere below. We use τ 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

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

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:

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

The quantity $\frac{1}{2}x \cdot n - H$ is denoted by ϕ. 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

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

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

1.4. Parabolic rescaling. Parabolic rescaling is best phrased 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.

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

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

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

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

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

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 Σ_t. RMCF is the negative L^2-gradient flow of F, i.e. along $\nabla_\Sigma^{d\mu} F = -\phi$. Furthermore $t \mapsto \lambda(\Sigma_t)$ decreases along a RMCF. The stationary points of RMCF are called shrinkers and are those for which $\phi = 0 \Leftrightarrow H = \frac{1}{2} x \cdot n$.

RMCF is obtained from MCF by setting

$$t = -\log(-\tau), \quad \Sigma_t = \frac{1}{\sqrt{-\tau}} M_\tau$$

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, λ_0 such that if

(1) Σ_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 $\mathcal{C} \in \mathcal{C}_k$ 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 $|F(\Sigma_t) - F(\mathcal{C}_k)|^{2-\tau} \leq K\big(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\big)$.

Remark. The finite dimensional Łojasiewicz 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 Σ_{t_1} is $\mathcal{C}^{2,\alpha}$-$(\varepsilon_0/2)$ 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 Σ_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 $\int |\nabla_\Sigma F| \, dt$, since the evolution is in terms of $\nabla_\Sigma F$. 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 $\big(\Sigma_s\big)_{s\in[t-1,t+1]}$ can all be written as a uniformly small cylindrical graph over a cylinder depending only on t. Then by 6.14, $|F(\Sigma_t) - F(\mathcal{C}_k)|^{2-\tau} \le K\big(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\big)$ 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, $|F(\Sigma_t) - F(\mathcal{C}_k)| \le C\,t^{-1/(1-\tau)}$, we get

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

At this point we can estimate

$$\int_1^{\infty}[\nabla_{\Sigma_t}F]_{L^1}\,dt = \int_1^{\infty}[\phi]_{L^1(\Sigma_t)}\,dt \le F(\Sigma_0)^{1/2}\sum_{j=1}^{\infty}\Big[F(\Sigma_j) - F(\Sigma_{j+1})\Big]^{1/2} < \infty$$

and therefore the total length of the flow is finite, so there is a unique limit point. □

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 $\mathcal{C}^{2,\alpha}$ close to a cylinder. For given $\varepsilon_0 > 0$ (small), $\ell \in \mathbb{N}$ (large), $C_\ell > 0$ (large), we define $R^{\varepsilon_0,\ell,C_\ell}_{\text{graph}}(\Sigma)$ to be the largest radius R for which:

(1) $|\nabla^\ell A| \le C_\ell$ on $\Sigma \cap B_R$, and
(2) $\Sigma \cap B_R$ is $\mathcal{C}^{2,\alpha}$ ε_0-close to a cylinder in $\mathcal{C}_k$.

Remark. The constant ε_0 will be chosen universally by two lemmas that come from pure functional analysis and will depend only on n. It is ℓ, C_ℓ that we have to be careful with.

Definition (Entropy radius)**.** Let Σ_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

$$\exp\Big(-\frac{R_{\text{entr}}(\Sigma_t)^2}{2}\Big) = F(\Sigma_{t-1}) - F(\Sigma_{t+1})$$

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:

$$\exp\Big(-\frac{R_{\text{entr}}(\Sigma_t)^2}{2}\Big) = F(\Sigma_{t-1}) - F(\Sigma_{t+1}) \approx |\delta F_t| \approx [\nabla_{\Sigma_t}F]^2_{L^2} \equiv [\phi]^2_{L^2(\Sigma_t)}$$
$$\Rightarrow \exp\Big(-\frac{R_{\text{entr}}(\Sigma_t)}{4}\Big) \approx [\phi]_{L^2}$$

We often run into expressions of the form:

$$[\phi]_{L^2(\Sigma_t\cap B_R)} + \exp\Big(-\frac{R^2}{4}\Big)$$

or powers thereof. Note the following points:

(1) The goal is to write everything in terms of ϕ. 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 $[\phi]^2_{L^2(\Sigma_t)} \approx \exp(-R_{\text{entr}}(\Sigma)^2_t/2)$, we have a problem when $R \ll R_{\text{entr}}(\Sigma_t)$: 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), $\ell_0 \ge 1$ (large), $C > 0$ (large) depending on n, λ_0, ε_0, such that if:

(1) Σ_t is a RMCF on $[-1, 1]$, and
(2) $\ell \ge \ell_0$,

then there exists $C_\ell > 0$ (large) depending on n, λ_0, ε_0, ℓ, such that

$$(1+\mu)R_{\mathrm{entr}}(\Sigma_0) \le \min_{-1/2\le t\le 1} R^{\varepsilon_0,\ell,C_\ell}_{\mathrm{graph}}(\Sigma_t) + C$$

2.3. **Łojasiewicz inequalities.** We will reach Theorem 6.14 through two Lojasiecz 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 $\ell_0 \ge 1$ (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 $\lambda(\Sigma) \le \lambda_0$,
(2) $\ell \ge \ell_0$, $C_\ell < \infty$, and
(3) $R < R^{\varepsilon_0,\ell,C_\ell}_{\mathrm{graph}}$,

then

$$\inf_{\mathcal{C}\in\mathcal{C}_k} [w_{\mathcal{C}} - \sqrt{2k}]^2_{L^2(\Sigma\cap B_R)} \le C\, R^\rho \Big\{ [\phi]^{c_{\ell,n}}_{L^1(\Sigma\cap B_R)} + \exp\Big(- c_{\ell,n}\cdot \frac{R^2}{4}\Big)\Big\}$$

for $C = C(n, \lambda_0, \varepsilon_0, \ell, C_\ell)$, where $w_{\mathcal{C}}$ denotes the distance function to the axis of the cylinder $\mathcal{C}$.

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 ϕ alongside an exponentially decaying noncompact error term.

The second Łojasiewicz inequality requires more effort, but it will follow from the first Łojasiewicz 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 $\ell_0 \ge 1$ (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 $\lambda(\Sigma) \le \lambda_0$,
(2) $\ell \ge \ell_0$, $C_\ell < \infty$,
(3) $R < R^{\varepsilon_0,\ell,C_\ell}_{\mathrm{graph}}$, and
(4) $\beta \in [0,1)$,

then

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

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

Remark. This corresponds to the "gradient Łojasiewicz" 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 $\ker L$ (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, λ_0 such that if

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

Then $|F(\Sigma_t) - F(\mathcal{C}_k)|^{2-\tau} \le K\big(F(\Sigma_{t-1}) - F(\Sigma_{t+1})\big)$.

Proof of: 0.26 + 5.3 ⇒ 6.14. By LI2 it is true that

$$|F(\Sigma_t) - F(\mathcal{C}_k)| \le C\, R^\rho \left\{ [\phi]^{c_{\ell,n}\cdot\frac{3+\beta}{2(1+\beta)}}_{L^2(\Sigma_s\cap B_R)} + \exp\Big(- c_{\ell,n}\frac{3+\beta}{2(1+\beta)}\cdot\frac{R^2}{4}\Big) + \exp\Big(- \frac{3+\beta}{4}\cdot\frac{R^2}{4}\Big)\right\}$$

for the maximal choice of radius, $R \triangleq R_{\text{graph}}^{\varepsilon_0,\ell,C_\ell}$. 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 $C_\ell > \infty$ is large enough depending on ℓ. The point is that μ, C are independent of ℓ, since we'll pick ℓ large in a little bit.

Then the last term can be bounded by

$$\exp\Big(-\frac{3+\beta}{4}\cdot\frac{R^2}{4}\Big) \leq C\,\exp\Big(-\frac{3+\beta}{4}\cdot(1+\mu)\cdot\frac{R_{\text{entr}}(\Sigma_t)^2}{4}\Big) \leq (\delta F_t)^{\frac{3+\beta}{4}\cdot\frac{1+\mu}{2}}$$

Picking $\beta \in (0,1)$ close enough to 1 makes the exponent larger than $1/2$. The first two terms on the RHS can be dealt with simultaneously, provided one observes a mean-value-type inequality saying that

$$[\phi]^2_{L^2(\Sigma_t\cap B_R)} \leq C'\,(F(\Sigma_{t-1}) - F(\Sigma_{t+1})) = C'\,\delta F_t$$

At this point we choose ℓ large enough such that

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

and the result follows. ☐

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

$$\nabla^{d\mu}_{\Sigma} F = -\phi = H - \frac{1}{2}x\cdot n.$$

If $\Sigma^n \subset \mathbb{R}^{n+1}$ is some hypersurface and $u : \Sigma^n \to \mathbb{R}$ is "small" then the graphical surface Σ_u has its own Euler-Lagrange functional of F with respect to Gaussian measure, ϕ_u. If we like, we can pull that back to Σ. 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 Σ is a cylinder,

$$\mathcal{M}u = \frac{\nu_u}{w_u}\Big(H_u - \frac{1}{2}\,\eta_u\Big)\,\exp\Big(-\frac{2\sqrt{2k}\,u + u^2}{4}\Big).$$

where

(1) ν_u is the relative volume element,
(2) H_u is the mean curvature scalar of the cylinder,
(3) η_u is the support function, $\langle \cdot + u(\cdot)n(\cdot), n_u(\cdot)\rangle$, 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,

$$L = \mathcal{L} + |A|^2 + \frac{1}{2} = \Delta - \frac{1}{2}\,\nabla_{x^t} + |A|^2 + \frac{1}{2}.$$

When Σ is a cylinder,

$$L = \mathcal{L} + 1 = \Delta - \frac{1}{2}\,\nabla_{x^t} + 1.$$

3. First Łojasiewicz Inequality

The first Łojasiewicz inequality states that ϕ correctly captures the L^2-closeness of Σ 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 $\mathcal{C}^{2,\alpha}$ 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 $\mathcal{C}^{2,\alpha}$ 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 $\mathcal{C}^1$ 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 $\mathbb{S}^k_{\sqrt{2k}}$ with $\|u\|_2 \le \varepsilon_0$,

then $\|u\|_{2,\alpha} \le C\ \|\phi\|_{0,\alpha}$.

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

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

On the other hand L is the linearization of $\nabla_\Sigma F = \phi$, so $\|\phi - Lu\|_{0,\alpha} \le C(k,\alpha)\,\|u\|_2\,\|u\|_{2,\alpha}$ provided we require $\|u\|_2 \le \varepsilon_0$ to be uniformly bounded (so the constant C is finite, and independent of ε_0). Therefore:

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

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

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 by a plane, then we get an almost-cylindrical almost-shrinker of one flat dimension lower.

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

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

then if ϕ_0 is the ϕ-function for Σ_0,

$$|\phi - \phi_0| + |\nabla^{\Sigma_0}(\phi - \phi_0)| \le C\,\varepsilon\,\big(1 + |\phi| + |\nabla\phi|\big)$$

for a universal constant $C > 0$.

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

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

where $v = \nabla^t x_{k+1}$ near x_0. Immediately we see $|\phi - \phi_0| \le C\,\varepsilon + C\,\varepsilon\,|\phi|$. By differentiating along Σ_0 and using Kato's inequality we get $|\nabla^{\Sigma_0}(\phi - \phi_0)| \le C\,\varepsilon + C\,\varepsilon\,|\phi| + C\,\varepsilon\,|\nabla\phi|$. □

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 ε_1, $\varepsilon_2 > 0$ (small) and $C > 0$ (large) depending on n, $\delta > 0$, $M > 0$, so that if

(1) $H \ge \delta > 0$, $|A| + |\nabla A| \le M$ on $\Sigma \cap B_R$,
(2) $\Sigma \cap B_{5\sqrt{2n}}$ is $\mathcal{C}^2$ ε_1-close to a cylinder in $\mathcal{C}_k$, $k \ge 1$, and
(3) $5\sqrt{2n} \le r < R$ is such that

$$r\,\|\phi\|_{1,B_{5\sqrt{2n}}} + r^5\,\|\nabla(A/H)\|_{1,B_r} \le \varepsilon_2$$

then $\Sigma \cap B_{\sqrt{r^2 - 3k}}$ is a a graph over a cylinder in $\mathcal{C}_k$ with

$$\|\cdot\|_1 \le C\Big\{r\,\|\phi\|_{1,B_{5\sqrt{2n}}} + r^5\,\|\nabla(A/H)\|_{1,B_r}\Big\}$$

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 $\mathcal{C}^{2,\alpha}$-close to $\mathbb{S}^k_{\sqrt{2k}}$ 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 $y \in \Sigma \in B_{\sqrt{r^2-3k}}$ and u the graph function over the cylinder,

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

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

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

Since $\Sigma \cap B_{5\sqrt{2n}}$ is $\mathcal{C}^2$ ε_1-close to a cylinder, by possibly shrinking ε_1 depending on n alone we can get an orthonormal basis of eigenvectors (flat:) $v_1(p)$, ..., $v_{n-k}(p)$, (spherical:) $z_1(p)$, ..., $z_k(p)$ with corresponding eigenvalues λ_1, ..., λ_{n-k}, σ_1, ..., σ_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 $n(p)$ to be the outer unit normal at p. By a general lemma of Colding and Minicozzi from a different paper, having λ_i and σ_j far apart forces a better yet bound on the flat eigenvalues λ_i in terms of $\|\nabla\tau\|_1$:

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

where the δ 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

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

and the tangent vectors

$$v_i(x) \triangleq \nabla^t f_i(x), \qquad i \leq n-k$$

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 $\Omega_r = \Sigma \cap B^\Sigma_{3r}(p)$, the latter ball being a geodesic ball. The v_i enjoy the following properties in Ω_r:

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

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

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

so

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

so

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

Since $\nabla^{\mathbb{R}^{n+1}}_{\dot\gamma} w = A(\dot\gamma, w)\, n$, 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 $|v_i - v_i(p)|$:

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

Furthermore by the bound on $|\nabla\tau|$,

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

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

Step 3. Dimension reduction. We seek to slice Σ 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 $\mathcal{C}^2$ close to $\mathbb{S}^k_{\sqrt{2k}}$, 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, ..., 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 $\Sigma_0 \approx \mathbb{S}^k_{\sqrt{2k}} \subset \Omega_{5\sqrt{2n}}$. Indeed, at any point x on the topological sphere:

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

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

$$\begin{aligned} \nabla_X \nabla^t f_i &= \nabla_X(\nabla f_i - (\nabla f_i \cdot n) n) \\ &= -X(\nabla f_i \cdot n)\, n - (\nabla f_i \cdot n) \nabla_X n \\ \Rightarrow \nabla^t_X \nabla^t f_i &= (\nabla f_i \cdot n)\, A(X) \\ &= \Big[(\nabla f_i - \nabla^t f_i) \cdot n\Big]\, A(X) \end{aligned}$$

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

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

We're writing $r_0 = 5\sqrt{2n}$ above. In view of the lower bound $r \ge 5\sqrt{2n}$, as long as we require ε_2 to be small enough depending on n, δ, M to make the constants above comparable to $1/2$, we can apply 2.11 and get that Σ_0 is $\mathcal{C}^{2,\alpha}$ $(\varepsilon_\phi + \varepsilon_\tau)$-close to $\mathbb{S}^k_{\sqrt{2k}}$. Therefore the radial axis distance function satisfies:

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

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 Σ_0 by v moves Σ_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 $|1 - |\nabla^t f|| \le \sum_i |v_i - v_i(p)| \le C\, r^2\, \varepsilon_\tau(r)$ on Ω_r. By possibly shrinking ε_2 depending on n, δ, M, we can control the right hand side by $1/2$ and thus get $|v| \le 2$ on Ω_r. Note that

$$\begin{aligned} \nabla_{v_i} \nabla g_j = 0 \Rightarrow \nabla_{v_i} \nabla^t g_j &= -\nabla_{v_i} \nabla^\perp g_j = \nabla_{v_i}\big((\nabla g_j \cdot n)\, n\big) \\ &= \nabla_{v_i}(\nabla g_j \cdot n)\, n + (\nabla g_j \cdot n)\, \nabla_{v_i} n \quad = (\nabla g_j \cdot \nabla_{v_i} n)\, n + (\nabla g_j \cdot n)\, \nabla_{v_i} n \end{aligned}$$

and therefore $|\nabla_{v_i} \nabla^t g_j| \le C(n)\, |A(v_j, \cdot)| \le C\, r^2\, \varepsilon_\tau(r)$. Since $|v| \le 2$, we conclude that

$$|\nabla_v \nabla^t g_j| \le C\, r^2\, \varepsilon_\tau(r)$$

in Ω_r and, consequently, that

$$|\nabla_v \nabla^t w^2| \le 2\,C(n) \sum_{j=1}^{k+1} \big|\nabla_v (g_j \nabla^t g_j)\big| \le 2\,C(n) \sum_{j=1}^{k+1} |\nabla_v g_j|\,|\nabla^t g_j| + |g_j|\,|\nabla_v \nabla^t g_j| \le C\,r^3\,\varepsilon_\tau(r)$$

in Ω_r, the extra r coming from the linear g_j on the right.

Let $\Omega_{r,f}$ be the region obtained by integrating Σ_0 along v for time $t^2 \le r^2 - 3k$, intersected with Ω_r. By the fundamental theorem of calculus,

$$|\nabla^t w^2| \le \sup_{\Sigma_0} |\nabla^t w^2| + C\,r \sup_{\Omega_{r,f}} |\nabla_v \nabla^t w^2| \le C\,\varepsilon_\phi(r_0) + C\,r^4\,\varepsilon_\tau(r)$$

as well as

$$|w^2 - 2k| \le C\,r\,\varepsilon_\phi(r_0) + C\,r^5\,\varepsilon_\tau(r)$$

in $\Omega_{r,f}$. Since w is bounded from below, these bounds can be turned into $\mathcal{C}^1$ bounds for $w - \sqrt{2k}$. These in turn extend to $\mathcal{C}^1$ bounds on the graph function u. □

3.4. **Cylindrical estimates.** In order to obtain the first Łojasiewicz inequality from Proposition 2.1 we need to absorb the $\|\nabla(A/H)\|_1$ term into a ϕ term. This is obtained by very generic estimates for surfaces that resemble cylinders in some sense.

The key operators are the drift Laplacian $\mathcal{L} = \Delta - \frac{1}{2}\nabla_{x^t}$ and $L = \mathcal{L} + |A|^2 + \frac{1}{2}$. Recall that L is the linearization of $\mathcal{M}$, the E-L functional of F. The second fundamental form and the function ϕ are related by a PDE:

$$LA = A + \nabla^2\phi + \phi A^2$$

from which we get as a corollary that:

$$\begin{aligned}
\mathcal{L}_{H^2}(A/H) &= \frac{\nabla^2\phi + \phi A^2}{H} + \frac{A\,(\Delta\phi + \phi\,|A|^2)}{H^2} \\
\mathcal{L}_{H^2}|A/H|^2 &= 2\,|\nabla(A/H)|^2 + 2\,\frac{\langle \nabla^2\phi + \phi\,A^2, A\rangle}{H^2} + 2\,\frac{|A|^2\,(\Delta\phi + \phi|A|^2)}{H^3}
\end{aligned}$$

where $\mathcal{L}_f\sigma = \mathcal{L}\sigma + \langle \nabla \log f, \nabla\sigma\rangle$ for any tensor (or function) σ.

Remark. One point is that every term involving A above is paired up with something linear in ϕ 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, $H \ge \delta > 0$ and $|A| \le M$, then for all $s \in (0, R)$ we can estimate

$$\big[\nabla(A/H)\big]^2_{L^2(\Sigma\cap B_{R-s})} \le C\Big\{\frac{1}{s^2}\,\mathrm{vol}(\Sigma \cap B_R)\,\exp\Big(-\frac{(R-s)^2}{4}\Big) + [\phi]_{L^1(\Sigma\cap B_R)} + [\nabla^2\phi]_{L^1(B_R)}\Big\}.$$

for $C = C(n, \delta, M)$.

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

Corollary (1.27)**.** If $\Sigma \cap B_R$ is smooth, $H \ge \delta > 0$, $|A| + |\nabla^\ell A| \le M$ for some $\ell \ge 2$, and $\lambda(\Sigma) \le \lambda_0$, then for all $|y| < R - 2$:

$$|\nabla(A/H)|(y) + |\nabla^2(A/H)|(y) \le C\,R^{2n}\Big\{[\phi]^{\frac{1}{2}c_{\ell,n}}_{L^1(B_R)} + \exp\Big(-\frac{1}{2}\,c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big\}\,\exp\Big(\frac{1}{2}\cdot\frac{|y|^2}{4}\Big)$$

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) $R < R_{\text{graph}}^{\varepsilon_0, \ell, C_\ell}$,

then

$$\inf_{\mathcal{C} \in \mathcal{C}_k} [w_{\mathcal{C}} - \sqrt{2k}]^2_{L^2(\Sigma \cap B_R)} \leq C\, R^\rho \Big\{ [\phi]^{c_{\ell,n}}_{L^1(\Sigma \cap B_R)} + \exp\Big(- c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big\}$$

for $C = C(n, \lambda_0, \varepsilon_0, \ell, C_\ell)$, where $w_{\mathcal{C}}$ denotes the distance function to the axis of the cylinder $\mathcal{C}$.

Proof. Suppose we are given n, λ_0, ε_0, ℓ, C_ℓ as above. Then by writing out the graphs and interpolation, $|A| + |\nabla A| \leq M(\varepsilon_0, \ell, C_\ell)$ and $H \geq \delta(\varepsilon_0)$ on $\Sigma \cap B_R$. In fact

$$\|\phi\|_{1, B_{5\sqrt{2n}}} \leq C(n, \ell) \Big[\|\phi\|_{L^1, B_{10\sqrt{2n}}} + \|\phi\|^{a_{\ell,n}}_{L^1} \|\nabla^\ell \phi\|^{1 - a_{\ell,n}}_{L^\infty, B_{10\sqrt{2n}}} \Big]$$

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

$$\|\phi\|_{1, B_{5\sqrt{2n}}} \leq C(n, \ell, C_\ell)\, \|\phi\|^{3/4}_{L^1, B_{10\sqrt{2n}}}$$

seeing as to how we may suppose $\|\phi\|_{L^1, B_{10\sqrt{2n}}} \leq 1$.

We may suppose that the quantity:

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

($\delta = \delta(\varepsilon_0)$, $M = M(\varepsilon_0, \ell, C_\ell)$ 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, λ_0, ε_0, ℓ, C_ℓ.

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

$$\begin{aligned}
&r^2\, \|\phi\|_{1, B_{5\sqrt{2n}}} + r^5\, \|\nabla(A/H)\|_{1, B_r} \\
&\leq C \Big\{ r^2\, \|\phi\|^{3/4}_{L^1, B_{10\sqrt{2n}}} + r^5 R^{2n} \Big[[\phi]^{\frac{1}{2} c_{\ell,n}}_{L^1, B_R} + \exp\Big(- \frac{1}{2}\, c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big] \exp\Big(\frac{1}{2} \cdot \frac{r^2}{4} \Big) \Big\} \\
&\leq C\, R^{2n+5} \Big[[\phi]^{\frac{1}{2} c_{\ell,n}}_{L^1, B_R} + \exp\Big(- \frac{1}{2}\, c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big] \exp\Big(\frac{1}{2} \cdot \frac{r^2}{4} \Big) \leq \varepsilon_1
\end{aligned}$$

and therefore by 2.1, $\Sigma \cap B_r$ is a graph over *some* cylinder $\mathcal{C}$, with $\|u\|_1 \leq$ the expression above prior to ε_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:

$$|u(y)| + |\nabla u(y)| \leq C\, R^{2n+5} \Big[[\phi]^{\frac{1}{2} c_{\ell,n}}_{L^1, B_R} + \exp\Big(- \frac{1}{2}\, c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big] \exp\Big(\frac{1}{2} \cdot \frac{|y|^2}{4} \Big)$$

for $y \in B_{R_1}$.

At this point we're done because by the bound on $|u(y)|$ in the remark above

$$\begin{aligned}
[u]^2_{L^2,B_{R_1}} &= \int_{B_{R_1}} |u(y)|^2 \exp\Big(-\frac{|y|^2}{4}\Big)\, dV(y) \\
&\le C \operatorname{vol}(B_{R_1})\, R^{4n+10} \Big[[\phi]^{c_{\ell,n}}_{L^1,B_R} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big] \\
&\le C\, R^{\rho(n)} \Big[[\phi]^{c_{\ell,n}}_{L^1,B_R} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big] \\
\Rightarrow [w-\sqrt{2k}]^2_{L^2,B_{R_1}} &\le C\, R^{\rho(n)} \Big[[\phi]^{c_{\ell,n}}_{L^1,B_R} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big]
\end{aligned}$$

and on $B_R \setminus B_{R_1}$ we have the crude estimate $|w-\sqrt{2k}| \le |x|$, so that

$$\begin{aligned}
[w-\sqrt{2k}]^2_{L^2,B_R\setminus B_{R_1}} &\le C\, R^{n+2} \exp\Big(-\frac{R_1^2}{4}\Big) \\
&\le C\, R^{5n+12} \Big[[\phi]^{c_{\ell,n}}_{L^1,B_R} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big]
\end{aligned}$$

the last inequality following from the definition of R_1. The result follows by adding these two inequalities. □

4. Second Łojasiewicz Inequality

The second Łojasiewicz 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 $\phi = \nabla_\Sigma F$, 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 $\ker L$ and $(\ker L)^\perp$.
(2) Using those estimates and by looking at what happens near $\ker L$ versus near $(\ker L)^\perp$ we will establish Proposition 4.1, which bounds the same quantity as LI2 ($F(\Sigma) - F(\mathcal{C})$) in terms of ϕ, 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 Σ refers to a round cylinder in $\mathcal{C}_k$, in which case L collapses to $\mathcal{L}+1$, seeing as to how $|A|^2 \equiv 1/2$. 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

$$\int_\Sigma u\,\mathcal{L}u \exp\Big(-\frac{|x|^2}{4}\Big) = -\int_\Sigma \langle \nabla u, \nabla v\rangle \exp\Big(-\frac{|x|^2}{4}\Big).$$

(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 $\mathcal{C}^\infty$ functions.

We can also show

Lemma (3.4)**.** There exists C depending on n such that

$$[|x|\, u]^2_{L^2} \le C\left([u]^2_{L^2} + [\nabla_{\mathbb{R}^{n-k}} u]^2_{L^2}\right) \le C\,[u]^2_{H^1}$$

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.

Sketch of proof of 3.4. 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 $y = x^t$ for the $\mathbb{R}^{n-k}$ factor coordinates, then $|x|^2 = y^2 + 2k$ on any cylinder in $\mathcal{C}_k$ it suffices to bound $[|y|\, u]_{L^2}^2$ instead of $[|x|\, u]_{L^2}^2$. We compute

$$\begin{aligned}\operatorname{div}_\Sigma \Big(u^2 y \exp\Big(- \frac{|x|^2}{4}\Big)\Big) \exp\Big(\frac{|x|^2}{4}\Big) &= 2u \,\langle \nabla u, y\rangle + (n-k)\, u^2 - u^2\, \frac{|y|^2}{4} \\ &\leq 4\, |\nabla_{\mathbb{R}^{n-k}} u|^2 + (n-k)\, u^2 - u^2\, \frac{|y|^2}{4}\end{aligned}$$

The result follows after integrating. $\square$

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).

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

$$\begin{gathered}\mu\, [u^\perp]_{H^2} \leq [Lu]_{L^2} \leq C\, [u]_{H^2} \\ [u_K]_{H^2} \leq C\, [u_K]_{L^2}\end{gathered}$$

Proof. The inequality $[Lu]_{L^2} \leq C\, [u]_{H^2}$ is a corollary of 3.4.

The inequality $\mu\, [u^\perp]_{H^2} \leq [Lu]_{L^2}$ 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

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

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

$$\begin{aligned}
&\operatorname{div}_\Sigma \Big\{ \Big[\nabla^2 u(\nabla u) - (\mathcal{L}u)\, \nabla u\Big] \exp\Big(- \frac{|x|^2}{4}\Big)\Big\} \exp\Big(\frac{|x|^2}{4}\Big) \\
&= \operatorname{div}_\Sigma \Big\{ \nabla^2 u(\nabla u)\Big\} - \operatorname{div}_\Sigma \Big\{ (\mathcal{L}u)\, \nabla u\Big\} - \langle \nabla^2 u(\nabla u) - (\mathcal{L}u)\, \nabla u, \frac{1}{2}\, x^t\rangle \\
&= \Big\{ \Delta\Big(\frac{1}{2}|\nabla u|^2\Big)\Big\} - \Big\{ \langle \nabla \mathcal{L}u, \nabla u\rangle + \mathcal{L}u\, \Delta u\Big\} - \Big\{ \langle \frac{1}{2} \nabla_{x^t} \nabla u, \nabla u\rangle - \mathcal{L}u\, \frac{1}{2} \nabla_{x^t} u\Big\} \\
&= \Delta\Big(\frac{1}{2}|\nabla u|^2\Big) - (\mathcal{L}u)^2 - \langle \frac{1}{2} \nabla_{x^t} \nabla u, \nabla u\rangle - \langle \nabla \mathcal{L}u, \nabla u\rangle \\
&= \Delta\Big(\frac{1}{2}|\nabla u|^2\Big) - (\mathcal{L}u)^2 - \frac{1}{2} \nabla_{x^t} \Big(\frac{1}{2}|\nabla u|^2\Big) - \langle \nabla \mathcal{L}u, \nabla u\rangle \\
&= \Delta\Big(\frac{1}{2}|\nabla u|^2\Big) - (\mathcal{L}u)^2 - \frac{1}{2} \langle \nabla_{x^t} \nabla u, \nabla u\rangle - \langle \nabla \mathcal{L}u, \nabla u\rangle \\
&= \Delta\Big(\frac{1}{2}|\nabla u|^2\Big) - (\mathcal{L}u)^2 - \frac{1}{2} \langle \nabla \nabla_{x^t} u, \nabla u\rangle - \langle \nabla \mathcal{L}u, \nabla u\rangle \\
&= |\nabla^2 u|^2 + \operatorname{Ric}(\nabla u, \nabla u) + \langle \nabla \Delta u, \nabla u\rangle - (\mathcal{L}u)^2 - \frac{1}{2} \langle \nabla_{x^t} \nabla u, \nabla u\rangle - \langle \nabla \mathcal{L}u, \nabla u\rangle \\
&= |\nabla^2 u|^2 + \operatorname{Ric}(\nabla u, \nabla u) - (\mathcal{L}u)^2 \\
&\geq |\nabla^2 u|^2 - (\mathcal{L}u)^2
\end{aligned}$$

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

(1) $\nabla_{x^t} \nabla u = \nabla \nabla_{x^t} u$ on the cylinder,
(2) the Bochner formula $\frac{1}{2}\Delta |\nabla u|^2 = |\nabla^2 u|^2 + \operatorname{Ric}(\nabla u, \nabla u) + \langle \nabla \Delta u, \nabla u\rangle$, and
(3) $\operatorname{Ric} \geq 0$ on the cylinder

in the last four steps. Integrating, we get $[\nabla^2 u]_{L^2}^2 \leq [\mathcal{L}u]_{L^2}^2$ as claimed.

By applying this elliptic estimate to $u^\perp$,

$$\begin{aligned}[u^\perp]_{H^2} &\leq C \Big[[u^\perp]_{L^2} + [\mathcal{L}u^\perp]_{L^2}\Big] = C \Big[[u^\perp]_{L^2} + [Lu - u^\perp]_{L^2}\Big] \\ &\leq C \Big[[u^\perp]_{L^2} + [Lu]_{L^2}\Big] \leq C\, [Lu]_{L^2},\end{aligned}$$

the last step holding true by ellipticity.

For the last claim observe that $Lu_K = 0 \Rightarrow \mathcal{L}u_K = -u_K$, and that by the previous application of Bochner's theorem this forces $[\nabla^2 u]_{L^2} \le [\mathcal{L}u]_{L^2} = [u_K]_{L^2}$. This bounds the second derivative. For the first derivative, integrating $\mathcal{L}u_K = -u_K$ by parts gives $[\nabla u_K]_{L^2} = [u_K]_{L^2}$. The result follows. □

Next we seek to understand $K = \ker L$. 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 $\mathbb{S}^k_{\sqrt{2k}}$ 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) $\mathcal{L}$-harmonic functions are constant. Indeed suppose $f \in \mathcal{C}^\infty(\Sigma) \cap L^2(\Sigma)$ be such that $\mathcal{L}f = 0$. If ζ is an arbitrary test function then $\mathcal{L}f = 0 \Rightarrow \langle \zeta^2 f, \mathcal{L}f\rangle = 0 \Rightarrow \langle \nabla(\zeta^2 f), \nabla f\rangle = 0 \Rightarrow [\zeta\, \nabla f]^2_{L^2} = -\langle f\, \nabla\zeta^2, \nabla f\rangle$. For appropriate cutoff functions ζ, $|\nabla \zeta|^2 \le C\,\zeta$, so

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

i.e., $[\zeta\, \nabla f]^2_{L^2} \le C\, [f]^2_{L^2}$. Since ζ were arbitrary, $f \in H^1$. Repeating this for the L^2 $\mathcal{L}$-harmonic function ∇f we deduce that $f \in H^2$ so $\mathcal{L}f \in L^2$. Therefore $[\nabla f]^2_{L^2} = -\langle f, \mathcal{L}f\rangle = 0$, i.e. f is constant.

The rest of the argument is clearly described in the paper. □

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

$$|v| \le C\,(1+|y|^2)\,[v]_{L^2}, \quad |\nabla v| \le C\,(1+|y|)\,[v]_{L^2}, \quad |\nabla^2 v(\cdot, \mathbb{R}^{n-k})| \le C\,[v]_{L^2}, \quad |\nabla^2 v| \le C\,(1+|y|)\,[v]_{L^2}.$$

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 $u \in H^2$,

$$\{u_K\}^2 \le C\,[u_K]^2_{L^2} \text{ and } \{u^\perp\}^2 \le C\,\|u\|_2\,[u^\perp]_{H^2}$$

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

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

Remark. Observe that $\{av\} = |a|\,\{v\}$. 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 $\bar{\varepsilon} > 0$ (small) and $C > 0$ (large) depending on n such that if $u : \mathcal{C} \to \mathbb{R}$ has $\|u\|_{2,\Sigma} \le \bar{\varepsilon}$, then

$$[\mathcal{M}u - Lu]_{L^2(\Sigma)} \le C\,\{u\}^2$$

and

$$\left|F(u) - F(\mathcal{C}_k) - \frac{1}{2}\,\langle u, Lu\rangle\right| \le C\,[u]_{L^2}\,\{u\}^2$$

Sketch of proof. One can compute $\mathcal{M}$ to be of the form

$$\mathcal{M}u(p) = f(u, \nabla u) + \langle p, V(u, \nabla u)\rangle + \langle \Phi(u, \nabla u), \nabla^2 u\rangle$$

where f, V, Φ depend smoothly on (s, y) for $|s|$ small. Since L is the linearization of $\mathcal{M}$ and cylinders are a critical point for $\mathcal{M}$, we have the standard nonlinear (pointwise) estimate

$$|\mathcal{M}u - Lu| \le C\,(1+|x|)\,(|u| + |\nabla u|)^2 + C\,(|u| + |\nabla u|)\,|\nabla^2 u|$$

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

$$[\mathcal{M}u - Lu]_{L^2} \le C\,[(1+|x|)\,u^2]_{L^2} + C\,[(1+|x|)\,|\nabla u|]_{L^2} + C\,[(1+|x|)^{-1}\,|\nabla^2 u|]_{L^2}$$

The first claim follows by Lemma 3.4. The second claim follows by integration. $\square$

Proposition (4.1)**.** There exists $C > 0$ (large) depending on n, λ_0, and $\overline{\varepsilon} > 0$ (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 $\mathcal{C} \cap B_R$, with $\|u\|_{2,\Sigma\cap B_R} \leq \overline{\varepsilon}$, and
(3) $\beta \in [0, 1)$,

then

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

Remark. The first two terms on the right are important when u_K is tiny, and the last term is important when $u^\perp$ 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 Σ_u denotes the graph of this entire function u then we have

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

for $\gamma \in (0,1)$ which we can just take to be $\frac{3+\beta}{4} < \gamma < 1$. Therefore we might as focus on bounding $F_0(u) \equiv F(\Sigma_u) - F(\mathcal{C})$. The point is to use the nonlinear comparison inequalities we've come up with:

(1) $|[\mathcal{M}u]_{L^2} - [Lu]_{L^2}| \leq C\,\{u\}^2$, and
(2) $|F_0(u) - \frac{1}{2}\langle u^\perp, Lu^\perp\rangle| \leq C\,[u]_{L^2}\,\{u\}^2$.

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

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

Indeed:

$$\begin{aligned}
|F_0(u)| &\leq \frac{1}{2}\,|\langle u^\perp, Lu^\perp\rangle| + C\,[u]_{L^2}\,\{u\}^2\\
&\leq \frac{1}{2}\,[u^\perp]_{L^2}\,[Lu^\perp]_{L^2} + C\,[u]_{L^2}\,\{u_K\}^2 + C\,[u]_{L^2}\,\{u^\perp\}^2\\
&\leq C\,[u]_{L^2}\,[u^\perp]_{H^2} + C\,[u]_{L^2}\,\{u_K\}^2 + \overline{\varepsilon}\,C\,[u]_{L^2}\,[u^\perp]_{H^2}\\
&\leq C\,[u]_{L^2}\,[u^\perp]_{H^2} + C\,[u]_{L^2}\,\{u_K\}^2
\end{aligned}$$

where we've used $[Lu^\perp] \leq [u^\perp]_{H^2}$ and $\{u^\perp\}^2 \leq C\,\|u\|_{2,\Sigma}\,[u^\perp]_{H^2}$. 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 $[u^\perp]_{H^2}$ that is sufficiently large to make up for our crude $[\mathcal{M}u]$ bound. In particular, suppose

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

Then

$$\{u\}^2 \leq 2\,\{u_K\}^2 + 2\,\{u^\perp\}^2 \leq 2\,\eta\,[u^\perp]_{H^2} + 2\,C\,\|u\|_{2,\Sigma}\,[u^\perp]_{H^2} \leq \frac{1}{2}\,\mu\,[u^\perp]_{H^2}$$

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

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

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

$$\begin{aligned}
|F_0(u)| &\leq C\,[u]_{L^2}\,[u^\perp]_{H^2} + C\,[u]_{L^2}\,\{u_K\}^2\\
&\leq C\,\big(\{u_K\} + [u^\perp]_{H^2}\big)\,[u^\perp]_{H^2} + C\,\big(\{u_K\} + [u^\perp]_{H^2}\big)\,\{u_K\}^2\\
&\leq C\,\{u_K\}^3 + C\,\{u_K\}\,[u^\perp]_{H^2} + C\,\{u_K\}^2\,[u^\perp]_{H^2} + C\,[u^\perp]_{H^2}^2\\
&\leq C\,\eta\,[u^\perp]^{\frac{3(1+\beta)}{2}} + C\,[u^\perp]^{\frac{3+\beta}{2}} + C\,[u^\perp]_{H^2}^2 \leq C\,[u^\perp]^{\frac{3+\beta}{2}}
\end{aligned}$$

because the middle term dominates for small u. But now we're done because

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

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

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

Then

$$|F_0(u)| \leq C\,[u]_{L^2}\,[u^\perp]_{H^2} + C\,[u]_{L^2}\,\{u_K\}^2$$
$$\leq C\,\eta^{-1}\,[u]_{L^2}\,\{u_K\}^{\frac{2}{1+\beta}}$$
$$\leq C\,[u]_{L^2}\,[u]_{L^2}^{\frac{2}{1+\beta}} = C\,[u]_{L^2}^{\frac{3+\beta}{1+\beta}}$$

and the proposition follows. □

Notice that we've been using $\bar\varepsilon$ instead of ε_0 in this subsection. This is because ε_0 is required to be much smaller to allow for some wiggle room to allow us to jump between cylindrical graphs:

Lemma (4.49)**.** There exists $\varepsilon_0 > 0$ (small) depending on $\bar\varepsilon$, n, such that if a surface Σ is

(1) an R_1-graph over a cylinder $\mathcal{C}_1$ with $\|\cdot\|_{1,B_{R_1}} \leq \varepsilon_0$, and
(2) an R_2-graph over another cylinder $\mathcal{C}_2$ with $\|\cdot\|_{2,\alpha,B_{R_2}}$

with $5\sqrt{2n} \leq R_1 < R_2$, then it is also an R-graph, $R = \min\{2R_1, R_2\}$, over $\mathcal{C}_1$ with $\|\cdot\|_{2,B_R} \leq \bar\varepsilon$.

Remark. This and LI1 on spheres are the two lemmas that determine the value of ε_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 Łojasiewicz inequality.** Recall the statement of LI2:

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) $R < R_{\text{graph}}^{\varepsilon_0,\ell,C_\ell}$, and
(4) $\beta \in [0,1)$,

then

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

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

Proof. Arguing as in LI1, we may assume that

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

is large enough that $R_1 > 5\sqrt{2n}$, for a choice of η that allows us to employ Proposition 2.1. Accordinng to 2.1, there exists a cylinder $\mathcal{C} \in \mathcal{C}_k$ over which $\Sigma \cap B_{R_1}$ is a graph of a function u with

$$\|u\|_{1,\mathcal{C}\cap B_{R_1}} \leq C\,R^\rho\,\Big[[\phi]_{L^2,B_{R_1}}^{\frac{1}{2}c_{\ell,n}} + \exp\Big(-\frac{1}{2}\,c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big]\,\exp\Big(\frac{1}{2}\cdot\frac{R_1^2}{4}\Big) \leq \varepsilon_0$$

(by a smaller choice of $\eta > 0$) and by LI1,

$$[u]^2_{L^2(\Sigma\cap B_{R_1})} \le C\,R^{2\rho}\,\Big[[\phi]^{c_{\ell,n}}_{L^2(\Sigma\cap B_R)} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big]$$

By the rotation lemma, 4.49, Σ is also a graph over $\mathcal{C}\cap B_{R_2}$, $R_2 = \min\{2R_1, R\}$, with $\|u\|_{2,\mathcal{C}\cap B_{R_2}} \le \overline{\varepsilon}$. Notice that the LI1 bound can be extended from B_{R_1} to B_{R_2} since

$$\begin{aligned}[u]^2_{L^2(\Sigma\cap B_{R_2}\setminus B_{R_1})} &\le C\,\overline{\varepsilon}^2\,R^n\,\exp\Big(-\frac{R_1^2}{4}\Big)\\ &\le C\,\overline{\varepsilon}^2\,R^{2\rho}\,\Big[[\phi]^{c_{\ell,n}}_{L^2(\Sigma\cap B_R)} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big]\end{aligned}$$

the latter by the very definition of R_1. That is, $\|u\|_{2,\Sigma\cap B_{R_2}} \le \overline{\varepsilon}$ and

$$[u]^2_{L^2(\Sigma\cap B_{R_2})} \le C\,R^{\rho}\,\Big[[\phi]^{c_{\ell,n}}_{L^2(\Sigma\cap B_R)} + \exp\Big(-c_{\ell,n}\cdot\frac{R^2}{4}\Big)\Big],$$

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

By 4.1 we can estimate

$$|F(\Sigma) - F(\mathcal{C})| \le C\,[\phi]^{\frac{3+\beta}{2}}_{L^2(\Sigma\cap B_R)} + C\,R_2^{n-1}\,\exp\Big(-\frac{3+\beta}{4}\cdot\frac{R_2^2}{4}\Big) + C\,[u]^{\frac{3+\beta}{1+\beta}}_{L^2(\Sigma\cap B_{R_2})}.$$

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

$$\begin{aligned}\exp\Big(-\frac{3+\beta}{4}\cdot\frac{R_2^2}{4}\Big) &= \exp\Big(-\frac{3+\beta}{4}\cdot R_1^2\Big)\\ &= C\,R^{\rho}\Big[[\phi]^{c_{\ell,n}\,(3+\beta)}_{L^2(\Sigma\cap B_R)} + \exp\Big(-c_{\ell,n}\,(3+\beta)\cdot\frac{R^2}{4}\Big)\Big]\end{aligned}$$

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 $\mathcal{C}^{2,\alpha}$ estimates on a slightly smaller ball $B_{(1-\tau)R}$. This is called the improvement scheme.
(2) As long as we have strong $\mathcal{C}^{2,\alpha}$ 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 $\mathcal{C}^1$ estimate obtained in Proposition 2.1 into a $\mathcal{C}^{2,\alpha}$ estimate using interpolation.

Theorem (2.54)**.** There exists $\ell_0 \ge 1$ (large) depending only on n, and there exist $R_0 > 0$ (large) and $\sigma_0 > 0$ (small) depending on n, λ_0, δ_0, μ, $\ell \ge \ell_0$, M such that if:

(1) $\Sigma^n \subset \mathbb{R}^{n+1}$ has $\lambda(\Sigma) \le \lambda_0$,
(2) $\Sigma \cap B_R$ has $|A| + |\nabla^\ell A| \le M$ and $H \ge \mu > 0$ for some $R > R_0$, $\ell \ge \ell_0$, and
(3) $\Sigma \cap B_{R_0}$ is $\mathcal{C}^2$ σ_0-close to a cylinder in $\mathcal{C}_k$,

then $\Sigma \cap B_{R_1}$ is $\mathcal{C}^{2,\alpha}$ δ_0-close to a possibly different cylinder in $\mathcal{C}_k$, where

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

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

$$\|\phi\|_{1,B_{5\sqrt{2n}}} \le C(n,\ell) \Big[\|\phi\|_{L^1, B_{10\sqrt{2n}}} + \|\phi\|^{a_{\ell,n}}_{L^1, B_{10\sqrt{2n}}} \|\nabla^\ell \phi\|^{1-a_{\ell,n}}_{L^\infty, B_{10\sqrt{2n}}} \Big]$$

and so we continue to choose $\ell_0 = \ell_0(n) \ge 1$ large enough so that $a_{\ell,n} \ge 3/4$ for all $\ell \ge \ell_0$, because then $\|\phi\|_{1,B_{5\sqrt{2n}}} \le C(n,\ell,M)\, \|\phi\|^{3/4}_{L^1, B_{10\sqrt{2n}}}$. We may suppose that

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

is $\ge 5\sqrt{2n}$. Here ε_2 is actually $\le \varepsilon_2 = \varepsilon_2(n,\mu,M)$ as in 2.1, and also squeezed further depending on n, λ_0, μ, ℓ, M, δ_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 σ_0 small depeding on n, λ_0, μ, ℓ, M. Of course we also need to have $\sigma_0 \le$ the $\varepsilon_1(n,\mu,M)$ of 2.1.

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

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

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

$$\begin{aligned} \|\cdot\|_{2,\alpha,B_{R_1}} &\le \eta\, \|\cdot\|_{\ell+2,B_{R_1}} + C(\eta,\ell)\, \|\cdot\|_{1,B_{R_1}} \\ &\le \eta\, C(n,M) + C(\eta,\ell)\, \varepsilon_2 \end{aligned}$$

Pick $\eta > 0$ small enough depending on n, M such that the first term is $\le \delta_0/2$. Now the choice of ε_2 is clear: pick it depending on η, ℓ, δ so that the second term is $\le \delta_0/2$. $\square$

Claim (Improvement step)**.** There exist $\ell_0 \ge 1$ (large) depending on n, τ, and $R_0 > 0$ (large) depending on n, λ_0, ε_0, δ_0, K, τ, ℓ, C_ℓ such that if:

1. Σ_s is RMCF with $\lambda(\Sigma_s) \le \lambda_0$ for $s \in [-1,1]$,
2. $\ell \ge \ell_0$, $C_\ell < \infty$,
3. $s_0 \in [-1,1]$, $\tau \in (0,1/2]$,
4. $R \ge R_0$, $R \le R_{\text{entr}}(\Sigma_0)$, $R \le R^{\varepsilon_0,\ell,C_\ell}_{\text{graph}}(\Sigma_{s_0})$, and
5. $|\nabla_{\Sigma_{s_0}} F|^2_{B_R} \le K\,(F(\Sigma_{-1}) - F(\Sigma_1))$,

then $\Sigma_{s_0} \cap B_{(1-\tau)R}$ is $\mathcal{C}^{2,\alpha}$ δ-close to a cylinder.

Proof. Observe that

$$\begin{aligned} [\phi]^{c_{\ell,n}}_{L^2(\Sigma_{s_0} \cap B_R)} = |\nabla_{\Sigma_{s_0}} F|^{c_{\ell,n}}_{B_R} &\le K^{\frac{1}{2} c_{\ell,n}} \Big[F(\Sigma_{-1}) - F(\Sigma_1) \Big]^{\frac{1}{2} c_{\ell,n}} \\ &\le K \exp\Big(-\, c_{\ell,n} \cdot \frac{R_{\text{entr}}(\Sigma_0)^2}{4} \Big) \end{aligned}$$

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

$$\begin{aligned} &R^\rho \Big[[\phi]^{c_{\ell,n}}_{L^2}(\Sigma_{s_0} \cap B_R) + \exp\Big(-\, c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big] \exp\Big((1-\tau)^2 \cdot \frac{R^2}{4} \Big) \\ &\quad \le K\, R^\rho \Big[\exp\Big(-\, c_{\ell,n} \cdot \frac{R_{\text{entr}}(\Sigma_t)^2}{4} \Big) + \exp\Big(-\, c_{\ell,n} \cdot \frac{R^2}{4} \Big) \Big] \exp\Big((1-\tau)^2 \cdot \frac{R^2}{4} \Big) \\ &\quad \le 2\, K\, R^\rho \exp\Big(((1-\tau)^2 - c_{\ell,n}) \cdot \frac{R^2}{4} \Big) \le 2\, K \exp\Big(-\, \tau \cdot \frac{R_0^2}{4} \Big) \end{aligned}$$

since $R \leq R_{\mathrm{entr}}(\Sigma_0)$, and since we may require ℓ_0 to be larger yet depending on n, τ so that $c_{\ell,n} \geq 1-\tau$ for all $\ell \geq \ell_0$. Additionally require ℓ_0 to exceed the ℓ_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 ℓ_0, Theorem 2.54 kicks in to give the required $\mathcal{C}^{2,\alpha}$ 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:

$$(4\pi\tau)^{-n/2} \int_{M_{-\tau}} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \leq 1+\sigma$$

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 τ 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, ε_0, M, such that if:

1. $R > \sqrt{2n}$,
2. M_τ 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_{-1} is $\mathcal{C}^{2,\alpha}$ δ_0-close to a cylinder $\mathcal{C} \in \mathcal{C}_k$,

then $M_\tau \cap B_R$ is $\mathcal{C}^{2,\alpha}$ ε_0-close to $\sqrt{-\tau}\,\mathcal{C}$ 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. Σ_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 $\mathcal{C}^{2,\alpha}$ δ_0-close to $\mathcal{C} \in \mathcal{C}_k$,
3. $R \geq R_0$,

then $\Sigma_{t+\eta} \cap B_{(1+\kappa)R}$ is $\mathcal{C}^{2,\alpha}$ ε_0-close to the same cylinder, for η, κ, δ_0, R_0 depending on ε, ε_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. Σ_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

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

Proof. One may compute that

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

If ζ 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:

$$\partial_t(\phi^2\zeta^2) = \phi^3\,\langle \nabla\zeta^2, n\rangle + \zeta^2\Big(\mathcal{L}\phi^2 + \phi^2(2\,|A|^2+1) - 2\,|\nabla\phi|^2\Big).$$

Then by direct computation we can estimate

$$\begin{aligned}
\partial_t \int_{\Sigma_t} \phi^2 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) &= -\int_{\Sigma_t} \phi^4 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad + \int_{\Sigma_t} (2\,|A|^2+1)\, \phi^2 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad - 2\int_{\Sigma_t} |\nabla \phi|^2 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad + \int_{\Sigma_t} \phi^3 \, \langle \nabla \zeta^2, n\rangle \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad - \int_{\Sigma_t} \langle \nabla \phi^2, \nabla \zeta^2\rangle \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\leq \int_{\Sigma_t} \Big((2\,|A|^2+1)\,\zeta^2 + 5\,|\nabla\zeta|^2\Big)\, \phi^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad - \int_{\Sigma_t} |\nabla \phi|^2 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x)
\end{aligned}$$

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 $s \in [t_1+\beta, t_2]$:

$$\begin{aligned}
&\int_{\Sigma_s} \phi^2 \, \zeta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad \leq \min_{t\in[t_1,t_1+\beta]} \int_{\Sigma_t} \phi^2 \, \eta^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) + C \int_{t_1}^{t_2} \int_{\Sigma_t} \phi^2 \, \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x) \\
&\quad \leq (C+\beta^{-1}) \int_{t_1}^{t_2} \int_{\Sigma_t} \phi^2 \exp\Big(-\frac{|x|^2}{4}\Big) \, dV(x)
\end{aligned}$$

as claimed. □

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) Σ_s is RMCF with $\lambda(\Sigma_s) \leq \lambda_0$, $s \in [-1, 1]$,
(2) $R \geq R_0 + 2$, $R \leq R_{\mathrm{entr}}(\Sigma_0)$, $x_0 \in B_{R-R_0}$, $\tau \in (0, 1/2]$, and
(3) we can control

$$(4\pi\tau)^{-n/2} \int_{\Sigma_1} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big) \, dV(x) \leq 1 + \frac{\sigma}{2}$$

then we can also control

$$(4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big) \, dV(x) \leq 1 + \sigma$$

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

$$(4\pi\tau)^{-n/2} \int_{\Sigma_s \setminus B_{r_0\sqrt{\tau}}(y)} \exp\Big(-\frac{|x-y|^2}{4\tau}\Big) \, dV(x) \leq \frac{\sigma}{4}$$

By use of a cutoff function, it's relatively simple to show that

$$\begin{aligned}
&\int_{\Sigma_s\cap B_R} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \\
&\qquad \le \int_{\Sigma_1\cap B_{R+2}} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) + \frac{R+2}{\tau}\int_{-1}^{1}\int_{\Sigma_t\cap B_{R+2}} |\phi|\, \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big) + |\phi|^2\, dV(x)\, dt
\end{aligned}$$

The second term on the last line can be estimated by Cauchy-Schwarz and our entropy bounds:

$$\begin{aligned}
&\frac{R+2}{\tau}\int_{-1}^{1}\int_{\Sigma_t\cap B_{R+2}} |\phi|\, \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \\
&\qquad \le \frac{R+2}{\tau}\sqrt{(4\pi\tau)^{n/2}\, 2\,\lambda_0}\, \exp\Big(\frac{(R+2)^2}{8}\Big)\Big[\int_{-1}^{1} [\phi]^2_{L^2(\Sigma_t)}\, dt\Big]^{1/2} \\
&\qquad \le \frac{R+2}{\tau}\sqrt{(4\pi\tau)^{n/2}\, 2\,\lambda_0}\, \exp\Big(\frac{(R+2)^2}{8}\Big)\, \exp\Big(-\frac{R_{\mathrm{entr}}(\Sigma_0)^2}{4}\Big) \le \mu(n,\lambda_0,R_0)\,\tau^{-1}
\end{aligned}$$

where $\mu(n,\lambda_0,R_0)$ can be made arbitrarily small by requiring that R_0 be large (depending on n, λ_0, and the smallness factor). Along the same lines,

$$\begin{aligned}
&\int_{-1}^{1}\int_{\Sigma_t\cap B_{R+2}} |\phi|^2\, dV(x) \\
&\qquad \le \exp\Big(\frac{(R+2)^2}{4}\Big)\int_{-1}^{1} [\phi]_{L^2(\Sigma_t)^2}\, dt \le \exp\Big(\frac{(R+2)^2}{4}\Big)\, \exp\Big(-\frac{R_{\mathrm{entr}}(\Sigma_0)^2}{2}\Big) \le \mu'(R_0)
\end{aligned}$$

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:

$$\begin{aligned}
&(4\pi\tau)^{-n/2}\int_{\Sigma_s} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \\
&\quad = (4\pi\tau)^{-n/2}\int_{\Sigma_s\cap B_R} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) + (4\pi\tau)^{-n/2}\int_{\Sigma_s\setminus B_R} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \\
&\quad \le (4\pi\tau)^{-n/2}\int_{\Sigma_1} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) + (4\pi\tau)^{-n/2}\,\mu(n,\lambda_0,R_0)\,\tau^{-1} + (4\pi\tau)^{-n/2}\,\mu'(R_0) + \frac{\sigma}{4} \\
&\quad \le 1 + \frac{3\sigma}{4} + (4\pi\tau)^{-n/2}\,\mu(n,\lambda_0,R_0)\,\tau^{-1} + (4\pi\tau)^{-n/2}\,\mu'(R_0)
\end{aligned}$$

This can be made $\le 1+\sigma$ 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 recularity theorem.

Proposition (5.6)**.** There exists $\delta_0 > 0$ (small) depending on n, and there exists $R_0 > 0$ (large) depending on n, λ_0, τ, and also $C_\ell > 0$ (large) depending on n, $\ell \ge 2$ such that if:

1. Σ_t is RMCF with $\lambda(\Sigma_s) \le \lambda_0$, $s \in [-1,1]$,
2. $R \ge R_0 + 2$, $R \le R_{\mathrm{entr}}(\Sigma_0)$, $x_0 \in B_{R-R_0}$, $\tau \in (0, 1/2]$, and
3. we can estimate

$$|A|^2 \le \delta_0\,\tau^{-1} \text{ on } \Sigma_1 \cap B_{R_0\sqrt{\tau}}(x_0)$$

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:

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

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

$$(4\pi\tau)^{-n/2}\int_{\Sigma_s\setminus B_{R_0\sqrt{\tau}}(y)} \exp\Big(-\frac{|x-y|^2}{4\tau}\Big)\, dV(x) \le \frac{\sigma}{4}$$

where σ is as in Brian White's local regularity theorem (referred to as ε 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 δ_0 is small enough depending on σ, 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.,

$$(4\pi\tau)^{-n/2} \int_{\Sigma_1} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \leq 1 + \frac{\sigma}{2}.$$

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

$$(4\pi\tau)^{-n/2} \int_{\Sigma_s} \exp\Big(-\frac{|x-x_0|^2}{4\tau}\Big)\, dV(x) \leq 1 + \sigma$$

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, λ_0, ε_0, and there exists $C_\ell > 0$ (large) depending on n, λ_0, ε_0, ℓ such that if:

(1) $R \geq R_0 + 2$, $R \leq R_{\text{entr}}(\Sigma_t)$,

(2) $\Sigma_s \cap B_R$ is $\mathcal{C}^{2,\alpha}$ δ_0-close to a cylinder (depending on s) for each $s \in [-1/2, 1]$,

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

$$R^{\varepsilon_0,\ell,C_\ell}_{\text{graph}}(\Sigma_s) \geq (1+\theta)\, R \qquad \text{and} \qquad |\nabla_{\Sigma_s} F|^2_{L^2(B_{(1+\theta)R})} \leq K\,(F(\Sigma_{-1}) - F(\Sigma_1))$$

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 $\mathcal{C}^{2,\alpha}$ δ_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 ℓ-th order derivative) bounds on a ball $B_{(1+\kappa)R}$ on a time interval $s \in [-1 - \log(1 - 7\tau/8), 1]$.

The crude curvature bounds immediately give the required estimate:

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

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

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

By interpolation,

$$\begin{aligned}
\|\phi\|_{2,\alpha,\Sigma_s\cap B_{(1+\kappa)R}} &\leq C\, \|\phi\|_{L^2(\Sigma_s\cap B_{(1+\kappa)R})} + C\, \|\phi\|_{\ell,\Sigma_s\cap B_{(1+\kappa)R}} \\
&\leq C \exp\Big(\frac{R^2}{4}\Big)\, [\phi]_{L^2(\Sigma_s\cap B_{(1+\kappa)R})} + C(C_3, \tau) \\
&\leq C(K, R_0, C_3, \tau) = C(n, \lambda_0, \varepsilon_0, \delta_0, C_3)
\end{aligned}$$

[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 $\mathcal{C}^{2,\alpha}$ δ_0-bounds on B_R become $\mathcal{C}^{2,\alpha}$ $2\delta_0$-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 $\mathcal{C}^{2,\alpha}$ ε_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 $\mathcal{C}^{2,\alpha}$ ε_0-tight bounds on $B_{(1+\theta)R}$ on the original time interval $[-1/2, 1]$, and curvature bounds $|A|^2 + \tau^\ell\, |\nabla^\ell A|^2 \leq \frac{C_\ell}{\tau}$. By replacing C_ℓ by $(\tau^{-\ell-1}\, C_\ell)^{1/2}$, this implies that $R^{\varepsilon_0,\ell,C_\ell}_{\text{graph}} \geq (1+\theta)\, R$ as claimed. □

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, λ_0, ε_0, such that if:

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

then there exists $C_\ell > 0$ (large) depending on n, λ_0, ε_0, ℓ, such that

$$(1+\mu) R_{\text{entr}}(\Sigma_0) \leq \min_{-1/2 \leq t \leq 1} R^{\varepsilon_0,\ell,C_\ell}_{\text{graph}}(\Sigma_t) + C$$

Proof. Let R_0, δ, θ, K be as in the extension step, depending on n, λ_0, ε_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, τ. Let $C_{\ell_0} > 0$ be as in the extension step, depending on n, λ_0, ε_0, λ_0. Finally, let $R_0 > 0$ be as in the improvement step depending on n, λ_0, ε_0, δ, K, ℓ_0, C_{ℓ_0}, quantities that have already been determined. The dependence of everything so far traces back down to n, λ_0, ε_0.

Notice that if the minimum graphical radius

$$\rho \triangleq \min_{s \in [t-1/2, t+1]} R^{\varepsilon_0,\ell_0,C_{\ell_0}}_{\text{graph}}(\Sigma_s)$$

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 $\mathcal{C}^{2,\alpha}$ close to a cylinder, for every $s \in [t-1/2, t+1]$. By the extension step,

$$\min_{s \in [t-1/2, t+1]} R^{\varepsilon_0,\ell_0,C_{\ell_0}}_{\text{graph}}(\Sigma_s) \geq (1+\theta)(1-\tau)\rho$$

We may iterate this until

$$\min_{s \in [t-1/2, t+1]} R^{\varepsilon_0,\ell_0,C_{\ell_0}}_{\text{graph}}(\Sigma_s) \geq (1+\theta)(1-\tau)\rho$$

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 $\ell \geq \ell_0$. □